PALGRAVE STUDIES IN PRAGMATICS,
LANGUAGE AND COGNITION

Conditionals

Logic, Linguistics and Psychology

Edited by
Stefan Kaufmann
David E. Over
Ghanshyam Sharma

MOREMEDIA

palgrave
macmillan

Palgrave Studies in Pragmatics, Language and Cognition

Series Editors
Richard Breheny, Psychology and Language Sciences, University College London, London, UK
Nicole Gotzner, University of Potsdam, Potsdam, Germany

Palgrave Studies in Pragmatics, Language and Cognition is a series of high quality research monographs and edited collections of essays focusing on the human pragmatic capacity and its interaction with natural language semantics and other faculties of mind. A central interest is the interface of pragmatics with the linguistic system(s), with the 'theory of mind' capacity and with other mental reasoning and general problem-solving capacities. Work of a social or cultural anthropological kind is included if firmly embedded in a cognitive framework. Given the interdisciplinarity of the focal issues, relevant research comes from linguistics, philosophy of language, theoretical and experimental pragmatics, psychology and child development. The series aims to reflect all kinds of research in the relevant fields - conceptual, analytical and experimental.

Stefan Kaufmann · David E. Over ·
Ghanshyam Sharma
Editors

Conditionals

Logic, Linguistics and Psychology

palgrave
macmillan

Editors
Stefan Kaufmann
Department of Linguistics
University of Connecticut
Storrs Mansfield, CT, USA

David E. Over
Department of Psychology
Durham University
Durham, UK

Ghanshyam Sharma
Department of South Asian Studies
National Institute of Oriental Languages
and Civilisations (INALCO)
Paris, France

ISSN 2946-2576 ISSN 2946-2584 (electronic)
Palgrave Studies in Pragmatics, Language and Cognition
ISBN 978-3-031-05681-9 ISBN 978-3-031-05682-6 (eBook)
https://doi.org/10.1007/978-3-031-05682-6

This Palgrave Macmillan imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

List of Contributors

Cariani Fabrizio University of Maryland, College Park, MD, USA

Copley Bridget SFL (CNRS/Université Paris 8/UPL), Paris, France

Cruz Nicole Institute for Psychology, University of Innsbruck, Innsbruck, Austria

Csipak Eva Department of Linguistics, University of British Columbia, Vancouver, BC, Canada

Douven Igor IHPST / CNRS / Panthéon–Sorbonne University, Paris, France

Edgington Dorothy Birkbeck, University of London, London, UK

Égré Paul Institut Jean-Nicod (CNRS/ENS/EHESS), Département de philosophie & Département d'études cognitives, Ecole normale supérieure, PSL University, Paris, France

Elqayam Shira School of Applied Social Sciences, De Montfort University, Leicester, UK

Haegeman Liliane Research Groups Ugent, DiaLing, University of Geneva, Geneva, Switzerland

Kaufmann Stefan University of Connecticut, Storrs, CT, USA

Krzyżanowska Karolina ILLC, University of Amsterdam, Amsterdam, Netherlands;
Arché, University of St Andrews, Scotland, UK

Mackay John Department of Philosophy, University of Wisconsin-Madison, Madison, WI, USA

Over David E. Department of Psychology, Durham University, Durham, UK

Pfeifer Niki Department of Philosophy, University of Regensburg, Regensburg, Germany

Rips Lance J. Psychology Department, Northwestern University, Evanston, IL, USA

Romero Maribel Department of Linguistics, University of Konstanz, Konstanz, Germany

Rossi Lorenzo Center for Logic, Language and Cognition (LLC), Department of Philosophy and Educational Science, Università degli Studi di Torino, Torino, Italy

Schönenberger Manuela University of Geneva, Geneva, Switzerland

Sharma Ghanshyam Department of South Asian Studies, National Institute of Oriental Languages and Civilisations (INALCO), Paris, France

Sikorski Michał Warsaw University of Technology, Warsaw, Poland

Sprenger Jan Center for Logic, Language and Cognition (LLC), Department of Philosophy and Educational Science, Università degli Studi di Torino, Torino, Italy

List of Figures

Experimenting with (Conditional) Perfection: Tests of the Exhaustivity Theory

Toward a Unified Linguistic Approach to Conditionals—Some Empirical Evidence

How Fake Is Fake Past?

List of Tables

The Logic and Pragmatics of Conditionals Under Uncertainty: A Mental Probability Logic Perspective

Rethinking the Acceptability and Probability of Indicative Conditionals

The External Syntax of Conditional Clauses

Toward a Unified Linguistic Approach to Conditionals—Some Empirical Evidence

How Fake Is Fake Past?

Introduction

Stefan Kaufmann, David E. Over, and Ghanshyam Sharma

Humans think and talk a great deal about hypothetical states of affairs. What will be, is, or would have been the case under certain suppositions informs many of our basic mental and social activities, such as planning, explaining, and deciding ("nothing will happen if I just have one

S. Kaufmann (✉)
Department of Linguistics, University of Connecticut, Storrs Mansfield, CT, USA
e-mail: stefan.kaufmann@uconn.edu

D. E. Over
Department of Psychology, Durham University, Durham, UK
e-mail: david.over@durham.ac.uk

G. Sharma
Department of South Asian Studies, National Institute of Oriental Languages and Civilisations (INALCO), Paris, France
e-mail: ghanshyam.sharma@inalco.fr

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_1

drink"), attitudes like relief and regret ("if only I had not been drinking"), and social relations to do with trust and power ("you will go to jail if you drink and drive"). It is therefore not surprising that hypotheticals permeate the way we talk about the world, ourselves, and each other. Conditional sentences—exemplified by the English *if–then* family of constructions and their counterparts in other languages—are the kind of expression most typically associated with suppositional reasoning, and the most widely studied, both as the overt expression of such reasoning and as a window into its underlying structures and mental processes. Part of what makes conditionals a fascinating research topic is their interdisciplinary appeal. We use them in both talking and theorizing about deep and thorny problems—knowledge and ignorance, causality and chance, fiction and fact, inferences and truth. Viewed as interpreted linguistic forms, conditionals exhibit highly intricate structural subtleties and interactions between grammatical categories (such as tense, aspect, modality and mood) and contextual factors.

In short, conditionals are of great interest to logicians (both philosophical and mathematical), linguists, and psychologists, among others. They also hold tremendous promise as a potential area for joint interdisciplinary research. However, for reasons that are hard to pin down, the history of research on conditionals in these diverse disciplines has been marked by long periods of relative isolation and parochialism, punctuated by occasional bursts of mutual interest and intellectual cross-fertilization. The record shows, however, that such bouts of interdisciplinary activity can lead to significant progress within the individual disciplines themselves.

In linguistics, the formal semantic study of conditionals and modality received a transformational impulse from the ideas of Robert Stalnaker, David Lewis and others, which were adapted for linguistic research primarily through the work of Angelika Kratzer in the late Seventies and early Eighties. After this, though, the basic ideas and tools underwent only small incremental changes until about the turn of the millennium. Meanwhile, the fine syntactic and morphological details that drive and constrain the interpretation of conditionals went largely unaddressed even by linguists working in formal semantics, and were appreciated even less within philosophy.

Until about the turn of the millennium, most psychological studies of conditionals uncritically assumed the correctness of the truth-functional, material conditional interpretation of the natural language conditional (Pfeifer this volume), according to which *if A then C* is logically equivalent to the disjunction, *not-A or C*. The most influential psychological study of conditionals is found in the *mental model* theory of Johnson-Laird and Byrne (1991), who presupposed the validity of inferring *if A then C* from *not-A* and from *C*, two jarringly counterintuitive consequences of the material conditional interpretation. They went so far as to claim that a conditional with a false antecedent "… can hardly be false, and so, since the propositional calculus allows only truth and falsity, it must be true" (p. 7). The propositional calculus specifies the logic of the material conditional, but there have long existed other logical systems in which conditionals with false antecedents could be false, as well as logical systems in which truth and falsity are not the only options (see the papers by Cariani and Rips; Égré et al.; Over and Cruz; and Pfeifer in this volume). Johnson-Laird and his collaborators have recently revised their position to make both conditionals and disjunctions equivalent to modal conjunctions (see, for example, Khemlani et al. 2018). This position has also been heavily criticized (for example, in Bringsjord and Govindarajulu 2020; Oaksford 2022; Over 2022). Williamson (2020) is a recent sophisticated defense of the material conditional analysis of the natural language conditional.

So, what happened around the turn of the millennium? Or rather we should ask, what *began* then? For, with hindsight, it is clear that we are still very much in the middle of a process that was initiated by what is arguably the most consequential burst of cross-disciplinary interactions yet. Multiple new developments contributed to this event. One that affected all three disciplines was the rise—or more accurately, the return—of causality as a respectable notion in statistics and the social and cognitive sciences. Most prominently associated with Pearl (2000), this development, with its attendant toolbox for the formalization of notions such as intervention and independence (see in this volume, Copley; Kaufmann; and Over and Cruz), provided the study of counterfactual conditionals in particular with a much-needed new perspective.

In another stimulating confluence, psychologists finally awoke to the hypothesis that the probability of the conditional *if A then C* is the conditional probability of *C* given *A* (see in this volume, Edgington; Over and Cruz; and Pfeifer). This hypothesis, sometimes simply called *The Equation* or *The Thesis* by philosophers, originated as a philosophical position in the works of Ernest Adams, Robert Stalnaker, and others. Its deep roots lie in the even earlier mathematical studies of subjective probability by de Finetti and Ramsey (Pfeifer this volume), but for a long time it had failed to attract the attention of psychologists. When it eventually did (Evans et al. 2003), psychologists soon uncovered a burgeoning stream of evidence in support of it and of the probabilistic, or Bayesian, approach to the study of conditionals in general (Oaksford and Chater 2007, 2020; Over and Cruz this volume). These findings in turn motivated philosophers and even linguists to show renewed interest in *The Thesis*, and the unique challenges involved in its formal implementation.

In linguistics, since the turn of the millennium, there has been an increasing interest in the long-neglected syntactic and morphological fine-print of conditionals and their related semantic subtleties, including the marking of counterfactuality, the semantic ramifications of tense and temporal reference, the role of contextual factors, and so forth. These findings in turn have fed into a better understanding of the relevant logical properties, as well as potentially into a better design of experimental stimuli.

This is an exciting time to be working on conditionals, for all the reasons already mentioned, plus one more: while we have made much progress in the last two decades or so, we are not much closer to a single, generally accepted theoretical framework. Proposals abound, each with its staunch defenders. An optimistic take on this situation, however, is that it may well be for the best: deepening our understanding of a phenomenon is not the same as, and does not require, agreeing on a single approach.

This optimistic stance is the one we take in presenting this collection of papers. Our goal in assembling them is to showcase both the range of current research questions and the variation in theoretical and methodological perspectives. We can do justice to neither in a single volume, but we hope to have succeeded in giving an impression of the multifaceted nature and truly interdisciplinary appeal of the phenomenon. Despite the diversity of perspectives and approaches on display, the individual essays, like guests at a lively dinner party, end up engaging with each other in a web of overlapping conversations. We would like to point out some of the larger themes at the center of these conversations.

One recurring theme involves the relation between truth and probability, which is particularly fraught in the case of conditionals. Many find *The Thesis* (that the probability of a conditional *if A, C* is the conditional probability of *C*, given *A*) intuitively appealing. Yet it does not sit well with a semantic framework in which conditionals have objective truth conditions: a conditional probability cannot in general be identified with the probability that a proposition is true, as Lewis (1976) was the first to point out, and as a wealth of subsequent work has reaffirmed. For those who seek to reconcile *The Thesis* with truth conditions, all known ways out of this quandary offend against some beliefs that people hold dear, *e.g.*, that conditionals are always either true or false.

The main challenge for truth-conditional accounts of conditionals lies in the case where the antecedent is false. Attempts to reconcile truth conditions with probabilities are not exempt from the difficulties this raises. Dorothy Edgington, one of the foremost experts on the topic, addresses precisely this problem in her contribution, asking: what is the truth value of a conditional *if A, C* when *A* is false? She argues for the view that conditionals have truth *conditions*, which may, however, fail to yield a determinate truth *value* when the antecedent is false. This allows her to argue that conditionals denote propositions, albeit ones whose values are not always fully determined by the facts.

Paul Égré, Lorenzo Rossi, and Jan Sprenger also address this problem in their chapter, but their starting point is an argument by Allan Gibbard that, under three assumptions which Gibbard considered uncontroversial, any semantic treatment of the conditional as a propositional operator implies that it is the material conditional. Most subsequent work has seen this unwelcome result as evidence against one of the three assumptions, namely the so-called *Import–Export Principle*, which maintains that *if A, then if B, C* is equivalent to *if A and B, then C*. Égré *et al.* instead give up a further assumption which Gibbard made and which has not been questioned as widely in subsequent work: that the conditional is a *bivalent* propositional operator that is always true or false. They assume instead that the conditional is a trivalent truth-functional operator, which takes on a third value (which could be interpreted as either undefined or void) when the antecedent is false. The move to a trivalent logic opens up a host of related theoretical choices. Égré *et al.* carefully chart these options and argue that they hold much promise for a more satisfactory account of conditionals.

Niki Pfeifer's chapter explores yet another line on the problem of false antecedents. The logical foundation of his account is *coherence-based probability logic*, a family of approaches based on de Finetti's (and Ramsey's) subjectivist definition of probability and his account of *coherence* based on the concept of a "Dutch book" (a series of bets that the bettor must lose). The semantic treatment of a conditional *if A then C* is based on the notion of a *conditional event* C|A, a trivalent entity whose value is that of *C* when *A* is true and which is *void* when *A* is false. According to this approach, the assertion of *if A then C* by a speaker is comparable to a conditional bet, *if A then I bet C*, which is called off, becoming void, when *A* is false. Unlike the trivalent truth tables explored in Égré *et al.*'s contribution, these conditional events are associated with conditional probabilities, which are taken to be formally primitive within this framework, i.e., not defined as the ratio of the probability of *A & C* to the probability of *A*. Pfeifer deploys this logic in a model of probabilistic inference, focusing specifically on the treatment of special cases, such as zero-probability antecedents, where the predictions of his framework differ sharply from those suggested by the material conditional analysis.

Michał Sikorski's paper takes a step back and reviews the challenges facing attempts to reconcile *The Thesis* with a truth-conditional approach under which probability is understood as the probability of truth in an objective sense. He examines *The Thesis* and proposals about when the assertion of a conditional is acceptable. He considers the experiments that confirm *The Thesis* and points out that some recent studies appear to cast doubt on its full generality. The probability of *if A then C* may only be the probability of *C* given *A* when *A* raises the probability of *C*, and not when A and C are independent or when A lowers the probability of C.

The psychological evidence for and against *The Thesis* is taken up in the chapters by David E. Over and Nicole Cruz on the one hand, and Igor Douven, Shira Elqayam, and Karolina Krzyżanowska on the other. Over and Cruz explore the relationship between indicative and counterfactual conditionals from a Bayesian perspective, arguing that when beliefs are updated with new information, a belief in a counterfactual can sometimes "collapse" to a belief in an indicative, and a belief in an indicative can sometimes "expand" to a belief in a counterfactual. In their Bayesian approach, the value of a conditional with a false antecedent is the conditional probability itself. They cover possible limitations to the conditional probability hypothesis and criticize *truth condition inferentialism*, which is the view that a "standard" conditional *if A then C* can only be true when there is a deductive, inductive, or abductive (or some other epistemic) relation between A and C. Pfeifer, in his chapter, also finds fault with inferentialism.

In contrast, Douven, Elqayam, and Krzyżanowska, with the support of philosophical arguments and appeals to experimental evidence, put forward their own version of truth condition inferentialism. In their account, a "standard" conditional *if A then C* can only be true when there is a compelling argument from A to C. They closely compare *if A then C* to an inference that leads from A as a premise to C as a conclusion, and hold that it must be a *belief bias* effect which causes people to endorse if *A then C* as "true"—when C is believed to be true—in the absence of a compelling argument from A to C. They also argue in support of the position that *modus ponens*, inferring C from *if A then C* and A, does not always preserve truth.

Cruz and Over reply to the points made by Douven *et al*. They define an "independence conditional" as a conditional *if A then C* that is used, in a particular context, as part of an argument that A and C are independent of each other. They argue that independence conditionals are of importance in themselves and should not be labeled "non-standard" conditionals, or "unconditionals". They also question the positions taken by Douven *et al*. on "belief bias" and *modus ponens*.

Fabrizio Cariani and Lance Rips study *conditional perfection*, the inference of the biconditional, *A if and only if C*, from an assertion of the conditional, *if A then C*. They interpret an assertion of a conditional as a statement that *C* is true when *A* is true across a range of contextually given possible worlds. This is a "strict" interpretation, and it allows *if A then C* to be false in some contexts when *A* is false. According to the *exhaustivity hypothesis*, there will be evidence for conditional perfection when there is an increased endorsement in experiments of the fallacies of affirming the consequent, inferring *A* from *if A then C* and *C*, and denying the antecedent, inferring *not-C* from *if A then C* and *not-A*. In seven experiments, Cariani and Rips find that only a speaker's explicit statement that *A* is the sole way to bring *C* about tends to induce perfection.

All the chapters mentioned so far are concerned in one way or another with issues relating to the truth conditions of conditionals and the reasoning underlying their interpretation and use. A host of additional questions open up, however, when we turn to conditionals *qua* linguistic expressions. As we mentioned above, linguists, especially those working in the formal semantic tradition that is heavily indebted to philosophical logic, were slow to start paying close attention to the finer details of the syntactic and morphological makeup of conditionals. One major exception to this is Angelika Kratzer's argument that there is no binary sentential operator in natural-language conditionals; rather, the consequent (the *matrix clause*, in linguistic terms) is just a standalone sentence headed by a (covert or overt) modal operator, while the *if*-clause is an optional modifier. This view, though articulated early on and now standard in linguistic theory, is not, however, a syntactic analysis, and does not do justice to the structural intricacies of the clauses involved. As in philosophy, the constituents were treated as the unanalyzed wholes *A* and *C*.

The linguistic chapters included in this collection show how far the field has moved on from that early view. From a syntactic perspective, the claim that conditional *if*-clauses are modifiers is not saying very much, as Liliane Haegeman and Manuela Schönenberger show in their contribution. They identify three classes of conditional adverbial clauses, distinguished by their syntactic properties, and with corresponding semantic differences. It is fair to say that the implications of this classification for semantic theory have yet to be fully explored (but see, for instance, Csipak and Romero this volume, on "biscuit conditionals").

With the aim of bridging the gap between logical and linguistic traditions, Ghanshyam Sharma's contribution reexamines Greenberg's Universal 14, and looks into the parameters laid down by Comrie (1986) for cross-linguistic typological research. In doing so, Sharma reassesses some much debated topics and terms, such as clause inversion, the consequent marker *then,* and the role of grammatical categories such as tense and aspect in producing a counterfactual construal. He argues that the role of aspect in counterfactuality still remains underexplored.

Another long-standing open issue in the linguistic expression of conditionals concerns the morphological marking of the class of "counterfactual" or "subjunctive" conditionals, as opposed to their "indicative" counterparts. Many languages use some form of Past or Perfect marking on counterfactuals, even when they refer to future states of affairs; English in fact has not one but two such forms. Just how exactly these "fake Pasts" contribute to the interpretation of conditionals is an unresolved question in linguistics and philosophy, with authors disagreeing on whether or to what degree fake Past retains at least some of the temporal meaning of the ordinary Past, or whether it is used as a purely modal marker indicating a kind of "remoteness" from reality. The chapters by John MacKay and Stefan Kaufmann discuss different facets of this open issue. Mackay argues that both camps are partly right and proposes a hybrid account: Past marking in conditionals can be temporal or modal, and which reading is available depends in part on certain properties of the context of the utterance. Kaufmann argues for a different kind of hybrid account: fake Past is always "modal" in that it signals a (hypothetical) *intervention* upon the actual course of events; in addition, though, which of the English forms is used depends on whether that intervention lies in the past of the speech time.

Eva Csipak and Maribel Romero are also concerned with the temporal morphology in counterfactual conditionals, but they focus on the special sub-class of so-called *biscuit conditionals*, so named after John Austin's (1956) example "There are biscuits on the sideboard, if you want some". Biscuit conditionals, unlike the more familiar *hypothetical* conditionals, are concerned with a connection between the truth of the antecedent and the relevance, rather than the truth, of the consequent. Csipak and Romero take a unified semantic approach to biscuit conditionals and hypothetical conditionals, but observe that unlike hypothetical counterfactuals, English biscuit counterfactuals do not allow fake Past in the consequent (similarly, Spanish biscuit counterfactuals have indicative mood in the consequent, unlike their hypothetical counterparts). Hypothetical counterfactuals exhibit the requisite marking (fake Past in English, subjunctive mood in Spanish) in both constituents, but there is a "mismatch" of sorts between the antecedent and the consequent in biscuit conditionals. Csipak and Romero view this difference as a modal analog of similar mismatches in the domain of tense and temporal reference, known as violations of *Sequence of Tense* (Sequence of Tense is observed in "Mary said that she *was* coming tomorrow"; it is broken in "Mary said that she *is* coming tomorrow").

Bridget Copley's chapter uses causal models to deal with the role of temporal semantics in the interpretation of conditionals. For some conditionals *If A, C* a reading under which *C* is a *causal* consequence of *A* is very salient; others lack this particular flavor. Copley argues that the aspectual properties of the constituents—in particular, the distinction between stative and eventive clauses—play a central role in determining this behavior.

The idea of this collection was hatched in the aftermath of the summer school *Conditionals in Paris—Logic, Linguistics, and Psychology*, at which many of the contributors taught for a highly stimulating week in June, 2019. The summer school was organized by Ghanshyam Sharma and Lucia Tovena and held under the auspices of the French National Research Center (CNRS) and the Laboratory of Excellence: Empirical Foundations of Linguistics (LABEX-EFL) in Paris. We want to use this opportunity to express our deepest gratitude to the organizers and sponsors of this event, which was conceived in the same spirit of interdisciplinarity that we aim to capture in this book. We are also most grateful to all contributors to the book and to the publishers for their generosity, patience, and endurance during the challenging times of a global pandemic. Finally, Stefan Kaufmann gratefully acknowledges support from the National Science Foundation (#2116972, "Research on conditional and modal language") and the University of Connecticut Humanities Institute.

References

Austin, J. L. (1956). Ifs and cans. In P*roceedings of the British Academy*, *42*, 109–132. Reprinted in Urmson, J. O. and G. J. Warnock, editors (1961). *J. L. Austin: Philosophical Papers*. Oxford: Clarendon Press.

Bringsjord, S., & Govindarajulu, N. S. (2020). Rectifying the mischaracterization of logic by mental model theorists. *Cognitive Science, 44,* e12898.

Comrie, B. (1986). Conditionals: A typology. *On Conditionals*, 77–99. Traugott, E. C., ter Meulen, A., Reilly, J. S. and C. A. Ferguson, editors. Cambridge: Cambridge University Press.

Evans, J. St. B. T., Handley, S. J., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*, 321–335.

Johnson-Laird, P. N., & Byrne, R. M. J. (1991). *Deduction*. Hove & Hammersmith: Erlbaum.

Khemlani, S. S., Byrne, R. M. J., & Johnson-Laird, P. N. (2018). Facts and possibilities: A model-based theory of sentential reasoning. *Cognitive Science*, *42(6)*, 1887–1924.

Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. *The Philosophical Review, 85*, 297–315.

Oaksford, M. (2022). Mental models, computational explanation and Bayesian cognitive science: Commentary on Knauff and Gazzo Castañeda (2022). *Thinking & Reasoning,* 1–12.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality*: *The probabilistic approach to human reasoning*. Oxford: Oxford University Press.

Oaksford, M., & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology*, *71*, 305–330.

Over, D. E. (2022). The new paradigm and massive modalization. *Thinking & Reasoning*. https://www.researchgate.net/publication/357433854

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge: Cambridge University Press.

Williamson, T. (2020). *Suppose and tell: The semantics and heuristics of conditionals*. Oxford: Oxford University Press.

Conditionals, Indeterminacy, Probability and Truth

Dorothy Edgington

1 Introduction

This paper defends a slightly different view of conditionals from the view I have advocated for a long time. Many of the ingredients remain in place: conditionals essentially involve suppositions; conditionals are often uncertain, and the best tool for handling uncertain conditionals is conditional probability—the probability of the consequent on the supposition of the antecedent; a conditional probability cannot be equated with the probability of the truth of a proposition. Hence, on this view, conditionals do not express propositions. I took this to mean that conditionals do not have truth conditions. One disadvantage of this is that we no longer have an account of conditionals embedded in wider contexts, either in terms of truth, or in terms of probability.

D. Edgington (✉)
Birkbeck, University of London, London, UK
e-mail: dorothy.edgington@philosophy.ox.ac.uk

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_2

I have been impressed by an idea of Richard Bradley's (2012), who argues that, although conditionals do not express (ordinary) propositions, they may express something to which truth conditions can be ascribed, such that the probability of truth is a conditional probability, after all; and thus the problem of embedded conditionals is solved. But in order to make Bradley's view acceptable, I have had to develop some ideas of my own. The most important is this: even if they can be given truth *conditions*, such that the probability of truth is the conditional probability, I argue that it is often indeterminate (when the antecedent is false) what the truth *value* of the conditional is. I link the treatment of this indeterminacy to another species of indeterminacy—that due to vagueness.

Another issue I develop is the relation between the indicative conditional and a certain reading of the counterfactual. Indicative conditionals are the main focus of this paper, but as we have to decide what to say about them should the antecedent be false, the counterfactual form of words comes into the story too.

The rest of this introduction is a summary of the advantages and disadvantages of the suppositional view.

Will she accept if she is offered the job? Will I be cured if I have the operation? To assess a conditional, you *suppose* the antecedent is true, and consider what you think about the consequent, on that basis. The idea goes back to Frank Ramsey: 'If two people are arguing "if *p*, will *q*?" and are both in doubt as to *p*, they are adding *p* hypothetically to their stock of knowledge, and arguing on that basis about *q*' (1929, in 1931 p. 247). Often, this procedure yields verdicts which fall short of certainty: 'I think she'll accept if she's offered the job, but she might not'; 'It's around 90% likely that you will be cured if you have the operation'. Here the notion of conditional probability is what is needed: the probability of *q* on the supposition that *p*. Ramsey's remark continues: 'they are fixing their degrees of belief in *q* given *p*' (ibid.).

This well-known and much-used notion, a conditional probability, cannot be equated with the probability that a proposition is true. It essentially involves a *supposition*. When considering how likely it is that A, you are comparing A and $\neg A$; you ask yourself which is more likely, and by how much, A or $\neg A$? When considering how likely it is that if A,

B, that is, B on the supposition that A, you ask yourself which is more likely, and by how much, $A\&B$, or $A\&\neg B$?

Ramsey realized that this treatment of conditionals was not equivalent to treating them as propositions: 'Many sentences express cognitive attitudes without being propositions. … This is even true of the ordinary hypothetical' (1929, in 1931, pp. 239–30). David Lewis (1976) was the first to prove that there is no proposition the probability of whose truth can be systematically equated with a conditional probability, and there have been many subsequent proofs. Rather than give proofs now, I shall illustrate the difference between the suppositional approach and various propositional accounts with some examples.

First example: we have an ordinary coin, and I'm not sure whether Jane is going to toss it or not. How likely is it that if she tosses, it lands heads? I reckon toss-and-heads is no more or less likely than toss-and-tails; so my answer is 50%. That is the suppositionalist approach. Let's now add that it's 50% likely that she tosses the coin. Here are the three relevant possibilities, together with their probabilities:

(1) No toss (50%)
(2) Toss and heads (25%)
(3) Toss and tails (25%).

For those who accept the truth-functional, material implication account, the conditional is equivalent to 'Either she won't toss it, or she will and it will land heads'. So their answer to the question is 75%: the conditional is true except in case (3), so it has a 75% probability of being true.

Let us look at David Lewis's truth conditions (1973). (Lewis restricted these to counterfactuals, but he applied them widely, for instance to the forward-looking conditionals used in decision making.) Our conditional is true if case (2) obtains. It is false if case (3) obtains. And it is false if case (1) obtains, because it is not the case that the consequent is true in *all* relevant (closest) antecedent worlds. So the conditional gets only 25%.

Lewis accepts centring: if antecedent and consequent are both true, the conditional is true. But there are also theories which do not accept centring [e.g. Lycan (2001), Gillies (2009)]: a conditional involves quantification over a set of relevant possibilities—it is true if and only if every relevant antecedent-world is a consequent-world. On this strict-conditional reading, our example, 'How likely is that if she tosses, it lands heads?' deserves 0: the consequent is not true in all relevant antecedent worlds, and is therefore certainly false.

We have above three theories which (a) disagree with the suppositional theory, and (b) give bad answers to the question. They are not well suited to uncertain conditionals. There is nothing special about the example. Here is another. Jane hasn't decided whether she will accept if she is offered the job; but she is leaning in the direction of accepting. Suppose I think it is only about 10% likely that she will be offered the job. That divides into two possibilities: offered and accepts; offered and declines. I evaluate these two at 9% and 1% respectively. That is, on the supposition that she is offered the job, I think it's only about 10% likely that she will decline. As above, according to the truth-functional theory, it's 91% likely that if she's offered, she'll decline (i.e. either she is not offered, or she is offered and declines). According to a Lewisian theory, it's only 1% likely; and according to a strict-conditional theory, it's 0% likely. Thus these truth conditions either underestimate or overestimate the probabilities we naturally assign to conditionals.

(I have left Robert Stalnaker's theory aside. Stalnaker's aim was to provide truth conditions for conditionals such that the probability of truth is the conditional probability of consequent given antecedent (1968, 1970). He did not succeed, but it is not surprising that there are no glaring, blatant, differences between Stalnaker's and the suppositional assignment of probabilities to conditionals. An amended version of Stalnaker's account will be defended later.)

Can we live with the idea that conditionals don't have truth conditions? One issue concerns the validity of arguments containing conditionals, such as those employing modus ponens or modus tollens. A valid argument is often defined as one such that the truth of the premises guarantees the truth of the conclusion. Fortunately, Ernest Adams (1975) found a good way to treat the validity of arguments in this context. Take an argument which just consists of propositions, no conditionals, which necessarily preserves truth. Then, demonstrably, it necessarily preserves probability, in the following sense: the uncertainty (i.e. 1—the probability) of the conclusion cannot exceed the sum of the uncertainties of the premises.[1] Applying this latter criterion of validity to arguments with conditionals (where the uncertainty of a conditional is 1—the conditional probability of consequent given antecedent) gives us a criterion of validity for arguments with conditionals.

There are other problems, however, if we give up truth conditions. What are we to say of complex sentences which have conditionals as parts? For instance, mother says 'If it rains tomorrow, we'll go to the cinema, and if it doesn't rain, we'll go to the beach'. If the two contained conditionals are assessed by conditional probability, and hence are not propositions, we have no established means of assessing their role embedded in a conjunction—not in terms of truth value, and not in terms of probability either, for probability theory does not embed conditional probabilities in wider contexts. There have been some proposals to deal with this, which I shall discuss later, but they have not been an unqualified success.

A remaining problem is that it is just counterintuitive to deny that some conditionals are plainly true, and some are plainly false, depending on how the world is, whether or not we know their truth value. This can be so even when the antecedent is false. For example: one of these two boxes contains the prize, I know not which, and I must choose only one box. 'If I choose the box on the left, I'll get the prize' is true if the prize is in the left box, false if not (irrespective of whether it has a false antecedent). In addition, there are the necessities, and impossibilities: 'If it's a square, it has four sides', 'If it's a pentagon, it has six sides', which, intuitively, are as true or as false as anything is.

Then there is the case when the antecedent turns out to be true. It is natural to think that in this case, the truth of the consequent *verifies* the conditional, the falsity of the consequent *falsifies* the conditional. (Not all agree: advocates of the 'strict conditional' approach, who deny centering, hold that the truth of antecedent and consequent is insufficient for the truth of the conditional. This is at odds with the suppositional approach:

---

[1] Suppose two premises, A and B, entail C. Then $\neg C$ entails ($\neg A$ or $\neg B$). So $p(\neg C) \leq p(\neg A$ or $\neg B) \leq p(\neg A) + p(\neg B)$. This two-premise case easily extends to any number of premises.

the strict-conditional approach has the consequence that one may be certain that *A*, certain that *B*, but uncertain about whether if *A*, *B*. But if the probability of *A* is 1, and the probability of *B* is 1, it follows that the probability of *B* given *A* is 1. The truth-functional and Lewis's truth conditions and the suppositional view all agree that that *A*&*B* entails if *A*, *B*).

It will do no great harm if the suppositionalist agrees that we can speak of truth and falsity in the true-antecedent case, but it is of little help unless we have an account of what to say about the false-antecedent case. Some have suggested that there is no truth value if the antecedent is false.[2] Then the probability of the conditional cannot be equated with the probability that it is true, for it is true if and only if *A* and *B* are true, and a conditional can be highly probable while *A*, and hence *A*&*B*, are highly improbable. It is not a fault in a conditional that it has a false antecedent, and hence, on this view, is not true. I say, 'If you press that button there will be an explosion'. You don't press it—a disaster is avoided—thanks to the fact that my conditional is not true (for it has a false antecedent). If we are to reinstate truth conditions in a useful way, the focus must be on the case where the antecedent is false.

Before looking at Bradley's solution to this problem, I shall discuss the two issues I think need to be in place if it is to succeed.

2 Indeterminacy

Suppose we are to assign truth conditions to conditionals. Then, I shall argue, it is often the case that conditionals whose antecedents turn out to be false are not merely uncertain, but indeterminate—neither determinately true not determinately false. Nevertheless, the probabilities we assign to them are in good order. Here are some examples.

(1) 90% of the red balls in the bag have a black spot. You are to shake the bag, put your hand in and select a ball. It's 90% likely that if you pick a red ball, it will have a black spot. In fact, the ball you

[2] See, for example, Bruno de Finetti (1936), Nuel Belnap (1970).

pick is not red. It's 90% likely that if you had picked a red ball, it would have had a black spot. This probability judgement is unassailable. But there is, typically at least, no fact of the matter as to which red ball you would have picked, had you picked a red ball.[3] This is so whether or not our world obeys deterministic laws. If some indeterminism is involved in the picking of balls, the result is obvious. But assuming determinism is no help. If determinism governs the case, the past and the laws rule out your picking a red ball. They say nothing about which red ball you would have picked, had you done so.[4] These are probabilities without outcomes.

(2) A dog almost always, but not quite always, attacks and bites when strangers approach. We can't tell the difference between the cases in which it does, and the few cases in which it does not. 'It's very likely that you will be bitten if you approach', I'm told. I don't approach. 'It's very likely that I would have been bitten if I had approached', I say. Assume determinism. We can distinguish two kinds of case: (a) whether the dog bites is determined by some standing feature of the person approaching, such as the shape of their nose, but we haven't figured this out. In that case, the uncertainty is mere ignorance. (b) Whether the dog bites depends on very fine details of the manner of approach. There's no fact of the matter about how exactly, down to this level of detail, I would have approached, had I done so. In this case, even under determinism, again we have not mere uncertainty, but indeterminacy. The probability judgement is in order, either way.

(3) 'It's about 90% likely that you will be cured if you have the operation' I'm told. I decline the operation. 'It's about 90% likely that you would have been cured if you had had the operation', I'm told later. Again, we can distinguish two scenarios: (a) the uncertainty is due to some feature of the present state of my body, which can't be discovered until I am opened up—this is the case of mere uncertainty; (b) the uncertainty depends, at least in part, on very fine details of the hand movements and cuts the surgeon would have made, had I

---

[3] An exception would be a case in which your hand hovers over two balls, such that if you don't pick one you will pick the other.

[4] That is why Lewis (1979a) required 'small miracles', relative to the laws of the actual world, in assessing counterfactuals, under the assumption of determinism.

gone ahead. In this case we have not mere uncertainty, but indeterminacy. This metaphysical difference has no effect on the probability judgements.

As we have seen, there are various sources of indeterminacy. Indeterminism is one. But even under determinism, our ordinary antecedents can typically be realized in many ways. (Which red ball would I have picked? How exactly would I have approached the dog? Exactly what would the surgeon's hand movements have been?) And the outcome depends on fine details not specified in these ordinary antecedents. Also, the vocabulary of the conditional might not be suitable for subsumption under deterministic laws of nature, if such there be. This may be true of the counterfactuals we accept and assert about our own and others' mental lives. Even if determinism is true, these are not the categories that belong with deterministic laws. On the other hand, when the antecedent together with the present state of the world is a sufficient condition for the truth (or falsity) of the consequent, the conditional is determinately true (or determinately false), although we may be ignorant of these facts—as in 'If I pick the left box I will win'.

One final example: a lottery is called off at the last minute. Had it gone ahead, one of the tickets would have won, each individual ticket would very probably have lost, but there is no determinate fact of the matter about which ticket would have won. Had it gone ahead, the world would have been this way or this way or …; but it is indeterminate which.

This counterfactual indeterminacy is, to my mind, similar in structure to the indeterminacy brought about by vagueness. One second after noon is noonish. 10,000 seconds after noon is not noonish. Therefore, it's not the case that for all n, if n seconds after noon is noonish, $n + 1$ seconds after noon is noonish. Therefore, for some n, n seconds after noon is noonish and $n + 1$ seconds after noon is not noonish. That is, there is a last noonish second.[5] But it is indeterminate which the last noonish second is. That is like: one of the tickets would have won—the world would have been this way or that way or …—but it is indeterminate which. I have argued that the indeterminacy due to vagueness is best

[5] This example is from Sorensen (2001), p. 58.

theorized using the notion 'degree of closeness to clearly true', and giving this notion probabilistic structure (Edgington 1996). For me this is a nice rapprochement. But whether or not I am right about vagueness-related indeterminacy, it is unavoidable to give counterfactual indeterminacy probabilistic structure.

It is also important. We use these judgements in inferences. 'She probably didn't approach, because she is unhurt, and it's very probable that she would have been bitten if she had approached'. 'You probably don't have the virus, because the test was negative, and if you had the virus, it probably would have shown up on the test'. It makes no difference to us in practice, whether the uncertainty is mere uncertainty, or involves some indeterminacy.

3 The Transition From the Indicative to the Counterfactual

We have seen examples above of indicative conditionals which, should the antecedent prove false, transform seamlessly into equivalent counterfactual conditional judgements. This is indeed very common. 'Don't go in there', I say, 'If you go in you will get hurt'. You look sceptical but stay outside, when there is a large crash as the ceiling collapses. 'You see', I say, 'If you had gone in you would have got hurt—just as I said'. 'If they arrive by eight, we'll eat at nine' is rephrased hungrily at ten, 'If they had arrived at eight, we would have eaten at nine'. But there are pairs which do not fit this pattern. The most famous example (Adams 1970):

> If Oswald didn't kill Kennedy, someone else did.
> If Oswald hadn't killed Kennedy, someone else would have.

One may accept the first, yet reject the second. Certainly this is so on a natural reading of the second. But counterfactuals, in different contexts, admit of different readings. Counterfactuals involve a change in perspective away from your present epistemic state. More than one change can be permissible, and there is always available the reading which matches the indicative. First consider this exchange: 'Why did you arrest Smith?'.

'We knew the crime was committed by either Smith or Jones. If it hadn't been Jones, it would have been Smith'.

Similarly, the police are asked: 'You already had Oswald; why did you round up other people from that part of the crowd?' Answer: 'We didn't know it was Oswald; if it hadn't been Oswald it would have been one of them'.

Here is another telling case: argument by modus tollens. Take any indicative conditional, if *A*, *B*. Suppose you believe it. You then discover that not *B*; you argue: not *A*, because if it had been the case that *A*, it would have been the case that *B*, and it isn't the case that *B*. You use the counterfactual form merely to re-express what you expressed in the indicative when you considered *A* to be a possibility. Thus, any indicative conditional judgement can be re-expressed as a counterfactual, should the antecedent turn out to be false.

4 Stalnaker's Truth Conditions

Stalnaker wrote:

> Consider a possible world in which *A* is true, and which otherwise differs minimally from the actual world. '*If A, then B*' is true (false) just in case *B* is true (false) in that possible world. (1968, p. 35)

There is no need for qualms about weird entities called possible worlds. Typically they are just a partition of possibilities, fine enough for the problem at hand, identified by a set of propositions, like the lines of a truth table. They are integral to probabilistic thinking, and do not involve a change in commitments.

Stalnaker's semantics assumes that there is a unique world minimally different from the actual world. He is aware that this is often unrealistic, and suggests that when this is not so we need to use supervaluations and say: true if true for all admissible selections, false if false for all admissible selections, otherwise indeterminate—neither true nor false, he says (Stalnaker 1981, pp. 87, 90).

I agree that when the antecedent is false, the conditional is often indeterminate, but that verdict is not very helpful as it stands. 'If you had tossed the coin ten times, you would have got at least one head' and 'If you had tossed the coin ten times, you would have got ten heads' are both indeterminate, but one is almost certain, the other incredible. We need finer distinctions.

Further, when something is indeterminate, 'neither true nor false' is not the correct verdict, in my view. Rather, we should say it is indeterminate whether it is true or false. This is different, for if it is not true, and not false, it is not indeterminate whether it is true or false. In the indeterminate examples of §2, if they were neither true nor false, the probability that they are true would be 0. This is wrong: it's 90% likely to be true that if you had picked a red ball it would have had a black spot, that the dog would have bitten if I had approached, that I would have been cured if I had had the operation, etc., even when the uncertainty is not mere uncertainty, but indeterminacy. (Similarly, I argue, for vague judgements: if it is indeterminate whether the patch is red, it is indeterminate whether it is true that the patch is red, rather than not true, and not false either).

This point is crucial in what follows. It is not an idiosyncratic point of view, but quite widely held in the philosophical literature on vagueness and indeterminacy. See, for example, Elizabeth Barnes and Robert Williams (2011). In the case of vagueness, Crispin Wright argues against the 'not true and not false' verdict, by saying that it 'cannot do justice to the … datum that in general borderline cases come across as hard cases where we are baffled to choose between conflicting verdicts about which polar verdict applies, rather than as cases which we recognize as enjoying a status inconsistent with both' (2001, pp. 69–70). Similar remarks are made by Stephen Schiffer (2003, p. 191).

5 Bradley's Amendments

Bradley's first amendment to Stalnaker's account is to abandon the notion 'minimally different from the actual world', or 'most similar to the actual world', in favour of a probability distribution over the relevant antecedent worlds. I think this is right. Consider this example: the short straws.

You pick a straw from a collection of 100 straws. From the angle you can see them—end on—they all look the same; and they are the same, except for the length. 90 are of length 10 cm, 1 is 11 cm, and 9 are 20 cm. The straw you pick is 10 cm. long. Consider the conditional:

If you had picked a longer straw, it would have been less than 15 cm. long.

This does not deserve much credence, as only one in ten of the longer straws are less than 15 cm, long. But if we go by minimal difference from the actual world, it should be clearly true.

Bradley's second amendment is to distinguish conditionals from propositions. Conditionals involve two propositions which play different roles, one a supposition, one a judgement within its scope. They cannot be represented by the set of worlds in which they are true. Indeed conditionals are not 'in' worlds—they are cross-world entities. Bradley proposes that the content of a conditional $A \rightarrow B$ can be represented by the set of *pairs* of worlds, $<w_i, w_j>$ such that, if w_i is actual, and w_j is the world that would be actual if *A* were true, the conditional would be true, because *B* is true at w_j. I shall call the world that would be actual if *A* were true, the *potential A*-world.[6] It may not be determinate which the potential *A*-world is, (just as it is not determinate which the last noonish second is).

It is not ad hoc or unheard-of to claim that some kinds of content cannot be represented by a set of worlds. Some examples: to capture the content of indexical thoughts using 'I', 'now', etc., we need the richer notion of a 'centred world'—an ordered triple of a world, an individual and a time [see Lewis (1979b)]. Allan Gibbard (1990) proposes that the content of a normative judgement can be represented by a set of ordered pairs $<w, n>$ where w is a world and n is a system of norms. Sarah Moss (2018) argues that the contents of probability judgements are not propositions but sets of probability spaces.

[6] Bradley continues to speak, in Stalnaker-Lewis terms, of the 'nearest' *A*-world; but as there is no ordering relation of worlds, this name is not apt.

It is important to be clear about the sense in which conditionals are not propositions. As I have said, they cannot be represented by the set of worlds in which they are true. Some argue that propositions are, in any case, more fine-grained than the set of worlds in which they are true—for instance, they deny that 'John is asleep' expresses the same proposition as 'John is asleep and $2 + 2 = 4$', despite their being true at the same set of worlds. This debate is not relevant to the present issue, for even the fine-grained propositions can be mapped (many-one) on to the set of worlds in which they are true.

One might be more liberal with the word 'proposition' and allow that the content of any thought is a proposition, whether or not it can be represented by the set of worlds in which it is true. Andrew Bacon (2018) calls the content of vague thought propositions, while insisting that they cannot be represented by a set of possible worlds. This dispute 'is just disagreement about which entities we should grant the honorific title "propositions" to', he says (p. 75). Perhaps there is a liberal sense in which conditionals do express propositions—what matters here is whether they can be mapped on to the set of possible worlds in which they are true.

6 Bradley's Theory

Two types of uncertainty, Bradley notes, are involved in assessing a conditional $A \rightarrow B$: uncertainty about the facts—about which world is actual; and uncertainty about what would be the case if some supposition were true. (Note: 'the facts' are construed as not containing conditionals. Note also: this second type of uncertainty may turn out to be not mere uncertainty but indeterminacy, but that does not fundamentally alter the picture, I have argued.) We have a probability distribution over the facts; and, for antecedent A, we have a probability distribution over the candidate A-worlds. They can be combined into a joint probability distribution over these ordered pairs.

We do not need to divide logical space any finer than is needed for a particular example, so for this example we say there are just three possible worlds (classes of worlds if you prefer, but I shall call them worlds): at

w_1, A and B are true; at w_2, A is true and B is false; at w_3, A is false. These generate the following four possibilities for the conditional $A \to B$.

			$A \to B$	
w_1	A,B	$<w_1,w_1>$	T	(1)
w_2	$A,\neg B$	$<w_2,w_2>$	F	(2)
w_3	$\neg A$	$<w_3,w_1>$	T	(3)
		$<w_3,w_2>$	F	(4)

The probabilities of these four lines sum to 1.

The first two lines are the cases in which the antecedent is true, so in those the 'potential' A-world is the actual world (Thus we have centering. As already mentioned, this is essential for the suppositional conditional).

If on the other hand w_3 is actual, that does not tell us whether the potential A-world is w_1, in which case the conditional is true, or w_2, in which case the conditional is false.

The crucial rule governing this non-propositional entity, Bradley proposes, is this: **the probability of $A \to B$ given A, is the same as the probability of $A \to B$ given $\neg A$; the conditional is probabilistically independent of its antecedent.**

This guarantees that $p(A \to B) = p(B|A)$ (using, as I shall henceforth, this standard notation for 'p(B given A)'). By the law of total probability,

$$p(A \to B) = p((A \to B)|A)p(A) + p((A \to B)|\neg A)p(\neg A).$$

If $p((A \to B)|A) = p((A \to B)|\neg A)$, it follows that

$$p(A \to B) = p((A \to B)|A)\cdot(p(A) + p(\neg A)) = p((A \to B)|A).$$

But $p((A \to B)|A) = p(A \& B|A) = p(B|A)$.

Therefore, $p(A \to B) = p(B|A)$.

It is not a new discovery that if there is a conditional the probability of whose truth is a conditional probability, it must be probabilistically independent of its antecedent. Bas van Fraassen (1976) was the first to show this, in the wake of Lewis's proof. But if we construe the conditional as an ordinary proposition, there are always some probability distributions in which it is not independent of its antecedent, as I shall now show.

Suppose we just redescribe the four lines above as four possible worlds, four ways the world might be, in two of which the conditional is true—as Stalnaker did:

		$A \to B$	
w_1	A, B	T	(1)
w_2	A, $\neg B$	F	(2)
w_3	$\neg A$, w_1 is the potential A-world	T	(3)
w_4	$\neg A$, w_2 is the potential A-world	F	(4)

Suppose we start off thinking each of the four is equally likely. Then we learn $\neg(A\&B)$: the first line goes out. We learn nothing other than that. $p(B|A)$ is now 0. But $p(A\to B)$ is not 0: the third line remains a possible way the world might be, and we haven't eliminated that. Indeed, if probabilities change by conditionalization, the third line now has a probability 1/3. But we need not assume anything as strong as conditionalization: it is built into the picture that these are four exclusive and exhaustive ways the world might be, a bit like a four-horse race. It is built into the picture that it is possible that (1) is false and (3) is true, so that we can eliminate (1), without eliminating (3). So, on the propositional picture, $p(A\to B)$ is not 0, although $p(B|A)$ is 0. (This is one illustration of the result that there is no *proposition* the probability of whose truth is the probability of consequent given antecedent in all probability distributions. Indeed, no two contingent propositions are probabilistically independent in all probability distributions.)

On Bradley's approach, if line (1) is eliminated, line (3) gets 0 as well, for $p(A\to B)$ must be the same in the $\neg A$-worlds as it is in the A-worlds. Having eliminated $A\&B$, and thus completely rejected B on the supposition that A, then, should A prove false, on my present evidence I also eliminate the corresponding counterfactual: if A were true, B would be true.

It's not as if you have to go through the latter, rather convoluted, thought process to evaluate 'if A, B'. You suppose that A, and judge how likely it is that B, on that supposition, and that's all you have to do. But, as we saw from examples in §§2 and 3, there is always available a reading of the corresponding counterfactual which you are committed to, to the same degree, should the antecedent prove false. It is a sort of reflection of your conditional belief on to the false-antecedent case.

Thus, the non-propositional nature of the conditional is essential here. And it is a new discovery that there is this non-propositional entity the probability of whose truth is a conditional probability.

Some remarks about this independence: first, this is synchronic independence, and does not imply diachronic independence. It does not have any consequences for what happens on changing, or updating your beliefs. Conditional beliefs do not change by conditionalization, on this view—as we saw in the argument above. Nor does it follow that acquiring the belief that if *A*, *B* has no effect on your attitude to *A*. After all, you might already be pretty sure that *B* is false, in which case you will infer that *A* is probably false. The claim is just that at a given time, in a given epistemic state, your attitude to if *A*, *B* is indifferent to whether *A* is true or false.

Second, it is generally agreed that conditionals are typically independent of their antecedents. Recall: 'If you go in you will get hurt'. The ceiling collapses. 'You see', I say, 'If you had gone in you would have got hurt—just as I said'. Countless conditionals have that pattern.

Some have argued that there are relatively rare cases where independence fails. Here is one example adapted from Brian Ellis (1979). A spy, Smith, has been captured. The questions are whether he will confess, and whether he will be killed if he confesses. These could go either way. But Smith has very skilled judgement on such matters: if he confesses, he almost certainly won't be killed. Now suppose he doesn't confess. Is it not now more likely that he would have been killed if he had confessed? Certainly that is one way, perhaps the most natural way, of reading the counterfactual, but it is not the only way. Suppose he doesn't confess. If he had confessed, that would have been because he had compelling evidence that he wouldn't be killed. So if he had confessed he wouldn't be killed. It is this latter reading that is relevant here. There is always available the reading of the counterfactual as a mere reflection from the indicative to the case in which the antecedent is false.

7 Truth Conditions

Returning to the diagram that displays Bradley's truth conditions, I want to stress that they do deserve that name. We could express them like this: the conditional is true iff either $A\&B$, or $\neg A$ and the world that would be actual if A were true is a B-world, false iff $A\&\neg B$, or $\neg A$ and the world that would be actual if A were true is a $\neg B$-world. And we have probabilities over these alternatives. The truth conditions are in order, I argue, even in the case when, A being false, it is not determinate whether a B-world or a $\neg B$-world would obtain if A had been true. We know, and can state, what has to be the case for the conditional to be true, but it may be indeterminate whether it is true. Think again of the lottery that is called off. Had it gone ahead, I know how the world would have to be for my ticket to have won, but it may be indeterminate whether the world would have been that way. Truth *conditions* are in order, even if truth *value* is indeterminate.

In my view, the same holds for vague statements. We can write a truth table in the usual way for, e.g. 'The ball is large and red, or it's heavy'. But it might be indeterminate which of the eight lines is the true one. (In my view they have 'verities' as I called degrees of closeness to clear truth, which sum to 1.) Also, we saw, there is a last noonish second, although it is indeterminate which second that is. Analogously, there is a way the world would have been if A had been true, but it may be indeterminate which way that would have been.

I should emphasize that this appeal to indeterminacy is not found in Bradley. It is rather something I had to think about before I could accept his framework. Bradley writes in hyper-realist vein, as though it is always mere ignorance when we don't know which world would be actual if A were true. His name for the ways things would be if some supposition were true is 'the counterfacts'. His official stance, however, is neutrality on this metaphysical issue: 'Realists can construe both the facts and the counterfacts … as different features of reality about which we can be uncertain …. Antirealists can construe the use of counterfacts to fix the truth conditions of conditionals as merely a formal device to support a compositional semantics' (2012, p. 560).

I don't accept either of these options. His 'realist' is what I called 'hyper-realist'—according to whom it is always determinate, e.g. which red ball I would have picked, if I had picked a red ball, although it is not fixed by any feature of the actual world. I find this view unbelievable.[7] But I am not an antirealist in Bradley's sense: I do not see his construction as 'merely a formal device', of only instrumental value, at best a kind of fictionalism. I am not an antirealist at all: I think counterfactual indeterminacy is how things really are; probabilities can be assigned to the various possible outcomes of a counterfactual supposition, even if it is indeterminate which would have obtained. And the truth conditions of a counterfactual are given in terms of the various ways things might have turned out, if A had been the case, even if it is indeterminate which of these possibilities would have obtained.

8 *Compounds of Conditionals*

With truth conditions in place, Bradley's framework readily extends to truth-functional compounds, such as conjunctions of conditionals, to which probabilities of truth can be assigned—always subject to the rule that for any contained conditional $A \to B$, the probability of $A \to B$ given A (which is just $p(B|A)$ is equal to the probability of $A \to B$ given $\neg A$. I will briefly discuss some of the difficulties with previous attempts to extend the suppositional-probabilistic approach to compounds of conditionals.[8]

First, the truth-value-gap approach: according to de Finetti (1936), $A \to B$ is true if $A \& B$, false if $A \& \neg B$, and has no truth value—is undefined—if $\neg A$. Its probability is not the probability of its truth, but the probability that it is true given that it has a truth value. He then gave 3-valued tables for, e.g. conjunctions of conditionals: the conjunction is true iff both conjuncts are true, false iff at least one conjunct is false, otherwise undefined. The probability of the conjunction is

[7] This hyper-realism has a history in Jesuit theology, and is sometimes called 'Molinism' after the sixteenth-century Jesuit theologian Luis de Molina.

[8] I discuss these alternative approaches in more detail in Edgington (2020).

again the probability of truth given either true or false. Now consider a conjunction of conditionals with incompatible antecedents, e.g.

(R) If it doesn't rain tomorrow we'll go to the beach, and if it rains we'll go to the cinema.

On this account, there is no way (R) can be true. On the other hand it might be false, due to some unlikely contretemps such as illness. So the probability that it is true, given that it has a truth value, is 0. One might be 99% confident of each conjunct, but one must have zero confidence in the conjunction. This is a bad result!

Second, the Expected Value approach, which is more specific about the third semantic value: $A \rightarrow B$ gets 1 (=true) if $A \& B$, 0 (=false) if $A \& \neg B$, and gets a semantic value equal to $p(B|A)$ if $\neg A$.[9] Its 'probability' is the expected value of this entity: a weighted average of the semantic values, the weights being their probability. That is, $p(A \rightarrow B) = p(A \& B)0.1 + p(A \& \neg B)0.0 + p(\neg A).p(B|A)$, which simplifies to $p(B|A)$.

Considering again a conjunction of conditionals with incompatible antecedents, $(A \rightarrow B)\&(\neg A \rightarrow C)$, we can prove that its probability is the product of the probabilities of the two conditionals.[10]

This is in the right ballpark for (R) above: if each conditional gets 0.9, the conjunction gets 0.81. But for other examples it yields curious answers. Return to this example: I must choose one of two boxes, call them Left and Right. One and only one contains a prize. It's 50–50 which contains the prize. Consider these two conjunctions:

(C_1) If I choose Left, I'll win, and if I choose Right, I'll win.
(C_2) If I choose Left I'll win, and if I choose Right I won't win.

[9] See e.g. Vann McGee (1989), Richard Jeffrey (1991), and Stalnaker and Jeffrey (1994).

[10] In the case of incompatible antecedents, the expected value of the conjunction turns out to be $p(A \& B)p(C|\neg A) + p(\neg A \& C)p(B|A)$ which simplifies to $p(B|A)p(C|\neg A)$.

Intuitively, (C_1) is definitely false, and deserves 0 probability. (C_2) deserves 0.5 probability. On the present proposal, both get the value 0.25.[11]

Let us turn to Bradley. When a sentence contains two conditionals, with two antecedents, A_1 and A_2—for instance, a conjunction of conditionals—its semantics requires not ordered pairs but ordered triples of worlds, $<w_i, w_j, wk>$ such that if w_i is actual and w_j is the potential A_1-world and w_k is the potential A_2-world, the sentence is true. We shall consider a sentence of the form: $(A\rightarrow B)\&(\neg A\rightarrow B)$. There are four possible worlds, w_1-w_4, $A\&B$, $A\&\neg B$, $\neg A\&B$ and $\neg A\&\neg B$, respectively. Now for a conditional of this form, one of the antecedents is true and the other false. Given centring, the truth value of the conditional with the true antecedent depends just on the truth value of B in the actual world; but the truth of the conditional with the false antecedent can go either way, depending on whether the potential antecedent-world is a consequent-world. Thus there are eight cases to consider[12]:

		$A\rightarrow B$	$\neg A\rightarrow B$	$(A\rightarrow B)\&(\neg A\rightarrow B)$	
w_1	$A\&B$	T	T	T	(1)
		T	F	F	(2)
w_2	$A\&\neg B$	F	T	F	(3)
		F	F	F	(4)
w_3	$\neg A\&B$	T	T	T	(5)
		F	T	F	(6)
w_4	$\neg A\&\neg B$	T	F	F	(7)
		F	F	F	(8)

Consider (C_1) 'If I choose Left I'll win, and if I choose Right I'll win'. Let A be 'I choose Left'. In the context, we can use $\neg A$ for 'I choose Right'. Let B be 'I'll win'. In this example, we know that the two conditionals can't have the same truth value, so we can eliminate, i.e. assign

[11] Mark Lance (1991) has a similar example: there is a werewolf. It's 50% likely to be in our area tonight. If it is, it kills everyone outside. 'If John went out, he was killed' gets 0.5. 'If John went out the front door, he was killed, and if John went out the back door, he was killed' still deserves 0.5; but it gets only 0.25 on the present proposal.

[12] I could have listed the eight ordered triples alongside the following eight lines (e.g. line 1 is $<w_1, w_1, w_1>$ and line 2 is $<w_1, w_1, w_2>$). But I thought this would be less useful than my explanation above.

probability 0 to, four of these lines: lines (1), (4), (5) and (8) get 0 probability. We are left with the four lines in which the two conditionals get opposite truth values. These each get equal value of 0.25. The probability of (C_1) is 0; the probability of each of its constituent conditionals is 0.5; and each conditional has the same probability given A as it does given $\neg A$. (C_2), If I choose Left, I'll win and if I choose Right, I won't win', gets probability 0.5.

Returning to (R)—rain, beach and cinema—on Bradley's approach, if you are 90% confident in each conditional, you must be at least 80% confident in the conjunction. Demonstrably, as there is an entailment of the conjunction from the two conditionals as premises, the uncertainty of the conclusion cannot exceed the sum of the uncertainties of the premises (see Note 1). I said 81% was in the right ballpark, but it doesn't have to be 81%. If the reasons for the uncertainty are things such as illness, or car breakdown, which would prevent both outings, the conjunction could be 90%. All we can say in general is that the probability of the conjunction is between 80 and 90% inclusive—the probability of a conjunction is not, in general, determined by the probabilities of the conjuncts.

9 Concluding Remarks

Bradley's theory is considerably more conservative than the rivals I have considered. Just one tweak to the classical notion of a proposition, and suppositional conditionals have truth conditions. Their probability is the probability of their truth. Validity is necessary preservation of truth, and hence demonstrably, validity preserves probability in Adams's sense (see Note 1). Some conditionals can be straightforwardly true or false, even when they have false antecedents, whether we know their truth value or not. Many with false antecedents are neither determinately true nor determinately false, I have argued, but can be given truth *conditions* nevertheless—we know what would make them true; and they still have probabilities. And the theory gives reasonable results for compounds of conditionals.[13]

The theory does not, I think, describe a method we use in assessing conditionals. But then, formal semantics is typically too abstract for that. For simple conditionals, the suppositional method is fine as it stands. For e.g. conjunctions, one will be guided by such facts as that if two contained conditionals are highly probable, the conjunction cannot be much less probable; and if the two contained conditionals have probabilities that sum to 1 or less than 1, their conjunction may have 0 probability. What Bradley provides is a valuable theoretical framework which justifies our use of 'true' and 'false' in suppositional contexts.

So I was wrong, all these years [see e.g. Edgington (1995)], in saying conditionals don't have truth conditions! I was not wrong to deny that they express classical propositions about how the world is, or to insist that they essentially involve suppositions, or to claim that they are assessed by conditional probability. It had not occurred to me that there was this slightly different entity to which truth and falsity apply. And I have argued here that although truth and falsity do apply, there are many cases in which the truth value of a conditional with a false antecedent is indeterminate.[14]

References

Adams, Ernest W. 1970. 'Subjunctive and Indicative Conditionals', *Foundations of Language*, 6, 89–94.

Adams, Ernest W. 1975. *The Logic of Conditionals*. Dordrecht: Reidel.

Bacon, Andrew 2018. *Vagueness and Thought*. Oxford: Oxford University Press.

Barnes, Elizabeth and Williams, J. Robert G. 2011. 'A Theory of Metaphysical Indeterminacy', in Karen Bennett and Dean W. Zimmerman (eds.), *Oxford Studies in Metaphysics*, vol. 6. Oxford: Oxford University Press, 103–148.

[13] Bradley makes brief remarks about nested conditionals, which will involve ordered pairs as members of ordered pairs, but the details have still to be worked out. Quantification should be unproblematic.

[14] My thanks to Scott Sturgeon and to two anonymous reviewers for helpful comments.

Belnap, Nuel. 1970. 'Conditional Assertion and Restricted Quantification', *Noûs*, 4, 1–13.

Bradley, Richard. 2012. 'Multidimensional Possible-World Semantics for Conditionals', *Philosophical Review*, 122(4), 539–571.

de Finetti, Bruno. 1936. 'La logique de la probabilité', translated as 'The Logic of Probability', *Philosophical Studies*, 77 (1995), 181–190.

Ellis, Brian. 1979. *Rational Belief Systems*. Oxford: Blackwell.

Edgington, Dorothy. 1995. 'On Conditionals', *Mind*, 104, 235–329.

Edgington, Dorothy. 1996. 'Vagueness by Degrees', in R. Keefe and P. Smith (eds.) *Vagueness: A Reader*. Cambridge MA: MIT Press, 294–316.

Edgington, Dorothy. 2020. 'Compounds of Conditionals, Uncertainty and Indeterminacy', in Shira Elqayam, Igor Douven, Jonathan St B. T. Evans and Nicole Cruz (eds.), *Logic and Uncertainty in the Human Mind*. London: Routledge, 42–56.

Gibbard, Allan. 1990. *Wise Choices, Apt Feelings*. Oxford: Clarendon Press.

Gillies, Anthony. 2009. 'On Truth Conditions for If (But Not Quite Only If)', *Philosophical Review*, 118(3), 325–349.

Jeffrey, Richard. 1991. 'Matter of Fact Conditionals', *Proceedings of the Aristotelian Society*, 65, 161–183.

Lance, Mark. 1991. 'Probabilistic Dependence among Conditionals', *Philosophical Review*, 100, 269–276.

Lewis, David. 1973. *Counterfactuals*. Oxford: Basil Blackwell.

Lewis, David. 1976. 'Probabilities of Conditionals and Conditional Probabilities', *Philosophical Review*, 85, 297–315.

Lewis, David. 1979a. 'Counterfactuals and Time's Arrow', *Noûs*, 13, 455–476.

Lewis, David. 1979b. 'Attitudes *De Dicto* and *De Se*', *Philosophical Review*, 88, 513–543.

Lycan, William. 2001. *Real Conditionals*. Oxford: Oxford University Press.

McGee, Vann. 1989. 'Conditional Probabilities and Compounds of Conditionals', *Philosophical Review*, 98, 485–542.

Moss, Sarah. 2018. *Probabilistic Knowledge*. Oxford: Clarendon Press.

Ramsey, Frank. 1929. 'General Propositions and Causality' in Ramsey, Frank 1931, *The Foundations of Mathematics*. London: Routledge and Kegan Paul, 237–255.

The Foundations of Mathematics. London: Routledge and Kegan Paul, 237–255.

Schiffer, Stephen. 2003. *The Things We Mean*. Oxford: Oxford University Press.

Sorensen, Roy. 2001. *Vagueness and Contradiction*. Oxford: Oxford University Press.

Stalnaker, Robert. 1968. 'A Theory of Conditionals', in *Studies in Logical Theory, American Philosophical Quarterly* (Monograph Series 2), Oxford: Blackwell, 98–112, reprinted in Frank Jackson (ed.). 1991, *Conditionals*. Oxford: Oxford University Press, 28–45. Page reference to the reprint.

Stalnaker, Robert. 1970. 'Probability and Conditionals', *Philosophy of Science*, 37, 64–80.

Stalnaker, Robert. 1981. 'A Defense of Conditional Excluded Middle', in W. L. Harper, R. Stalnaker and G. Pearce (eds.), *Ifs*. Dordrecht: Reidel, 87–103.

Stalnaker, Robert and Richard Jeffrey. 1994. 'Conditionals as Random Variables', in E. Eells and B. Skyrms (eds.), *Probability and Conditionals*. Cambridge: Cambridge University Press, 31–46.

Van Fraassen, Bas. 1976. 'Probabilities of Conditionals', in W. Harper and C. Hooker (eds.), *Foundations of Probability Theory, Statistical Inference, and Statistical Theories of Science*, Volume I. Dordrecht: Reidel.

Wright, Crispin. 2001. 'On Being in a Quandary', *Mind*, 110, 45–98.

Gibbardian Collapse and Trivalent Conditionals

Paul Égré, Lorenzo Rossi, and Jan Sprenger

1 Introduction

In conditional logics, the law of Import-Export denotes the principle that a right-nested conditional of the form $A \rightarrow (B \rightarrow C)$ is logically equivalent to the simple conditional $(A \wedge B) \rightarrow C$ where both antecedents are united by conjunction. The law holds in classical logic for the material conditional, and several reasons can be given for its plausibility in the case of indicative conditionals. For instance, to use an example from (Cooper 1968, 300), the sentences "If Smith attends and Jones attends, then a quorum will be present", and "if Smith attends, then if Jones

P. Égré (✉)
Institut Jean-Nicod (CNRS/ENS/EHESS), Département de philosophie & Département d'études cognitives, Ecole normale supérieure, PSL University, Paris, France
e-mail: paul.egre@ens.fr

L. Rossi (✉) · J. Sprenger (✉)
Center for Logic, Language and Cognition (LLC), Department of Philosophy and Educational Science, Università degli Studi di Torino, Torino, Italy
e-mail: lo.rossi@unito.it

J. Sprenger
e-mail: jan.sprenger@unito.it

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_3

attends, a quorum will be present" appear to convey the same hypothetical information. The same appears to hold more generally, at least when A, B, and C themselves are non-conditional sentences, and the equivalence has been described as "a fact of English usage" (McGee 1989). Some psycholinguistic experiments thus indicate that pairs of sentences of the form "if A, then if B, C" and "if A and B, C" get similar answer profiles both regarding truth evaluations and regarding probabilistic assignments (see van Wijnbergen-Huitink et al. 2015, 799).[1]

In a celebrated paper, however, Allan Gibbard (1980) showed that a binary conditional connective "$\rightarrow$" collapses to the material conditional of classical logic "$\supset$" if the following conditions hold: (i) the conditional connective satisfies Import-Export, (ii) it is at least as strong as the material conditional ($A \rightarrow C \models_L A \supset C$), where $\models_L$ is the consequence relation of the target logic of conditionals, (iii) it is supraclassical in the sense that it reproduces the valid inferences of classical logic in conditional form ($\models_L A \rightarrow C$ whenever $A \models_{\mathsf{CL}} C$). From (i)–(iii) and some natural background assumptions, Gibbard infers $A \supset C \models_L A \rightarrow C$. Given (ii), $\rightarrow$ and $\supset$ are thus logically equivalent, according to the logic of conditionals ($\models_L$) under consideration. Prima facie, the conditional then needs to support all inference schemes validated by the material conditional in classical logic. However, various inferences based on the material conditional enjoy little plausibility in ordinary reasoning with conditionals (for instance, the inference from $\neg(A \rightarrow C)$ to $A \wedge \neg C$).

Gibbard's result poses a challenge for theories that compete with the material conditional as an adequate analysis of the indicative conditional.[2] For example, Stalnaker's logic **C2** (Stalnaker 1968) and Lewis's logic **VC** (Lewis 1973) are both supraclassical and make the conditional stronger than the material conditional, but they invalidate Import-Export for that matter.

---

[1] Import-Export has been challenged on linguistic grounds, see for instance Khoo and Mandelkern (2019), drawing on examples from Fitelson. The alleged counterexamples are subtle, however, and even Khoo and Mandelkern accept a version of the law. See also Appendix A. We return to some objections to the principle in the conclusion.

[2] Notable defenders of the material conditional analysis are Lewis (1976), Jackson (1979), and Grice (1989).

Not all theories make that choice, however. In this paper, we explore how certain *trivalent* logics of conditionals address Gibbard's challenge. These logics analyze an indicative conditional of the form "if A then C" as a *conditional assertion* that is indeterminate, or *void*, if the antecedent turns out to be false, and that takes the truth value of the consequent C if A is true (Reichenbach 1935; de Finetti 1936; Quine 1950; Belnap 1970, 1973). This analysis assigns a third truth value ("neither true nor false") to such "void" assertions, and gives rise to various logics that combine a truth-functional conditional connective with existing frameworks for trivalent logics (e.g., Cooper 1968; Farrell 1979, 1986; Milne 1997; Olkhovikov 2002 [2016]; Cantwell 2008; Baratgin et al. 2013; Égré et al. 2021a, b).

By retaining truth-functionality, these trivalent logics allow for simple algorithmic calculation of the truth value of even the most complex nested conditionals. Modal analyses and non truth-functional analyses are, by contrast, far more complex. Truth-functional analyses in particular avoid the problem of determining the right context for evaluating the conditional; there is much less inter-speaker variety than, for example, in Kratzer (2012)'s approach where it is variable to what degree (and in what way) the possible worlds need to be restricted for evaluating the consequent.[3] In the logics we consider, there is a fact of the matter as to the truth value of a conditional, and so, after getting to know the truth values of the components, there is no disagreement about whether the conditional is true, false, or indeterminate. Yet, as discussed in Égré et al. (2021a), these logics avoid some of the problems that have plagued two-valued truth-functional accounts.

We begin our investigation with a precise explication of Gibbard's result, including a more formal version of his original proof sketch (Sect. 2). Then we present two trivalent logics of indicative conditionals, paired with Strong Kleene semantics for conjunction and negation, and we examine how they deal with Gibbardian collapse (Sects. 3 and 4).

---

[3] See in particular Kratzer (2015) for a reassessment of the benefits of truth-functionality to evaluate certain embedded conditionals.

We then turn to trivalent logics that replace the Strong Kleene operators with Cooper's quasi-connectives where the conjunction of the True and the third truth value is the True (Sect. 5). Specifically, we show why rejecting supraclassicality—and retaining both Import-Export and a stronger-than-material conditional—is a viable way of avoiding Gibbardian collapse.

In the second part of the paper, we consider a recent strengthening of Gibbard's result due to Branden Fitelson (2013) and apply it to the above trivalent logics (Sects. 6, 7, and 8). From this analysis it emerges that Gibbard's result may be better described as a *uniqueness result*: we cannot have two conditional connectives that satisfy Import-Export as well as Conjunction Elimination, where one is strictly stronger than the other, and where the weaker (already) satisfies Modus Ponens. We conclude the paper with a response to an objection raised by David E. Over against the Export principle in relation to the paradoxes of the material conditional.

We also provide three appendices: Appendix A rebuts a recent attempt at a *reductio* of Import-Export, Appendix B provides the proofs of various lemmata stated in the paper, and Appendix C gives a more constrained derivation of Gibbardian collapse than his original proof, of particular relevance for the first trivalent system we discuss. For more in-depth treatment of trivalent logics of conditionals, we refer the reader to our comprehensive survey and analysis in Égré et al. (2021a, b).

2 Gibbard's Collapse Result

The Law of Import-Export is an important bridge between different types of conditionals: it permits to transform right-nested conditionals into simple ones. Import-Export is of specific interest in suppositional accounts of indicative conditionals that assess the assertability of a conditional by the corresponding conditional probability (as per Adams' thesis, viz. Adams 1965). Import-Export is then an indispensable tool for providing a probabilistic analysis of embedded conditionals. However, when Adams' Thesis, originally limited to conditionals with Boolean antecedent and consequent, is extended to nested conditionals, Import-Export creates unexpected problems.[4] Thus, a famous result by David Lewis (1976) shows that combining this latter equation with the usual laws of probability and an unrestricted application of Import-Export trivializes the probability of the indicative conditional.[5] Gibbard establishes a second difficulty with Import-Export, namely that any conditional satisfying Import-Export in combination with other intuitive principles collapses to the material conditional.

Gibbard's original proof (Gibbard 1980, 234–235)—in reality more of an outline—was based on semantic considerations and left various assumptions implicit. Here we provide a more formal derivation. In particular, Gibbard only stressed conditions (i)–(iii) below, but implicitly assumed two further conditions, here highlighted as (iv) and (v), as well as structural constraints on the underlying consequence relation. In what follows we use $\models_{\mathsf{CL}}$ for classical consequence, and $\equiv_L$ for the conjunction of $\models_L$ and its converse. Under (v) we mean that $\supset$ obeys classical principles, that is classical logical laws, inferences, and metainferences.[6]

Theorem 1 (Gibbard) *Suppose L is a logic whose consequence relation $\models_L$ is at least transitive, with $\supset$ and $\rightarrow$ two binary operators, obeying principles (i)–(v) for every formulae A, B, C. Then $\rightarrow$ and $\supset$ are provably equivalent in L.*

(i)	$A \rightarrow (B \rightarrow C) \equiv_L (A \wedge B) \rightarrow C$	*Import-Export*
(ii)	$A \rightarrow B \models_L (A \supset B)$	*Stronger-than-Material*

[4] The unrestricted version of Adams' equation is often called Stalnaker's Thesis (going back to Stalnaker 1970) or simply "The Equation", with the latter name being prevalent in the psychological literature. Adams defends it in his 1975 monograph, too.

[5] On the reasons to defend Import-Export in relation to probabilities of conditionals, see McGee (1989) and Arló-Costa (2001). For some objections linked to Lewis's triviality result, see for example (Sanfilippo et al. 2020, 151), whose theory assigns distinct expected values to a right-nested conditional and its simplified counterpart. A discussion of the links between Gibbardian collapse and Lewisian triviality lies beyond the scope of this paper, but we refer to Lassiter (2020) for a survey of Lewisian triviality results and their treatment in a trivalent framework.

[6] An *inference* is a relation between (sets of) formulae: for instance, the relation between $(A \supset B) \wedge A$ and $A \wedge B$; a *metainference* is a relation between inferences, for example, the relation between $A \models B$ and $\models A \supset B$. See for instance Ripley (2013) and Dicher and Paoli (2019) about the distinction and its underpinnings.

(iii)	If $A \models_{\mathsf{CL}} B$, *then* $\models_L A \to B$	*Supraclassicality*
(iv)	*If* $A \equiv_L A'$ *then* $A \to B \equiv_L A' \to B$	*Left Logical Equivalence*
(v)	$\supset$ *obeys classical principles in L*	*Classicality of* $\supset$

Proof

1.	$(A \supset B) \to (A \to B) \equiv_L ((A \supset B) \wedge A) \to B$	by (i)
2.	$((A \supset B) \wedge A) \equiv_L (A \wedge B)$	by (v) (classical inferences)
3.	$(A \supset B) \wedge A) \to B \equiv_L (A \wedge B) \to B$	by 2 and (iv)
4.	$(A \wedge B) \to B \equiv_L (A \supset B) \to (A \to B)$	1, 3 and the transitivity of $\models_L$
5.	$A \wedge B \models_{\mathsf{CL}} B$	Conjunction Elimination
6.	$\models_L (A \wedge B) \to B$	5 and (iii)
7.	$\models_L (A \supset B) \to (A \to B)$	4, 6, and the transitivity of $\models_L$
8.	$(A \supset B) \to (A \to B) \models_L (A \supset B) \supset (A \to B)$	by (ii)
9.	$\models_L (A \supset B) \supset (A \to B)$	7, 8 and the transitivity of $\models_L$
10.	$A \supset B \models_L A \to B$	by 9 and (v) (classical metainference)

□

This is not the only proof of Gibbard's result. In particular, Fitelson (2013) and Khoo and Mandelkern (2019) give more parsimonious derivations. But our presentation closely matches the structure of his original argument: first Gibbard shows that $(A \supset B) \to (A \to B)$ is a theorem of L (step 1–7), from that he derives $\models_L (A \supset B) \supset (A \to B)$ (step 8–9) and finally, he infers $A \supset B \models_L A \to B$ (step 10).

With Gibbard we can grant that the assumptions (ii) and (iii) introduced alongside Import-Export are fairly weak.[7] Stronger-than-Material is shared by all theories that classify an indicative conditional with true antecedent and false consequent as false.[8] Supraclassicality, a restricted

[7] Our reconstruction of Gibbard's proof requires the two "directions" of the Import-Export principle (i.e., importing and exporting the outer conditional operator, see Kaufmann and Kaufmann (2015) for more on the two directions in a modal setting). For a variant of Gibbard's proof by Khoo and Mandelkern (2019) that requires only the Export direction, see Appendix C. Whether we can get a collapse from Import alone is an open question.

[8] We use "stronger" in a reflexive sense (as opposed to "strictly stronger"). The name MP is sometimes used for this principle, see Unterhuber and Schurz (2014), or Khoo and Mandelkern (2019) who call it Modus Ponens. We find more appropriate to use "Stronger-than-Material" since Modus Ponens is strictly speaking a two-premise argument form. The two principles are not necessarily equivalent: in the system DF/TT introduced below, for instance, Stronger-than-Material holds but not Modus Ponens (in the form $A \to B, A \models B$).

version of the principle of Conditional Introduction, means that deductive relations are supported by the indicative conditional. Even that could be weakened by just assuming the conditional to support conjunction elimination as in step 6.

Assumptions (iv) and (v), on the other hand, are stronger than meets the eye. While the substitution rule LLE was taken for granted by Gibbard, likely on grounds of compositionality, it raises issues in relation to counterpossibles and other forms of hyperintensionality (see Nute 1980; Fine 2012). However, even if one is inclined to give up principle (iv), one may not find fault with applying it in this particular case. Similarly, (v) implies that the material conditional supports classical absorption laws (step 2 of the proof) and (meta-inferential) Modus Ponens (step 10) in L—two properties not necessarily retained in non-classical logics.

Gibbard's result also leaves a number of questions unanswered. One of them concerns the implication of the mutual entailment between $\rightarrow$ and $\supset$. Does the collapse imply that the two conditionals can be replaced by one another in all contexts, for example? The answer to this question is in fact negative, as we proceed to show using trivalent logic in the next section.

3 The Trivalent Analysis of Indicative Conditionals

From his result, Gibbard drew the lesson that if we want the indicative conditional to be a propositional function, and to account for a natural reading of embedded indicative conditionals, then the function must be "$\supset$", namely the bivalent material conditional. We disagree with this conclusion: trivalent truth-functional accounts of the conditional can satisfy Import-Export and yield a reasonable account of embeddings without collapsing to the material conditional. We now explain why one may want to adopt such an approach, and then, in the next two sections, how they deal with Gibbard's result.

Reichenbach and de Finetti proposed to analyze an indicative conditional "if A, then C" as an assertion about C upon the supposition that

Table 1 Truth tables for the de Finetti conditional (left) and the Cooper conditional (right)

$f_{\to_{\mathsf{DF}}}$	1	1/2	0
1	1	1/2	0
1/2	1/2	1/2	1/2
0	1/2	1/2	1/2

$f_{\to_{\mathsf{CC}}}$	1	1/2	0
1	1	1/2	0
1/2	1	1/2	0
0	1/2	1/2	1/2

A is true. Thus the conditional is true whenever A and C are true, and false whenever A is true and C is false. When the supposition (=the antecedent A) turns out to be false, there is no factual basis for evaluating the conditional statement, and therefore it is classified as neither true nor false. This basic idea gives rise to various truth tables for $A \to C$. Two of them are the table proposed by Bruno de Finetti (1936), and the one first proposed by William Cooper (1968), later rediscovered by Nuel Belnap (1973), by Grigory Olkhovikov (2002 [2016]), and by John Cantwell (2008) (see Table 1).[9] In both of them the value 1/2 can be interpreted as "neither true nor false", "void", or "indeterminate". There is moreover a systematic correspondence and duality between those tables: whereas de Finetti treats "not true" antecedents (<1) in the same way as false antecedents ($=0$), Cooper and Cantwell treat "not false" antecedents (>0) in the same way as true ones ($=1$) (and likewise Belnap 1973). Thus in de Finetti's table the second row copies the third, whereas in Cooper's table it copies the first.

One way to define the other logical connectives is via the familiar Strong Kleene truth tables (see Table 2). Conjunction corresponds to the "minimum" of the two values, disjunction to the "maximum", and negation to inversion of the semantic value. In particular, beside the indicative conditional $A \to C$, the trivalent analysis also admits a Strong

[9] In Égré et al. (2021a, b), we call it the Cooper–Cantwell conditional, but there we actually missed Belnap's 1973 contribution, as well as Olkhovikov's. In Belnap (1970), published under nearly the same title, Belnap originally worked with the de Finetti conditional instead. Belnap himself underemphasized the change in his 1973 paper ("I now choose to treat the case when A is nonassertive as on the side of the case when A is true rather than on the side of the case when A is false", p. 51), despite the centrality of this move. Henceforth we find simpler to call this conditional "Cooper's conditional", but we still name it the **CC** conditional to avoid confusion. We are indebted to Hitoshi Omori for bringing Olkhovikov's work to our notice.

Table 2 Strong Kleene truth tables for negation, conjunction, and the material conditional

	$f_\neg$
1	0
1/2	1/2
0	1

$f_\wedge$	1	1/2	0
1	1	1/2	0
1/2	1/2	1/2	0
0	0	0	0

$f_\supset$	1	1/2	0
1	1	1/2	0
1/2	1	1/2	1/2
0	1	1	1

Kleene "material" conditional $A \supset C$, definable as $\neg(A \wedge \neg C)$ (see again Table 2).

To have a logic, however, we also need a definition of validity. This question is non-trivial in a trivalent setting since preservation of (strict) truth is not the same as preservation of non-falsity. Like Cooper and Cantwell, but with independent reasons,[10] we opt for a tolerant-to-tolerant (TT-) consequence relation where non-falsify is preserved: an inference $A_1, \ldots, A_n \models C$ is valid if, for any valuation function (of the appropriate kind) v from the sentences of the language to the values $\{0, 1/2, 1\}$, whenever $v(A_i) \in \{1/2, 1\}$ for all $i \leq n$, then also $v(C) \in \{1/2, 1\}$. This choice yields two logics depending on how the conditional is interpreted: the logic **DF/TT** based on de Finetti's truth table, and the logic **CC/TT** based on Cooper's table.[11]

Both logics make different predictions, but they agree on a common core, and they give a smooth treatment of nested conditionals. In particular, both **DF/TT** and **CC/TT** satisfy the Law of Import-Export. We now investigate how they deal with Gibbardian collapse.

[10] Other trivalent consequence relations come with problematic features (Fact 3.4 in Égré et al. 2021a): they either fail the Law of Identity (i.e., $\not\models A \rightarrow A$), or they license the inference from a conditional to its converse (i.e., $A \rightarrow C \models C \rightarrow A$). Belnap (1973) too compares several options for validity in a trivalent setting, including the nontransitive notion **ST** (going from true premises to a non-false conclusion, see Cobreros et al. (2012), Belnap calls it *weak implication*). Like for Cooper, Belnap's favored notion of validity is the preservation of non-false values, alternatively the preservation of non-false and true values (i.e., the schemes we call **TT** and **SS** $\cap$ **TT** in Égré et al. 2021a). We return to the **SS** $\cap$**TT** notion of validity in the conclusion.

[11] The system **CC/TT** actually matches Cantwell's system. Cooper's logic, called **OL**, rests on a different choice of truth tables for conjunction and disjunction, and restricts valuations to two-valued atoms.

4 Gibbardian Collapse in DF/TT and CC/TT

We first consider Gibbard's triviality result in the context of DF/TT with its indicative and material conditionals. DF/TT is contractive, reflexive, monotonic, and transitive. An inspection of the principles (i)–(v) in Theorem 1 shows that:

- Assumption (i) holds. In particular, both sides of the Law of Import-Export receive the same truth value in any DF-valuation.
- Assumption (ii) also holds: if there is a DF-valuation v such that $v(A \supset B) = 0$, then $v(A) = 1$ and $v(B) = 0$, but then $v(A \to B) = 0$ as well, thus failing to make $A \to B$ tolerantly true.
- Assumption (iii) holds in DF/TT. We prove this in Appendix B.
- Assumption (iv) *fails* in DF/TT. In fact, $A \models_{\mathsf{DF/TT}} B$ and $B \models_{\mathsf{DF/TT}} A$ if, for any DF-valuation v, one of the following is given:

  (a) $v(A) = 1 = v(B)$ (c) $v(A) = 1;\ v(B) = 1/2$
  (b) $v(A) = 1/2 = v(B)$ (d) $v(A) = 1/2;\ v(B) = 1$

  Therefore, letting $v(C) = 0$, cases (c) and (d) provide counterexamples since either $A \to C \not\models_{\mathsf{DF/TT}} B \to C$ or $B \to C \not\models_{\mathsf{DF/TT}} A \to C$. A concrete example is the following:

  $$p \vee \neg p \models_{\mathsf{DF/TT}} (p \to \neg p) \vee (\neg p \to p)$$
  $$(p \to \neg p) \vee (\neg p \to p) \models_{\mathsf{DF/TT}} p \vee \neg p$$

  but

  $$[(p \to \neg p) \vee (\neg p \to p)] \to (p \wedge \neg p) \not\models_{\mathsf{DF/TT}} (p \vee \neg p) \to (p \wedge \neg p)$$

- Assumption (v) *fails* in general of $\supset$ in DF/TT. In particular, step 2 of Gibbard's proof fails: $(A \supset B) \wedge A \not\models_{\mathsf{DF/TT}} A \wedge B$, assuming $v(A) = 1/2$ and $v(B) = 0$.

The failure of Gibbard's conditions (iv) and (v) may seem to make DF/TT irrelevant for the discussion of his result. But this is not so: despite assumptions (iv) and (v) failing for DF/TT's indicative conditional and material conditional, the two conditionals turn out to be equivalent. More precisely, DF/TT validates the equivalence of $A \supset B$ and $A \to B$, as a reciprocal entailment ($\equiv_{\mathsf{DT/TT}}$), as a material biconditional (denoted by $\supset\subset$), and as an indicative biconditional (denoted by $\leftrightarrow$).

Lemma 2 *For every* $A, B \in \mathsf{For}(L)$:

$$\begin{aligned} A \supset B &\equiv_{\mathsf{DF/TT}} A \to B \\ &\models_{\mathsf{DF/TT}} (A \supset B) \supset\subset (A \to B) \\ &\models_{\mathsf{DF/TT}} (A \supset B) \leftrightarrow (A \to B) \end{aligned}$$

This result is not a coincidence. As it turns out, Gibbard's result can be derived using only principles (i), (ii), (iii), (v) and structural assumptions on logical consequence, in such a way that all uses of (v) are DF/TT sound. This result directly follows from the version of Gibbard's result established by Khoo and Mandelkern (2019), as we prove in Appendix C. We also give a sequent-style proof of the collapse in Appendix B, making use of the system presented in our Égré et al. (2021b).

However, such an extended form of equivalence between the indicative and the material conditional in DF/TT does not mean that the two conditionals are identified with each other or indistinguishable. In fact, they obey very different logical principles, such as the following connexive law:

$$A \to B \models_{\mathsf{DF/TT}} \neg(A \to \neg B) \text{ but } (A \supset B) \not\models_{\mathsf{DF/TT}} \neg(A \supset \neg B).$$

This shows that indicative and material conditional cannot be validly replaced in complex formulae in DF/TT. Put differently, DF/TT fails the classical principle of replacement of equivalents.

What is, then, the import of DF/TT's equivalences between different conditionals? Not much, one might argue. A look at the DF semantics and the status of the premises of Gibbard's Theorem in DF/TT shows that such equivalences are largely a byproduct of (i) the fact that the DF truth table assigns value 0 to an indicative conditional in the same cases in which it assigns value 0 to a material conditional, and (ii) the fact that the tolerant–tolerant consequence relation does not distinguish between value 1 and 1/2.

Notably, things are different when we move to CC/TT, keeping the tolerant-tolerant notion of consequence fixed, but moving to a truth table for the conditional which assigns value 0 to the indicative conditional in more cases. Like DF/TT, CC/TT is contractive, reflexive, monotonic, and transitive. Moreover:

- Assumption (i) and (ii) hold in CC/TT for the same reasons as DF/TT.
- Assumption (iii) *fails* in CC/TT. For example, $A \wedge \neg A \models_{\mathsf{CL}} B$, but $\not\models_{\mathsf{CC/TT}} (A \wedge \neg A) \rightarrow B$. A CC-valuation v s.t. $v(A) = 1/2$ and $v(B) = 0$ provides a counterexample.
- Assumption (iv) holds in CC/TT. As in the DF/TT case, we have that $A \models_{\mathsf{CC/TT}} B$ and $B \models_{\mathsf{CC/TT}} A$ if, for any CC-valuation v, one of the following is given:

  (a) $v(A) = 1 = v(B)$ (c) $v(A) = 1;\ v(B) = 1/2$

  (b) $v(A) = 1/2 = v(B)$ (d) $v(A) = 1/2;\ v(B) = 1$

  However, the row of value 1 is identical to the row of value 1/2 in CC-truth tables of the indicative conditional. Therefore, whenever one of (a)–(d) holds, for every formula C, we have that $v(A \rightarrow C) = v(B \rightarrow C)$, proving the claim.
- Assumption (v) *fails* in CC/TT, for the same reason it fails in DF/TT.

One of (i)–(iv) thus fails for CC/TT as it does for DF/TT, and (v) fails in both. The failure of assumption (iii), supraclassicality, is irrelevant for blocking the proof since the only classically valid inference required for the proof is Conjunction Elimination ($A \wedge B \models A$). This inference is also validated by CC/TT. The proof is thus blocked exclusively by the failure of assumption (v): $\supset$ does not behave classically in CC/TT (i.e., step 2 in our reconstruction of Gibbard's proof fails). Unlike DF/TT, CC/TTavoids Gibbardian collapse: it declares both conditionals materially equivalent, but neither logically equivalent nor equivalent according to the indicative biconditional:

Lemma 3 *For every* $A, B \in \mathsf{For}(L)$:

$$A \to B \models_{\mathsf{CC/TT}} A \supset B \text{ but } A \supset B \not\models_{\mathsf{CC/TT}} A \to B$$
$$\models_{\mathsf{CC/TT}} (A \supset B) \supset\subset (A \to B)$$
$$\not\models_{\mathsf{CC/TT}} (A \supset B) \leftrightarrow (A \to B)$$

In general, the indicative conditional of CC/TT is *strictly stronger* than its material counterpart: $A \to B$ entails $A \supset B$, but is not entailed by it. Altogether, DF/TT and CC/TT avert Gibbardian triviality in different ways. In both of them the material conditional is not fully classical, but an extensional collapse takes place in DF/TT anyway; this, however, does not make the material conditional always replaceable by the indicative in DF/TT. On the other hand, the indicative conditional of CC/TT is more remote from its material counterpart: not only does it validate different conditional principles (removing the most pressing paradox of the material implication), but it is also extensionally distinct from the material conditional within CC/TT itself.

Summing up, while Gibbardian collapse is avoided more markedly in CC/TT than in DF/TT, in neither logic does it constitute a form of "triviality": even when indicative and material conditionals are declared to be equivalent, they are firmly set apart by their inferential behavior. This concludes our study of Gibbard's original collapse result in trivalent logics based on Strong Kleene connectives. In the next section, we expand the scope of our analysis and look at trivalent logics of conditionals with a different semantics for the standard logical connectives.

5 Gibbardian Collapse in QCC/TT

The logics DF/TT and CC/TT solve a large set of problems related to the indicative conditional, but they also have important limitations. First, both CC/TT and DF/TT validate the Linearity principle $(A \to B) \vee (B \to A)$ for arbitrary A and B. This schema was famously criticized by

Table 3 Truth tables for trivalent quasi-conjunction and quasi-disjunction and the material conditional based on quasi-disjunction, as advocated by Cooper (1968)

$f'_\wedge$	1	1/2	0
1	1	1	0
1/2	1	1/2	0
0	0	0	0

$f'_\vee$	1	1/2	0
1	1	1	1
1/2	1	1/2	0
0	1	0	0

$f'_\supset$	1	1/2	0
1	1	0	0
1/2	1	1/2	0
0	1	1	1

MacColl (1908): one may reject both "if John is red-haired, then John is a doctor" and "if John is a doctor, then he is red-haired". So it is unclear on which basis we should accept, or declare as true, the disjunction of both sentences.

In a similar vein, some highly plausible conjunctive sentences can never be true on **DF/TT** or **CC/TT**. The schema $(A \to A) \wedge (\neg A \to \neg A)$ ("if A, then A; and if ¬A, then ¬A") is always classified as neither true nor false, although each of the conjuncts is a **DF/TT**- and **CC/TT**-theorem.[12] Likewise, an ensemble of conditional predictions of the form $(A \to B) \wedge (\neg A \to C)$ will always be indeterminate or false (see Belnap 1973, 60–61; Bradley 2002, 368–370). However, a sentence such as:

(1) If the sun shines tomorrow, Paul will go to the office by bike; and if it rains, he will take the metro.

seems to be true (with hindsight) if the sun shines tomorrow and Paul goes to the office by bike.

A principled reply to these challenges consists in modifying the truth tables for trivalent conjunction and disjunction, as proposed by Cooper (1968) (see also Belnap 1973; Dubois and Prade 1994; Calabrese 2002). In these truth tables, reproduced in Table 3, the conjunction of value 1 and value 1/2 is value 1, and similarly the disjunction of the value 0 with 1/2 is 0. This is coherent with the idea that a conditional assertion with two components (e.g., in Bradley's examples) should be classified

[12] We are indebted to Paolo Santorio for this example.

as true if one of the assertions came out true, and the other one void. Notably, the (quasi-)material conditional $A \supset C$ (definable as $\neg A \vee B$ or as $\neg(A \wedge \neg B)$) of a TT-logic based on these quasi-connectives blocks the paradoxes of the material conditional ($\neg A \not\models A \supset C$, $C \not\models A \supset C$), reflecting the fact that quasi-disjunction fails to validate Disjunction Introduction ($A \not\models A \vee B$).

Adopting "quasi-conjunction" and "quasi-disjunction" (the terminology is due to Adams 1975) invalidates Linearity and gives non-trivial truth conditions for ensembles or partitions of conditional assertions. In particular, $(A \to A) \wedge (\neg A \to \neg A)$ is always true, and so is $(A \to B) \wedge (\neg A \to C)$ when one of its conjuncts is true, irrespective of whether $\to$ is the DF or CC conditional. We call the resulting logics QDF/TT and QCC/TT.[13] However, when paired with DF/TT, quasi-conjunction leads to a violation of Import-Export, but not so in CC/TT. So the system of interest for us in this section is QCC/TT.

How does QCC/TT then fare with respect to the five premises of Gibbard's proof?

- Assumption (i) holds since both sides of the Law of Import-Export receive the same truth value in any QCC-valuation.
- Assumption (ii) *fails* since the (quasi-)material conditional is strictly stronger than the indicative conditional. The valuation $v(A) = 1$ and $v(B) = 1/2$ is a model of $A \to B$, but not of $A \supset B$, which takes the same truth values as $\neg A \vee B$.
- Assumptions (iii) and (v) *fail* with the same countermodels as in CC/TT.
- Assumption (iv) holds: it is independent of the interpretation of the standard connectives and the proof for CC/TT can be transferred.

In QCC/TT, two steps of Gibbard's proof are blocked, corresponding to the failure of assumptions (ii) and (v). As before, the failure of (iii) is inessential since the proof just requires Conjunction Elimination instead of the more general property of Supraclassicality (Table 4).

[13] QCC/TT is almost identical to Cooper's logic OL, except that Cooper requires valuations to be bivalent on atomic formulae.

Table 4 Overview of which premises of Gibbard's proof are satisfied by the logics DF/TT, CC/TT and QCC/TT. CE = conjunction elimination (=a sufficient surrogate for (iii)), TRM = transitivity, monotonocity and reflexivity of the logic. $\equiv, \leftrightarrow, \supset\subset$ concern whether logical, indicative, or material equivalence holds between $\supset$ and $\rightarrow$

Condition	(i)	(ii)	(iii)	CE	(iv)	(v)	TRM	$\equiv$?	$\leftrightarrow$?	$\supset\subset$?
DF/TT	✓	✓	✓	✓	✗	✗	✓	✓	✓	✓
CC/TT	✓	✓	✗	✓	✓	✗	✓	✗	✗	✓
QCC/TT	✓	✗	✗	✓	✓	✗	✓	✗	✗	✗

Since the material conditional is strictly stronger than the indicative in QCC/TT, Gibbardian collapse does not happen, and moreover, neither the material nor the indicative conditional declares the two connectives equivalent:

Lemma 4 *For every* $A, B \in \mathsf{For}(L)$:

$$A \supset B \models_{\mathsf{QCC/TT}} A \rightarrow B \text{ but } A \rightarrow B \not\models_{\mathsf{QCC/TT}} A \supset B$$
$$\not\models_{\mathsf{QCC/TT}} (A \supset B) \supset\subset (A \rightarrow B)$$
$$\not\models_{\mathsf{QCC/TT}} (A \supset B) \leftrightarrow (A \rightarrow B)$$

In QCC/TT, the connectives are thus more distinct than in DF/TT (where they are logically and materially equivalent) and CC/TT (where they are not logically, but still materially equivalent). The way out provided by QCC/TT is notable for another reason, too. Most theorists react to Gibbardian collapse either by giving up or restricting Import-Export (e.g., Stalnaker, Kratzer), or by endorsing a material conditional analysis of the indicative conditional (e.g., Grice, Lewis, Jackson). Denying that $\supset$ satisfies the classical laws in a logic of conditionals—the road taken by CC/TT—is already less common. However, Cooper's original approach is probably unique in entertaining the possibility of an indicative conditional that is strictly weaker than the material conditional. The explanation is probably that bivalent logic has been the default framework for formal work on conditionals and the material conditional represents, in that framework, the weakest possible conditional connective. The logic QCC/TT thus shows an original and surprising way of defining the relationship between the two connectives.

6 Fitelson's Generalized Collapse Result

Our rendition of Gibbard's original argument has revealed that one of the premises—namely that $\models A \to C$ whenever A classically implies C—is stronger than needed: we only require that $(A \wedge C) \to C$ be a logical truth. On the other hand, Gibbard's argument uses some properties of classical logic and the material conditional, such as the fact that $A \wedge (A \supset C)$ is logically equivalent to $A \wedge C$. Gibbard's result can thus be generalized along two dimensions: first, use premises only as strong as we need them for the proof of the collapse result; second, make explicit the classicality assumptions (compare Sect. 2) and extend the result to other logics than just classical logic with the material conditional.

Branden Fitelson (2013) has provided one such generalized result. It concerns the relation between two binary connectives represented by the symbols $\to$ and $\rightsquigarrow$ in an arbitrary logic L, whose consequence relation we denote with $\models_L$. Letting A, B, and C stand for arbitrary formulae of L, and $\models_L$ for some consequence relation defined for the language of L, Fitelson states eight independent conditions sufficient to derive a general collapse result:[14]

(1) $\models_L (A \wedge B) \rightsquigarrow A$ (Conjunction Elimination for $\rightsquigarrow$)
(2) $\models_L (A \wedge B) \to B$ (Conjunction Elimination for $\to$)
(3) $\models_L A \rightsquigarrow (B \rightsquigarrow C)$ if and only if $\models_L (A \wedge B) \rightsquigarrow C$ (Import-Export for $\rightsquigarrow$)
(4) $\models_L A \to (B \to C)$ if and only if $\models_L (A \wedge B) \to C$ (Import-Export for $\to$)
(5) If $\models_L A \to B$, then $\models_L A \rightsquigarrow B$ ($\to$ implies $\rightsquigarrow$)
(6) If $\models_L A \rightsquigarrow B$, then $A \models_L B$ (Conditional Elimination for $\rightsquigarrow$)

[14] Our notation swaps the meaning of the symbols $\to$ and $\rightsquigarrow$ in Fitelson's work to make it consistent with the rest of our paper. We should note that Fitelson's result assumes something strictly speaking weaker than Conjunction Elimination, but we do not go into the details for reasons of simplicity.

(7) If $A \equiv_L B$ and $\models_L A \to C$, then also $\models_L B \to C$ (Left Logical Equivalence)

(8) If $A \models_L B$ and $A \models_L C$, then $A \models_L B \wedge C$ (Conjunction Introduction)

In short, Fitelson's result concerns the relationship between two conditionals which satisfy both Conjunction Elimination (1+2) and Import-Export (3+4), and of which one is stronger than the other one (5). The stronger conditional, represented by the normal arrow $\to$, is supposed to represent the indicative conditional. Moreover, it is assumed that the weaker connective $\rightsquigarrow$ satisfies Conditional Elimination[15] relative to the logic $\models_L$ (6), which means—given certain structural assumptions on L, such as transitivity and monotonicity—that $\rightsquigarrow$ satisfies (meta-inferential) Modus Ponens. The final conditions concern the substitution of $\models_L$-equivalents in the premises of $\to$-validities (7) and Conjunction Introduction (8): if two propositions follow from a third, then so does their conjunction.

Fitelson shows that these axioms are logically independent from each other and that they are sufficient to show that the two connectives $\to$ and $\rightsquigarrow$ are logically equivalent:

Theorem (Fitelson 2013): From conditions (1)–(8) it follows that

$$A \rightsquigarrow B \models_L A \to B \qquad \text{and} \qquad A \to B \models_L A \rightsquigarrow B$$

As Fitelson emphasizes, this result should not be taken to imply that the connective $\to$ collapses to the *material* conditional, or that the indicative conditional "If A, then C" should be interpreted as "not A or C".[16]

[15] What we call Conditional Elimination is the converse of Conditional Introduction; the two properties together are known as the Deduction Theorem.

[16] Fitelson (p.c.) has recently sharpened his result, showing that Import-Export for $\to$ is actually *equivalent* to the collapse, over (1)–(3) and (5)–(8). We thank Branden Fitelson for bringing this to our attention. While technically and conceptually interesting, this sharpening does not affect the trivalent analysis of Fitelson's generalized collapse: the ways in which trivalent conditionals

Fitelson's result is interpretation-neutral and concerns *any* two connectives with the said properties; specifically, it does not presuppose that the weaker connective $\rightsquigarrow$ corresponds to the material conditional $\supset$. Whether the material conditional $A \supset C$ (i.e., $\neg A \vee C$) satisfies the properties of $\rightsquigarrow$ (i.e., conditions (1), (3), (5), and (6)) will depend on which logic we choose to interpret $\models_L$, and we will soon see that it need not in a trivalent setting. What Fitelson shows is rather that if a conditional connective satisfies Conjunction Elimination, Import-Export, and (meta-inferential) Modus Ponens, then in any logic with Conjunction Introduction, there cannot be a strictly stronger conditional connective that satisfies these conditions as well as axiom (7)—the substitution of equivalents in the premises of its theorems. In this sense, Fitelson proves the existence of an upper bound for the strength of a conditional that satisfies these intuitively desirable logical properties. Moreover Fitelson shows that such a connective must also validate some central intuitionistic principles.

7 Fitelson's Result and Trivalent Logic

What does Fitelson's result mean for trivalent logics when his two connectives $\rightarrow$ and $\rightsquigarrow$ are identified with the indicative and the material conditional? Keeping the tolerant-to-tolerant character of the logical consequence relation fixed (see Sect. 4 for why), we have to assign values to the following parameters:

- the truth table for the indicative conditional (de Finetti or Cooper–Cantwell);
- the truth table for conjunction and disjunction (Strong Kleene operators or Cooper's quasi-conjunction and disjunction);
- which connective in Fitelson's result represents the indicative conditional, and which connective represents the material conditional.

block the weaker version of Fitelson's result immediately extend to the sharper version. For this reason, we will focus on the 2013 version in what follows.

This leaves us with eight different logics, characterized by the choice of the truth table for the indicative conditionals (DF or CC), the truth tables for conjunction and disjunction (Strong Kleene or Cooper), and the assignment of conditionals to Fitelson's connectives ($\to$ and $\leadsto$). Fitelson suggests that the stronger connective $\to$ stands for the indicative conditional. However, the properties of $\leadsto$, which include Modus Ponens, Conjunction Elimination, and Import-Export, could also square well with the indicative conditional. Moreover, the indicative conditional can be *weaker* than the material conditional in QCC/TT. Thus, we have to carefully examine all ways of distributing Fitelson's connectives to truth tables.

As noticed in the previous section, QDF/TT does not satisfy Import-Export for the indicative conditional and so we set it aside (either condition (3) or condition (4) will fail). All the other logics satisfy conditions (1)–(4) and also condition (8). Thus our discussion will be limited to those logics and the more controversial properties (5), (6), and (7). Actually, we see that none of our trivalent logics satisfies all of these principles:

DF/TT with $\to = \to_{\mathsf{DF}}$ Satisfies (5)—material and indicative conditional are DF-equivalent—, but neither (6) nor (7). For (6), consider a valuation v such that $v(A) = 1/2$, $v(B) = 0$, and for (7), consider $v(A) = 1/2$, $v(B) = 1$, and $v(C) = 0$.

DF/TT with $\to = \supset$ Satisfies (5), but neither (6) and (7). Consider the same examples as above.

CC/TT with $\to = \to_{\mathsf{CC}}$ Satisfies (5) and (7), but not (6). Consider again v such that $v(A) = 1/2$ and $v(B) = 0$.

CC/TT with $\to = \supset_Q$ Satisfies (6), but neither (5) nor (7). For (5), consider v such that $v(A) = 1/2$ and $v(B) = 0$; for (7) consider v such that $v(A) = 1/2$, $v(B) = 1$, and $v(C) = 0$.

QCC/TT with $\to = \to_{\mathsf{CC}}$ Satisfies (6) and (7), but not (5). Consider v such that $v(A) = 1$ and $v(B) = 1/2$.

QCC/TT with $\to = \supset_Q$ Satisfies (5) and (6), but not (7). The counterexample is given by some v such that $v(A) = v(C) = 1/2$ and $v(B) = 1$.

Table 5 summarizes our findings. As we see, none of our trivalent candidate logics for the indicative conditional obeys all of these axioms. Since there are no obvious alternatives to the (various forms of the) material

Table 5 Overview of the satisfaction/violation of Fitelson's conditions (5)–(7) in different trivalent logics

Condition/Logic		DF/TT		CC/TT		QCC/TT	
Assignment of Symbols	Indicative=?	$\rightarrow$	$\rightsquigarrow$	$\rightarrow$	$\rightsquigarrow$	$\rightarrow$	$\rightsquigarrow$
	Material=?	$\rightsquigarrow$	$\rightarrow$	$\rightsquigarrow$	$\rightarrow$	$\rightsquigarrow$	$\rightarrow$
(5): $\rightarrow$ implies $\rightsquigarrow$		✓	✓	✓	✗	✗	✓
(6): Conditional Elimination for $\rightsquigarrow$		✗	✗	✗	✓	✓	✓
(7): Substitution of Equivalents ($\rightarrow$)		✗	✗	✓	✗	✓	✗
Collapse strongly blocked?		✓	✓	✗	✗	✗	✗

conditional as the second connective in Fitelson's theorem, Gibbardian collapse is blocked for the entire range of trivalent logics that we study. In particular, since at least one of the axioms fails for all configurations we have looked at, the connective $\rightarrow$ must also fail one of the principles of the intuitionistic conditional (this is, as mentioned above, a consequence of satisfying conditions (1)–(8)).

8 Blocking Fitelson's Collapse Strongly and Weakly

In order to better assess the distinct ways in which Fitelson's collapse is blocked in trivalent logics, we introduce a useful distinction. We say that a logic of indicative conditionals L blocks the collapse *strongly* if at least one of conditions (1)–(8) is not satisfied by letting $\rightarrow = \rightarrow_{\text{ind}}$, where $\rightarrow_{\text{ind}}$ is the connective that, in L, is taken to model the indicative conditional. We say that the L blocks the collapse *weakly* if $\rightarrow = \rightarrow_{\text{ind}}$ *and* $\rightsquigarrow = \supset$, where $\supset$ is the material conditional in L. In other words, L blocks Fitelson's collapse strongly if some of Fitelson's premises fails in L once $\rightarrow$ is interpreted as L's candidate for the indicative conditional, regardless of how the other conditional $\rightsquigarrow$ is interpreted. On the other hand, L blocks Fitelson's collapse only weakly if some of Fitelson's premises fails in L once $\rightarrow$ is interpreted as L's candidate for the indicative conditional *and* $\rightsquigarrow$ is interpreted as L's material conditional. In the former case, L's indicative conditional is non-trivial (in the sense of the collapse) *by itself*, whereas in the latter case it is non-trivial only if we assume (at least some of) the features of $\supset$ in L for the other conditional.

A glance at our findings shows that Fitelson's collapse result is blocked strongly for the DF/TT-logics, and only weakly for all (Q)CC/TT-logics. The failure of collapse in the (Q)CC/TT-logics is due to both features of the Cooper–Cantwell conditional in a TT-consequence relation *and* the choice of the material conditional as the interpretation of the weaker connective $\rightsquigarrow$. Does this show that the indicative conditional of the (Q)CC/TT-logics is "trivial", or in some sense uninteresting? Not really. All Fitelson's result can be used to argue for is that, given (1)–(8), the indicative conditional of (Q)CC/TT-logics is L-equivalent to (i.e., inter-L-inferrable with) an unspecified conditional which: (i) cannot be the material conditional of L (since (Q)CC/TT-logics weakly block the collapse), and (ii) satisfies conditions (1), (3), (5), and (6), over a background logic which satisfies (8).[17] Now, not only are these properties unproblematic—by themselves, they do not give rise to any paradox of implication—, they are indeed desirable. Hence, it should actually be a welcome result that an indicative conditional is equivalent to a conditional with such properties.

In summary, since the trivalent logics we have examined block Fitelson's collapse result systematically, we do not find ourselves in the dilemma of having to sacrifice Import-Export, or another plausible condition to avoid triviality. To us, the most reasonable construal of Fitelson's theorem is as a *uniqueness result*: it is impossible to have two conditional connectives both satisfying Import-Export and Conjunction Elimination, such that one is strictly stronger than the other and where the weaker one satisfies Conditional Elimination. This leaves Left Logical Equivalence (condition (7)) out of the picture, but as Table 5 shows, this condition is only required to prevent collapse in one case, namely QCC/TT, in which the material conditional is stronger than the indicative conditional. For all other combinations there is a tension between the relative strength of the connectives (as codified by (5)) and the

[17] Respectively: Conjunction Elimination (1), Import-Export (3), being entailed by indicative conditionals (5), Conditional Elimination/Modus Ponens (6), and Conjunction Introduction (8).

fact that the weaker connective should satisfy Conditional Elimination (namely (6)).

9 Conclusion

This paper has given a precise reconstruction of Gibbard's informal argument that any indicative conditional that satisfies Import-Export and is supraclassical and stronger than the material conditional must collapse to the material conditional. Specifically, we have seen that Gibbard's argument requires additional premises (e.g., structural assumptions on the underlying logic L) and that the premises are not tight either (e.g., Supraclassicality can be replaced without loss of validity by Conjunction Elimination).

We have then explored how a family of trivalent logics, all based on the idea that a conditional is void when its antecedent turns out false, fare with respect to Gibbardian collapse. The logics we have examined all block an important premise of Gibbard's proof, namely the classical behavior of the material conditional $\supset$, as well as one additional premise (different for each logic). Nonetheless, in **DF/TT**—the tolerant-to-tolerant logic based on de Finetti's truth table for the indicative conditional—Gibbardian collapse occurs, but this does not mean that both conditionals obey the same logical principles. In contrast, Cantwell's logic **CC/TT** and Cooper's logic **QCC/TT**, based on their common truth table for the indicative, avoid Gibbardian collapse altogether. This shows us that the apparent lesson from Gibbard's result—that one has to give up Import-Export or endorse the material analysis of the conditional—is mistaken.

We confirmed that diagnosis by looking at these logics in the context of the strengthening of Gibbard's result proposed by Fitelson (2013). Specifically, we have re-interpreted Fitelson's result as showing the impossibility of having two distinct connectives that both satisfy a set of characteristic properties (Conjunction Elimination, Import-Export), where the weaker one also satisfies Conditional Elimination. A logic of indicative conditionals does not have to choose between forswearing Import-Export and embracing the material conditional analysis: trivalent logics of conditionals offer a simple, yet articulate and fully truth-functional alternative that avoids both problems.

Some challenges remain for the logics we presented, however. One issue we have set aside is whether the Export principle is as fundamental as the Import principle toward the intuition that the law of Import-Export is valid. As stressed by David E. Over, from the Export principle, and from the validity of $A \wedge B \to A$, it follows that $A \to (B \to A)$ is valid. If $\to$ obeys Modus Ponens, then this implies that $B \to A$ follows from A, a version of the paradox of material implication. In a system like **CC/TT**, which supports both MP and Export, this prediction is borne out. But even in **DF/TT**, in which MP fails, $B \to A$ follows from A.

In order to avoid this feature, one option is to make validity more demanding, for instance, by requiring the preservation of both non-false values and of true values, namely using so-called **SS ∩ TT**-validity (see Belnap 1973; Dubois and Prade 1994; McDermott 1996). In **DF/SS∩TT** as in **CC/SS∩TT**, the law of Import-Export remains valid, but $B \to A$ does *not* follow from A. In Égré et al. (2021a), we did not select this notion of validity because it rules out sentential validities such as $A \to A$ (see Footnote 10). However, back then we also admitted that systems like **CC/TT** and **DF/TT** fall prey to the paradoxes of the material conditional. Obviously, systems like **CC/SS∩TT** or **DF/SS∩TT** also block Gibbardian collapse, being subsystems of the logics we discussed. Of the Gibbardian assumptions (i), (ii), (iii) and (v), **CC/SS∩TT** too preserves (i), (ii), and fails (iii) and (v). And **DF/SS∩TT** likewise preserves (i) and (ii), but this time it fails (iii) beside (v).

The upshot is that Export need not be viewed as more suspect than Import in this context.[18] The issue, instead, is whether it matters to preserve sentential validities based on the conditional (formulae such as $A \wedge B \to A$), or whether argument forms suffice to capture conditional reasoning. We think that this is an important foundational issue. We shall not pursue it here, but we conclude that this makes the choice

[18] Setting aside complexity considerations: D. Over points out that Import corresponds to a simplifying assumption, unlike Export. We agree, but our point here is merely about the logical problem raised by Export.

between trivalent options for validity more complex than acknowledged in our previous work.

Funding Funding for this research was provided by grants ANR-17-EURE-0017 (program FrontCog), and ANR-19-CE28-0004-01 (program Probasem) (P.E.) and by the European Research Council (ERC) through Starting Grant No. 640638 (J.S.). P.E. also thanks the University of Torino for a research invitation in June 2021. We are indebted to David E. Over, to Stefan Kaufmann, and to an anonymous reviewer for valuable comments and criticisms. Thanks also to Colin Caret, Michael De, Mario Günther, Branden Fitelson, Dan Lassiter, Matt Mandelkern, Hitoshi Omori, Hans Rott, and to audiences in Regensburg and in Utrecht for very helpful exchanges.

Appendices

Appendix A: Import-Export Revisited

Our case study on trivalent logics shows that it is possible to have Import-Export without restriction in a conditional logic without running into undesirable results of collapse to the material conditional, or to other connectives that are clearly too weak. Specifically, even if a conditional connective $\rightarrow$ validates Import-Export, the schema $(A \wedge B) \rightarrow A$, and is stronger than the material conditional, it need not be logically equivalent to the latter.

This observation raises the suspicion that the scope of Gibbardian collapse results may have to do with the absence or presence of bivalence. Note that the third truth value has been essential to constructing suitable counterexamples to Fitelson's conditions (1)–(8), and to blocking a generalized collapse theorem. In other words, we conjecture that Gibbardian collapse is a characteristic feature of conditional connectives with Import-Export in *bivalent logic*.

This conjecture shall now be probed by studying a recent reductio argument against Import-Export. Matthew Mandelkern (2020) argues that Import-Export, when conjoined with other plausible principles, leads to absurd conclusions (compare also Mandelkern 2019). Specifically, for a logic $(L, \models_L)$ with formulae A, B, and C and a connective $\rightarrow$ representing the indicative conditional, Mandelkern considers (and defends) the following three principles:

If	$A \models_L B$	then	$\models_L A \rightarrow B$	(Conditional Introduction)
If	$\models_L A \rightarrow B$	then	$A \rightarrow (B \rightarrow C) \equiv_L A \rightarrow C$	(Nothing Added)
If	$A \rightarrow C \equiv_L B \rightarrow C$	then	$A \equiv_L B$	(Equivalence)

where $\equiv_L$ means, as before, that both $\models_L$ and its converse hold. Conditional Introduction is valid in all trivalent logics we considered, whereas Nothing Added and Equivalence hold in **(Q)CC/TT**, but not in **(Q)DF/TT**.[19] Mandelkern requires another premise, restricted to atom-classical formulae A (i.e., such that all propositional variables have a classical value), but without restrictions on B:

$$\text{For atom-classical } A: \qquad \models_L (A \wedge \neg A) \rightarrow B \qquad \text{(Quodlibet)}$$

Quodlibet too holds in the trivalent logics we surveyed. From these four principles Mandelkern derives the following intermediate result:

$$\text{For atom-classical } A: \qquad A \models_L \neg A \rightarrow B \qquad \text{(Intermediate)}$$

Intermediate also holds in **CC/TT**, and plausibly so: If A holds then any conditional assertion with $\neg A$ as a premise is void, and thus valid in a logic with a tolerant-to-tolerant consequence relation. Intermediate is equivalent to $\neg A \models_L A \rightarrow B$, from which Mandelkern derives:

$$\text{For atom-classical } A: \qquad \neg(A \rightarrow B) \models_L A \qquad \text{(Ex Falso)}$$

The lesson Mandelkern takes from this is:

> [Intermediate] is clearly false [...]. For this conclusion entails that the falsity of $\neg A \rightarrow B$ entails the falsity of A; more succinctly (given classical

---

[19] Countermodel for Nothing Added in **(Q)DF/TT**: $v(A) = 1$, $v(B) = 1/2$, $v(C) = 0$. Countermodel for Equivalence in **(Q)DF/TT**: $v(A) = v(C) = 0$, $v(B) = 1/2$.

> negation, which is not in dispute here), the falsity of $A \to B$ entails the truth of A. (Mandelkern 2020, symbolic notation changed)

Ex Falso is definitely an unacceptable principle for a theory of indicative conditionals. As it turns out, it is invalidated in our trivalent logics, including CC/TT (Consider $v(A) = 0$.) What happened in the step from Intermediate to Ex Falso? As hinted by Mandelkern's parenthetical remark, the step is blocked in CC/TT because trivalent negation is no longer classical. In particular, TT-consequence does not obey Contraposition. This feature suggests a tradeoff: the trivalent logics of conditionals we considered validate Import-Export without restriction, and they do not fall prey to Mandelkern's reduction. However, they no longer validate Contraposition without restriction, and because CC/TT satisfies the full Deduction Theorem, the associated conditional fails Contraposition too. For indicatives as well as for counterfactuals, Contraposition is moot, however, in that regard the way in which Mandelkern's reductio is blocked here does not appear problematic.[20]

Appendix B: Proofs

In this appendix, we first prove that assumption (iii) of Gibbard's Theorem holds in DF/TT. Then, we give a syntactic proof of the mutual DF/TT-entailments of $A \to B$ and $A \supset B$ (cf. Lemma 2), in the three-sided sequent calculus for DF/TT from (Égré et al. 2021b). The remaining claims of the Lemma are then immediate. The calculus is sound and complete for DF/TT, so the proof immediately establishes the corresponding semantic claims, but we believe that a syntactic proof provides a good illustration of how one can, rather naturally, reason in trivalent logics. Similar proofs are available for the corresponding claims in CC/TT.

[20] Mandelkern does not dispute the validity of Import-Export for simple right-nested conditionals where it looks very compelling; he just thinks that Import-Export has less than general scope. Specifically, he has doubts about the application of Import-Export to compound conditionals with left-nesting, such as $A \to ((B \to C) \to D)$. Naturally, it is very difficult to find reliable empirical data or expert intuitions on how such sentences are, or should be, interpreted.

Lemma 5 *Supraclassicality holds in* DF/TT.

Proof We prove the contrapositive. Suppose $\not\models_{\mathsf{DF/TT}} A \to B$. Then there is a DF-evaluation $v : \mathsf{For}(L) \longmapsto \{0, 1/2, 1\}$ s.t. $v(A) = 1$ and $v(B) = 0$. We then claim that, in this case, then there is always a *classical* evaluation $v_{\mathsf{cl}} : \mathsf{For}(L) \longmapsto \{0, 1\}$ s.t. for every $C \in \mathsf{For}(L)$, if $v(C) = 1$, then $v_{\mathsf{cl}}(C) = 1$ and if $v(C) = 0$, then $v_{\mathsf{cl}}(C) = 0$, thus showing that $A \not\models_{\mathsf{CL}} B$. We prove this by induction on the logical complexity (cp) of A and B:

- $\mathsf{cp}(A) = \mathsf{cp}(B) = 0$. Then, $A \to B$ has the form $p \to q$, and $v(p) = 1$, $v(q) = 0$. v_{cl} is any classical evaluation which agrees with v on p and q, so clearly $p \not\models_{\mathsf{CL}} q$.
- $\mathsf{cp}(A) = m + 1$ and $\mathsf{cp}(B) = 0$. Then $A \to B$ has the form $C \to q$, for C a logically complex sentence. We assume the claim as IH up to m, and reason by cases:
  - C is $\neg D$. Then $v(\neg D) = 1$ and $v(q) = 0$, and $v(D) = 0$. By IH, then, there is a classical evaluation v_{cl} s.t. $v_{\mathsf{cl}}(D) = 0$ and $v(q) = 0$, so that $C \not\models_{\mathsf{CL}} q$.
  - C is $D \vee E$. Then $v(D \vee E) = 1$ and $v(q) = 0$. There are several cases, all similar between them, where at least one of the disjunct receives value 1:
    $v(D) = 1$ and $v(E) = 1$
    $v(D) = 1$ and $v(E) = 1/2$
    $v(D) = 1$ and $v(E) = 0$
    $v(D) = 1/2$ and $v(E) = 1$
    $v(D) = 0$ and $v(E) = 1$
    Let X be the (or "a") disjunct which receives value 1 by v. By IH, $v_{\mathsf{cl}}(X) = 1$, and then $v_{\mathsf{cl}}(D \vee E) = 1$ and $v_{\mathsf{cl}}(q) = 0$, hence $C \not\models_{\mathsf{CL}} q$
  - The case where C has the form $D \wedge E$ is similar to the above one.
  - C is $D \to E$. Then $v(D \to E) = 1$ and $v(q) = 0$, and therefore $v(D) = v(E) = 1$. By IH, then, $v_{\mathsf{cl}}(D) = v_{\mathsf{cl}}(E) = 1$, hence $C \not\models_{\mathsf{CL}} q$.

- The cases where $\mathsf{cp}(A) = 0$ and $\mathsf{cp}(B) = n+1$, and where $\mathsf{cp}(A) = m+1$ and $\mathsf{cp}(B) = n+1$ are dealt with similarly.

$\square$

Notice that, in this proof, a DF-valuation for the language including the conditional is mapped to a *classical* valuation for the same language, i.e., a classical valuation which also interprets formulae of the form $A \to B$. However, the proof does not specify how formulae of the form $A \to B$ are classically interpreted—that is, $A \to B$ may or may not be interpreted as a classical material conditional. We also note that an attempted proof along the lines of the above one would fail for CC/TT exactly because the conditions under which an indicative conditional receives value 0 under a CC-evaluation strictly exceed the conditions under which a material conditional receives value 0 under a classical evaluation, unlike in a DF-evaluation.

We now present a formal derivation of the interderivability of the material conditional and the conditional in DF/TT. Readers unfamiliar with the three-sided sequent calculus are invited to look at the sequents semantically. From a semantic point of view, a three-sided sequent of form $\Gamma \mid \Delta \mid \Sigma$ is satisfied by a three-valued valuation v provided either some element of Γ gets the value 0 under v, or some element of Σ gets the value $1/2$ or some element of Δ gets the value 1. When $\Delta = \Sigma$, this matches the definition of TT consequence.

Lemma 6 *Let $\Gamma \vdash_{\mathsf{DF/TT}} \Delta$ indicate that there is a derivation of the three-sided sequent $\Gamma \mid \Delta \mid \Delta$ in the calculus developed in Égré et al. (2021b), §§3.1–3.2. Then, for every $A, B \in \mathsf{For}(L)$:*
$A \supset B \vdash_{\mathsf{DF/TT}} A \to B$ and $A \to B \vdash_{\mathsf{DF/TT}} A \supset B$

Proof We write $\neg(A \wedge \neg B)$ for $A \supset B$, as the two formulae are definitionally equivalent in DF/TT. The following derivation establishes that $A \supset B \vdash_{\mathsf{DF/TT}} A \to B$:

$$
\dfrac{
  \dfrac{
    \dfrac{}{A \mid A, B \mid A \to B, A}\,\textsf{SRef}
  }{\varnothing \mid A \to B \mid A \to B, A}\,{\to}\text{-}1/2
  \qquad
  \dfrac{
    \dfrac{
      \dfrac{
        \dfrac{}{A, B \mid A, B \mid A}\,\textsf{SRef}
        \quad
        \dfrac{}{A, B \mid A, B \mid B}\,\textsf{SRef}
      }{A, B \mid A, B \mid A \to B}\,{\to}\text{-}1
    }{B \mid A \to B \mid A \to B}\,{\to}\text{-}1/2
  }{\varnothing \mid A \to B \mid A \to B, \neg B}\,\neg\text{-}1
}{
  \dfrac{\varnothing \mid A \to B \mid A \to B, A \wedge \neg B}{\neg(A \wedge \neg B) \mid A \to B \mid A \to B}\,\neg\text{-}0
}\,\wedge\text{-}1
$$

We now show that $A \to B \vdash_{\mathsf{DF/TT}} A \supset B$. First, let $\mathcal{D}_0$ be the following derivation:

$$
\dfrac{
  \dfrac{}{A, \neg B \mid A \mid A, A}\,\textsf{SRef}
  \quad
  \dfrac{}{A, \neg B \mid \neg B \mid A, \neg B}\,\textsf{SRef}
  \quad
  \dfrac{}{A, \neg B \mid A, \neg B \mid A}\,\textsf{SRef}
}{A, \neg B \mid A \wedge \neg B \mid A}\,\wedge\text{-}1/2
$$

Second, let $\mathcal{D}_1$ be the following derivation:

$$
\dfrac{
  \dfrac{}{A, \neg B, B \mid A \mid A}\,\textsf{SRef}
  \quad
  \dfrac{}{A, \neg B, B \mid \neg B \mid \neg B}\,\textsf{SRef}
  \quad
  \dfrac{
    \dfrac{
      \dfrac{}{A, B \mid A, B \mid B}\,\textsf{SRef}
    }{A, B \mid A, \neg B \mid B}\,\neg\text{-}1/2
  }{A, \neg B, B \mid A, \neg B \mid \varnothing}\,\neg\text{-}1
}{A, \neg B, B \mid A \wedge \neg B \mid \varnothing}\,\wedge\text{-}1/2
$$

Finally, combining $\mathcal{D}_0$ and $\mathcal{D}_1$ yields the desired result:

$$
\dfrac{
  \dfrac{
    \dfrac{
      \dfrac{
        \begin{matrix}\mathcal{D}_0 \\ A, \neg B \mid A \wedge \neg B \mid A\end{matrix}
        \qquad
        \begin{matrix}\mathcal{D}_1 \\ A, \neg B, B \mid A \wedge \neg B \mid \varnothing\end{matrix}
      }{A, \neg B, A \to B \mid A \wedge \neg B \mid \varnothing}\,{\to}\text{-}0
    }{A \wedge \neg B, A \to B \mid A \wedge \neg B \mid \varnothing}\,\wedge\text{-}0
  }{A \to B \mid A \wedge \neg B \mid \neg(A \wedge \neg B)}\,\neg\text{-}1
}{A \to B \mid \neg(A \wedge \neg B) \mid \neg(A \wedge \neg B)}\,\neg\text{-}1/2
$$

□

Appendix C: Gibbardian Collapse Without Left Logical Equivalence

Khoo and Mandelkern (2019, 489) prove Gibbard's collapse result using Reasoning by Cases. They do not use Left Logical Equivalence (as in our reconstruction of Gibbard's original proof) and explicitly refer to principles (i)–(iii) only (i.e., Import-Export, Stronger-than-Material, and Supraclassicality). However, like Gibbard, they actually make use of more assumptions, in particular (v): the classicality of $\supset$. Their proof can be formalized thus:

Theorem 7 *Let L be a reflexive, monotonic, and transitive consequence relation, with $\vee$ satisfying Reasoning by Cases. Then if (i), (ii), (iii), and (v) hold in L, $\supset$ entails $\rightarrow$, that is, for any $A, B \in \mathsf{For}(L)$, $A \supset B \models_L A \rightarrow B$.*

1. $\neg A \wedge A \models_{\mathsf{CL}} B$, classical logic
2. $\models_L (\neg A \wedge A) \rightarrow B$, by 1 and (iii)
3. $\models_L \neg A \rightarrow (A \rightarrow B)$, by 2 and (i)
4. $\neg A \rightarrow (A \rightarrow B) \models_L \neg A \supset (A \rightarrow B)$, by (ii)
5. $\models_L \neg A \supset (A \rightarrow B)$, by 3, 4 and Transitivity
6. $\neg A \models_L \neg A \supset (A \rightarrow B)$, by 5 and Monotonicity
7. $\neg A \models_L \neg A$, by Reflexivity
8. $\neg A \models_L A \rightarrow B$, by 6, 7 and (v), using (meta) Modus Ponens for $\supset$
9. $B \wedge A \models_{\mathsf{CL}} B$, classical logic
10. $\models_L (B \wedge A) \rightarrow B$, by 9 and (iii)
11. $\models_L B \rightarrow (A \rightarrow B)$, by 10 and (i)
12. $B \rightarrow (A \rightarrow B) \models_L B \supset (A \rightarrow B)$, by (ii)
13. $\models_L B \supset (A \rightarrow B)$, by 11, 12, and Transitivity
14. $B \models_L B \supset (A \rightarrow B)$, by 13 and Monotonicity
15. $B \models_L B$, by Reflexivity
16. $B \models_L A \rightarrow B$, by 14, 15, (v), using (meta) Modus Ponens
17. $\neg A \vee B \models_L A \rightarrow B$, by 8, 16 and Reasoning by Cases
18. $A \supset B \models_L A \rightarrow B$, by 17 and (v)

This version does not use the replacement principle (iv) of Gibbard's original proof, making it particularly interesting, notably in relation to DF/TT. Indeed, Reasoning by Cases is valid in DF/TT and CC/TT, as are structural assumptions on logical consequence. We know that CC/TT fails Supraclassicality and so steps 2 and 10 of the proof are blocked. Interestingly, however, all steps of the proof here are *sound in* DF/TT. Although principle (v) does not hold of $\supset$ in full generality in DF/TT, all instances of (v) are sound in this case, unlike in Gibbard's original proof. Readers may observe that the proof of $A \supset B \vdash_{\mathsf{DF/TT}} A \rightarrow B$ produced in the sequent-system of Appendix B also mirrors Reasoning by Cases (see Lemma 6): on the third line from the root of the tree, the left branch of the derivation tree actually establishes that $\neg A \vdash_{\mathsf{DF/TT}} A \rightarrow B$, while the right branch establishes that $B \vdash_{\mathsf{DF/TT}} A \rightarrow B$.

References

Adams, Ernest W. (1965). The logic of conditionals. *Inquiry 8*, 166–197.

Adams, Ernest W. (1975). *The logic of conditionals*. Dordrecht: Reidel.

Arló-Costa, Horacio (2001). Bayesian epistemology and epistemic conditionals: On the status of the Export-Import laws. *The Journal of Philosophy 98*, 555–593.

Baratgin, Jean, David Over, and Guy Politzer (2013). Uncertainty and the de Finetti tables. *Thinking & Reasoning 19*, 308–328.

Belnap, Nuel D. (1970). Conditional assertion and restricted quantification. *Noûs*, 1–12.

Belnap, Nuel D. (1973). Restricted quantification and conditional assertion. In Hugues Leblanc (ed.), *Truth, syntax and modality*, Volume 68, pp. 48–75. Studies in Logic and the Foundations of Mathematics: Elsevier.

Bradley, Richard (2002). Indicative conditionals. *Erkenntnis 56*, 345–378.

Calabrese, Philip (2002). Deduction with uncertain conditionals. *Information Sciences 147*, 143–191.

Cantwell, John (2008). The logic of conditional negation. *Notre Dame Journal of Formal Logic 49*, 245–260.

Cobreros, Pablo, Paul Égré, David Ripley, and Robert van Rooij (2012). Tolerant, classical, strict. *Journal of Philosophical Logic 41*, 347–385.

Cooper, William S. (1968). The propositional logic of ordinary discourse. *Inquiry 11*, 295–320.

de Finetti, Bruno (1936). La logique de la probabilité. In *Actes du congrès international de philosophie scientifique*, Volume 4, Paris, pp. 1–9. Hermann Editeurs.

Dicher, Bogdan and Francesco Paoli (2019). ST, LP and tolerant metainferences. In Can Baskent and Thomas Ferguson (eds.), *Graham Priest on dialetheism and paraconsistency*, pp. 383–407. Springer.

Dubois, Didier, and Henri Prade (1994). Conditional objects as nonmonotonic consequence relationships. *IEEE Transactions on Systems, Man, and Cybernetics 24*, 1724–1740.

Égré, Paul, Lorenzo Rossi, and Jan Sprenger (2021a). De Finettian logics of indicative conditionals. Part I: Trivalent semantics and validity. *Journal of Philosophical Logic 50*, 187–213.

Égré, Paul, Lorenzo Rossi, and Jan Sprenger (2021b). De Finettian logics of indicative conditionals. Part II: Proof theory and algebraic semantics. *Journal of Philosophical Logic 50*, 215–247.

Farrell, Robert J. (1979). Material implication, confirmation, and counterfactuals. *Notre Dame Journal of Formal Logic 20*, 383–394.

Farrell, Robert J. (1986). Implication and presupposition. *Notre Dame Journal of Formal Logic 27*, 51–61.

Fine, Kit (2012). Counterfactuals without possible worlds. *The Journal of Philosophy 109*, 221–246.

Fitelson, Branden (2013). Gibbard's collapse theorem for the indicative conditional: An axiomatic approach. In Maria Paola Bonacina and Mark E. Stickel (eds.), *Automated reasoning and mathematics: Essays in memory of William W. McCune*, pp. 181–188. Beriln: Springer.

Gibbard, Allan (1980). Two recent theories of conditionals. In William Harper (ed.), *Ifs*, pp. 211–247. Springer.

Grice, H. Paul (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Jackson, Frank (1979). On assertion and indicative conditionals. *Philosophical Review 88*, 565–589.

Kaufmann, Magdalena, and Stefan Kaufmann (2015). Conditionals and modality. In Shalom Lappin and Chris Fox (eds.), *The handbook of contemporary semantic theory*, pp. 237–270. Wiley Online Library.

Khoo, Justin, and Matthew Mandelkern (2019). Triviality results and the relationship between logical and natural languages. *Mind 128*, 485–526.

Kratzer, Angelika (2012). *Modals and conditionals: New and revised perspectives*, Volume 36. Oxford University Press.

Kratzer, Angelika (2015). Chasing hook: Quantified indicative conditionals. In Lee Walters and John Hawthorne (eds.), *Conditionals, probability, and paradox: Themes from the philosophy of Dorothy Edgington*, pp. 40–56. Oxford University Press Oxford.

Lassiter, Daniel (2020). What we can learn from how trivalent conditionals avoid triviality. *Inquiry*, *63*(9–10), 1087–1114.

Lewis, David (1973). *Counterfactuals*. Oxford: Basil Blackwell.

Lewis, David (1976). Probabilities of conditionals and conditional probabilities. *Philosophical Review 85*, 297–315.

MacColl, Hugh (1908). 'If' and 'Imply'. *Mind 17*, 453–455.

Mandelkern, Matthew (2019). Crises of identity. In Julian J. Schlöder, Dean McHugh, and Floris Roelofsen (eds.), *Proceedings of the 22nd Amsterdam Colloquium*, pp. 279–288.

Mandelkern, Matthew (2020). Import-Export and 'and'. *Philosophy and Phenomenological Research 100*, 118–135.

McDermott, Michael (1996). On the truth conditions of certain 'if'-sentences. *The Philosophical Review 105*, 1–37.

McGee, Vann (1989). Conditional probabilities and compounds of conditionals. *The Philosophical Review 98*, 485–541.

Milne, Peter (1997). Bruno de Finetti and the logic of conditional events. *The British Journal for the Philosophy of Science 48*, 195–232.

Nute, Donald (1980). *Topics in conditional logic*, Volume 20. Springer Science & Business Media.

Olkhovikov, Grigory (2002 [2016]). On a new three-valued paraconsistent logic. *IFCoLog Journal of Logic and Applications 3*, 317–334. Originally published in Russian in 2002, in *Logic of Law and Tolerance*, pp. 96–113, Yekaterinburg. English translation by T. Ferguson.

Quine, W. V. O. (1950). *Methods of logic*. Cambridge, MA: Harvard University Press.

Reichenbach, Hans (1935). *Wahrscheinlichkeitslehre*. Leiden: Sijthoff.

Ripley, D. (2013). Paradoxes and failures of cut. *Australasian Journal of Philosophy 91*, 139–164.

Sanfilippo, Giuseppe, Angelo Gilio, David E Over, and Niki Pfeifer (2020). Probabilities of conditionals and previsions of iterated conditionals. *International Journal of Approximate Reasoning 121*, 150–173.

Stalnaker, Robert (1968). A Theory of Conditionals. In *Studies in Logical Theory: American Philosophical Quarterly Monograph Series, No. 2*. Oxford: Blackwell.

Stalnaker, Robert (1970). Probability and conditionals. *Philosophy of Science 37*, 64–80.

Unterhuber, Matthias and Gerhard Schurz (2014). Completeness and correspondence in Chellas-Segerberg semantics. *Studia Logica 102*, 891–911.

van Wijnbergen-Huitink, Janneke, Shira Elqayam, and David E. Over (2015). The probability of iterated conditionals. *Cognitive Science 39*, 788–803.

The Logic and Pragmatics of Conditionals Under Uncertainty: A Mental Probability Logic Perspective

Niki Pfeifer

1 Introduction

In the last two decades, a new research paradigm has emerged within the psychology of reasoning. It is characterised by the use of probability theory instead of bivalent truth-functional logic[1] as the gold standard

[1] In the psychology of reasoning, if not specified otherwise, "logic" refers traditionally to standard bivalent truth-functional logic, i.e., either to the standard propositional calculus, or—in the context of quantifiers—to the predicate calculus. For extensions of logic or alternative logics see, e.g., the various volumes of the *Handbook of Philosophical Logic*, edited by Dov M. Gabbay and Franz Guenthner. For a recent overview, specifically on conditional logic, see, e.g., Égré and Rott (2021).

N. Pfeifer (✉)
Department of Philosophy, University of Regensburg, Regensburg, Germany
e-mail: niki.pfeifer@ur.de
URL: https://go.ur.de/np

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_4

for evaluating the rationality of human inference (see, e.g., Elqayam et al. 2020; Evans and Over 2004; Oaksford and Chater 2007, 2020; Pfeifer and Kleiter 2005; Pfeifer 2013a). The new paradigm is driven by the observation that human reasoning about the truth values of conditionals (i.e., *if A, then C* constructions) violates the bivalent semantics implied by the material conditional $A \supset C$, which is the conditional of classical (bivalent truth-functional) logic. The material conditional is at the core of this logic: it is closely related to the logical consequence relation (via the deduction theorem) and to logical validity. However, for formalising common sense arguments, the material conditional is inappropriate. One may even doubt that the material conditional is a genuine conditional since it is logically equivalent to the disjunction *not-A or C*, i.e., $(A \supset C) \equiv (\neg A \vee C)$ is a tautology. The material conditional is *false* if $A \wedge \neg C$ is true and *true* in each of the other three possible truth table cases.

It has already been observed in truth table tasks that people's judgements about the truth values of *if A, then C* agree with the first two cases of the truth table of $A \supset C$ (i.e., $A \wedge C$ and $A \wedge \neg C$, respectively) but diverge in the third and fourth cases (i.e., $\neg A \wedge C$ and $\neg A \wedge \neg C$, respectively). People judge false antecedent cases to be "irrelevant", "undetermined", or in similar terms, but not—as implied by the material conditional—as true (see, e.g., Wason 1966; Evans 1982). The corresponding response pattern was pejoratively called "defective truth table pattern". With the probabilistic turn in the psychology of reasoning, however, this response pattern was rehabilitated. Specifically, this response pattern corresponds to the semantics of the conditional object which is measured by the conditional probability: namely, the conditional event $C|A$, which is what is measured by the conditional probability, $p(C|A)$. The conditional event $C|A$ is *true* when $A \wedge C$ is true, *false* when $A \wedge \neg C$ is true, and *void* (or undetermined) when $\neg A$ is true (see de Finetti 1936/1995, who used the term *triévénement* in the *Actes du Congrès International de Philosophie Scientifique* [Sorbonne, Paris] in 1935 but preferred later the term *conditional event*). Thus, while human reasoning about the truth conditions of conditionals appears to be irrational by the standards of bivalent truth functional logic, it is rational by the standards of probability theory. Instead of "defective truth table pattern", the term "de Finetti table" became popular in honour of Bruno de Finetti's work on conditional events.

Among various probabilistic approaches to the psychology of reasoning, I advocate Mental Probability Logic (MPL; see, e.g., Pfeifer 2006, 2013a, 2014, 2021; Pfeifer and Kleiter 2005, 2009). The research focus of MPL is on reasoning competence. Descriptive and normative research questions are how people solve (and how they should solve, respectively) epistemic reasoning and argumentation problems. The key assumptions are that everyday life reasoning is based on incomplete information and on uncertain premises. Moreover, conclusions are defeasible. People interpret the uncertainty of conditionals (*if A, then C*) by conditional probabilities ($p(C|A)$). The computational level problem description (in the sense of Marr 1982) is first to reason to an interpretation of the premises and second to draw a rational conclusion. This process requires us to

1. make any implicit assumptions and logical relations explicit,
2. assign uncertainty to the premises, and
3. transmit the uncertainty from the premises to the conclusion.

The long term goal of MPL is the development of a unified, normatively, and descriptively adequate theory of human reasoning under uncertainty.

The underlying rationality framework of MPL is coherence-based probability logic (see, e.g., Coletti and Scozzafava 2002; Gilio 2002; Gilio et al. 2016; Pfeifer and Kleiter 2009). "Coherence" refers to the subjective interpretation of probability which originated with Bruno de Finetti (see, e.g., de Finetti 1931, 1970/1974). The coherence principle was generalised, for example, to conditional probability and to imprecise probability (see, e.g., Berti et al. 1998; Biazzo et al. 2005; Capotorti et al. 2003, 2007; Coletti and Scozzafava 2002; Gilio 1990, 1996; Gilio and Sanfilippo 2013; Gilio et al. 2016; Holzer 1985; Lad 1996; Petturiti and Vantaggi 2017; Regazzini 1985; Walley et al. 2004) as well as to nested conditionals in terms of previsions in conditional random quantities (see, e.g., Gilio and Sanfilippo 2014, 2019, 2021; Gilio et al. 2020; Sanfilippo et al. 2020, 2017). In the framework of coherence, probability is interpreted as a *degree of belief*. Of course, degrees of belief are naturally of more interest to psychologists compared, for example, to frequentistic interpretations of probability, since psychologists are interested in belief formation and reasoning about beliefs and not in the measurement of probability in the ontological sense of an objective quantity belonging to the outside world. In betting terms, a degree of belief (i.e., a subjective probability assessment) is *coherent* if and only if in any finite combination of n bets, it cannot happen that—after discarding the cases where the bet is called off—the remaining values of the random gain are all positive or all negative. Intuitively, coherence requires that *no Dutch Book* should be constructible. For an example of a Dutch Book consider that you pay 100 EUR for receiving either 50 EUR (when you "win") or lose the 100 EUR (when you lose): after discarding the case of calling off the bet, you will lose money for sure.

Some approaches (e.g., Kolmogoroff 1933), require a complete algebra and define conditional probability by the fraction of the probability of the conjunction of the conditioned and the conditioning event (i.e., $p(A \wedge C)$), divided by the probability of the conditioning event (i.e., $p(A)$):

$$p(C|A) =_{\text{def.}} \frac{p(A \wedge C)}{p(A)}, \quad \text{provided that} \quad p(A) > 0.$$

In the coherence approach, however, conditional probability is *primitive* and properly managed even if $p(A) = 0$, while the above-mentioned definition in terms of a fraction is undefined in case the denominator ($p(A)$) is equal to zero. The ability to deal with zero-probability antecedents is not only important from theoretical points of views (e.g., they allow for reducing the complexity of the probabilistic inference). Practical implications are that different results are obtained depending on whether zero-probability antecedents are allowed (and properly managed!) or excluded by assumption. Two illustrative examples are given below.

Coherence allows for imprecise probabilities, which is relevant in cases under incomplete probabilistic knowledge, i.e., when only interval-valued probabilities are available. In everyday contexts, we are usually faced with partial and incomplete knowledge. Thus, premises may often be evaluated by lower and/or upper probability bounds only, instead of (precise) point-valued probabilities. Moreover, even if you start reasoning from a premise set consisting of precise-probability assessments, you will often end up with concluding interval-valued probability assessments. In a further reasoning step, such conclusions can function as a(n imprecise) premise of a further argument. For an example, consider the following instance of the probabilistic modus ponens[2]:

p(I take the train at five) $= .7$
p(I am home in time for dinner | I take the train at five) $= .9$
Therefore, p(I am home in time for dinner) $\in [.63, .93]$ is coherent.[3]

The conclusion, p(I am home in time for dinner) $\in [.63, .93]$, can then function in a further argument as a premise which is, naturally, imprecise (if, of course, coherence does not force it to a point-value in the light of other premises of the new argument).

Probability logic refers to the study of uncertain argument forms, where the assigned uncertainty of the premises is transmitted to the conclusion. As the probability propagation rules are derived deductively within coherence-based probability theory, this transmission of the uncertainties is deductive. Hence, the consequence relation between the premise set and the conclusion remains deductive in coherence-based probability logic.

In this chapter we will use coherence-based probability logic as a rationality framework for studying selected semantic and alleged pragmatic phenomena. Specifically, we will look at a paradox of the material conditional and the role of zero-probability antecedents in the interpretation of

---

[2] The probability propagation rule for the probabilistic modus ponens is as follows: from $p(A) = x$ and $p(C|A) = y$ infer $p(C) \in [xy, xy + 1 - x]$ is coherent (see, e.g., Pfeifer and Kleiter 2006, 2009).

[3] The values in the square brackets $[x, y]$ indicate the lower (x) and the upper probability bound (y). Here and in what follows the given bounds are always understood as the tightest possible coherent bounds, i.e., all assessments inside these bounds are coherent, whereas assessments outside these bounds are not coherent. If the tightest bounds coincide with the unit interval for all probability assessments of the premises, the corresponding argument form is called *probabilistically non-informative*, since the premise probabilities do not constrain the coherent bounds on the conclusion.

conditionals. What is semantics and what is pragmatics? Carnap (1948) gives a classic definition:

> In an application of language, we may distinguish three chief factors: the speaker, the expression uttered, and the designatum of the expression, i.e., that to which the speaker intends to refer by the expression [...] An investigation of a language belongs to *pragmatics* if explicit reference to a speaker is made; it belongs to *semantics* if designata but not speakers are referred to; it belongs to *syntax* if neither speakers nor designata but only expressions are dealt with. (Carnap 1948, p. 8)

Some psychologists of reasoning appeal to pragmatics when predictions from semantic considerations lead to counterintuitive consequences; likewise, when normative predictions turn out to be unreliable descriptions of actual reasoning behaviour. As an example, consider the following paradox of the material conditional:

C; therefore, *if* A*, then* C.

This inference is called a paradox of the material conditional, since if the conditional is interpreted by the material one (i.e., by $A \supset C$), the resulting argument

C; therefore $A \supset C$

is logically valid. Thus, classical truth-functional semantics suggests that it is impossible that the premise C is true and the conclusion $A \supset C$ is false; hence, normatively, the inference should hold. However, it is easy to find instantiations for A and C, where this inference appears counterintuitive: e.g., according to this argument form, from *Ludwig is rich* it follows that *if Ludwig donates all his money, then Ludwig is rich*. To block such paradoxes, for example, Johnson-Laird and Byrne (2002) use a principle of "pragmatic modulation", which "can modulate the core interpretation of a conditional" (p. 659). The danger of such pragmatic considerations is that anything can be explained away by pragmatics.

From a probabilistic point of view, the above-mentioned paradox is blocked on purely semantic grounds since the inference is probabilistically non-informative (Pfeifer 2014). I.e., for all probability values x:

$p(C) = x$; therefore $p(C|A) \in [0, 1]$ is coherent.[4]

If an inference is probabilistically non-informative, it does not allow for constraining the conclusion probability, as all you can infer is the unit interval. Hence, probabilistically non-informative inferences are generally not probabilistically valid (p-valid; see Pfeifer and Kleiter 2009). Therefore, if an inference is probabilistically non-informative, the corresponding inference pattern is blocked.

The probabilistic analysis of the previous paradox allows us also to illustrate the importance of zero-probability antecedents. When the conditional probability is defined by the above-mentioned fraction, then $p(C|A) = 1$ (or undefined if $p(A) = 0$) in the special case when $p(C) = 1$. It appears strange to say that the inference is probabilistically non-informative for all probability assessments except for probability one in the premise: why should a premise-probability infinitesimally close to one block the paradox but not a premise-probability equal to one? To explain away this unwanted consequence Bonnefon and Politzer (2010) proposed to use pragmatic considerations. In the coherence approach, however, the probability propagation rule states that the coherence of $p(C|A) \in [0, 1]$ follows from $p(C) = x$, even if $p(C) = 1$ (for a proof see Pfeifer 2014). Thus, within coherence-based probability logic, by exploiting and properly managing zero-probability antecedents, this paradox is blocked on semantic grounds without the need of employing pragmatic considerations. Experimental data suggest that people understand that this paradox is probabilistically non-informative even if the premise-probability is equal to one (Pfeifer and Kleiter 2011).

Let me stress that I do not want to argue that pragmatic considerations are never useful. Rather, I argue that if we can deal with the problem

[4] If not otherwise stated, A and C are logically independent, i.e., they are contingent and any logical relation needs to be made explicit.

at hand within semantics alone, we do not need to use pragmatic considerations to explain away unwanted consequences.

Another example, which highlights the importance of zero-probability antecedents, is transitivity. Probabilistic transitivity is probabilistically non-informative (see, e.g., Gilio et al. 2016); for all probability values x and y:

$p(B|A) = x$
$p(C|B) = y$
Therefore, $p(C|A) \in [0, 1]$ is coherent.

Blocking (unrestricted) transitivity is a good thing, since validity of transitivity would imply antecedent strengthening, which is a psychologically implausible inference scheme.[5] If, however, the conditioning events are assumed to be positive, then, in the extreme case when $x = y = 1$, the probability of the conclusion of transitivity jumps to one. If conditioning events are not assumed to be positive (i.e., zero-probability antecedents are left as a possibility), probabilistic transitivity is probabilistically non-informative even if $x = y = 1$ within the framework of coherence. Similar to the above case of the paradox of the material conditional, approaches which have to assume positive conditioning events to avoid fractions over zero, end up with probability one in the conclusion in the special case when $x = y = 1$.

Having considered the relevance of zero-probability antecedents in the context of probabilistic argument forms, let us now turn to the semantics of conditionals. Frank P. Ramsey's (1929/1994) famous footnote inspired qualitative and quantitative accounts of conditionals:

[5] Antecedent strengthening would, for example, validate inferences like *if Tweety is a bird, then Tweety can fly*, therefore, *if Tweety is a bird and a penguin, then Tweety can fly*. Therefore, while first order logic is monotonic and hence validates antecedent strengthening, we do not want to validate (unrestricted) monotonicity. Intuitively, monotonicity means that you cannot simply retract conclusions (e.g., Tweety can fly) in the light of new evidence (that Tweety is a penguin). Coherence-based probability logic is nonmonotonic and allows in principle to revise conclusions when new premises are available.

> If two people are arguing 'If A will C?' and are both in doubt as to A, they are *adding A hypothetically to their stock of knowledge and arguing on that basis about C*; so that in a sense *'If A, C' and 'If A, $\neg C$' are contradictories.* We can say they are fixing their degrees of belief in C given A. *If A turns out false, these degrees of belief are rendered void*. (Ramsey 1929/1994, p. 155, adjusted notation; my emphasis)

Let me highlight three important points in this quote.

First, Ramsey's footnote makes an intuitive justification, why it does not make sense to conditionalise on contradictions, i.e., why the conditioning event must not be equivalent to $\perp$. Ramsey talks about adding the antecedent A (or the conditioning event) hypothetically to a stock of knowledge. On traditional accounts of knowledge, knowledge implies truth. Even if it is unknown whether A is true or false, it must *in principle* be possible that A is true. However, if A is a contradiction (i.e., $A \equiv \perp$), then A is by definition false under all interpretations. Thus, $\perp$ can never, not even hypothetically, be consistently assumed to be true (as it is impossible). However, for an $A \not\equiv \perp$ it could very well be that $p(A) = 0$, which can be added to the stock of knowledge. In other words, conditionalising on a non-contradictory conditional event A with probability zero makes sense. As mentioned above, zero-probability antecedents are properly managed in the coherence approach to probability. Note that of course if $A \equiv \perp$, then $p(A) = 0$, but the converse does not hold: $p(A) = 0$ does not imply that $A \equiv \perp$. Consider, for example, the following coherent assessment on a coin toss: $p(\text{heads}) = .5$, $p(\text{tails}) = .5$, and $p(\text{coin lands standing on its edge}) = 0$. The latter event with probability zero is of course contingent and not a logical contradiction. In this example it is also natural to conditionalise on a zero-probability antecedent, like $p(\text{heads in the second toss} \mid \text{coin lands standing on its edge in the first toss}) = .5$.

Second, the idea that $A \to C$ and $A \to \neg C$ (where "$\to$" denotes an uninterpreted conditional) cannot both be true is intuitively plausible and a key principle in connexive logic (for an overview on connexive logic see, e.g., Wansing 2020), namely Abelard's First Principle, i.e.,

$\neg((A \to C) \wedge (A \to \neg C))$. However, under the material conditional interpretation, Abelard's First Principle, i.e., $\neg((A \supset C) \wedge (A \supset \neg C))$ is contingent, since it is logically equivalent to A. Abelard's First Principle is thus not a theorem of binary truth functional logic, but can be validated in probability logic (Pfeifer and Sanfilippo 2021). Furthermore, experimental data suggest that people think that the negation of a self-contradictory conditional is true, i.e., the connexive principle called Aristotle's thesis holds. This is invalid in binary truth-functional logic, but holds in probability logic (see, e.g., Pfeifer 2012, 2021).

Third, Ramsey's term "void" can be understood within coherence-based probability logic in a qualitative and in a quantitative sense. Under the qualitative interpretation of void, consider the trivalent conditional event $C|A$, which is *void* if $\neg A$ is true, *true* if $A \wedge C$ is true, and *false* if $A \wedge \neg C$ is true. Under the quantitative interpretation, consider the corresponding conditional probability assessment, $p(C|A) = x$. In betting terms, the probability assessment $p(C|A) = x$ (for $x \in [0, 1]$) means that, for every real number s (for stakes $s \neq 0$), you agree

$$\text{to pay the amount } sx \text{ in order to receive } \begin{cases} s, & \text{if } A \wedge C \text{ is true,} \\ 0, & \text{if } A \wedge \neg C \text{ is true,} \\ sx, & \text{if } \neg A \text{ is true.} \end{cases}$$

In this case, when s is positive, "you" refers to a person who is paying the price sx for the bet in order to receive, according to the conditions specified above, s, 0, or sx; when s is negative, "you" refers to a person who agrees to receive the amount $-sx$ in order to pay $-s$, 0, or $-sx$, when $A \wedge C$ is true, $A \wedge \neg C$ is true, and when $\neg A$ is true, respectively. Intuitively, this means that the person who is buying a betting ticket (the bettor) and the person who is selling a betting ticket (the bank) can in principle change roles. You (in the role of the bettor) *win*, or *lose*, or *get your money back*, when $A \wedge C$ is true, or $A \wedge \neg C$ is true, or when $\neg A$ is true, respectively. You (in the role of the bank) *lose*, or *win*, or *return the money paid*, when $A \wedge C$ is true, or $A \wedge \neg C$ is true, or when $\neg A$ is true, respectively.

Let us now turn to an experiment on a generalised probabilistic truth table task, which involves imprecise probabilities, to investigate how people interpret conditionals.[6] The experiment uses the probabilistic truth table task material for incomplete probabilistic knowledge as in Pfeifer (2013a) but measures additionally how sure participants are in the correctness of their solutions (for similar tasks involving complete probabilistic knowledge see Fugard et al. 2009, 2011, 2011; Kleiter et al. 2018). Moreover, two additional tasks are used to investigate whether previously observed low frequencies of material conditional responses may be due to difficulties in calculating the respective values. Contrary to the current sample, the sample reported by Pfeifer (2013a) involved naive participants in the sense that they did not receive a special training in logic or statistics.

2 Method

2.1 Participants and Procedure

Twenty Carnegie Mellon University students (ten female and ten male) in the age between 18 and 40 years old (mean: 22.95) participated in the experiment. The students received \$10 for their participation. Most participants had a formal background like mathematics, logic, probability, statistics, or decision theory. They were administered the paper and pencil tasks in the Laboratory for Empirical Approaches to Philosophy (LEAP) at the Department of Philosophy (CMU). To ensure optimal experimental conditions, the participants were tested individually and not in groups. The experimenter was present in the background and asked each participant to re-read the instructions in case of questions.

2.2 Method

The instructions informed the participants that they were taking part in a study on reasoning about uncertainty, that they should read the instructions carefully, and that they could ask the experimenter if they

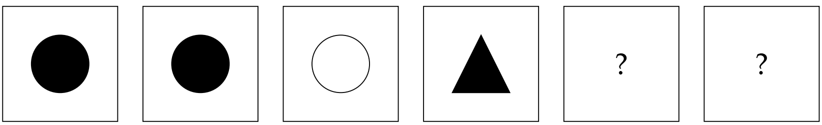
[6] Originally, the probabilistic truth table tasks were formulated in terms of complete probabilistic knowledge (Evans et al. 2003; Oberauer and Wilhelm 2003).

had questions. They were given as much time as they needed to solve the tasks and were also told that most of the participants needed about half an hour. Four introductory examples explained the task material as follows: Example 1 explained the response format for point-valued responses and its meaning on a scale. Example 2 explained that question marks indicated covered sides. In such cases of incomplete probabilistic knowledge, participants were told that the response could be given in terms of *at least* and *at most*. Example 3 explained how to respond with a point-value in the response format which also allows for interval-valued responses (e.g., *at least* 2 out of 6 and *at most* 2 out of 6). Example 5 introduced the question "How confident are you that the answer you just gave is correct?", which asks for a mark on the confidence scale.

After 23 target tasks, participants were asked to rate the overall clearness of the instructions and the difficulty of the tasks.

The target tasks were formulated on separate pages. Here is a sample task:

Illustrated here are all sides of a six-sided die. The sides have two properties: a color (*black* or *white*) and a shape (*circle*, *triangle*, or *square*). Question marks indicate covered sides.



Imagine that this die is placed in a cup. Then the cup is randomly shaken. Finally, the cup is placed on the table so that you cannot see what side of the die shows up.

Question: How sure can you be that the following sentence holds?

If the side facing up shows a *circle*, **then** the side shows *white*.

The depicted/covered sides of the die contain the information of the premise set of the probabilistic truth table task. For example, they allow the construction of the following (imprecise) probabilities of the truth table cases[7]:

$$p(\text{circle} \wedge \text{white}) \in \left[\frac{1}{6}, \frac{3}{6}\right] = \left[\frac{1}{6}, \frac{1}{2}\right]$$

$$p(\text{circle} \wedge \neg\text{white}) \in \left[\frac{2}{6}, \frac{4}{6}\right] = \left[\frac{1}{3}, \frac{2}{3}\right]$$

$$p(\neg\text{circle} \wedge \text{white}) \in \left[\frac{0}{6}, \frac{2}{6}\right] = \left[0, \frac{1}{3}\right]$$

$$p(\neg\text{circle} \wedge \neg\text{white}) \in \left[\frac{1}{6}, \frac{3}{6}\right] = \left[\frac{1}{6}, \frac{1}{2}\right]$$

Because only intervals can be derived from the above-mentioned premises, we are in a situation of incomplete probabilistic knowledge. The shapes and colours (black/white) which were depicted in the premises and mentioned in the conditional in the conclusion as well as the number of question marks were manipulated systematically, such that the predictions concerning the target interpretations differed for each task. The three target interpretations are (i) the conditional probability interpretation, (ii) the probability of the conjunction, and (iii) the probability of the material conditional. For the above-mentioned die, the predictions according to the target interpretations are as follows:

- **(i) Conditional probability:** $p(\text{white}|\text{circle}) = p\left(\frac{\text{circle} \wedge \text{white}}{\text{circle}}\right) \in \left[\frac{1}{5}, \frac{3}{5}\right]$
- **(ii) Probability of the conjunction:** $p(\text{circle} \wedge \text{white}) \in \left[\frac{1}{6}, \frac{3}{6}\right] = \left[\frac{1}{6}, \frac{1}{2}\right]$

[7] Of course, coherence requires that the probabilities of all truth table cases must add up to one. Therefore, when a coherent (point or interval-valued) probability assessment on $n-1$ truth table cases is given, the coherent probability assessment on the remaining truth table case is determined.

(iii) Probability of the material conditional: $p(\text{circle} \supset \text{white}) = p(\text{circle} \wedge \text{white}) + p(\neg\text{circle} \wedge \text{white}) + p(\neg\text{circle} \wedge \neg\, \text{white}) = 1 - p(\text{circle} \wedge \neg\text{white}) \in \left[\frac{2}{6}, \frac{4}{6}\right] = \left[\frac{1}{3}, \frac{2}{3}\right]$

Predictions concerning other interpretations are (iv) the probability of the material equivalence (or material bivalence) and (v) the probability of the biconditional event interpretation. The latter was proposed, for example, in a developmental psychological context (see, e.g., Gauffroy and Barrouillet 2009):

(iv) Probability of the material equivalence: $p(\text{circle} \equiv \text{white}) = p(\text{circle} \wedge \text{white}) + p(\neg\text{circle} \wedge \neg\, \text{white}) \in \left[\frac{2}{6}, \frac{4}{6}\right] = \left[\frac{1}{3}, \frac{2}{3}\right]$

(v) Probability of the biconditional event: $p(\text{white} \mid\mid \text{circle}) = p\left(\frac{\text{circle} \wedge \text{white}}{\text{circle} \vee \text{white}}\right) \in \left[\frac{1}{5}, \frac{3}{5}\right]$

The latter two interpretations ((iv) and (v)) overlap with target interpretations ((i)–(iii)) in the current example. However, we will see later that—depending on the configuration of the premises—these two biconditional interpretations do not overlap with the same target interpretations in all tasks. This will allow us to differentiate the target interpretations from other interpretations.

Of course, if instead of question marks, the figures behind the covered sides were known, we would be in a situation of complete probabilistic knowledge, which would result in point-valued assessments in the truth table cases. Then, the predictions concerning the probability of the conditional would be point-valued.

Probabilistic truth table tasks under incomplete probabilistic knowledge allow also for investigating three interpretations, which were proposed by Pfeifer and Tulkki (2017a). They can be seen as a "halfway conditional event strategy", as they are characterised by ignoring covered sides. Specifically, when faced with covered sides, some people may take a mental shortcut by computing the conditional probability based on the visible sides only (thereby ignoring possible cases marked by the question mark). This may impose a lower burden on the cognitive computational load compared to computing the best possible bounds on the conditional probabilities, where the possibilities entailed by the covered sides have to be considered. Such a strategy can be used for (vi) the lower bound only (indicated by $\bar{l}$), (vii) the upper bound only ($\overline{u}$), or (viii) for both bounds together ($\overline{lu}$). For the above-mentioned premises, the corresponding predictions are:

(vi) Conditional probability $\bar{l}$: $p(\text{white}|\text{circle})_{\bar{l}} \in \left[\frac{1}{3}, \frac{3}{5}\right]$
(vii) Conditional probability $\overline{u}$: $p(\text{white}|\text{circle})_{\overline{u}} \in \left[\frac{1}{5}, \frac{1}{3}\right]$
(viii) Conditional probability $\overline{lu}$: $p(\text{white}|\text{circle})_{\overline{lu}} = \frac{1}{3}$

The previous three strategies ((vi)–(viii)) simplify the tasks by reducing them partially (in case of $\bar{l}$ or $\overline{u}$) or completely (in case of $\overline{lu}$) to tasks under complete probabilistic knowledge. In principle also other halfway response strategies can be constructed (for example, based on the material conditional, conjunction, or the biconditional interpretation). However, we will see that the strategies (i)–(viii) already explain roughly 90% of the data. Therefore, other strategies can be neglected.

The response format was designed to keep mental arithmetics to a minimum and was depicted as follows:

Answer:

at least								*at most*							
	☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6		☐ 0	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6
out of		☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6	out of		☐ 1	☐ 2	☐ 3	☐ 4	☐ 5	☐ 6

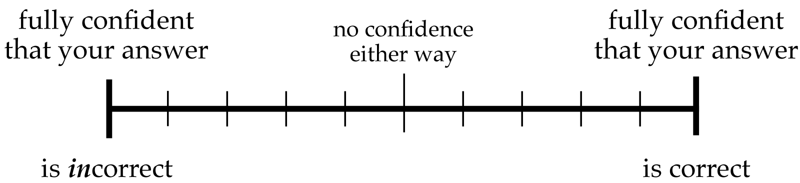
(please tick the appropriate boxes)

The confidence ratings were coded from −5 ("fully confident that your answer is **in**correct") to 5 ("[...] is correct").

Since both lower and upper bound responses needed to match for the classification according to the predictions, the *a priori* chance for guessing any specific interpretation is very low (i.e., $1/(42^2) = 0.0006$; see Pfeifer and Tulkki 2017a).

Then, the target task was separated by a dotted horizontal line from the following question:

How confident are you that the answer you just chose is correct?



The latter question was intended to investigate the participants' meta-cognitive feeling about the rightness of their responses (see, e.g., Thompson et al. 2011).

Inferentialist accounts of conditionals claim that some inferential connection between the antecedent A and the consequent C is needed in order to support a belief in the conditional *if A, then C*, which could, for example be

- deductive, inductive, or abductive (see, e.g., Douven et al. 2020); or
- the relevance of A for C; which can be measured, e.g., by Δp (see, e.g., Skovgaard-Olsen et al. 2016).

Δp emerged in the context of causal reasoning and is defined as follows:

$$\Delta p(\text{If } A, \text{ then } C) =_{def.} p(C|A) - p(C|\neg A)$$

For probabilistic truth table tasks under complete probabilistic knowledge, Δp can be calculated straightforwardly. However, for those tasks under incomplete probabilistic knowledge, it is not straightforward. For two covered sides, for instance, 16 possible configurations need to be considered (see Table 1). The median (or mean) can be used as a proxy for Δp (see Table 2 for summary statistics for the example in the text). Table 3 presents summary statistics of proxies for Δp values for tasks 1–20. Positive Δp values would imply for inferentialists high degrees of belief in conditionals, whereas zero or negative values should decrease the belief in conditionals as the inferential strength is smaller.

Table 1 Δp of all 16 possibilities in the sample task (i.e., a die with two covered sides and two black circles, one white circle, and one black triangle; the conditional in the conclusion is: if circle, then white)

c1	c2	w1	w2	#wc	#c	#w¬c	#¬c	*p*(*w*|*c*)	$p(w\|\neg c)$	Δp
no	no	no	no	1	3	0	3	0.33	0.00	0.33
no	no	no	yes	1	3	1	3	0.33	0.33	0.00
no	no	yes	no	1	3	1	3	0.33	0.33	0.00
no	no	yes	yes	1	3	2	3	0.33	0.67	−0.33
no	yes	no	no	1	4	0	2	0.25	0.00	0.25
no	yes	no	yes	2	4	0	2	0.50	0.00	0.50
no	yes	yes	no	1	4	1	2	0.25	0.50	−0.25
no	yes	yes	yes	2	4	1	2	0.50	0.50	0.00
yes	no	no	no	1	4	0	2	0.25	0.00	0.25
yes	no	no	yes	1	4	1	2	0.25	0.50	−0.25
yes	no	yes	no	2	4	0	2	0.50	0.00	0.50
yes	no	yes	yes	2	4	1	2	0.50	0.50	0.00
yes	yes	no	no	1	5	0	1	0.20	0.00	0.20
yes	yes	no	yes	2	5	0	1	0.40	0.00	0.40
yes	yes	yes	no	2	5	0	1	0.40	0.00	0.40
yes	yes	yes	yes	3	5	0	1	0.60	0.00	0.60

Note "c1" (resp., "c2") denotes whether the first (resp., second) question mark covers a circle and whether the corresponding figure is white (w1 and w2, respectively). "#wc" (resp., "#w¬c") denotes the number of white circles (resp., not-white circles), #c (resp., #¬c) denotes the number of circles (resp., not-circles), and $\Delta p = p(w|c) - p(w|\neg c)$. For summary statistics (see Table 2)

Table 2 Summary statistics of Δp for the sample task (see Table 1)

Δp	
Minimum:	−0.3333
1st quartile:	0.0000
Median:	0.2250
Mean:	0.1625
3rd quartile:	0.4000
Maximum:	0.6000

Table 3 Summary statistics of Δp values for tasks 1–20 (T1–T20)

	T1	T2	T3	T4	T5	T6	T7
# covered	0	0	0	1	2	3	2
Minimum	0.4000	0.3333	0.7500	0.0000	−0.8000	−0.8000	−0.5000
1st quartile	0.4000	0.3333	0.7500	0.3000	−0.6167	−0.5000	−0.2500
Median	0.4000	0.3333	0.7500	0.4500	−0.4500	−0.2500	0.0000
Mean	0.4000	0.3333	0.7500	0.3750	−0.4208	−0.2344	0.0000
3rd quartile	0.4000	0.3333	0.7500	0.5250	−0.2500	0.0000	0.2500
Maximum	0.4000	0.3333	0.7500	0.6000	0.0000	0.3333	0.5000
	T8	T9	T10	T11	T12	T13	T14
# covered	1	2	1	3	0	2	3
Minimum	0.2500	0.0000	0.0000	−0.6000	−0.5000	−0.6000	−0.3333
1st quartile	0.3125	0.2500	0.1875	−0.2125	−0.5000	−0.4000	0.0000
Median	0.4167	0.3333	0.2917	0.0000	−0.5000	−0.2250	0.2500
Mean	0.4375	0.3542	0.2708	0.0719	−0.5000	−0.1625	0.2344
3rd quartile	0.5417	0.5000	0.3750	0.3333	−0.5000	0.0000	0.5000
Maximum	0.6667	0.7500	0.5000	0.7500	−0.5000	0.3333	0.8000

	T15	T16	T17	T18	T19	T20
# covered	1	3	2	1	0	3
Minimum	−0.5000	−0.7500	−0.4000	−0.5000	0.7500	−0.5000
1st quartile	−0.3750	−0.3333	−0.2000	−0.1250	0.7500	−0.1750
Median	−0.2917	0.0000	0.0000	0.0000	0.7500	0.2500
Mean	−0.2708	−0.0719	0.0958	−0.0750	0.7500	0.2844
3rd quartile	−0.1875	0.2125	0.3333	0.0500	0.7500	0.6875
Maximum	0.0000	0.6000	0.6667	0.2000	0.7500	1.0000

Note "#covered" denotes the number of covered sides

2.3 Results

On the average, 89.75% of the response frequencies are explained by the categorisation given by the predictions in Table 4. The conditional probability interpretation is the best predictor for the responses: 54.75% of the responses coincide exactly with the best possible coherent lower and upper probability bounds of the conditional event. Figure 1 shows the conditional probability responses as a function of the item position. The frequency drops after the first three tasks (T1–T3 are relatively easy, since they do not involve covered sides), but then increases again during the course of the experiment. The shift towards conditional probability interpretation is not as pronounced as reported in Pfeifer (2013a). This can be explained by the different samples: in Pfeifer (2013a) no students of mathematics, philosophy, computer science, or psychology were included in the sample, whereas in the current sample, participants had some formal background.

Two participants gave almost consistently probability of the material conditional responses. Probability of the conjunction responses were a bit more frequent but still rare. Likewise, both biconditional interpretations receive high frequencies only if they overlap with the conditional probability interpretation (which is an artefact). Thus, they occurred also relatively seldom. Cases in which conditional probability responses occurred just under 50% can be explained by halfway conditional event strategies. In those tasks, which allow for unambiguously differentiating among the halfway strategies, no participant used the shortcut of ignoring covered sides for both bounds at once: the strategy was used either for the lower or for the upper bound alone. The frequencies of unambiguous halfway response strategies were higher in the first half of the tasks (T1–T10) compared to the last half (T11–T20).

Overall, the responses appear not to depend on Δp, which poses a challenge to inferentialist accounts of conditionals. According to inferentialist accounts, to hold a high degree of belief in a conditional some additional inferential connection between the antecedent and the consequent is required, beyond conditional probability. Thus, for negative Δp values (or $\Delta p = 0$), conditional probability should not be the best predictor. However, as can be seen in Table 4, conditional probability

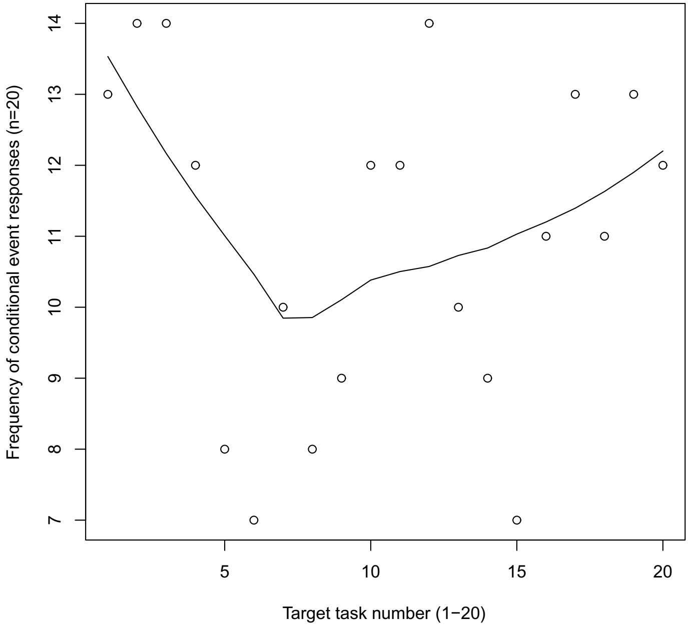


Fig. 1 Conditional event responses ($n = 20$). The solid line was generated using the locally weighted scatter plot smoother method (`lowess`, see Cleveland [1981]; implemented in `R`)

remains the best predictor for the data even in cases where Δp is negative or equal to zero. Pfeifer (2013b, Table 1) presents eight candidates for measures of argument strength, which are based on prominent measures of confirmation: Δp is one of them. Without going into the details of these measures, note that except for tasks 7, 17, 18, and 21, each sign ($+$, $-$, or 0) of their medians coincides with the respective sign of the median of Δp. Thus, replacing Δp by similar other measures does not yield a better predictive validity compared to the conditional probability interpretation of conditionals.[8]

Task 20 was pragmatically special as it showed a die with three white circles and three question marks while it asked for the conditional

[8] For this chapter a publicly available `R` subroutine for calculating the predictions in the probabilistic truth table task and an interactive `R Shiny App` have been developed (see Schöppl 2021).

Table 4 Response frequencies, classified by interpretation, and mean (and SD) confidence in correctness responses for tasks 1–20 (T1–T20; $n = 20$)

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15	T16	T17	T18	T19	T20
$P(C\|A)*$	13	14	14	12	8	7	10	8	9	12	12	14	10	9	7	11	13	11	13	12
$P(A \supset C)*$	2	1	1	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2
$P(A \wedge C)*$	2	2	2	2	4	4	2	3	3	2	3	4	3	3	2	4	4	4	4	5
P(C‖A)	[13]	[14]	[14]	0	[4]	[4]	[2]	[8]	[9]	0	[12]	[4]	[3]	2	[2]	[11]	[13]	[4]	[13]	[12]
$P(A \equiv C)$	[2]	[1]	[1]	0	[4]	[4]	[2]	[2]	[2]	0	[2]	[4]	[3]	0	[2]	[2]	[2]	[4]	[2]	[2]
$P(C\|A)_{\overline{l}}$	–	–	–	1	2	3	2	3	4	1	[12]	–	3	1	2	[11]	[13]	2	–	[12]
$P(C\|A)_{\overline{u}}$	–	–	–	[12]	3	2	[10]	4	2	[12]	0	–	1	[9]	2	0	0	[11]	–	0
$P(C\|A)_{\overline{lu}}$	–	–	–	[1]	0	0	[2]	0	0	[1]	[0]	–	0	[1]	0	[0]	[0]	[2]	–	[0]
Other	3	3	3	3	1	2	4	0	0	3	3	0	1	3	5	3	1	1	1	1
Mean	4.20	4.25	4.10	3.20	3.10	3.05	3.35	3.40	3.15	3.70	3.70	4.35	3.65	3.55	3.80	3.85	3.25	3.70	4.00	2.65
SD	1.24	1.41	1.52	2.48	2.07	2.28	2.21	2.16	2.18	1.89	2.05	1.27	1.95	1.85	1.77	1.42	2.05	1.92	1.45	2.43
$Md.(\Delta p)$	+	+	+	+	–	–	0	+	+	+	0	–	–	+	–	0	0	0	+	+

Note " $*$" denotes target interpretations. Square brackets denote response frequencies consistent with multiple predictions. Halfway interpretations are undefined for complete probabilistic knowledge tasks (i.e., those without covered sides)

"$Md.(\Delta p)$" denotes whether the median value of Δp is positive ($+$), negative ($-$) or zero (0; see also Table 3) involving different patterns (**If** the side facing up shows a *square*, **then** the side shows *black*). The visible patterns validate the material conditional interpretation, since *circle* validates the disjunct $\neg$ square, which in turn makes the material conditional $\neg$ square $\vee$ black true. For the conditional probability interpretation, however, the visible sides are irrelevant and the participant has to reason about the possible patterns behind the question marks. This could lead to some pragmatic oddness of this task, which is reflected by a slight decrease of the mean confidence rating.

After Task 20, participants were presented with a die consisting of six white circles. The target conditional was the same as in Task 20, i.e., its antecedent was inconsistent with all visible patterns. In this task, 14 participants responded with zero, three with one, two with the unit interval ([0, 1]), and one participant responded with .5. If participants interpret the visible sides as the only logically possible side-patterns (i.e., there can only be a white circle), then the antecedent of the conditional is logically impossible and hence conditional probability is undefined. As there was no opt-out option, except for three probability 1 responses, all other responses (0, [0, 1], and .5 as a kind of guessing) can be interpreted as rejecting such conditionals. The mean confidence in the correctness of the responses in Task 21 was 3.20 (SD = 2.09).

To investigate in general whether there are any difficulties with calculating the probability of the material conditional, tasks 22 (all sides visible) and 23 (one covered side) formulated a material conditional in terms of disjunctions (e.g., the side facing up does **not** show *black* **or** the side shows a *circle*). In Task 22, 17 out of 20 participants inferred correctly the probability of the material conditional (mean confidence in correctness: 3.45, SD = 1.82). In Task 23, which involves incomplete probabilistic knowledge, 11 out of 20 participants made correct inferences (mean confidence in correctness: 2.35, SD = 3.15). Thus, the probability of the material conditional is not too hard to calculate for most participants; rather, people interpret conditionals differently, namely as conditional probabilities.

After the target tasks, the participants answered two questions *Were the instructions clear to you?* and *How difficult were the tasks?* on rating scales with the extreme poles *very* **un***clear* (*very difficult*, resp.; coded as

0) and *very clear* (*very easy*, resp.; coded as 10). The overall clearness of the tasks was rated by 8.75 (SD = 1.48) on the average. The overall difficulty was rated by 6.10 (SD = 2.31) on the average. This indicates that the task formulation was clear and that the tasks had an appropriate discriminatory power (i.e., they were neither too easy nor too hard).

3 Concluding Remarks

I presented Mental Probability Logic as a competence theory of human inference, which uses coherence-based probability logic as the rationality framework. Zero-probability antecedents are important not only for theoretical reasons, but also for practical reasons. If properly managed, they allow for explaining some alleged pragmatic effects in purely semantic terms. I also presented an experiment on a generalised probabilistic truth table task. The main result is that, like naive participants, most mathematically minded participants interpret degrees of beliefs in conditionals in terms of conditional probabilities. Some make cognitive shortcuts by ignoring some possibilities when reasoning about lower or upper probability bounds on conditionals. Participants' ratings appear independent of Δp, which poses a challenge for inferentialist accounts of conditionals which propose Δp (or similar measures) of inferential strength.

Under inferentialism, one would expect differences in the probabilistic truth table tasks which involve neutral material (like circle/square and black/white) compared to causal or abductive task material (like *if a drug is administered, then symptoms are diminished* or *if symptoms are diminished, then a drug was administered*, respectively). However, the data of probabilistic truth table tasks which compared such task material suggest that there are no such differences and conditional probability is the best predictor (Over et al. 2007; Pfeifer and Stöckle-Schobel 2015; Pfeifer and Tulkki 2017a).

Another challenge for inferentialism is Centering (*from A and C infer if A, then C*), which does not hold in inferentialist accounts (e.g., Douven 2016, p. 40). However, Centering holds in probability logic (Sanfilippo et al. 2018). Consistently with the formal predictions of probability logic, experimental data suggest that most responses are in line with the predictions of Centering (see, e.g., Cruz et al. 2015, 2016; Pfeifer and Tulkki 2017b; Politzer and Baratgin 2016). Thus, Centering challenges inferentialist accounts of conditionals from both, formal and descriptive points of views (see also the chapter by Over and Cruz in press, in this volume).

Frequencies play a communicative role in probabilistic truth table tasks: they allow for clearly communicating uncertainties to the participants. This does not mean, however, that here a frequentistic interpretation of probability is more appropriate compared to the subjective interpretation of probability. First, the task is about a singular event. Specifically, it is about the particular outcome of throwing a specific die, which is made clear to the participants in the instruction ("imagine that this die is placed in a cup"). It is well-known that frequentistic approaches cannot deal with probabilities of singular events, while coherence can properly manage the uncertainty of singular events. Second, the task explicitly asks participants to infer how sure they can be that the respective sentence holds, i.e., the task asks for the participants' degrees of belief. Thirdly, using frequencies is one of several possibilities for forming degrees of belief. The resulting beliefs do not imply the ontological existence of an objective probability.[9] Epistemically, however, there are correct solutions to the tasks.

The probabilistic truth table task paradigm is a fair and clear litmus test for how people interpret various kinds of conditionals. It is not restricted to precise-probability assessments. Future research could be devoted to investigating other logical connectives with the truth table task (and looking also at corresponding halfway interpretations), to combining truth table tasks with neuroscientific methods, or to studying to what extent mental disorders modulate reasoning about conditionals. The latter strand of future research is inspired by the pioneer of the experimental psychology of deductive reasoning, Gustav Störring, who argued that "in the pathological cases nature makes experiments for us" (Störring 1900, p. 11; my translation): studying reasoning processes with people

[9] This is in line with de Finetti's famous and provocative slogan "Probability does not exist" (1970/1974, p. vii, Vol. 1): probability is not an objective frequentistic quantity existing in the outside world; rather, probability is subjective and coincides with coherent degrees of belief.

suffering from mental disorders can give us supplementary insight for the understanding of normal reasoning processes.

Acknowledgements This work is supported by the BMBF project 01UL1906X. The Department of Philosophy at the Carnegy Mellon University funded the author's research stay at the Center for Formal Epistemology (CMU) and the reimbursement of the participants. David Danks, Angelo Gilio, Stefan Kaufmann, Kevin T. Kelly, Tim Kraft, David E. Over, Katrin Pfeifer, Hans Rott, Giuseppe Sanfilippo, Romina Schmid, and participants of the seminar "Current Topics in Theoretical Philosophy" (Summer term, 2021) at the Department of Philosophy at the University of Regensburg provided useful comments to an earlier draft. Leon Schöppl assisted in the data analysis.

References

Berti, P., Regazzini, E., & Rigo, P. (1998). Well calibrated, coherent forecasting systems. *Theory of Probability & Its Applications, 42*(1), 82–102.

Biazzo, V., Gilio, A., Lukasiewicz, T., & Sanfilippo, G. (2005). Probabilistic logic under coherence: Complexity and algorithms. *Annals of Mathematics and Artificial Intelligence, 45*(1–2), 35–81.

Bonnefon, J.-F., & Politzer, G. (2010). Paradoxes, mental models and one paradox of the material conditional. *Mind & Language, 26*(2), 141–155.

Capotorti, A., Galli, L., & Vantaggi, B. (2003). Locally strong coherence and inference with lower-upper probabilities. *Soft Computing, 7*(5), 280–287.

Capotorti, A., Lad, F., & Sanfilippo, G. (2007). Reassessing accuracy rates of median decisions. *American Statistician, 61*(2), 132–138.

Carnap, R. (1948). *Introduction to semantics*. Chicago: Harvard University Press.

Cleveland, W. S. (1981). LOWESS: A program for smoothing scatterplots by robust locally weighted regression. *The American Statistician, 35*(1), 54.

Coletti, G., & Scozzafava, R. (2002). *Probabilistic logic in a coherent setting*. Dordrecht: Kluwer.

Cruz, N., Baratgin, J., Oaksford, M., & Over, D. E. (2015). Bayesian reasoning with ifs and ands and ors. *Frontiers in Psychology, 6*, Article 192.

Cruz, N., Baratgin, J., Oaksford, M., & Over, D. E. (2016). Centering and the meaning of conditionals. In A. Papafragou, D. Grodner, D. Mirman, &

J. Trueswell (Eds.), *Proceedings of the 38th annual meeting of the cognitive science society* (pp. 1104–1109). Austin, TX: The Cognitive Science Society.

de Finetti, B. (1931). Sul significato soggettivo della probabilitá. *Fundamenta Mathematicae, 17*, 298–329.

de Finetti, B. (1936/1995). The logic of probability. *Philosophical Studies, 77*, 181–190.

de Finetti, B. (1970/1974). *Theory of probability* (Vols. 1, 2). Chichester: Wiley.

Douven, I. (2016). *The epistemology of indicative conditionals: Formal and empirical approaches*. Cambridge: Cambridge University Press.

Douven, I., Elqayam, S., Singmann, H., & Wijnbergen-Huitink, J. (2020). Conditionals and inferential connections: Toward a new semantics. *Thinking & Reasoning, 26*, 311–351.

Égré, P., & Rott, H. (2021). The logic of conditionals. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Fall 2021 ed.). Metaphysics Research Lab, Stanford University. https://plato.stanford.edu/entries/logic-conditionals/.

Elqayam, S., Douven, I., Evans, J. S. B. T., & Cruz, N. (Eds.). (2020). *Logic and uncertainty in the human mind: A tribute to David E. Over*. London: Routledge.

Evans, J. St. B. T. (1982). *The psychology of deductive reasoning*. London: Routledge.

Evans, J. St. B. T., Handley, S. J., & Over, D. E. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(2), 321–355.

Evans, J. St. B. T., & Over, D. E. (2004). *If*. Oxford: Oxford University Press.

Fugard, A. J. B., Pfeifer, N., & Mayerhofer, B. (2011). Probabilistic theories of reasoning need pragmatics too: Modulating relevance in uncertain conditionals. *Journal of Pragmatics, 43*, 2034–2042.

Fugard, A. J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. D. (2009). How people interpret an uncertain If. In T. Kroupa & J. Vejnarová (Eds.), *Proceedings of the 8th workshop on uncertainty processing* (pp. 80–91). Liblice, CZ.

Fugard, A. J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. D. (2011). How people interpret conditionals: Shifts towards the conditional event. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*(3), 635–648.

Gauffroy, C., & Barrouillet, P. (2009). Heuristic and analytic processes in mental models for conditionals: An integrative developmental theory. *Developmental Review, 29*, 249–282.

Gilio, A. (1990). Criterio di penalizzazione e condizioni di coerenza nella valutazione soggettiva della probabilità. *Bollettino dell'Unione Matematica Italiana, 4–B*, 645–660.

Gilio, A. (1996). Algorithms for conditional probability assessments. In D. A. Berry, K. M. Chaloner, & J. K. Geweke (Eds.), *Bayesian analysis in statistics and econometrics: Essays in honor of Arnold Zellner* (pp. 29–39). New York: John Wiley.

Gilio, A. (2002). Probabilistic reasoning under coherence in System P. *Annals of Mathematics and Artificial Intelligence, 34*, 5–34.

Gilio, A., Pfeifer, N., & Sanfilippo, G. (2016). Transitivity in coherence-based probability logic. *Journal of Applied Logic, 14*, 46–64.

Gilio, A., Pfeifer, N., & Sanfilippo, G. (2020). Probabilistic entailment and iterated conditionals. In S. Elqayam, I. Douven, J. S. B. T. Evans, & N. Cruz (Eds.), *Logic and uncertainty in the human mind: A tribute to David E. Over* (pp. 71–101). London: Routledge.

Gilio, A., & Sanfilippo, G. (2013). Quasi conjunction, quasi disjunction, t-norms and t-conorms: Probabilistic aspects. *Information Sciences, 245*, 146–167. https://doi.org/10.1016/j.ins.2013.03.019.

Gilio, A., & Sanfilippo, G. (2014). Conditional random quantities and compounds of conditionals. *Studia Logica, 102*(4), 709–729.

Gilio, A., & Sanfilippo, G. (2019). Generalized logical operations among conditional events. *Applied Intelligence, 49*(1), 79–102.

Gilio, A., & Sanfilippo, G. (2021). Compound conditionals, Fréchet-Hoeffding bounds, and Frank t-norms. *International Journal of Approximate Reasoning, 136*, 168–200.

Holzer, S. (1985). On coherence and conditional prevision. *Bollettino dell'Unione Matematica Italiana, 4*(6), 441–460.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review, 109*(4), 646–678.

Kleiter, G. D., Fugard, A. J. B., & Pfeifer, N. (2018). A process model of the understanding of uncertain conditionals. *Thinking & Reasoning, 24*(3), 386–422.

Kolmogoroff, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Berlin, Heidelberg: Springer.

Lad, F. (1996). *Operational subjective statistical methods: A mathematical, philosophical, and historical introduction*. New York: Wiley.

Marr, D. (1982). *Vision. A computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford: Oxford University Press.

Oaksford, M., & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology, 71*, 305–330.

Oberauer, K., & Wilhelm, O. (2003). The meaning(s) of conditionals: Conditional probabilities, mental models and personal utilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(4), 680–693.

Over, D. E., & Cruz, N. (in press). Indicative and counterfactual conditionals in the psychology of reasoning. In S. Kaufmann, D. E. Over, & G. Sharma (Eds.), *Conditionals—Logic, linguistics, and psychology.* London: Palgrave Macmillan. https://link.springer.com/book/9783031056819

Over, D. E., Hadjichristidis, C., Evans, J. St. B. T., Handley, S. J., & Sloman, S. (2007). The probability of causal conditionals. *Cognitive Psychology, 54*, 62–97.

Petturiti, D., & Vantaggi, B. (2017). Envelopes of conditional probabilities extending a strategy and a prior probability. *International Journal of Approximate Reasoning, 81*, 160–182.

Pfeifer, N. (2006). *On mental probability logic*. Unpublished doctoral dissertation, Department of Psychology, University of Salzburg. (The abstract is published in *The Knowledge Engineering Review*, 2008, 23, 217–226.)

Pfeifer, N. (2012). Experiments on Aristotle's Thesis: Towards an experimental philosophy of conditionals.*The Monist, 95*(2), 223–240.

Pfeifer, N. (2013a). The new psychology of reasoning: A mental probability logical perspective. *Thinking & Reasoning, 19*(3–4), 329–345.

Pfeifer, N. (2013b). On argument strength. In F. Zenker (Ed.), *Bayesian argumentation. The practical side of probability* (pp. 185–193). Dordrecht: Synthese Library (Springer).

Pfeifer, N. (2014). Reasoning about uncertain conditionals. *Studia Logica, 102*(4), 849–866.

Pfeifer, N. (2021). Probability logic. In M. Knauff & W. Spohn (Eds.), *Handbook of rationality* (pp. 277–284). Cambridge, MA: The MIT Press.

Pfeifer, N., & Kleiter, G. D. (2005). Coherence and nonmonotonicity in human reasoning. *Synthese, 146*(1–2), 93–109.

Pfeifer, N., & Kleiter, G. D. (2006). Inference in conditional probability logic. *Kybernetika, 42*, 391–404.

Pfeifer, N., & Kleiter, G. D. (2009). Framing human inference by coherence based probability logic. *Journal of Applied Logic, 7*(2), 206–217.

Pfeifer, N., & Kleiter, G. D. (2011). Uncertain deductive reasoning. In K. Manktelow, D. E. Over, & S. Elqayam (Eds.), *The science of reason:*

A Festschrift for Jonathan St. B.T. Evans (pp. 145–166). Hove: Psychology Press.

Pfeifer, N., & Sanfilippo, G. (2021). Interpreting connexive principles incoherence-based probability logic. In J. Vejnarová & J. Wilson (Eds.), *Symbolic and quantitative approaches to reasoning with uncertainty (ECSQARU 2021)* (LNAI, Vol. 12897, pp. 672–687). Cham: Springer.

Pfeifer, N., & Stöckle-Schobel, R. (2015). Uncertain conditionals and counterfactuals in (non-)causal settings. In G. Airenti, B. G. Bara, & S. G. (Eds.), *Proceedings of the EuroAsianPacific Joint Conference on Cognitive Science (4th European Conference on Cognitive Science; 10th International Conference on Cognitive Science)* (Vol. 1419, pp. 651–656). Aachen: CEUR Workshop Proceedings. Retrieved from http://ceur-ws.org/Vol-1419/paper0108.pdf.

Pfeifer, N., & Tulkki, L. (2017a). Abductive, causal, and counterfactual conditionals under incomplete probabilistic knowledge. In G. Gunzelmann, A. Howes, T. Tenbrink, & E. Davelaar (Eds.), *Proceedings of the 39th Cognitive Science Society Meeting* (pp. 2888–2893). Austin, TX: The Cognitive Science Society.

Pfeifer, N., & Tulkki, L. (2017b). Conditionals, counterfactuals, and rational reasoning. An experimental study on basic principles. *Minds and Machines, 27*(1), 119–165.

Politzer, G., & Baratgin, J. (2016). Deductive schemas with uncertain premises using qualitative probability expressions. *Thinking & Reasoning, 22*(1), 78–98.

Ramsey, F. P. (1929/1994). General propositions and causality (1929). In D. H. Mellor (Ed.), *Philosophical papers by F. P. Ramsey* (pp. 145–163). Cambridge: Cambridge University Press.

Regazzini, E. (1985). Finitely additive conditional probabilities. *Rendiconti del Seminario Matematico e Fisico di Milano, 55*, 69–89.

Sanfilippo, G., Gilio, A., Over, D. E., & Pfeifer, N. (2020). Probabilities of conditionals and previsions of iterated conditionals. *International Journal of Approximate Reasoning, 121*, 150–173.

Sanfilippo, G., Pfeifer, N., & Gilio, A. (2017). Generalized probabilistic modus ponens. In A. Antonucci, C. L., & O. Papini (Eds.), *Lecture notes LNAI* (Vol. 10369, pp. 480–490). Dordrecht: Springer.

Sanfilippo, G., Pfeifer, N., Over, D. E., & Gilio, A. (2018). Probabilistic inferences from conjoined to iterated conditionals. *International Journal of Approximate Reasoning, 93*, 103–118.

Schöppl, L. (2021). *Dice task subroutine (v1.0) [computer software].* Retrieved from https://github.com/leon-schoeppl/ptt-calculator/releases/tag/DTSv1.0/ on September 2, 2021.

Skovgaard-Olsen, N., Singmann, H., & Klauer, K. C. (2016). The relevance effect and conditionals. *Cognition, 150*, 26–36.

Störring, G. W. (1900). *Vorlesungen über Psychopathologie in ihrer Bedeutung für die normale Psychologie mit Einschluss der psychologischen Grundlagen der Erkenntnistheorie*. Leipzig: Wilhelm Engelmann.

Thompson, V. A., Prowse-Turner, J. A., & Pennycook, G. (2011). Intuition, reason, and metacognition. *Cognitive Psychology, 63*, 107–140.

Walley, P., Pelessoni, R., & Vicig, P. (2004). Direct algorithms for checking consistency and making inferences from conditional probability assessments. *Journal of Statistical Planning and Inference, 126* (1), 119–151.

Wansing, H. (2020). Connexive logic. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Spring 2020 ed.). https://plato.stanford.edu/archives/spr2020/entries/logic-connexive/.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New horizons in psychology I* (pp. 106–137). Harmandsworth: Penguin.

Rethinking the Acceptability and Probability of Indicative Conditionals

Michał Sikorski

1 Introduction

Indicative conditionals, like:

(1) If you press this button, the fire alarm goes off.

are an important part of our language. We use them, for example, to express our prediction or generalizations. Partly because of their importance, conditionals are interesting for philosophers and psychologists. They are interested, for example, in truth conditions of conditionals or updating our beliefs with them. Two other issues which received a lot of attention are the probability and acceptability of indicative conditionals.

M. Sikorski (✉)
Warsaw University of Technology, Warsaw, Poland
e-mail: michalpsikorski@gmail.com

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_5

In the case of probability, the reasons for all this attention are clear. For instance, if we were able to define the probabilities of conditionals, we could incorporate reasoning with conditionals into the popular and successful framework of Bayesian epistemology.

In the case of acceptability, the attention is a bit harder to explain. The acceptability conditions of other complex expressions are not so widely discussed. They are, to be sure, studied as a part of pragmatics or epistemology, but it seems that there is not, for example, a special problem of the acceptability conditions for conjunction. What is different in the case of conditionals? It seems to me that it is an influence of a very popular philosophical position called the non-truth value view (NTV). It claims that conditionals do not have truth values.[1] The proponents of NTV have to deal with at least two problems. Firstly, we systematically recognize that some conditionals are appropriate to utter in some situations while others are not. In the case of other sentences, it can be often explained by the difference between truth and falsity. So how can we explain that without postulating truth values for conditionals? Secondly, if we claim that conditionals are not truth-apt, it seems natural to assume that they are not probability-apt. The probability of a sentence is the probability of that sentence being true and if a sentence is not truth-apt (think for example about commands or questions), it makes little sense to ask about its probability. If it is so, we even in principle cannot incorporate the conditionals to the Bayesian framework. The answer to both challenges is provided by the notion of acceptability. We can use graded acceptability as a substitute for probability and categorical acceptability as a substitute for truth.

The discussion concerning the probability and the acceptability of conditional $(A \rightarrow B)$, is mainly organized around two influential theses. The first of them is so fundamental for the currently dominant paradigm of thinking about conditionals (see, e.g. Over and Cruz 2018) that it usually just called the Equation:

Equation $P(A \rightarrow B) = P(B|A)$

[1] For the details and motivation of the view see, e.g., Bennett (2003) or Edgington (1995). For the critical discussion see: Douven (2015).

The second thesis is called Adams' thesis:

AT $ac(A \rightarrow B) = P(B|A)$

where $P(B|A)$ indicates conditional probability and "$ac()$" indicates acceptability. AT in this form is not a good substitute for truth conditions. It does not provide us with a threshold of acceptability above which a conditional would be acceptable. Such a threshold is provided by another version of AT, the Qualitative Adams' Thesis:

(QAT) An indicative conditional "If A, B" is assertable for/acceptable to a person if and only if the person's conditional degree of belief, $P(B|A)$, is high.[2]

All three theses were evaluated from both empirical and theoretical perspectives. In this chapter, I will examine both of these perspectives and show that there are no convincing reasons to accept any of them, and therefore we should rethink their role in the future study of conditionals. In the second section, I will discuss the experiments dedicated to all three theses. Then I will discuss the theoretical considerations for and against them. In the last section, I will conclude and point to some alternative conceptualizations of the probability of conditionals.

2 Empirical Support

In this section, I will discuss empirical experiments concerning the three theses. Before that, I will make a distinction useful in this context.

Conditionals can be divided into positively relevant, irrelevant, and negatively relevant. The positively relevant conditionals are conditionals whose antecedents are positively probabilistically relevant for their consequents. If a sentence is positively probabilistically relevant for another one, then the truth of the first sentence makes the second one more probable. The negatively relevant conditionals are conditionals

[2] The source of this formulation is Douven and Verbrugge (2012).

whose antecedents are negatively probabilistically relevant for their consequents, which means that the truth of the antecedent decreases the probability of the consequent. Irrelevant conditionals are the conditionals whose antecedents are probabilistically irrelevant for their consequents. The concept of relevance can be mathematically represented in at least two ways. Firstly, we can use $\Delta P = P(B|A) - P(B|\neg A)$ as a measure of relevance, as proposed in Spohn (2012). If the value of ΔP is 0 the corresponding conditional is irrelevant; when it is higher, then it is positively relevant, and when it is lower, the conditional is negatively relevant. Secondly, relevance can be conceptualized as the difference measure, $P(B|A) - P(B)$. As in the case of ΔP, when the value of difference measure is 0, the conditional is irrelevant; if it is lower, it is negatively relevant; and if it is higher, it is positively relevant. Both conceptualizations classify conditionals in the same way, but the exact level of relevance will differ in some cases.[3] Both notions have been used in experiments on conditionals, and the difference will not matter for our conclusions.

An example of an intuitively irrelevant conditional is:

(2) If I eat an apple today, I will not inherit 1,000,000$ today.

And a negatively relevant one is:

(3) If he smokes, he will not develop lung cancer.

Going back to our three theses, all of them have been traditionally regarded as descriptively true.[4] Philosophers generally found all of them confirmed by their introspective case-by-case studies. Many such cases

[3] For a detailed discussion of the difference between the two notions and an experiment indicating that ΔP predicts intuitive relevance better than the difference measure, see Skovgaard-Olsen et al. (2016a).

[4] For example McGee, 1989: "Ernest Adams (1965, 1975) has advanced a probabilistic account of conditionals, according to which the probability of a simple English indicative conditional is the conditional probability of the consequent given the antecedent. The theory describes what English speakers assert and accept with unfailing accuracy, yet the theory has won only limited acceptance."

were presented, for example, in Bennett (2003), Edgington (1995), or Jackson (1987).

More systematic experimental studies were, firstly, directed toward the Equation. The results of most of these experiments support it. For example, Evans et al. (2003), Over et al. (2007) or Oberauer and Wilhelm (2003) found significant correlation between participants' responses concerning the probability of conditionals and conditional probability while using different types of conditionals. Over et al. (2007) used "causal" conditionals, i.e., conditionals justified by causal relations, while Oberauer and Wilhelm (2003) uses conditionals that describe relations between frequency distributions. Results of those, and many similar studies (e.g., Fugard et al. 2011; Barrouillet and Gauffiroy 2015; Evans et al. 2007; Cruz et al. 2016), support the Equation. They convinced many philosophers and psychologists that the Equation is a correct description of how people reason with conditionals and made it, and probabilistic theories based on it, a dominant paradigm for thinking about conditionals.[5]

Both AT and QAT did not receive so much attention. AT was first tested in Douven and Verbrugge (2010). In the experiment, the authors used inferential conditionals divided into inductive, abductive, and deductive conditionals. Inferential conditionals are conditionals that express inferences. Inductive conditionals express inductive inferences, deductive conditionals express deductive inferences, and abductive conditionals express abductive inferences. In the first experiment, the authors tested Adams' Thesis and four weaker versions of it:

(WAT1) $Ac(A \to B) \approx Pr(B|A)$
(WAT2) $Ac(A \to B)$ is high/middling/low iff $Pr(B|A)$ is high/middling/low.
(WAT3) $Ac(A \to B)$ highly correlates with $Pr(B|A)$.
(WAT4) $Ac(A \to B)$ at least moderately correlates with $Pr(B|A)$.

[5] See, e.g., Over and Cruz (2018) or Evans and Over (2004).

The theses were tested by comparing their prediction with responses given by participants to questions concerning the acceptability and probability of a given conditional.

Surprisingly, only a weak correlation between the conditional probability and the acceptability of conditionals was found. The correlation was especially weak in the case of inductive conditionals. It was not enough to support AT or even two weaker versions of it. Just the weakest version (WAT4) was supported for all kinds of conditionals (inductive, deductive, and abductive). In the third experiment presented in the paper, participants were asked to judge the conditional probability of the consequent given the antecedent and the probability of the conditional. The results of the first experiment and the third experiment were compared. The comparison showed a significant difference between participants' judgments concerning the acceptability and the probability of conditionals. I will discuss this issue later on.

QAT was, also, tested the first time by Igor Douven and Sara Verbrugge. The experiment was presented in Douven and Verbrugge (2012). The authors tested the predictions of QAT and the so-called Evidential Support Theory presented in Douven (2008):

EST An indicative conditional $(A \rightarrow B)$ is assertable/acceptable if and only if $Pr(B|A)$ is not only high but also higher than $Pr(B)$.

The idea behind EST is that a high conditional probability is not enough for a conditional to be acceptable, and positive relevance has to be included as an additional condition. Results show that QAT predicted judgments of speakers worse than EST, and especially poorly in the case of irrelevant and negatively relevant conditionals. This result was replicated in Krzyżanowska et al. (2017).

A similar idea, of using irrelevant and negatively relevant conditionals, was adopted by Skovgaard-Olsen et al. (2016b). The authors tested the Equation and AT. The items include positively relevant and, crucially, irrelevant and negatively relevant conditionals. The results showed a significant correlation between the conditional probabilities and the probabilities of the positively relevant conditionals. At the same time, this was not the case for irrelevant and negatively relevant conditionals.

There the probabilities of conditionals were much lower than the conditional probabilities. The results for acceptability were almost the same. The failure of AT is not that surprising if we take into consideration the failure of its qualitative version and the results from Douven and Verbrugge (2010), but the poor performance of the Equation is unexpected given the rich history of experiments that supported it. This result was replicated in experiments with different experimental designs. For example, the results of Krzyżanowska et al. (2017), Skovgaard-Olsen et al. (2016a), Vidal and Baratgin (2017), and Fugard et al. (2011) all suggest the Equation (by itself) does not correctly predict the probability of conditionals in the case of irrelevant and negatively relevant conditionals. This interpretation of the results is controversial. First, it is not clear how it squares with the earlier results, and second, there is an alternative interpretation of the effect.

How should we explain this discrepancy between the results presented in Skovgaard-Olsen et al. (2016b) and earlier experiments supporting the Equation? The authors claim that previous studies do not include irrelevant or negative relevant conditionals and therefore cannot support the unrestricted version of the Equation. For example, all conditionals considered in Over et al. (2007) seem to be intuitively positively relevant one.[6] The case of Oberauer and Wilhelm (2003) is similar. The successful replications and the lack of irrelevant and negatively relevant conditionals in the stimuli used in the earlier experiments strongly suggest that the effect of the relevance on the assessment of the probability or acceptability is robust, and the support for the Equation provided by those experiments should be re-evaluated.

A defender of the Equation may claim that the effect of the relevance of conditionals is pragmatic, and therefore the unrestricted version of the Equation can still be preserved. This solution is somewhat supported by the results of Skovgaard-Olsen et al. (2017) which suggests that the effect of relevance on the assessment of truth is much weaker than its effect on the acceptability or probability of conditionals. This suggests that the effect of relevance is pragmatic in nature. On the other hand,

[6] E.g., "If Adidas get more superstars to wear their new football boots then the sales of these boots will increase" or "If the cost of petrol increases then traffic congestion will improve."

results of different experiments do suggest that relevance influence truth assessments, for example, Krzyżanowska et al. (2017) or Douven et al. (2017).[7] The hypothesis that the effect is pragmatic was also tested directly in Skovgaard-Olsen et al. (2019). The authors tested three hypotheses describing different pragmatic mechanisms generating the reason-relation part of the content of indicative conditionals responsible for the effect. Firstly, they checked if it is cancelable in the way conversational implicatures are, secondly, they tested if its projection behavior resembles that of presuppositions, and finally, they tested if it is treated as not-at-issue content which is believed to be one of the characterizing features of the conventional implicature. Surprisingly, the results of all three experiments were negative, which suggests that the reason-relation part of the content is not conversational implicature, presupposition, nor conventional implicature, and therefore, likely, not pragmatic content. The authors in discussing their results point out that the features of conventional implicature (including it being not-at-issue content) are still very controversial and therefore, given the results of Skovgaard-Olsen et al. (2017) it is likely that the reason-relation part of the content of conditionals is conventional implicature. This, in the opinion of the authors, does not necessarily make it a part of the pragmatic content. Conventional implicature has been classified both as part of pragmatic and semantic content by different authors in the relevant literature. In light of that, it seems that the pragmatic origin of the effect of relevance on probability or acceptability of conditionals is not supported by the existing evidence.

Finally, we may wonder if it is possible to restrict the Equation to make it consistent with the available evidence? It seems possible. A version of the Equation restricted to the positively relevant conditionals seems to be in line with the results of all the mentioned experiments. Such a version can look, for example, like this:

Equation+ If $\Delta P > 0$ then $P(A \to B) = P(C|A)$

[7] For the discussion see: Douven et al. 2017.

All this seems to weaken the position of the unrestricted Equation. At the same time, it puts all the theses in a somehow similar position. All of them were initially regarded as intuitive and supported by introspective case-by-case examination. In light of the available empirical evidence, both QAT and AT seem to be empirically inadequate. QAT performs poorly (Douven and Verbrugge 2012) in comparison to an EST. AT was disconfirmed by results of Skovgaard-Olsen et al. (2016b) which show that it fails in the case of the irrelevant conditionals, and by the results of Douven and Verbrugge (2010), which show that it is not supported in the case of the inductive conditionals. Similarly, the results which were considered to be evidence for the Equation are to some degree undermined by the results of Skovgaard-Olsen et al. (2016b) and considerations concerning the conditionals used in the studies.

3 Theoretical Arguments

The theoretical studies concerning the Equation, AT and QAT have a longer history than the empirical ones. Still, it seems that there is not much theoretical justification for the three theses. Even some of their defenders seem to agree. For example, Douven (2015) says about the Equation:

> While there is no known argument for this thesis showing that it has any normative force, to many the proposal does ring true, at least prima facie.

In this section, I will discuss the theoretical considerations presented for and against the Equation, ST, and QAT. I will start by discussing the Ramsey Test, which is commonly used to argue for the Equation or AT. Then I will move to trivialization proofs. I will discuss them with special attention dedicated to the two most popular ways to block them: denying that conditionals are propositions and postulating that the meaning of a conditional depends on the beliefs of the speaker. Finally, I will discuss the relationship between the semantics of conditionals and their probability.

3.1 Ramsey Test

The Ramsey test was presented by Ramsey (1990):

> If two people are arguing 'If p will q' and both are in doubt as to p, they are adding p hypothetically to their stock of knowledge and arguing on that basis about q; so that in a sense 'If p, q' and 'If p, $\bar{q}$' are contradictories. We can say that they are fixing their degrees of belief in q given p. Ramsey (1990, p. 155)

The test is very popular among philosophers and psychologists,[8] and it is typically interpreted as the procedure for evaluating acceptability or probability of indicative conditionals (see, e.g., Gibbard 1981; Edgington 1995; Bennett 2003). and many cases in which its predictions are correct were considered and discussed.[9] Because of this intuitiveness, but also simplicity, the procedure served as a direct inspiration for three successful research programs: belief revision theory, possible world semantics for conterfactuals, and suppositional theories of indicative conditionals. The theories from the last group are typically committed to the Equation or AT. The Equation is a probabilistic reinterpretation of Ramsey test, and, therefore, the argument from the one to the other is straightforward: If you accept the Ramsey test and conditionalization as a rule for belief revision, which is typically accepted in this context (see, e.g., Pettigrew 2020), then you have to accept the Equation which is just its probabilistic reformulation.[10]

There are two problems with this argument. Firstly, the intuition behind the plausibility of both Ramsey test and the Equation seems to be exactly the same. The second is merely a reformulation of the first, and in all cases in which Ramsey test delivers a correct result, the Equation will give us just as satisfying an answer. Therefore, it seems that by

[8] E.g., "Most theorists of conditionals accept the Ramsey test thesis for indicatives" Bennett (2003).

[9] See, e.g., Evans and Over (2004, pp. 21–22).

[10] See, e.g., Bennett (2003) or Evans and Over (2004).

appealing to the test we do not provide any independent evidence for the Equation.

Secondly, the close parallel between the Equation and the Ramsey test, and the empirical results which established limits of the Equation, point toward possible limits of the test. As we have seen in the previous section, the Equation seems to fail for the irrelevant and negatively relevant conditionals. The situation seems to be similar in the case of the Ramsey test; considers once again a negatively relevant conditional:

(4) If he smokes, he will not develop a lung cancer.

Let us say that the lifestyle of the person in question is perfect and he does not have any genetic predispositions to developing cancer, so even in the case he smokes the probability that he will develop cancer is really low, for example 1%. In such a case, if we conduct the Ramsey test on (4) we will get the conditional probability of 99% and therefore we should believe in (4). Still, because antecedent of (4) is negatively relevant for its consequent, (4) is hard to accept. The intuition that negatively relevant indicative conditionals are defective is supported by the results of experiments that test acceptability and probability of negatively relevant conditionals (e.g., Douven and Verbrugge 2012; Skovgaard-Olsen et al. 2016b; Douven et al. 2017). This deficiency of the Ramsey test was considered, and the revised version of the test was proposed in Rott (1986).

To sum up, it seems that the intuitions behind the Ramsey test are the same intuitions that underline the Equation; therefore, appealing to the former does not provide any independent justification for the latter. Secondly, the plausibility of the Ramsey test may be restricted to positively relevant conditionals.

3.2 Triviality Proofs

Triviality proofs show that accepting the Equation leads to unacceptable conclusions. For example, the first proof from Lewis (1976) showed that we can infer from the Equation that $P(A \rightarrow B) = P(B)$ which is generally false:

(5) $P(A \rightarrow B)$
(6) $P(A \rightarrow B|B)P(B) + P(A \rightarrow B|\neg B)P(\neg B)$
(7) $P(B|A, B)P(B) + P(B|A, \neg B)P(\neg B)$
(8) $P(B)$

As we have already mentioned, the conclusion is clearly unacceptable. The two most popular ways to block the proof is to deny that conditionals are propositions (e.g., Bennett 2003; Edgington 1995) or to postulate that the meaning of conditionals depends on the beliefs of the speakers (e.g., Douven 2015; van Fraassen 1976).

The first option involves accepting NTV: that the conditionals are not propositions and are therefore not truth-apt. If conditionals are not propositions, they cannot occur in Boolean combinations; therefore, for example, we cannot use the law of total probability on conditionals, and therefore, Lewis' proof is blocked.

But how plausible is NTV? Several arguments for this view have been presented, I will discuss one of them later on and all of them were, in my opinion convincingly, countered in Douven (2015). On the oher hand, the rejection of the propositional view seems to be costly, and these costs are rarely acknowledged.

First of all, one of the consequences of NTV is that conditionals no longer have a probability. The probability of a sentence is typically understood as the probability of this sentence being true; therefore if a sentence is not truth-apt, it is also not probability-apt. Because of that, we have to replace the Equation with AT. It describes the acceptability of conditionals, and therefore, does not require them to have probabilities.

Secondly, the NTV has a problem with explaining the way conditionals are regularly used as premises in reasoning. Typically, we understood the validity of reasoning as the preservation of truth. If one of the premises is not truth-apt, there is nothing to be preserved. Therefore, NTV makes reasoning involving conditionals unexplainable, if one understands validity as truth preservation. This is an instance of the so-called Frege-Geach problem.[11] In general, the problem consists in the fact that a view that denies that expressions of a given class are truth-apt, has to explain possible occurrences of such expressions in truth-functional contexts (see, e.g., Schroeder 2008). To solve the problem one would have to propose an alternative, revisionary way of understanding the validity of reasoning. One such proposal, p-validity, was presented in Adams (1975) in which AT was also defended:

> ...an inference to be *probabilistically valid* (abbreviated p-valid) if and only if the uncertainty of its conclusion cannot exceed the sum of the uncertainties of its premises. (Adams 1998, p. 131)

This proposal on its own will not help us with our problem. As we have seen above, one of the consequences of NTV is that conditionals cannot have probability, or at least not in the sense the truth-apt sentences do,[12] therefore p-validity cannot be directly used to assess the validity of arguments with mixed conditional and unconditional promises. Perhaps we can use some proxy-quantity, in place of the probability of conditionals, to compute p-validity? There seem to be two natural candidates, acceptability and conditional probability, but neither of them is unproblematic. As quoted above p-validity is defined in terms of uncertainty. Uncertainty of a sentence, according to Adams, equals 1—probability of the sentence. In light of that, the acceptability cannot be used in computing p-validity as we have no idea if and how it relates to uncertainty. Additionally, acceptability is typically believed to have different properties than probability (therefore it can be used to avoid Lewis' trivialization), so it is not clear if we can extend the p-validity framework to incorporate acceptability. What about conditional probability? According to one of

[11] see e.g., Kölbel (1997).

[12] In fact Adams (1975) claims that this natural interpretation of probability is not applicable to conditionals. He seems to be aware of how problematic the consequences of NTV are, for example:

> "The author's very tentative opinion on the 'right way out' of the triviality argument is that we should regard the inapplicability of probability to compounds of conditionals as a fundamental limitation of probability, on a par with the inapplicability of truth to simple conditionals."

Adams (1975, p. 35).

the interpretations of the theory presented in (Adams 1975), the conditional probability differs significantly from (unconditional) probability. In light of that, someone may assume that conditionals have conditional probabilities, without having truth values or unconditional probabilities. Adams (1975) seems to be using this assumption, when analyzing cases of inference with mixed premises (e.g., antecedent restriction). His framework delivers many plausible results concerning the validity of such inferences (e.g., he shows that contraposition is not generally valid). At the same time, this approach seems to be based on questionable foundations. As discussed in (Hájek 2012), the Adams' conditional probability is in many respects dissimilar to (unconditional) probability. For example, in contrast to probability, Adams' conditional probabilities do not attach to the Boolean combination of sentences. As we have seen, p-validity was defined in terms of (unconditional) probability and, as it stands, conditional probability cannot be used when we calculate it. Additionally, given the discussed differences, it is not clear if p-validity can be easily generalized to be able to incorporate acceptability or conditional probability. This problem can be seen as a probabilistic version of the Frege-Geach problem, a probabilistic framework (e.g., Bayesianism or p-validity) that cannot accommodate conditionals that do not have a probability. Using p-validity to understand reasoning with mixed conditional and non-conditional premises is questionable if conditionals do not have truth values, and therefore probabilities.

Thirdly, accepting NTV makes it hard to make sense of conditionals embedded in truth-functional contexts like disjunction or conjunction, for example:

(9) Either he is in Rome, if he is in Italy, or he is in Bordeaux, if he is in France.[13]

According to NTV, conditionals are not the type of things that can occur in such contexts. The evaluation of the whole sentence requires its arguments to be true or false but according to NTV conditionals are neither. The defenders of AT developed elaborate ways of explaining away such

[13] Example from Kölbel (2000).

sentences (see, e.g., Edgington 1995); at the same time, others come up with new examples harder to explain away (see, e.g., Kölbel 2000). The other way to solve this problem is to provide an alternative, non-truth-functional analysis of contexts like disjunction or conjunction. Perusing this strategy may be challenging. In doing so, one not only goes against a well-entrenched understanding of logical connectives, but also for sake of completeness will have to provide a similar analysis for other truth-functional contexts in which conditionals can occur (e.g., It is true that $if A \rightarrow B$ etc.).

All these problems seem to suggest that conditionals behave as truth-apt propositions. It is also suggested by the reaction of participants of the experiment asked to assess truth values or probabilities of conditionals. They perfectly well understand both questions about truth values (see, e.g., Douven et al. 2020; Krzyżanowska et al. 2017) and probabilities of conditionals (e.g., all the articles which test the Equation) and do not seem to be confused by either of them. This is, once again, unexpected if conditionals are not propositions, consider for example asking somebody about the truth value of a question. In light of that, denying that the conditionals are propositions is both unintuitive and costly.

The second popular way to dodge triviality was explored in Douven (2015) (after van Fraassen 1976). The prove uses a generalized version of the Equation, GSH[14]:

$$\text{GSH} \quad P(A \rightarrow B|C) = P(B|A, C)$$

It was used to infer (7) from (6). Lewis derives GSH from three assumptions. The first assumption claims that the considered class of probability functions is closed under conditionalization. The second assumption is the Equation, and the third is that the interpretation of the natural language indicative conditionals does not depend on the belief states of the speaker. I will refer to this assumption as the independence assumption or IA. Both Douven (2015) and van Fraassen (1976) argue against the assumption in order to save the Equation.

[14] The Equation is sometimes called Stalnaker Hypothesis(SH), therefore its generalized version is called Generalized Stalnaker Hypothesis (GSH).

Van Fraassen believes that the source of Lewis' assumption is his metaphysical view, so-called modal realism. According to modal realism, possible worlds are real and objective in the sense in which the actual world is. If we combine modal semantics, which defines the meanings of conditionals in terms of the properties of possible worlds, with modal realism, the meanings of conditionals do not depend on our beliefs but on the objective properties of possible worlds. Van Fraassen claims that, if we adopt a less realistic notion of possible worlds, the assumption loses its appeal. If possible worlds are not objective and in some sense depend on our beliefs, then the meanings of conditionals will also depend on them.

Douven (2015) discusses the IA in more detail. He gives three arguments against it, and attacks some of the arguments, which were presented for it. I will start by discussing his three arguments:

Firstly, some of the popular and promising semantic theories proposed for conditionals suggest that IA is false. The two theories mentioned by the author are Stalnaker style modal semantics which uses the notion of similarity between possible worlds and inferentialist semantics.

Stalnaker semantics can also be interpreted in a way in which it supports IA. The realistic interpretation held, according to Van Fraassen, by Lewis is an example of such interpretation. More importantly, Stalnaker semantics is inconsistent with the Equation (see, e.g., Stalnaker 1976). Therefore appealing to it in order to attack IA and defend the Equation is not a convincing strategy.

The inferentialist semantics presented in Krzyżanowska et al. (2014) seems to be a very promising theory. Its main claim is:

Definition 1 "A speaker S's utterance 'If p, q' is true iff (i) q is a consequence-be it deductive, abductive, inductive, or mixed-of p in conjunction with S's background knowledge, (ii) q is not a consequence-whether deductive, abductive, inductive, or mixed-of S's background knowledge alone but not of p on its own, and (iii) p is deductively consistent with S's background knowledge or q is a consequence (in the broad sense) of p alone" (Krzyżanowska et al. 2014, p. 5).

If we consider this formulation, it is not clear why inferentialist semantic supports rejection of IA. The meanings of conditionals are here relative to the knowledge but not to the beliefs of the speaker. The authors explain that it would be counter-intuitive to treat as true conditionals whose consequences were inferred from antecedents with the use of false beliefs.

Douven (2015) presents a different version of the theory (see also Douven et al. 2020; Douven, Elqayam, and Krzyżanowska, this volume):

Definition 2 "A conditional is true in a given context iff the consequent follows via a number of steps from the antecedent, possibly in conjunction with contextually accepted background premises where, first, the steps are valid in deductive, inductive or abductive sense, and second the consequents does not follow (in the same generalized sense) from the premises alone" (Douven 2015, p. 38).

According to him the belief sensitivity of conditionals is imposed by this version of the semantics because the acceptability of potential background premises depends on the beliefs of the speaker or evaluator. This dependence causes the second formulation of inferentialist semantics to collide with IA, but it also makes the proposal vulnerable to the problem which motivated the phrasing of the first formulation.

If the speaker or the evaluator is liberal in accepting the background premises, for example, he accepts as premises all beliefs of the speaker, then his false beliefs can be a basis for true conditionals.

For example, let us assume that I believe that the moon is made of cheese and all my beliefs are acceptable premises for my conditionals. It is known to all of my interlocutors that I share this preposterous belief. It is easy to see that according to Definition 2 a conditional:

(10) If we bring the moon to the surface of the earth, we will end the world hunger.

Uttered by myself is true. Still, it seems to me that none of my sane interlocutors would agree to it. The fact that they know that I believe that the moon is made of cheese seems to make no difference for their assessment of (10) uttered by me. This seems to suggest that Definition

2 is too permissible in the way it relates the truth of a conditional to the beliefs of the speaker or evaluator.

Secondly, Lindström (1996) proposed rejecting IA as a way out of the so-called Gärdenfors' Paradox (Gärdenfors 1986). The paradox shows that no non-trivial belief system can at the same time satisfies both the Ramsey Test and the following Preservation Condition:

(P) If a proposition *B* is accepted in a given state of belief *K* and *A* is consistent with the beliefs in *K*, then *B* is still accepted in the minimal change of *K* needed to accept *A*. (Gärdenfors 1986, p. 82)

(P) seems to be a very natural assumption while the Ramsey Test, as we have seen, is a popular procedure for testing conditionals. Lindström shows that we can have both if we drop IA. As we have already noted, appealing to the Ramsey test, of which the Equation is a probabilistic reformulation, to defend the Equation seems not to give us a lot of additional independent evidence. Secondly, the empirical evidence concerning the effects of relevance on the probability of conditionals suggests that the intuitiveness of the Ramsey test may be limited, so despite its popularity, it may not be worth preserving.

As an independent justification for the rejection IA, Lindström presents the certeris paribus cases. These are cases in which we cease to accept a conditional after we have learned some additional evidence. An example of such a case is:

(11) If I pass today's exam, I will go for a beer afterward.

Which is true, or at least acceptable, about me. But it ceases to be the case if I learn that I have another, very hard exam tomorrow. Lindström claims that when I learn about the second exam, (11) changes its meaning. If (11) conveys the second meaning it is false while if it has the first meaning (the meaning it had before I learned about the second exam), it is, still, true. This explanation of the ceteris paribus cases seems to have an unintuitive consequence. Let us consider a discussion between me and my friend: she knows about the second exam of which I am still unaware. We disagreed about (11). According to Lindström's proposal, we talk past each other, because each of us means different things by (11). This is unintuitive.

Finally, Douven (2015) points out that similar proposals were made for different expressions (e.g., taste predicates, modal operators). This is undoubtedly true but as far as I know, neither of these proposals is uncontroversial (see, e.g., Hirvonen et al. 2019). Even if it was the case that these proposals were uncontroversial, it is not clear why their success should tell us anything about conditionals. It is possible, and maybe even plausible, that IA may be false, for example, in the case of taste predicates for reasons absent in the case of conditionals.

It seems that the postulated relativity should be reflected in the way we use conditionals. As far as I know, the only reported phenomenon which can suggest it is the so-called Gibbard phenomenon. Consider the following story:

> Sly Pete and Mr. Stone are playing poker on a Mississippi riverboat. It is now up to Pete to call or fold. My henchman Zack sees Stone's hand, which is quite good, and signals its content to Pete. My henchman Jack sees both hands, and sees that Pete's hand is rather low, so that Stone's is the winning hand. At this point, the room is cleared. A few minutes later, Zack slips me a note which says 'If Pete called, he won,' and Jack slips me a note which says 'If Pete called, he lost.' I know that these notes both come from my trusted henchmen, but do not know which of them sent which note. I conclude that Pete folded. (Gibbard 1981, p. 231)

Now according to Gibbard, if both conditionals are true, they would together with the so-called conditional non-contradiction rule:

CNC $\neg((A \rightarrow \neg B) \wedge (A \rightarrow B))$

lead to inconsistency. Both conditionals are based on true beliefs and the support for them seems to be symmetrical. Therefore, there is no reason why we should ascribe to them different truth values or judge either of them false. Gibbard concludes that both conditionals are acceptable, and the existence of such pairs is an argument for NTV. There seems to be a problem with this argument. The observation that in this situation both conditionals are acceptable is in tension with the Equation (and even more so with QAT).[15]

It is easy to see that according to the Equation, it cannot be the case that both $(A \rightarrow B)$ and $(A \rightarrow \neg B)$ are highly probable at the same time. Therefore, it is the case that two acceptable conditionals of these forms cannot have, at the same time, a high probability (>50%). That seems to show that using the example to argue for NTV to defend the Equation or AT is misguided.

The phenomenon is very controversial; many different interpretations were proposed. For example, Lycan (2003) denies that the support for both conditionals is symmetrical and therefore claims that just one of them is true. Finally, following Krzyżanowska et al. (2014), one can claim that the meaning of conditionals depends on the beliefs of the speaker. In the case described by Gibbard, it is clear that both Zack and Jack based their conditionals on different beliefs based on different evidence. Because of that, both conditionals, despite their superficial form, are not in any tension and therefore not inconsistent even when combined with CNC; they are based on different beliefs and therefore they express different relations. This interpretation of the phenomenon, in fact, supports rejections of IA.

It seems to me that it is unclear if natural language speakers are willing to accept the Gibbard-like pairs of conditionals. Even If they were, it is even less clear how to interpret this phenomenon. In light of that, this argument does not make IA significantly less plausible.

At the same time, it should be noted that rejection of IA can have potentially unwelcome consequences. For example, as noted by Lewis (1976), it is not clear whether we can explain a disagreement about conditionals if their meaning is relative in the proposed way (in line with

---

[15] It is also discussed in Jackson (1987): "When A is consistent, there is something quite generally wrong with asserting both $(A \rightarrow B)$ and $(A \rightarrow \text{not-}B)$. We cannot assert in the one breath 'If it rains, the match will be cancelled' and 'If it rains, the match will not be cancelled'. This conforms nicely with [AT]; for, by it, we have $As(A \rightarrow B) = 1 - As(A \rightarrow \text{not-}B)$, from the fact that $P(B/A) = 1 - P(\text{not-}B/A)$. Thus, the fact that $(A \rightarrow B)$ and $(A \rightarrow \text{not-}B)$ cannot be highly assertible together when A is consistent is nicely explained by [AT] as a reflection of the fact that $P(B/A)$ and $P(\text{not-}B/A)$ cannot both be high when A is consistent. Indeed, [AT] explains the further fact that $(A \rightarrow B)$ and $(A \rightarrow \text{not-}B)$ have a kind of 'see-saw' relationship. As the assertibility of one goes up, the assertibility of the other goes down."

our discussion of (11) above). It was countered by Douven (2015) that it is not necessary for the disagreement that the arguing parties interpret the proposition in question in exactly the same way. On the other hand, it seems that we should agree with Lewis that it may be hard to account for disagreement on the basis of the theory which makes the meaning of conditionals relative to opaque features of the speaker (her beliefs). As we have seen, in the case of Definitions 1 and 2 it is not clear if such explanation which does not run into other problems is available.

Finally, it seems that rejecting IA would be in tension with the Equation. The Equation claims that the probability of a conditional depends just on the conditional probability of its antecedent given its consequent and not on any other factors. If we reject IA, we claim that the meaning of a conditional and therefore its truth condition depends on some other factors, namely the beliefs of the user. If we assume that the probability of a sentence is determined by its truth condition, which seems to be a natural assumption, then it seems that meaning relativized to beliefs does not correspond well to a probability which is not explicitly relativized.

A number of other triviality proofs were proposed, for example, Carlstrom and Hill (1978), Milne (2003) or Fitelson (2015).[16] As far as I know, all of these proofs are blocked by NTV but not by rejecting IA. For example, in order to block a triviality proof from Hájek (1989), Douven has to claim that no finite model can represent a rational agent belief state (Douven 2015). Discussing the plausibility of this assumption goes beyond this scope of the paper.

It is hard to consider the triviality proofs conclusive arguments against the Equation. The two discussed ways to block the proofs, despite their problematic consequences, are available, and they are hardly the only ones (see, e.g., Bradley 2000 or Sanfilippo et al. 2020 which I will briefly discuss in the next subsection). On the other hand, as far as I know, none of these ways can be considered especially attractive and therefore the triviality proofs show, at the very least, that sticking to the Equation is costly.

Hájek (2012) argued that AT is also susceptible to a triviality proof analogous to one he presented in Hájek (1994) against the Equation.

[16] For discussion see: Hájek and Hall (1994).

He points there that a plausible conceptualization of the acceptability has to share features with probability which made it susceptible to his argument.

3.3 Truth Conditions and Probability

What is the relation between the truth conditions of a sentence and its probability? Let us start by considering sentences that are not truth-apt and therefore have no truth conditions. In such cases attributing probability to such sentences seems to be a category mistake. As we have already seen, it seems nonsensical to ascribe probabilities to questions (e.g., "Should I open the window?") or commands (e.g., "Open the window!"), uncontroversial and prototypical examples of non-truth-apt sentences. If a sentence S in question is truth-apt, as I already hinted, a natural and straightforward interpretation seems to be:

SP The probability of S is the probability of it being true.

This interpretation of the relation between semantics and probability seems to be uncontroversial to the point that, as far as I know, no alternative has been explicitly proposed.[17] SP captures the relation between the probabilities of complex sentences and their components, for example, the general probability rule for disjunction: $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$ reflects its truth conditions: (A or B) is true iff (A) is true or (B) is true.

Is the relation the same in the case of conditionals? It seems so. If we adopt the NTV view we are in the first case and, as we have already shown, we have to retreat from the Equation to AT, which does not claim anything about the probability of conditionals. Therefore SP is trivially fulfilled; no truth and no probability. Otherwise, we have to explain how it is possible that conditionals do not have truth values but have probabilities.

[17] Adams (1975) reject SP for conditionals but as far as I understand, he does not provide an alternative. At the same time, his theory is usually interpreted as describing the acceptability of conditionals rather than their probability.

Propositional semantics also adheres to SP. For example, the authors of Johnson-Laird and Byrne (2002) defend the mental model theory according to which the truth conditions of natural language conditionals are those of material implication: $(A \rightarrow B)$ is true iff (A) is false or (B) is true. Consequently, they propose a fitting probability definition: $P(A \rightarrow B) = P(\neg A \text{ or } B)$. So, the relation between semantic properties and probability of conditionals conforms to SP, and therefore the theory, despite its other well-described shortcomings (see, e.g., Bennett 2003), provides a coherent picture of truth and probability.

In light of that, it is interesting to see if there is a semantic theory that can provide a basis for the Equation, or conversely what semantic properties are suggested by it.

The best candidate seems to be trivalent semantics proposed by de Finetti. The theory is part of a more general subjective Bayesian theory of reasoning. In his de Finetti (1980)[18] he divided knowledge into three levels. *Level 0* describes the objective knowledge and is well described by the bivalent logic. *Level 1* describes categorical knowledge as possessed by humans and therefore it includes the third logical value *uncertain*, which represents a given individual being uncertain about a given sentence. Finally, *Level 2* is human knowledge represented in a graded numerical way. De Finetti's three-valued semantics for conditionals is a part of a description of *Level 1*. According to it, a conditional is true if both antecedent and consequent are true, is false if the antecedent is true and consequent is false, and it is uncertain or void if the antecedent is false. The semantics is often justified by the analogy between the conditionals and conditional bets (for more details, see Egré, Rossi, and Sprenger, this volume and Over and Cruz, this volume). A conditional bet is called off if its condition is not satisfied, similarly a conditional is void if its antecedent is false (see Table 1). The semantics is supported by the results of experiments in which participants tend to produce so-called defective truth tables, that is ones in which conditionals with false antecedents are judged to be devoid of value (see, e.g., Douven et al. 2020; Over and Baratgin 2017). On the other hand, the semantics performed poorly

[18] See de Finetti (n.d.) and Baratgin et al. (2018) for discussion.

Table 1 An example of a conditional bet and the corresponding conditional

	A	B	Conditional bet (B if A)	$(A \rightarrow B)$	$P(w^n)$
w^1	1	1	win	1	0.25
w^2	1	0	loss	0	0.25
w^3	0	1	called off	v	0.25
w^4	0	0	called off	v	0.25

in other experiments, for example, Skovgaard-Olsen et al. (2019) or Douven et al. (2020).[19]

What do these truth conditions tell us about the probability of conditionals? In the words of Over and Cruz (2018):

> The probability of the conditional *if p then q* for de Finetti is the probability that $p\&q$ holds given that the conditional makes a non-void assertion, that p holds, and this probability is of course the conditional probability of q given p, $P((p\&q)|p) = P(q|p)$.

As we see the semantics implies the Equation. But there seems to be a hidden assumption used in the derivation of probability. Consider the following example:

The probability of each of the situations (w^1, ..., w^4) is 0.25. If we use the trivalent truth conditions to calculate the probability of $(A \rightarrow B)$, we will get 0.25. The conditional is true just in w^1, it is false in w^2 and void in w^3 and w^4. So the probability that $(A \rightarrow B)$ is true equals 0.25. At the same time, $P(B|A)$ in the described situation will be 0.5. In order to equate the probability of $(A \rightarrow B)$, derived by means of the truth conditions with $P(B|A)$, we have to condition on the conditional not being void or, in other words, ignore the cases in which antecedent is false and therefore the conditional is void during the assessment of probability. Is this assumption justified? w^3 and w^4 seem to be just as legitimate cases as w^1 or w^2 and it is not clear why we should ignore them.[20]

In light of that, at the very least, it is not clear if the assumption necessary for connecting trivalent semantics and Equation is justified. Perhaps

---

[19] See e.g., Egré et al. (2019). For discussion of defective truth tables see: Baratgin et al. (2018) or Over and Baratgin (2017).

[20] See e.g., van Wijnbergen-Huitink et al. 2015.

the "void" value can be interpreted in a way that implies that a conditional does not have an objective truth value in false antecedent cases, and therefore these cases should not contribute to the calculation of its probability (see Over and Cruz 2018; Over and Cruz, this volume).

If the assumption is granted, the resulting theory has many attractive features. An example of such theory is a recent version of the trivalent semantics combined with Equation presented in Sanfilippo et al. (2020). The theory does not validate the import-export principle:

IE $P(B \to (A \to C)) = P((A \wedge B) \to C)$

assumed in Lewis' proof, and because of that, is not susceptible to this version of trivialization. Additionally, the authors show that their theory can be generalized to deal with iterated and nested conditionals. Because of these features, it is clearly a promising proposal (see also Over and Cruz, this volume, and Pfeifer, this volume). On the other hand, IE is often regarded to be plausible and therefore wanted (see, e.g., Egré et al., this volume). Secondly, as we have seen there are versions of triviality arguments that do not use the import-export principle; an example of such proof was proposed in Hájek (1989).

It seems worthwhile to consider how those theoretical considerations square with the results of psychological experiments. As we have seen, there is growing empirical evidence suggesting that the Equation holds only for the positively relevant conditionals. De Finetti semantics, combined with the discussed assumption, supports the unrestricted Equation and therefore accepting it commits us to the pragmatic explanation of results of, for example, Skovgaard-Olsen et al. (2016b). At the same time, it is unclear if and how the semantics can be modified in order to support the qualified version of the Equation. Perhaps combining the truth conditions defined by de Finetti's truth tables with the additional requirement of positive relevance would be a way to construct such a theory. As far as I know, this step has not been taken in the literature. Therefore it seems that we are dealing here with a curious situation in which empirical and theoretical considerations pull in opposite directions. The unrestricted version of the Equation is theoretically justified by the corresponding semantics, but not supported by the totality of empirical results, while it is not clear if the restricted version supported by the empirical evidence can be supported by any semantics theory.

The situation is a bit more complicated in the cases of QAT and AT. That is so because it is not clear what the relation is between the truth and the acceptability of a given sentence. In light of that, it seems that if we are to have any theoretical justification for QAT or AT, it will come from their relation to the Equation.

3.4 Probability and Acceptability

In this section, I will discuss the possible conceptual relation between all three theses.

The relation between probability and acceptability is a well-discussed topic in philosophy. The most straightforward way to relate the two notions is the Lockean Thesis[21]:

LT A proposition φ is acceptable iff the probability of φ is high.

From the Equation and LT we can deduce QAT. The intuition behind LT seems, also, to support AT. If categorical acceptance coincides with high probability then, it seems natural that, if there is something like graded acceptability, it will coincide with probability. But what if we accept the NTV and therefore deny that conditionals have probabilities? It seems that in such a case we have to reject LT in order to be still able to claim that conditionals have acceptability at all. If we endorse any other theory of acceptability[22] it seems that we are losing the theoretical basis for QAT and AT. In this place, we should also point out another controversial issue, namely the differences in our intuitions concerning the acceptability and the probability of conditionals. Results from Skovgaard-Olsen et al. (2016b) found no significant differences between assignments of acceptability and probability to conditionals made by participants. This suggests that $P(A \rightarrow B) = ac(A \rightarrow B)$. On the other hand, Douven and Verbrugge (2010) found a significant difference in the case

[21] LT seems to be quite popular, see, e.g., Foley (2009).

[22] Alternative theories are usually more complex see, e.g., Proust (2012).

of inductive and abductive conditionals. A possible explanation is that Skovgaard-Olsen et al. (2016b) used causal, non-inferential conditionals while Douven and Verbrugge (2010) used inferential conditionals. If so, it may be the case that there is a difference in intuitions concerning acceptability and probability is restricted to the inferential conditionals. It seems that more evidence should be collected in order to settle this issue. Replicating both experiments may be a good first step.

4 Conclusion

I will conclude by judging how the theses stand against the presented evidence, then I will discuss the proposed and possible alternatives to the three theses.

How do the three theses (the Equation, AT, and QAT) stand against the presented evidence? Let us start with the theoretical considerations. All three seem to be in a similar situation. There seem to be no strong theoretical arguments for any of them. The intuitions behind the Ramsey test seem to be the same intuitions that initially make the theses plausible. Therefore appealing to the test does not give us additional reasons to believe it. The Equation is supported by de Finetti's three-valued semantics, if we ignore the void cases when we consider the probability of conditionals. QAT is supported by the Equation if we accept LT and unsupported otherwise. AT seems to be, to some degree, supported by QAT.

At the same time, we have strong arguments against the Equation in the form of triviality proofs. Neither of the proofs is conclusive, given the possible ways to dodge them. On the other hand, they convinced some philosophers to abandon the Equation (e.g., Stalnaker 1976) and showed that sticking to it is costly. For example, we have to abandon IA which, as I tried to show in the third section, is plausible. A triviality argument of similar strength was also presented against AT. I am not aware of any comparable theoretical arguments against QAT.

As we have seen, all three theses were traditionally regarded as descriptively true, but the results of the empirical studies seem to paint a different picture. The situation is more complicated in the case of the

Equation than in the case of AT and QAT. QAT and AT attracted much less attention than the Equation but, as far as I know, they were not supported by the results of any of the relevant studies. AT was disconfirmed by Skovgaard-Olsen et al. (2016b) which showed that it fails in the case of the irrelevant and negatively relevant conditionals, and Douven and Verbrugge (2010) which showed that it is not supported in the case of inductive conditionals. QAT performs poorly (Douven and Verbrugge 2012) in comparison to EST.

The Equation has a long tradition of good performance in empirical studies. On the other hand, the results of Skovgaard-Olsen et al. (2016b) strongly suggest, that it fails in the cases of irrelevant and negatively relevant conditionals. The result was conceptually replicated by a few subsequent studies. At the same time, as is pointed out in Skovgaard-Olsen et al. (2016b), the experiments which confirmed the Equation did not include irrelevant on negatively relevant conditionals and therefore did not use a representative sample of conditionals. This seems to undermine them and together with results of Skovgaard-Olsen et al. (2016b) suggests that overall the unrestricted Equation is not empirically adequate. There is some evidence suggesting that the effect of relevance is pragmatic in nature (e.g., Skovgaard-Olsen et al. 2017) but different studies suggest that it is not the case (e.g., Krzyżanowska et al. 2017; Douven et al. 2017). In light of all that, it seems that we have neither theoretical nor empirical reasons for accepting the theses beyond their initial intuitiveness. Therefore, it seems that their role in the future study of indicative conditionals should be rethought.

On the other hand, I did not show that any of the theses is false. Conclusive arguments against them, as far as I know, do not exist and maybe never will. Specifically, someone impressed with the intuitiveness of any of the theses may treat it as a desideratum to be satisfied by a successful theory of conditionals. Even in such cases, the tension between them and some of the empirical findings and involved theoretical costs should remain clear.

Now we can discuss alternative proposals. I will start with the Evidential Support Theory proposed by Douven (2008). As we have seen, the core of the theory is the Evidential Support Thesis (EST):

EST An indicative conditional 'If *A*, *B*' is assertable/acceptable if and only if $Pr(B|A)$ is not only high but also higher than $Pr(B)$. (Douven and Verbrugge 2012, p. 484)

This is a counterproposal to QAT. In Douven and Verbrugge (2012), it was shown that EST predicts intuitions of natural language users much better than QAT. This is a clear advantage of EST and a good reason to prefer it over QAT. On the other hand, as it stands now, this approach also lacks theoretical justification.

EST is not supported by the Equation in a way in which QAT is and, as far as I know, it is not supported by any proposed semantics for conditionals. Perhaps further work on inferentialist semantics can provide a theoretical basis for EST.

As we have seen, EST is empirically more successful than QAT because it classifies irrelevant and negatively relevant conditionals as not acceptable. Consequently, it seems natural that users of language will judge the acceptability and the probability of conditionals as lower in such cases. Skovgaard-Olsen et al. (2016b) showed that this is true. If so, maybe we can restrict the Equation and AT to be more in line with this finding. As we have seen a restricted version of both may look for example:

Equation+/AT+ If $\Delta P > 0$ then $P/ac(A \to B) = P(C|A)$

The Equation+ and AT+ are more consistent with the available empirical evidence than the original theses. Because of the restriction, they are not undermined by the results of Skovgaard-Olsen et al. (2016b), but AT+ is still undermined by the results of Douven and Verbrugge (2010).

What about their theoretical position? Once again we lack any theoretical motivation for both theses. The situation is even worse in the case of the Equation+. There is nothing in it which would block a triviality proof analogous to Lewis' restricted to the positively relevant conditionals. The result of the proof will be that for all positively relevant conditionals $P(A \to B) = P(B)$. This is just as unacceptable as the original unrestricted result. The bottom line here seems to be that if the Equation is proposed for any kind of conditionals we can make Lewis-like argument for these conditionals. $P(A \to B) = P(B)$ is true for irrelevant conditionals, but the Equation restricted just to them would be both uninteresting and empirically inadequate (as suggested by the results of Skovgaard-Olsen et al. 2016b).

Let us move to theoretical considerations concerning conditionals. Can they point us toward a new definition of probability (or acceptability)? Triviality proofs do not give us clear help concerning the probability and acceptability of conditionals. They provide us with a purely negative lesson concerning the Equation (and AT), and it seems hard to predict which of the alternative proposals will be susceptible to analogous triviality proofs.

Perhaps a more promising and natural approach is to start with the truth conditions proposed by some of the plausible semantics, and on the basis of that, work out corresponding probability conditions. Most of the popular semantic theories postulate complex and subtle truth conditions which translate into similarly complex definitions of probability.[23] For example, if we combine, the already presented inferentialist semantics (for more on this semantics and the debate about it, see Cruz and Over, this volume, Douven et al, this volume, Over and Cruz, this volume, and Pfeifer, this volume), with SP we will get:

IP The probability of "If *p*, *q*" uttered by a speaker *S* is the probability that (i) *q* is a consequence-be it deductive, abductive, inductive, or mixed-of *p* in conjunction with S's background knowledge, (ii) *q* is not a consequence-whether deductive, abductive, inductive, or mixed-of *S*'s background knowledge alone but not of *p* on its own, and (iii) *p* is deductively consistent with *S*'s background knowledge or *q* is a consequence (in the broad sense) of *p* alone.

It is easy to see that IP is less elegant and harder to test than the Equation. At the same time, it is directly justified by the inferentialist semantics. That alone puts IP in a better theoretical position than the Equation and perhaps it is enough to make it worth further studies.

[23] As we have seen the material implication theory is an exception. It provides us with truth conditions that can be easily translated into the definition of probability. Sadly, both the definition of probability and truth conditions proposed by the material implication theory are unintuitive.

Can it accommodate the existing evidence concerning the probability of conditionals? Can we construct trivialization arguments against it or perhaps show that it is impossible? Answering those questions goes well beyond the scope of this paper. On the other hand, I hope that this example shows that there are promising alternatives to the Equation and further investigation of such alternative proposals is justified.

Acknowledgments I would like to thank Jan Sprenger, Rafał Urbaniak, Vincenzo Crupi, Andrea Iacona, Matteo Colombo, Karolina Krzyżanowska, Igor Douven, and the anonymous referees for their useful comments. I would also thank the editors of this volume, David E. Over and Stefan Kaufmann for the many suggestions which helped me to improve the paper.

Funding Information The research was supported by Starting Investigator Grant No. 640638 ("OBJECTIVITY—Making Scientific Inferences More Objective") of the European Research Council (ERC).

References

Adams, E. W. (1975). *The logic of conditionals: An application of probability to deductive logic*. D. Reidel Pub. Co.

Adams, E. W. (1998). *A primer of probability logic*. Stanford: Csli Publications.

Baratgin, J., Politzer, G., Over, D., & Takahashi, T. (2018). The psychology of uncertainty and three-valued truth tables. *Frontiers in Psychology, 9*, 1479.

Barrouillet, P., & Gauffiroy, C. (2015). Probability in reasoning: A developmental test on conditionals. *Cognition, 137*.

Bennett, J. (2003). *A philosophical guide to conditionals*. Oxford University Press.

Bradley, R. (2000). A preservation condition for conditionals. *Analysis, 60*(3), 219–222.

Carlstrom, I. F., & Hill, C. S. (1978). Book review: The logic of conditionals ernest w. adams. *Philosophy of Science, 45*(1), 155–158.

Cheng, P. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367–405.

Cruz, N., Over, D., Oaksford, M., & Baratgin, J. (2016). *Centering and the meaning of conditionals*.

de Finetti, B. (1980). *Probabilità [probability]*. Encyclopedia.

de Finetti, B. (n.d.). La logique de la probabilité. In *Actes du congrés international de philosophie scientifique*.

Douven, I. (2008). The evidential support theory of conditionals. *Synthese, 164*(1), 19–44.

Douven, I. (2015). *The epistemology of indicative conditionals: Formal and empirical approaches*. Cambridge University Press.

Douven, I., Elqayam, S., Singmann, H., & van Wijnbergen-Huitink, J. (2020). Conditionals and inferential connections: Toward a new semantics. *Thinking and Reasoning, 26*(3), 311–351.

Douven, I., Elqayam, S., Singmann, H., & Wijnbergen-Huitink, J. (2017). Conditionals and inferential connections: A hypothetical inferential theory. *Cognitive Psychology, 101.*

Douven, I., & Verbrugge, S. (2010). The adams family. *Cognition, 117*(3), 302–318.

Douven, I., & Verbrugge, S. (2012). Indicatives, concessives, and evidential support. *Thinking and Reasoning, 18*(4), 480–499.

Edgington, D. (1995). On conditionals. *Mind, 104*(414), 235–329.

Egré, P., Rossi, L., & Sprenger, J. (2019). *De finettian logics of indicative conditionals*.

Eva, B., Hartmann, S., & Rad, S. R. (2019). Learning from conditionals. *Mind*. fzz025.

Evans, J., Handley, S., Neilens, H. L., & Over, D. (2007). Thinking about conditionals: A study of individual differences. *Memory & Cognition, 35*, 1772–1784.

Evans, J., Handley, S., & Over, D. (2003). Conditionals and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition.*

Evans, J., & Over, D. (2004). *If*. Oxford University Press.

Fitelson, B. (2015). The strongest possible lewisian triviality result. *Thought: A Journal of Philosophy, 4*(2), 69–74.

Foley, R. (2009). Beliefs, degrees of belief, and the lockean thesis. In F. Huber & C. Schmidt-Petri (Eds.), *Degrees of belief* (pp. 37–47). Springer.

Fugard, A., Pfeifer, N., & Mayerhofer, B. (2011). Probabilistic theories of reasoning need pragmatics too: Modulating relevance in uncertain conditionals. *Journal of Pragmatics, 43*, 2034–2042.

Fugard, A., Pfeifer, N., Mayerhofer, B., & Kleiter, G. (2011). How people interpret conditionals: Shifts toward the conditional event. *Journal of Experimental Psychology Learning Memory and Cognition, 37*, 635–648.

Gärdenfors, P. (1986). Belief revisions and the ramsey test for conditionals. *Philosophical Review, 95*(1), 81–93.
Gibbard, A. (1981). Two recent theories of conditionals. In W. Harper, R. C. Stalnaker, & G. Pearce (Eds.), *Ifs* (pp. 211–247). Reidel.
Hájek, A. (1989). Probabilities of conditionals—Revisited. *Journal of Philosophical Logic, 18*(4), 423–428.
Hájek, A. (1994). Triviality on the Cheap? In E. Eells, B. Skyrms, & E. W. Adams (Eds.), *Probability and conditionals: Belief revision and rational decision* (pp. 113–140). Cambridge University Press.
Hájek, A. (2012). The Fall of "Adams' Thesis"? *Journal of Logic, Language and Information, 21*(2), 145–161.
Hájek, A., & Hall, N. (1994). The hypothesis of the conditional construal of conditional probability. In E. Eells, B. Skyrms, & E. W. Adams (Eds.), *Probability and conditionals: Belief revision and rational decision* (p. 75). Cambridge University Press.
Hirvonen, S., Karczewska, N., & Sikorski, M. P. (2019). On hybrid expressivism about aesthetic judgments. *Grazer Philosophische Studien, 96*(4).
Jackson, F. (1987). *Conditionals*. Blackwell.
Johnson-Laird, P., & Byrne, R. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review, 109*, 646–678.
Kölbel, M. (1997). Expressivism and the syntactic uniformity of declarative sentences. *Crítica; revista hispanoamericana de filosofía, 29*, 3–51.
Kölbel, M. (2000). Edgington on compounds of conditionals. *Mind, 109*(433), 97–108.
Krzyżanowska, K., Collins, P. J., & Hahn, U. (2017). Between a conditional's antecedent and its consequent: Discourse coherence vs. probabilistic relevance. *Cognition, 164*, 199–205.
Krzyżanowska, K., Wenmackers, S., & Douven, I. (2014). Rethinking Gibbard's Riverboat Argument. *Studia Logica, 102*(4), 771–792.
Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. *Philosophical Review, 85*(3), 297.
Lindström, S. (1996). The ramsey test and the indexicality of conditionals: A proposed resolution of gärdenfors' paradox. In A. Fuhrmann & H. Rott (Eds.), *Logic, action and information*. de Gruyter.
Lycan, W. G. (2003). Real conditionals. *Philosophical Quarterly, 53*(210), 134–137.
Milne, P. (2003). The simplest lewis-style triviality proof yet? *Analysis, 63*.

Oberauer, K., & Wilhelm, O. (2003). The meaning (s) of conditionals: Conditional probabilities, mental models, and personal utilities. *Journal of Experimental Psychology: Learning Memory and Cognition, 29*, 680–693.

Over, D., & Baratgin, J. (2017). The “defective” truth table: Its past, present, and future. In *The thinking mind: A festschrift for ken manktelow*.

Over, D., & Cruz, N. (2018). *Probabilistic accounts of conditional reasoning*.

Over, D., Hadjichristidis, C., Evans, J., Handley, S., & Sloman, S. (2007). The probability of causal conditionals. *Cognitive Psychology, 54*, 62–97.

Pettigrew, R. (2020). What is conditionalization, and why should we do it? *Philosophical Studies, 177*(11), 3427–3463.

Proust, J. (2012). The norms of acceptance. *Philosophical Issues, 22*(1), 316–333.

Ramsey, F. P. (1990). General propositions and causality. In D. H. Mellor (Ed.), *Philosophical Papers*. Cambridge University Press.

Rott, H. (1986). Ifs, though, and because. *Erkenntnis, 25*(3), 345–370.

Sanfilippo, G., Gilio, A., Over, D. E., & Pfeifer, N. (2020). Probabilities of conditionals and previsions of iterated conditionals. *International Journal of Approximate Reasoning, 121*, 150–173.

Schroeder, M. (2008). What is the frege-geach problem? *Philosophy Compass, 3*(4), 703–720.

Skovgaard-Olsen, N., Collins, P., Krzyżanowska, K., Hahn, U., & Klauer, K. C. (2019). Cancellation, negation, and rejection. *Cognitive Psychology, 108*, 42–71.

Skovgaard-Olsen, N., Kellen, D., Krahl, H., & Klauer, K. (2017). Relevance differently affects the truth, acceptability, and probability evaluations of ’and’, ’but’, ’therefore’, and ’if then’. *Thinking and Reasoning, 23*.

Skovgaard-Olsen, N., Singmann, H., & Klauer, K. C. (2016a). Relevance and reason relations. *Cognitive Science, 41*(S5), 1202–1215.

Skovgaard-Olsen, N., Singmann, H., & Klauer, K. C. (2016b). The relevance effect and conditionals. *Cognition, 150*(1), 26–36.

Spohn, W. (2012). *The laws of belief: Ranking theory and its philosophical applications*. Oxford University Press.

Sprenger, J., & Hartmann, S. (2019). *Bayesian philosophy of science*. Oxford University Press.

Stalnaker, R. (1976). Stalnaker to Van Fraassen. In W. H. C. Hooker (Ed.), *Foundations of probability theory, statistical inference, and statistical theories of science*.

Talbott, W. (2016). Bayesian epistemology. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy* (Winter 2016). Metaphysics Research Lab, Stanford University.

van Fraassen, B. (1976). Probabilities of conditionals. In W. H. C. Hooker (Ed.), *Foundations of probability theory, statistical inference, and statistical theories of science*.

van Wijnbergen-Huitink, J., Elqayam, S., & Over, D. (2015). The probability of iterated conditionals. *Cognitive Science, 39*(4), 788–803.

Vidal, M. & Baratgin, J. (2017). A psychological study of unconnected conditionals. *Journal of Cognitive Psychology, 29*(6), 769–781.

Indicative and Counterfactual Conditionals in the Psychology of Reasoning

David E. Over and Nicole Cruz

1 Introduction

Our objective in this chapter is to introduce the study of indictive and counterfactual conditionals in the psychology of reasoning, and to discuss how psychology can make progress in understanding the relationship between these conditionals in people's thinking and reasoning. There is a great deal of psychological research on both indicative and counterfactual conditionals, but to a large extent, the two research topics have not been closely connected in the psychology of reasoning. For

D. E. Over (✉)
Psychology Department, Durham University, Durham, UK
e-mail: david.over@durham.ac.uk

N. Cruz
Institute for Psychology, University of Innsbruck, Innsbruck, Austria
e-mail: Nicole.Cruz@uibk.ac.at

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_6

example, Byrne (2017) reviews recent psychological research on counterfactuals, but does not mention indicative conditionals, even in passing, while doing so (see Egré and Cozic 2016, for a logical and philosophical review that covers both conditionals). We can illustrate the importance of relating the two topics using the classical example of the difference between these two conditionals (Adams 1970):

(1) If Oswald did not kill Kennedy, then someone else did.
(2) If Oswald had not killed Kennedy, then someone else would have.

The indicative (1) and counterfactual (2) can appear to be quite different from each other, with (1) probably true, and (2) probably false. Someone who uses (2) also suggests that its antecedent is false, and that is not so of (1). We will focus in this chapter on uses of the counterfactual form that suggest, at least initially, that the antecedent is false (see Edgington 2014, on other uses). But suppose new scientific forensic evidence suddenly makes it probable that Oswald did not kill Kennedy. We could then infer, by *modus ponens* (MP), the believable conclusion that someone else did. For in this case, (2) would no longer seem improbable to us, and we would have a high level of confidence in the MP conclusion. After learning that the antecedent of (2) is probable, we would increase our confidence in (2), but what then would be the difference between (1) and (2)?

When we no longer believe that the antecedent of a counterfactual conditional is probably false, we will say that it *collapses* to an indicative conditional (Over 2020). It is an empirical question whether a given counterfactual, like (2), collapses to a given indicative, like (1), in people's judgments in circumstances like those we have described. Our example occurs in a context of dynamic inference and belief updating (Oaksford and Chater 2013, 2020). But our position is that psychologists of reasoning can throw light on the topic of how indicative and counterfactual conditionals are related to each other, by examining contexts in which people come to believe that the antecedent of the counterfactual is probably true and, more generally, update their beliefs about conditionals, changing counterfactuals into indicatives, and indicatives into counterfactuals.

Our topic is far too substantial to be thoroughly covered in a single chapter, but we can get insight into it by considering MP and other inferences in conditional reasoning that have been, and could be, studied in the psychology of reasoning. Experiments have already been run on MP and counterfactuals, which we will come to below, but these can be greatly extended, and experiments developed for other forms of conditional reasoning. Psychology should tell us, if any subject does, something about people's conditional beliefs, conditional reasoning from those beliefs, and belief updating. We will introduce, later below, the new *Bayesian approach* in the psychology of reasoning, which has the precise goal of understanding people's degrees of belief, as reflected in their subjective probability judgments, and how these beliefs are updated in dynamic reasoning (Oaksford and Chater 2013, 2020). But we will begin by making more points about how counterfactuals and indicatives are related to each other as our beliefs change.

2 The Rerunning History Hypothesis

We have given an example above of how a counterfactual can collapse, as we have put it, to an indicative conditional, as a result of learning that the antecedent is probably true. This learning happens as we go forward in time. But much previous research has focused on a hypothesis about evaluating counterfactuals by going back in time to earlier beliefs. Adams (1975) termed this proposal the "epistemic past tense" view. It is more intuitively called the *rerunning history hypothesis* (Kaufmann 2005 this volume; Pearl 2013). According to it, a counterfactual like (2) can be evaluated by returning to a corresponding indicative conditional before we came to believe that its antecedent is probably false. For the classic example, that indicative would be:

(3) If Oswald does not kill Kennedy, then someone else will.

We will say that an indicative conditional *expands* to a counterfactual when we come to believe that its antecedent is probably false. In this sense, (3) expands to become (2) as we acquire a high degree of belief that Oswald did kill Kennedy. The rerunning history hypothesis goes in the other direction. It is the claim that, to assess (2), we go back to a time before we became convinced that Oswald killed Kennedy. At that earlier time, we would judge (3) to be improbable, just as we later judge (2) to be improbable. There is some experimental support for the result of formulating this claim as a psychological hypothesis and restricting it to a type of counterfactual that could be justified by referring to a possible causal relationship (Over et al. 2007, Experiment 3; Over 2017).

Nevertheless, the rerunning history hypothesis has some intuitive counterexamples if it is not carefully qualified (Adams 1975; Edgington 2008, 2014; Kaufmann 2005). We must be careful about the beliefs we retain when we mentally rerun history. A simple example can illustrate the problem. Suppose a fair coin is being tossed, and a friend of ours is going to call "heads" or "tails" for the outcome. Consider this indicative conditional:

(4) If she calls "heads", then she will be right.

The intuitive probability of (4) is 0.5. But assume further that, in fact, she decided to call "tails", and that the coin does come up tails. Our friend has turned out to be right in calling "tails", but we make a judgment about this counterfactual:

(5) If she had called "heads", then she would have been right.

By the rerunning history hypothesis, the probability of (5), after the coin toss, is supposed to equal the probability of (4), 0.5, before the coin toss. But intuitively, the probability of (5) is 0. We could try to save the hypothesis by taking the knowledge that the coin will come up tails with us when we evaluate (4) at the earlier time, as that will make the probability of (4) intuitively 0 as well (see Kaufmann 2005, for a normative analysis designed to solve the problems with the rerunning history hypothesis). But with any such qualification of the hypothesis, we are not fully returning, in a mental simulation, to an earlier time before we came to believe the antecedent of (5) is false. People may do that as a way of processing some counterfactuals, as Over et al. (2007) confirm to some extent. But in the present example, we are apparently constructing a representation in which the antecedent of (5) is uncertain, while retaining the belief that the result of the coin toss was tails.

We might get the same effect in an experiment by asking participants to assume, at a given time, that the antecedent of (5) has become uncertain, and then asking them to evaluate:

(6) If she did call "heads", then she was right.

In (6), we have an indicative conditional with the same probability, 0, as the counterfactual (5). In the experiment, we could ask participants for their confidence in a counterfactual like (5), and to take (5) as a premise along with the supposition that the antecedent of (5) is uncertain, rather than definitely false. On this basis, we could ask them whether the indicative (6) followed in an inference, and what their confidence in (6) was as a conclusion. This kind of inference could be another way to collapse a counterfactual to an indicative conditional. It could provide further evidence that people go mentally back in time when evaluating counterfactuals, and that differences in the intuitive probability of counterfactuals just come from the different pieces of information people carry with them when they go into this mental process. An experiment like this could allow us to narrow down the factors that determine whether a piece of information will be retained or left aside when going back in time and give us a better understanding of the scope of the rerunning history analysis as a psychological hypothesis.

Going in the other direction, we could also see whether the indicative (6) expands to the counterfactual (5) by using (6) as a premise, along with the supposition that its antecedent is probably false, and then observing whether participants in the experiment inferred that the counterfactual (5) followed with the same probability as (6). These experiments could help us discover how people process the two types of conditional and relate them to each other, either collapsing or expanding one into the other, as they update their beliefs.

3 Lewis' Dialogue Technique and MP

We referred above to a form of MP in which the conditional premise is a counterfactual, but the categorical premise states that the antecedent of the conditional is in fact true or probable, rather than false. One might ask at this point how such inferences can be made sense of in a psychological experiment, when the use of the counterfactual form often suggests that the antecedent is false, and yet the categorical statement affirms it to be true. Lewis (1973, 1.7) gave us a way to deal with this problem. He used a thought experiment to support his position that MP is logically valid for counterfactuals (and not merely because its premises are inconsistent). He described a dialogue in which one speaker asserts a counterfactual, suggesting pragmatically, as Lewis held, that its antecedent is false. An apparently more knowledgeable second speaker then corrects the suggestion, by reporting that the antecedent is actually true. At this point, Lewis argued, the premises for MP fit well together pragmatically, and the conclusion of the inference intuitively follows.

One can question whether Lewis' dialogue technique really solves the normative problem he was worried about. He was trying to argue that MP is valid for counterfactuals, justifying its use as a rule of inference in his formal system. He did not illustrate his thought experiment with (1) and (2), but as we have seen above, a dialogue can cause (2) to collapse to (1). At least awkwardly for Lewis, he did use the difference between (1) and (2) to try to argue that indicative conditionals and counterfactuals are essentially different from each other (Lewis 1973, 1.1). He also used possible worlds for the semantics of counterfactuals, and not subjective probabilities or degrees of belief (see Jeffrey 1991, on de Finetti). The semantic contents of Lewis' indicative and counterfactual conditionals do not change when they have different subjective probabilities, and so it is unclear whether he could explain how (2) can become equivalent to the supposedly totally different (1) in the course of a dialogue, where degrees of belief are updated.

In any event, Thompson and Byrne (2002) used Lewis-style dialogues in experiments on counterfactuals and found that their participants happily endorsed the resulting MP inferences. An important conclusion not drawn by Thompson and Byrne from their experiments is that people do not understand a counterfactual to imply logically that its antecedent is false. We can infer from their studies of dialogues between two people that the first speaker's suggestion that the antecedent of a counterfactual is false is easily canceled by the second speaker in the dialogue, collapsing the counterfactual to the equivalent of an indicative conditional. If that were not so, the experimental participants would find the dialogue inconsistent. It is logically valid to infer any conclusion from an inconsistency in many logical systems for conditionals (Adams 1998; Lewis 1973; Stalnaker 1968), but it seems likely that ordinary people would be uncertain what to infer from a logical inconsistency in the premises for MP, and there does not appear to be evidence of this in the results of Thompson and Byrne (2002).

Some psychologists have defined a counterfactual in a way that makes it inconsistent with the truth of its antecedent. For them, it supposedly represents the negation of its antecedent as a definite fact, unqualified by a degree of belief (Johnson-Laird and Byrne 2002). The data on MP and counterfactuals shows what is wrong with doing that. We can profitably study conditionals like (2) in logic, linguistics, and psychology for some time without being able to define their class as precisely as we would hope to in the end (Adams 1975, Ch. 4; Lewis 1973, 1.1), but our first step should not be a definition of "counterfactual" that causes problems for that study. Moreover, using and understanding a counterfactual often depends on a degree of belief less than certainty that the antecedent is false. To represent such beliefs, and how they are updated over time, we can adopt the new Bayesian approach in psychology of reasoning, in which degrees of belief, and equivalently subjective probabilities, are fundamental.

4 The New Bayesian Paradigm in the Psychology of Reasoning

In traditional psychology of reasoning, the study of reasoning was usually focused on inferences from arbitrary assumptions, which were sometimes highly unbelievable (Evans and Over 2004). This traditional approach was also binary, in classifying propositions only as true or false, and inferences as only valid or invalid. In contrast, the new Bayesian approach in the psychology of reasoning (Elqayam and Over 2013; Oaksford and Chater 2007, 2020; Over 2009, 2020; Over and Cruz 2018) takes account of the fact that real-world reasoning, in everyday contexts and in science, is primarily from degrees of belief, or hypotheses that have reasonable subjective probability. In this reasoning, we do not usually accept the premises and conclusions as definitely true, but rather think of them as more or less probable. We judge the inferences to be relatively strong or weak, and not only valid or invalid, and our reasoning is often dynamic and aimed at belief updating, changing our degrees of belief, as new evidence and information is acquired and processed. Taking these points into account, the new Bayesian paradigm is developing radically different theories of reasoning, and of conditionals, than were found in the traditional binary paradigm (Oaksford and Chater 2013, 2020; Over 2020; Over and Cruz 2018).

The new approaches have taken versions of a mental process called the *Ramsey test* as a basis for evaluating conditionals. Ramsey's original proposal was restricted to indicative conditionals (Ramsey 1929/1990), but Stalnaker (1968) extended the test to counterfactuals like (2). In this version, we evaluate a conditional, *if p then q*, by hypothetically supposing *p*, while making minimal changes to preserve consistency in our beliefs, and then judging our degree of belief in *q* under this supposition of *p*. The clause about minimal changes to preserve consistency is what allows the test to be applied to conditionals with antecedents that are known, or believed to be, false. The result is that our degree of belief in the conditional, $P(\textit{if p then q})$, its believability, is the conditional subjective probability of *q* given *p*, $P(q|p)$. The Ramsey test has long been of great significance in logic and philosophy (Edgington 1995), and it is now fundamental to probabilistic accounts of conditionals (Evans and Over 2004; Over 2020; Over and Cruz 2018; Pearl 2013).

The consequences of the Ramsey test are so important in Bayesian accounts of conditionals and conditional reasoning that $P(\textit{if p then q}) = P(q|p)$ has simply been called *the Equation* (Edgington 1995; Oaksford and Chater 2007). A conditional that satisfies the Equation has been called a *probability conditional* in the literature (Adams 1998), and also a conditional event (de Finetti 1936, 1937, 1964, 1995; Pfeifer this volume). The material conditional of binary extensional logic, logically equivalent to *not-p or q*, is not a probability conditional, since P(*not-p or q*) only coincides with $P(q|p)$ in extreme cases (Gilio and Over 2012). Lewis (1976) also proved that conditionals like his (Lewis 1973) and Stalnaker's (Stalnaker 1968), which are objectively true or false at every possible world, are not probability conditionals. The probabilities of their conditionals will not often be the conditional probability (on Lewis' proof, see Douven and Dietz 2011, Edgington 1995, Evans and Over 2004, pp. 27–30; Sanfilippo et al. 2020).

In a research line going back to de Finetti's work, a *Jeffrey semantics* (Jeffrey 1991) can be given for the probability conditional *if p then q* by specifying that it is true in the case in which *p & q* is true, and false in the case in which *p & not-q* is true. In the cases in which *not-p* is true, it has the subjective conditional probability $P(q|p)$ as its value, which can vary from person to person, with different beliefs, and context to context, as in the kind of dialogues we have used as examples above. In what we are calling Jeffrey semantics, $P(q|p)$ is the *expected value*, or prevision in de Finetti's terms, of *if p then q* as a conditional assertion (Over 2020; Over and Baratgin 2017; Over and Cruz 2018; Pfeifer and Kleiter 2009; Sanfilippo et al. 2018, 2020). To take this Bayesian line is to think of truth and falsity as having epistemic utility, and $P(q|p)$ is then the expected epistemic utility of an assertion of *if p then q*, expressing someone's conditional belief. Epistemic utility, like subjective probability, can vary from person to person, and from context to context. Truth has high utility for dedicated scientists, and low utility for propagandists, and such differences could potentially be measured in psychological studies.

Do people understand conditionals in natural language as probability conditionals? In psychological research, the Equation becomes *the conditional probability hypothesis* that people will judge the probability of a natural language conditional, P(*if p then q*), to be $P(q|p)$. The first experiments on the conditional probability hypothesis used rather abstract indicative conditionals about arbitrary frequency distributions. Evans et al. (2003) asked participants to imagine a pack of cards with shapes of different colors printed on them. There were 4 cards with yellow diamonds, 1 card with a yellow circle, 16 with red circles, and 16 with red diamonds. Participants were then given a conditional like:

(7) If the card is yellow (y), then it has a circle printed on it (c).

To ensure that participants interpreted (7) as singular rather than general conditional, they were told that the conditional was about one card that was going to be randomly drawn from the pack, rather than about a set of cards (see Cruz and Oberauer 2014, on general conditionals). Participants were asked how likely (7) was to be true. By the conditional probability hypothesis, they were predicted to respond that $P(\textit{if } y \textit{ then } c)$ was $P(c|y)$, which in the example was 1/5, or 0.2. The hypothesis was confirmed for the majority of participants in this and similar experiments (Oberauer and Wilhelm 2003). There was also a minority response that $P(\textit{if } y \textit{ then } c)$ was the probability of the conjunction, $P(y \,\&\, c)$, 1/37 in our example. Such a minority conjunctive response is typically found in experiments of this type, about invented frequency distributions, but it tends to be replaced by the conditional probability response as the participants make more and more probability judgments (Fugard et al. 2011).

There is a deep theoretical correspondence between a probability conditional, which satisfies the Equation, $P(\textit{if } p \textit{ then } q) = P(q|p)$, and a conditional bet on q given p (de Finetti 1936/1995; Ramsey 1926/1990; Sanfilippo et al. 2018, 2020). Consider again the pack of 37 cards that we described above for example (7). Suppose that we randomly pick a card from this pack and assert as a conditional bet, “If the card is yellow then we bet it has a circle printed on it”. For a clear analysis, let us look at this conditional bet more formally as a kind of “lottery”. Let us also say that we know the above frequency distribution in the pack and are prepared to pay 0.2 of a Euro, 20 cents, for a “ticket” that will have the following payout conditions. We will win 1 Euro when $y \,\&\, c$ is the outcome of the random selection of the card, and we get 0 Euro when $y \,\&\, \textit{not-c}$ is the outcome and lose the 20 cents. If the card is red, and *not-y* holds, then the bet will be called off, becoming “void”, and we will get our 20 cents back. It is easy to see that the expected value of this bet is 20 cents, and that the probability of our winning the bet is the conditional probability of the *y & c* outcome given the bet is non-void, that *y* holds, which is $P((y \,\&\, c)|y) = P(c|y) = 0.2$.

If (7) is a probability conditional, we can say that (7) is "true" when *y & c* is the outcome, is "false" when *y & not-c* is the outcome, and is in some sense "void" when *not-y* holds, having no objective truth value. These non-binary truth conditions for a probability conditional parallel the non-binary payout conditions for a conditional bet. In an exact parallel with the conditional bet, the probability of (7) is the probability of the *y & c* outcome given the assertion is non-void, that y holds, which is $P((y \,\&\, c)|y) = P(c|y) = 0.2$ Psychologists have confirmed that there is this close parallel between indicative conditionals and conditional bets in people's judgments (Baratgin et al. 2013; Oberauer and Wilhelm 2003; Politzer et al. 2010). Using Jeffrey semantics, and representing "true" with 1 and "false" with 0, we can derive (Jeffrey 1991) that the expected value of the assertion of (7) in this context is 0.2.

Notice that there is a strong case for describing (7) as in some sense "void" when not-y holds (de Finetti 1936/1995; Ramsey 1929/1990). We would not look directly at a red card and, referring to it, begin a conditional by saying, "If the card is yellow …". We would instead switch to a counterfactual beginning, "If the card had been yellow, …" We can predict that, in the *not-y* outcome, the probability of (7), 0.2, would be transferred to the corresponding counterfactual (Baratgin et al. 2013, p. 312). This "switch" prediction has not yet been tested in an experiment, but it could be, using:

(8) If the card had been yellow, then it would have had a circle printed on it.

Probability judgments about (8) could be compared with probability judgments about (7) as expressed before the draw. There are also experimental techniques for making the outcomes of random draws uncertain (Baratgin et al. 2013). Using these techniques would allow us to study the relation between (7), (8), and this past tense indicative:

(9) If the card was yellow, then it had a circle printed on it.

Such possible extensions to the original experiments could give us some insight into the relation between indicatives and counterfactuals and the rerunning history hypothesis about the relation between the two kinds of conditional, when these are about frequency distributions.

5 Conditionals and Relations

Over et al. (2007) studied "causal" indicative and counterfactual conditionals that could be assessed for probability on the basis of evidence about causal relations, rather than by given frequency distributions, as in the example about cards above. A counterfactual that could be justified by referring to a causal relation between unemployment and homelessness is:

(10) If unemployment had risen, then homelessness would have increased.

For each conditional *if p then q* in the experiments, participants gave their judgment about $P(\textit{if } p \textit{ then } q)$, and their judgments about the four logical possibilities, $P(p \mathbin{\&} q)$, $P(p \mathbin{\&} \textit{not-}q)$, $P(\textit{not-}p \mathbin{\&} q)$, and $P(\textit{not-}p \mathbin{\&} \textit{not-}q)$, which were to sum to 100%. In Experiment 3 of this article, focused on counterfactuals, participants gave their judgments about the four logical possibilities at a time five years in the past. An implicit conditional probability judgment, $P(q|p)$, could be derived for each participant by using $P(p \mathbin{\&} q)/(P(p \mathbin{\&} q) + P(p \mathbin{\&} \textit{not-}q))$, and that implicit $P(q|p)$ compared with the explicit $P(\textit{if } p \textit{ then } q)$. The results confirmed the conditional probability hypothesis for counterfactual conditionals like (10), and there were no significant conjunctive responses. There was also initial evidence for the rerunning history hypothesis about counterfactuals, which we referred to above, because the judgments about the four possibilities concerned a time five years earlier than the counterfactual assertion.

Over et al. (2007) also found support for the conditional probability hypothesis in their first two experiments, for present tense indicative conditionals like, "If global warming continues then London will be flooded". But they did not study past tense indicative conditionals, such as:

(11) If unemployment rose last year, then homelessness increased.

We would expect that (7) to (9) would not be significantly different in probability, and similarly for (10) and (11). We would also expect a reply to the assertions of (8) and (10), in dialogue, with the minor premise for MP, not to have a significant effect on the probabilities of (8) and (10). Whether these intuitions would match those of participants in an experiment is an open question, but studying the possible differences between conditionals like (7), (8), (9), (10), and (11), on the one hand, and (1) and (2) on the other, should lead to a greater understanding of the similarities and differences between indicatives and counterfactuals.

There is evidence, in some recent psychological research, that people's judgments of the probability of indicative conditionals, *if p then q*, deviates from the corresponding conditional probability when there is no causal or epistemic relation between *p* and *q* (but see Pfeifer, this volume). Skovgaard-Olsen et al. (2016a, b) have results supporting this conclusion about a possible qualification, or limitation, of the conditional probability hypothesis. They distinguish between positive relevance, when $P(q|p) > P(q|not\text{-}p)$, and *p* is a reason to increase belief in *q*, negative relevance, when $P(q|not\text{-}p) > P(q|p)$, and *p* is a reason to decrease belief in *q*, and irrelevance, when $P(q|p) = P(q|not\text{-}p)$. Irrelevance produces a *missing-link conditional* (Douven 2016, 2017). Without positive relevance, people tended to judge that $P(if\ p\ then\ q)$ was lower than $P(q|p)$. That is, in these experiments, people assigned a lower probability to *if p then q* when *p* did not raise the probability of *q*. The underlying criterion here is delta-*p*: $P(q|p) - P(q|not\text{-}p)$. When delta-*p* is positive, *p* raises the probability of *q*, and there is positive relevance. When delta-*p* is 0, there is irrelevance and a missing-link, and when delta-*p* is negative, there is negative relevance. There is a *relevance effect* when *p* does not have positive relevance for *q*, and $P(if\ p\ then\ q)$ is less than $P(q|p)$.

It is plausible that people would expect a speaker to be able to support a counterfactual like (10) by reference to a causal relation, which would make delta-*p* positive. They would interpret the speaker as intending to convey that such a causal relation exists, and their probability could be about the existence of this relation. For example, people who doubt that unemployment causes homelessness might see (10) and (11) as missing-link conditionals. They might infer as well, pragmatically, that a speaker who asserts (10) or (11) has an unlikely, as they believe, causal link in mind, and their low probability judgment could be about the existence of that link.

We have just suggested a pragmatic explanation of why missing-link conditionals can be assigned a relatively low probability in experiments, producing a relevance effect. Cruz et al. (2016) and Lassiter (2022) make further pragmatic points about missing-link conditionals. We could make similar suggestions about conditionals where there is negative relevance, but we will mostly restrict our points here to missing-link conditionals. Consider the stronger theory, which we will call *truth condition inferentialism*, that a missing-link conditional is not only pragmatically misleading when asserted, but not true, because the truth conditions for *if p then q* require a deductive, or a sufficiently strong inductive or abductive, connection between *p* and *q* (Douven et al. 2020; Mirabile and Douven 2020). These inferentialists are uncertain whether a missing-link conditional is false or neither true nor false (Douven et al. 2020). But in any case, some of the examples of missing-link conditionals used to try to support the theory are so pragmatically bizarre, "If raccoons have no wings they cannot breathe under water" (Krzyżanowska et al. 2017), that it is hard to see what relevance they can have to understanding ordinary discourse, where non-bizarre missing-link conditionals are not uncommon. To investigate these uses more fully, we suggest studying a wider range of linguistic forms, conditional and non-conditional, and including more realistic examples.

Cruz et al. (2016) found that participants' probability judgments for missing-link indicative conditionals were reduced to a degree similar to their probability judgments for conjunctions and disjunctions with missing links. This indicates that the effect of a missing-link on probability judgments cannot be specific to conditionals. And if it is not specific to conditionals, then how can it be a central, let alone the sole, criterion for the meaning of conditionals? The finding that the effects of missing links are not specific to conditionals suggests that the role of a missing-link is not to determine the truth conditions of a statement, conditional or not, but that its impact on probability judgments lies elsewhere. One could hardly argue, for example, that a missing-link conjunction, *p & q*, was false, or neither true nor false, when *p* and *q* were both true. Future studies could also include missing-link counterfactuals, and other conditional speech acts such as conditional promises (see also Over 2020, on missing-link disjunctions).

Consider an example of a conditional promise. A father promises his daughter that, if she gets a distinction in her piano exam, he will buy her a new bike, when he intends to buy her the bike in any event, because she has already been working hard for the exam. He always rewards hard work, and not merely getting the highest grades, and for participants, in an experiment, who knew these facts about him, delta-*p* would be 0. Suppose the daughter gets the distinction, and the father buys her the bike. Then the father would seem to be "as good as his word", and it appears highly unlikely to us that the participants would judge his original promise to have been untrue and a lie, because of their confidence in the perfectly good, in our view, missing-link counterfactual that, if his daughter had not got the distinction, he would have bought her the bike anyway.

There is strong experimental evidence that the problem with a missing-link conditional is not that it is false, or neither true nor false (Wang et al. 2022). Skovgaard-Olsen et al. (2017) ran a truth table study in which participants were asked, for cases of positive relevance, irrelevance, and negative relevance, whether *if p then q* was true or false when *p* and *q* were both true. For all three degrees, they responded that *if p then q* is "true" in that truth table row (Skovgaard-Olsen et al. 2017, Fig. 1). Truth condition inferentialism implies, contrary to these findings, that the participants should have responded, for that row, with "false" for *if p then q* when there was negative relevance, and with "false", or "neither true nor false", when there was irrelevance.

Truth condition inferentialism allows *if p then q* to be "true" when *p* is true and *q* is false, making MP an "invalid" inference form. It proposes that a conditional *if p then q* is true if there is a deductive, inductive, or abductive relation between *p* and *q*, but *p* can be true and *q* false when there is an inductive or abductive relation between *p* and *q*. As Mirabile and Douven (2020) remark, their "… account is consistent with there being true conditionals with a true antecedent and a false consequent …" (see also Krzyżanowska et al. 2014). Inductive and abductive arguments provide only inconclusive grounds for a conclusion, and consequently such grounds can exist when *p* is true and *q* is false. However, the results of Skovgaard-Olsen et al. (2017) show that the participants judge *if p then q* false when *p* is true and *q* is false for all degrees of relevance (Skovgaard-Olsen et al. 2017, Fig. 1). More generally, MP is among the most highly endorsed inferences in the psychology of reasoning (Evans and Over 2004), and supporters of truth condition inferentialism have not produced an example in which *if p then q* is intuitively true when *p* is true and *q* false, where neither *p* nor *q* contain conditionals.

Let us consider:

(12) If Lincoln was assassinated, then John Wilkes Booth did it.

The evidence that Booth caused Lincoln's death by shooting him is extremely strong, and so by the inferentialist truth conditions, (12) could be "true" even if Lincoln was assassinated and yet Booth did not do it. In this way, MP would not always preserve truth and so would be "invalid". But we do not see how (12) can be true if Booth was innocent, and someone else was responsible for the assassination.

There are contexts in which missing-link conditionals are pragmatically felicitous, such as when a fair coin is tossed to determine who will serve first in a tennis match. Consider (4) again, about calling "heads" while a fair coin is tossed. This is a missing-link conditional: there is no relation between calling "heads" and the coin landing heads. Truth condition inferentialism implies that (4) should be as unacceptable as any other missing-link conditional. Yet intuitively the probability of (4) is 0.5, and we would expect participants in an experiment to agree with that. There would be no relevance effect in this case, and truth condition inferentialists themselves point out that explaining people's probability judgments about conditionals is a challenging problem for them (see

Douven et al. 2020; Mirabile and Douven 2020, who acknowledge some of the problems and open questions in their theory).

The results of Skovgaard-Olsen et al. (2019) provide evidence, for their materials, that the relevance requirement in conditionals is a conventional implicature, which can be seen as an aspect of semantic meaning over and above truth conditions (see also Skovgaard-Olsen 2020). But conventional inferentialism is like truth condition inferentialism in having a problem with conditionals like (4) and with *non-interference* conditionals, such as the following example from Douven (2016):

(13) If we triple her salary, Bettie will leave the Department.

Supporters of conventional inferentialism also must explain why their account does not apply to conditionals like (13). Some inferentialists have gone so far as to deny that non-interference conditionals are conditionals, dismissing them as "unconditionals" (Douven et al. 2020). But (13) looks like an acceptable conditional to us, as does this linked counterfactual:

(14) If we had tripled her salary, Bettie would have left the Department.

Research on inferentialism is further limited by the paucity of studies in which the *polarity* of the conditionals is manipulated, with AA (affirmative antecedent, affirmative consequent), AN (affirmative antecedent, negated consequent), NA, and NN examples (but see Over et al. 2007). Such a polarity manipulation can produce many missing-link examples that are hard to dismiss as "unconditionals". Here is an NA example:

(15) If our friends are not at Westminster Abbey (*not-w*), then they are in London (l).

A speaker using (15) might believe that the friends have reason to be in London that has nothing to do with being, or not being, at the Abbey. Note as well that delta-p is (slightly) negative for this use of (15), with $P(l|w) = 1$ greater than $P(l|not\text{-}w)$, which we can imagine is high but less than 1. Yet (15) is not at all infelicitous or odd, and we would make the same points about the counterfactual corresponding to (15). An AN example would be, "If you turn the key, the car will not start", used when the car is out of fuel. Inferentialists have to rule out these examples as "unconditionals", but that stipulation decreases the falsifiability of their theory, if it does not remove it completely (Cruz et al. 2016; Lassiter 2022). They have to say more than that their theory applies to all uses of "if" that do not falsify it.

A related point on falsifiability concerns the finding that the believability of a conditional is affected by the believability of its consequent, regardless of the presence of a deductive, inductive, or abductive relation between the antecedent and consequent (Douven et al. 2018, 2020). The conditional probability hypothesis implies this finding, because $P(q|p)$ will be high when $P(q)$ is high and q is independent of p. But puzzlingly, this absence of an effect of an inferential connection on people's judgments about conditionals was interpreted by some inferentialists (Douven et al. 2018, 2020) as evidence in favor of their theory. Their argument is that the absence of the effect of an inferential connection in conditionals is like *belief bias* in inferences, which is the tendency to be affected by the believability of the conclusion of an inference when making a judgment about its deductive validity (Evans and Over 2004). But it is a mystery to us how this evidence against an effect of inferential connections—deductive, inductive, or abductive—can be recast as evidence for the essential place of these connections in people's understanding of conditionals. The conditional probability hypothesis is directly supported by the "belief bias" results without auxiliary hypotheses and analogies.

A further important problem for inferentialism, which limits the extent to which it can be subjected to empirical test, is that it does not specify a logic, either for its "real" conditionals or its "unconditionals", although truth conditional inferentialism seemingly implies the highly counterintuitive (as we see it) position that MP is "invalid" for the former, but valid for the latter. The role of relations in conditionals has a clear formal specification in other accounts. For example, Stalnaker's extended Ramsey test can be applied to (13), (14), and (15). These are not "unconditionals" for him, but perfectly good conditionals, as they are for Lewis (1973). Crupi and Iacona (in press) and Rott (2019) analyze conditionals in the context of theories of explanatory power, without committing to a particular theory of the semantics of conditionals. But the probability conditional has a fully specified formal theory of conditional reasoning, as we will explain in the next section. It has its own logical system that is sometimes called *System P* (Gilio 2002), and Adams (1998) proves that it is sound, complete, and decidable.

6 The Probability Conditional and Inference

As we have emphasized above, the new probabilistic approach in the psychology of reasoning recognizes that most ordinary and scientific reasoning is from degrees of belief. In our decision making, we reflect on possible actions that we have some confidence we can put into effect. Counterfactual conditionals are evaluated by supposing what we do not believe, but they enable us to infer what follows from the supposition and our background beliefs, and to improve our decision making in similar antecedent conditions in the future. We also assume what we do not believe for reductio ad absurdum arguments, with the aim of rejecting the assumption by deriving an inconsistency from it, but this derivation will usually depend on background beliefs. The central place of belief in human reasoning is fully recognized by the new Bayesian approach. To develop it in the psychology of reasoning, we need to generalize the binary logical notions of consistency and validity, and this is accomplished with the concepts of coherence and probabilistic validity (see Over and Cruz 2018; Over 2020, on how these concepts are used in the psychology of reasoning).

Degrees of beliefs are said to be *coherent* if and only if they conform to probability theory. For example, it is binary inconsistent to hold that $p \,\&\, q$ is true (of value 1) but p is false (of value 0), but more generally, it is incoherent to judge that $P(p \,\&\, q) > P(p)$. A one-premise inference is binary valid when it preserves truth, and it is *probabilistically valid*, *p-valid*, if and only if its conclusion cannot coherently have a lower probability than its premise. Consider &-elimination, inferring p (or q)

from $p \mathbin{\&} q$. It is binary valid to infer p from $p \mathbin{\&} q$ because p cannot be false (value 0) and $p \mathbin{\&} q$ true (value 1), and it is p-valid to infer p from $p \mathbin{\&} q$ because $P(p)$ cannot be coherently lower than $P(p \mathbin{\&} q)$. It is binary valid and p-valid to infer the material conditional, *not-p or q*, from *not-p*, but it is not p-valid to infer the probability conditional, *if p then q*, from *not-p*, because $P(q|p)$ can be coherently lower than $P(\textit{not-p})$.

To define p-validity for more than one premise, let the uncertainty of p be one minus its probability, $1 - P(p)$. An inference is then p-valid if and only if the uncertainty of its conclusion cannot be coherently greater than the sum of the uncertainties of its premises (Adams 1998). More informally, a p-valid inference is one that cannot possibly increase our uncertainty. System P is a precise set of inference rules for p-validity. Two further examples of p-valid inferences are MP and the and-to-if inference, which is inferring *if p then q* from $p \mathbin{\&} q$. We shall call the and-to-if inference *one-premise centering* (see Lewis 1973, on centering as a semantic condition). That one-premise centering is p-valid for the probability conditional, $P(\textit{if } p \textit{ then } q) = P(q|p)$, follows immediately from the fact that $P(p \mathbin{\&} q) = P(p)P(q|p)$, and there is experimental evidence that people treat this inference as valid, at least for indicative conditionals (Cruz et al. 2015, 2016).

Precise coherence intervals can also be derived for p-valid and p-invalid inferences (Cruz 2020; Evans et al. 2015; Pfeifer and Kleiter 2009). For &-elimination, $P(p)$ must fall in the interval between $P(p \mathbin{\&} q)$ and 1 to be coherent, and for the p-invalid inference of inferring $p \mathbin{\&} q$ from p, $P(p \mathbin{\&} q)$ must be in the interval between 0 and $P(p)$ to be coherent. Two coherence intervals of special importance for our discussion of indicative and counterfactual conditionals are those for MP and *two-premise centering*, inferring *if p then q* from p and q as separate premises. As an example, consider the following as a probability conditional:

(16) If Booth smoked cigars (s), then he damaged his lungs (d).

We do not know whether Booth smoked cigars or not, but cigar smoking was popular in his day, and we judge that $P(s) = 0.5$. We have a causal model relating cigar smoking to lung damage and infer from it that $P(d|s) = 0.8$. From these two judgments, we can derive a coherence interval for MP, inferring d from *if s then d* and s, by using the total probability theorem from probability theory:

$$P(d) = P(s)P(d|s) + P(not\text{-}s)P(d|not\text{-}s)$$

To be coherent, we must judge that $P(not\text{-}s) = 1 - P(s) = 0.5$. But $P(d|not\text{-}s)$ could be completely open for us. It might be that $P(d|not\text{-}s)$ is high because, if Booth did not smoke cigars, then he smoked a pipe, or it might be that $P(d|not\text{-}s)$ is low because, if Booth did not smoke cigars, he did not smoke at all. In any event, $P(d|not\text{-}s)$ cannot be coherently higher than 1 or lower than 0. Therefore, for coherence, the conclusion of MP, $P(d)$, cannot be lower for us than $P(s)P(d|s)$, which is 0.4, or higher for us than $P(s)P(d|s) + P(not\text{-}s)$, which is 0.9. And that is the coherence interval for MP in this case, assuming that (16) is a probability conditional.

For two-premise centering, we observe that this reverses the direction of MP. Instead of beginning with the two premises, *if s then d* and s, to infer d using MP, we begin with the two premises, s and d, to infer *if s then d* by two-premise centering. Suppose then we start off with the judgments that $P(s) = 0.5$ and $P(d) = 0.6$. We now wish to infer, coherently, a judgment about $P(d|s) = P(s \;\&\; d)/P(s)$. The minimum overlap of $P(s)$ and $P(d)$ is $P(s) + P(d) - 1$, which is 0.1. This is the minimum value that $P(s \;\&\; d)$ can have in this example, and so the minimum value of $P(d|s) = P(s \;\&\; d)/P(s)$ is $0.1/0.5 = 0.2$. If $P(s) + P(d) - 1$ had been a negative number, leaving open the possibility that $P(s)$ and $P(d)$ did not overlap at all, then the minimum value of $P(d|s)$ would have been 0. The maximum value of $P(d|s)$ in this example is clearly 1, since $P(d)$, at 0.6, is greater than $P(s)$, at 0.5. If $P(d)$ had been less than $P(s)$, then the maximum value of $P(d|s)$ would have been $P(d)/P(s)$. In sum, the coherence interval for two-premise centering, in the example, is a minimum value of 0.2 and a maximum value of 1, still assuming of course that $P(\textit{if s then d}) = P(d|s)$. There is a formal account of two-premise centering in Sanfilippo et al. (2018), and a psychological study of it in Cruz et al. (2015), showing that participants endorse it as a valid inference for indicative conditionals. A current limitation of this work is that it has yet to specify the psychological mechanisms that allow people to make coherent judgments and to distinguish between those inferences that preserve truth and probability, from their premises to their conclusion, and those that do not.

Clearly, MP and two-premise centering are very closely related to each other for the probability conditional. Staying with our last example, suppose that we made a two-premise centering inference with $P(s) = 0.5$ and $P(d) = 0.6$, but with the incoherent conclusion that $P(\textit{if } s \textit{ then } d) = P(d|s) = 0.1$. Now turn this round and see it as an MP inference, with premises $P(d|s) = 0.1$ and $P(s) = 0.5$, and the conclusion $P(d) = 0.6$. The maximum coherent value that $P(d)$ can have, given those premises, and the coherence interval for MP, is $(0.1)(0.5) + 0.5 = 0.55$. Since our MP conclusion is that $P(d) = 0.6$, we are incoherent. This incoherent outcome illustrates a point about the probability conditional. If the conclusion of a two-premise centering inference is incoherent, then the corresponding MP conclusion will be incoherent. In general, given three judgments, $P(\textit{if } p \textit{ then } q) = P(q|p) = x$, $P(p) = y$, and $P(q) = z$, we can test their coherence either by arranging them as an MP inference, or as a two-premise centering inference, and then checking whether the conclusion, $P(q) = z$, or $P(\textit{if } p \textit{ then } q) = P(q|p) = x$, is in the MP, or in the two-premise centering, coherence interval, respectively. There are other ways to check the coherence of the three judgments, but we can exploit the connection between MP and two-premise centering to extend experiments on MP for indicative and counterfactual conditionals to studies of two-premise centering, for the purpose of understanding the relationship between the two conditionals.

People may switch back and forth between centering and MP to achieve coherence in their conditional reasoning, with indicative conditionals and counterfactuals replacing each other in a dynamic process of belief updating. Suppose for example that people acquire some confidence that Booth did not smoke cigars. They may then expand the indicative (16) to the counterfactual:

(17) If Booth had smoked cigars, then he would have damaged his lungs.

Let us now imagine that this dynamic process continues. They learn that that Booth probably did smoke cigars, and he had damaged lungs, and (17) collapses back to (16). Given high values for $P(s)$ and $P(d)$, people may use a procedure grounded in centering to infer an even higher probability value for (16). But because the coherence interval for two-premise centering is relatively complex, they might then go the other way and check the coherence of this high value by relying on the coherence interval for MP, with (16) as the major premise, *s* as the minor premise, and *d* as the conclusion. If they were within this MP coherence interval, they could be happy with the even higher probability for (16).

For a more complete picture of the relation between indicative and counterfactual conditionals, there must be an account of how people make probability judgments about counterfactuals like (17), and those in Over et al. (2007), which can be justified by referring to causal relations. Some recent research has employed causal Bayes nets for this purpose, and the next section will be about this research.

7 Bayesian Networks

We have described how people may switch back and forth between MP and centering inferences to make their judgments of $P(p)$, $P(q)$, and $P(q|p)$ coherent. We described how, to evaluate a counterfactual like (2), people may sometimes "rerun history" and engage in a mental simulation in which they go back in time to when it was uncertain whether Kennedy would be assassinated, allowing them to reduce the evaluation of the counterfactual (2) to that of (3), the corresponding indicative conditional. And we have pointed out how a counterfactual can collapse to an indicative conditional because of a dialogue in which the degree of belief in the minor premise of MP is updated. These assessments and changes in subjective probability judgments over time depend on dynamic reasoning, and the advent of the Bayesian approach has greatly extended the tools available in psychology for modeling this reasoning (Oaksford and Chater 2013, 2020).

Bayes' theorem relates a set of probabilities to one another at a fixed point in time, here at time 0:

$$P_0(q|p) = (P_0(p|q)P_0(q))/P_0(p)$$

But suppose that, in a process of belief revision, we learn p is true as we move from the initial state to time 1, $P_1(p) = 1$. We can then use *simple Bayesian conditionalization*, in dynamic reasoning, to infer that $P_1(q) = P_0(q|p)$, provided that $P_0(q|p)$ has remained invariant in this process, $P_0(q|p) = P_1(q|p)$. If at time 1 we do not learn p for certain but only with some lower probability, we can generalize simple Bayesian conditionalization to *Jeffrey conditionalization*:

$$P_1(q) = (P_0(q|p)P_1(p) + P_0(q|not\text{-}p)P_1(not\text{-}p)$$

This generalization can be justified by the total probability theorem with, again, the added assumption that invariance holds in the conditional probabilities. How closely people conform to these principles of Bayesian dynamic reasoning has been of interest to psychologists, with a growing number of findings showing people follow the principles more often than expected by chance. They also sometimes deviate from them when, for instance, this requires the integration of larger pieces of information or the processing of negations (Bramley et al. 2017; Cruz et al. 2020a, b; Hadjichristidis et al. 2014; Madsen et al. 2020; Oaksford and Chater 2013; Pilditch et al. 2019; Zhao and Osherson 2010).

The Bayesian approach can be powerfully extended by the use of *Bayesian networks* (Pearl 1988) to represent causal relations and the counterfactuals that can be justified by reference to them. Examples of basic building blocks of such networks, in which there are three variables that are connected in different ways, are shown in Fig. 1.

Bayesian networks are directed acyclic graphs, which means that their arrows go in only one direction: an arrow from p, to q, to r, and so on, will not form a loop leading back to p. The *variables*, or *nodes*, p, q, and r above, can take a range of values. We will assume here that these nodes are propositions that take the values of truth, 1, or falsity,

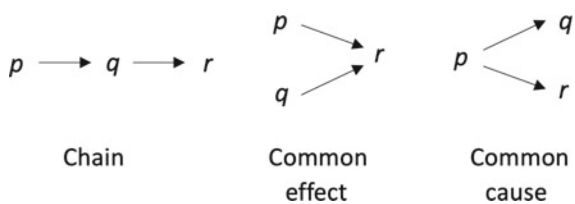


Fig. 1 Basic building blocks of Bayesian network structures

0. Nodes from which arrows start are called parent nodes, and nodes to which arrows are directed are called child nodes. Every node is associated with a probability table. For instance, consider a minimal network $p \rightarrow q$, which could be used to represent a causal relation between p and q and for the evaluation of the counterfactual, *if p were the case, then q would be the case*. There is one node here without a parent, p; the table for it would simply give the probability of p, $P(p)$. There is also one node here with a parent, q, and the table for it would list the conditional probabilities of the truth of the node given each possible combination of the truth or falsity of its parent. For $p \rightarrow q$, these probabilities would be $P(q|p)$ and $P(q|not\text{-}p)$.

An important underlying assumption in Bayesian networks is the *Markov condition*. It states that a given node is conditionally independent of nodes further upstream in the network, given the values of its parent nodes. In other words, a node p can only influence the value of another node q via the nodes that have direct incoming connections to q, so that if the nodes with direct connections to q are held constant, any nodes further upstream from q cease to have an influence. A related assumption is that a network makes explicit all relevant probabilistic dependencies between nodes, so that an arrow connects two nodes if and only if there is a probabilistic dependency between them. These assumptions make it easy to keep track of the pattern of probabilistic dependencies between variables, and they enable us to make precise quantitative predictions about how a change in the value of one variable "propagates" through the network affecting the values of other variables it is connected to. People's causal reasoning has sometimes been found to violate the Markov condition. But there is evidence that this is because people elaborate on the given causal information, based on knowledge about causal mechanisms, and that adherence to the Markov condition is stronger when this causal knowledge is taken into account (Park and Sloman 2013; Rehder and Burnett 2005).

Bayesian networks make it possible to represent causal relations between events, and the associated counterfactuals, in a formally precise way (Spirtes et al. 1993; Pearl 2000). The *strength* of these causal relations can be quantified using the relative values of $P(q|p)$ and $P(q|not\text{-}p)$. Psychologists have been interested in how far they can use these networks to model people's causal reasoning (Cheng 1997; Fernbach et al. 2010; Hayes et al. 2014; Holyoak and Cheng 2011; Griffiths and Tenenbaum 2009; Oaksford and Chater 2017; Pilditch et al. 2019; Sloman and Lagnado 2015; Waldmann 2017). There has also been psychological interest in whether the networks provide tools to help individuals and organizations make decisions about complex real-world causal scenarios (Cruz et al. 2020a, b).

The information in Bayesian networks, as directed acyclic graphs, and the associated tables can also be given in structural equations, where the relationships between the variables are formally written out (Pearl 2000). It is more plausible psychologically to hypothesize that people have mental representations more like Bayesian networks than structural equations, but the two formal representations, the graphs and the equations, are formally equivalent. Pearl (2013) argues for the use of structural equations to implement the Ramsey test (which he calls Ramsey's idea). But let us, returning to (17) as an example, assume that a simple network with an arrow from cigar smoking to lung damage can model people's knowledge of how the former can cause the latter.

For the counterfactual (17), people might well not have a view on whether Booth was a cigar smoker or not, and they could implement the Ramsey test in the following way. They could set the s node, or variable, at 1 or "true" in their Bayesian network representation and take the associated $P(d|s)$ as their confidence in (17). As Pearl (2013) would put it, this dynamic progress would be like rerunning history, under different conditions, to work out the effects of cigar smoking on Booth's lungs. This process would also be very similar to, if not identical with, the evaluation of the indicative (16). In both cases, we could interpret "rerunning history" as making an *intervention* to assign one or more nodes of the network a specific value and seeing what would unfold over time by inferring the consequences of those network changes (Kaufmann 2005; Sloman and Lagnado 2005; Rips 2010; Rips and Edwards 2013).

People's use of a Bayesian network for (2) could not be so simple. They would have to intervene to cut off arrows leading into the node representing Oswald's killing of Kennedy, changing the value of this variable from 1, or true, to 0, or false, while considering also other possible causes of Kennedy's assassination (Cruz et al. 2018; Fernbach et al. 2010). Yet people can make judgments about counterfactuals, like (2), embedded in complex causal scenarios. They would also have to go through a complex process to revise their actual beliefs if, as we imagined above, new forensic evidence demonstrated that Oswald did not in fact kill Kennedy, and (2) collapsed to (1).

There is active research in cognitive psychology on whether the Bayesian network analysis of conditional reasoning accurately describes people's actual conditional reasoning (Ali et al. 2011; Cruz et al. 2018; Fernbach and Erb 2013; Lucas and Kemp 2015; Sloman and Lagnado 2005; Rips 2010; Rips and Edwards 2013). More generally, Bayesian network representations provide a general framework that could help connect different findings about indicative and counterfactual conditionals and probabilistic reasoning, which were previously studied almost entirely in isolation from each other (Adams 1998; de Zoete et al. 2019).

8 Concluding Remarks on Truth

The probability conditional and its Jeffrey semantics fit well with the use of a Bayesian network representation (compare Oaksford and Chater 2020). The expected value of the assertion of a probability conditional, *if p then q*, is the conditional subjective probability, $P(q|p)$, which can be inferred from a Bayesian network. Judgments about $P(q|p)$ can vary from context to context, and person to person, and such changes could be represented in corresponding networks. Not all conditional assertions have causal justifications, and not all Bayesian networks can be usefully interpreted in a causal manner. For example, a network representation for indicative (7) and counterfactual (8), about a random card from a frequency distribution, could be purely probabilistic. There is not a causal relation between a card being yellow in the pack and having a circle on it: there is only the frequency distribution.

As we have noted above, Lewis (1976) proved that the probability of a conditional that is objectively true or false at every possibility cannot, in general, be the corresponding conditional probability. But a probability conditional *if p then q* is not objectively true or false at every possibility, as its value is the subjective conditional probability, $P(q|p)$, when *not-p* holds. Lewis' proof does not apply to a probability conditional (Sanfilippo et al. 2020). There is a long history of findings, in studies of the so-called "defective" truth table, that people do not judge conditionals like (7) to be true or false when their antecedents are false (Evans and Over 2004; Over and Baratgin 2017). There does not indeed appear to be an objective fact to make (7) and (8) true or false, rather than more or less probable, when a red card is randomly selected out of a distribution that contains some yellow cards with circles on them and some yellow cards with diamonds on them (compare Stalnaker 2020). People do call other examples of conditionals "true" when their antecedents are false, but the question is whether these are objective uses of "true", or whether they are simply based on high degrees of belief.

People apparently use "true" pragmatically at times just to express a high degree of belief in, or endorsement of, an utterance, even one that is, not only deontic, but an expression of purely subjective preference or taste. For example, after we say, "If you visit Oaxaca on your trip to Mexico, then you should try chapulines", one listener might reply, "That is true", and another, "That is false", without there being an objective way to decide between them. But there are also uses that would have a ground in objective logical or in causal relations. An example of the former would be *if p & q then q*, which we think that people will call "true" even when *p & q* is false. An example of the latter might be (17), which people might judge to be "true" when they were sure that there is a causal relation between smoking and lung damage. We would predict that the people who made these truth judgments would also assess $P(q|(p \,\&\, q))$ at 1, and of lung damage given smoking as at or near 1. These cases would be covered by Jeffrey semantics. The expected values of these conditionals would be at or near 1 for these people, and that would be the Jeffrey value of the conditionals when their antecedents were false.

Some philosophers have taken the view that a natural language conditional, *if p then q*, is never objectively true, or false, not even when *p & q*, or *p & not-q*, is true (Bennett 2003). But not even they have claimed that people do not apply the words "true" and "false", in some senses, to conditionals, and they have referred to pleonastic (Edgington 2003) and pragmatic notions of truth (Adams 1998). We have followed de Finetti in closely comparing conditionals and conditional bets, and our position is that the outcomes *p & q*, and *p & not-q*, are objective conditions for settling conditional bets and making indicative and counterfactuals conditionals true or false. But we also want to explain people's other uses of "true" and "false" as applied to conditionals in natural language when their antecedents are false. Unfortunately, the psychology of reasoning has not yet systematically studied those uses (but see Over 2020; Pfeifer 2012; Skovgaard-Olsen 2020; Wang and Yao 2018, for some relevant points). This must change before we can get a deeper understanding of how people use indicative and counterfactual conditionals and relate them to each other as their beliefs change. In this investigation, it is certain that psychologists will benefit greatly from what philosophical logicians and linguists say about "true" and "false".

References

Adams, E. (1970). Subjunctive and indicative conditionals. *Foundations of Language, 6*, 89–94.

Adams, E. (1975). *The logic of conditionals: An application of probability to deductive logic*. Dordrecht: Reidel.

Adams, E. (1998). *A primer of probability logic*. Standford: CLSI Publications.

Ali, N., Chater, N., & Oaksford, M. (2011). The mental representation of causal conditional reasoning: Mental models or causal models. *Cognition, 119*, 403–418.

Baratgin, J., Over, D. E., & Politzer, G. (2013). Uncertainty and de Finetti tables. *Thinking & Reasoning*, *19*, 308–328.

Bennett, J. (2003). *A philosophical guide to conditionals*. Oxford: Oxford University Press.

Bramley, N. R., Dayan, P., Griffiths, T. L., & Lagnado, D. A. (2017). Formalizing Neurath's ship: Approximate algorithms for online causal reasoning. *Psychological Review, 124*(3), 301–338.

Byrne, R. M. J. (2017). Counterfactual thinking: From logic to morality. *Current Directions in Psychological Science*, *26*(*4*), 314–322.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367–405.

Crupi, V., & Iacona, A. (in press). Three ways of being non-material. *Studia Logica.*

Cruz, N. (2020). Deduction from uncertain premises? In S. Elqayam, I. Douven, J. S. B. T. Evans, & N. Cruz (Eds.), *Logic and uncertainty in the human mind* (pp. 27–41). Abingdon: Routledge.

Cruz, N., & Oberauer, K. (2014). Comparing the meanings of "if" and "all". *Memory &. Cognition, 42*, 1345–1356.

Cruz, N., Baratgin, J., Oaksford, M., & Over, D. E. (2015). Bayesian reasoning with ifs and ands and ors. *Frontiers in Psychology, 6*, 192.

Cruz, N., Connor, D., Dewitt, S., Hahn, U., Lagnado, D., Liefgreen, A., Phillips, K., Phildитch, T., & Tešić, M. (2020a). Widening access to Bayesian problem solving. *Frontiers in Psychology*. https://doi.org/10.3389/fpsyg.2020.00660

Cruz, N., Hahn, U., Fenton, N., & Lagnado, D. (2020b). Explaining away, augmentation, and the assumption of independence. Frontiers in Psychology, 11, 502751

Cruz, N., Over, D., & Oaksford, M. (2018, July). Dynamic reasoning with counterfactuals. In M. Oaksford, V. Thompson, & N. Adams (Chairs), *The 11th London Reasoning Workshop*. London: Birkbeck, University of London.

Cruz, N., Over, D., Oaksford, M., & Baratgin, J. (2016). Centering and the meaning of conditionals. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 1104–1109). Austin: Cognitive Science Society.

de Finetti, B. (1936/1995). The logic of probability [translation of 1936 original by R. B. Angell]. *Philosophical Studies, 77*, 181–190.

de Finetti, B. (1937/1964). Foresight: Its logical laws, its subjective sources [translation of 1937 original]. In H. E. Kyburg, & H. E. Smokler (Eds.), *Studies in subjective probability* (pp. 55–118). New York: Wiley.

de Zoete, J., Fenton, N., Noguchi, T., & Lagnado, D. (2019). Resolving the so-called "probabilistic paradoxes in legal reasoning" with Bayesian networks. *Science & Justice*. https://doi.org/10.1016/j.scijus.2019.03.003

Douven, I. (2016). *The epistemology of indicative conditionals*. Cambridge: Cambridge University Press.

Douven, I. (2017). How to account for the oddness of missing-link conditionals. *Synthese*, *194*, 1541–1554.

Douven, I., & Dietz, R. (2011). A puzzle about Stalnaker's hypothesis. *Topoi, 30*, 31–37.

Douven, I., Elqayam, S., Singmann, H., & van Wijnbergen-Huitink, J. (2018). Conditionals and inferential connections: A hypothetical inferential theory. *Cognitive Psychology*, *101*, 50–81.

Douven, I., Elqayam, S., Singmann, H., & van Wijnbergen-Huitink, J. (2020). Conditionals and inferential connections: Toward a new semantic. *Thinking & Reasoning*, *26*(*3*), 311–351.

Edgington, D. (1995). On conditionals. *Mind, 104,* 235–329.

Edgington, D. (2003). What if? Questions about conditionals. *Mind and Language*, *18*, 380–401.

Edgington, D. (2008). Counterfactuals. *The Presidential Address, Proceedings of the Aristotelian Society, CVIII*(1)*,* 1–21.

Edgington, D. (2014). Estimating conditional chances and evaluating counterfactuals. *Studia Logica*, *102*, 691–707.

Elqayam, S., & Over, D. E. (2013). New paradigm psychology of reasoning: An introduction to the special issue edited by Elqayam, Bonnefon, & Over. *Thinking & Reasoning, 19*, 249–265.

Egré, P., & Cozic, M. (2016). Conditionals. In M. Aloni & P. Dekker (Eds.), *The Cambridge handbook of formal semantics* (pp. 490–524). Cambridge: Cambridge University Press.

Evans, J. St. B. T., Handley, S. J., & Over, D. E. (2003). Conditional and conditional probability. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 321–335.

Evans, J. St. B. T., & Over, D. E. (2004). *If*. Oxford: Oxford University Press.

Evans, J. St. B. T., Thompson, V., & Over, D. E. (2015). Uncertain deduction and conditional reasoning. *Frontiers in Psychology, 6*, 398.

Fernbach, P. M., Darlow, A., & Sloman, S. A. (2010). Neglect of alternative causes in predictive but not diagnostic reasoning. *Psychological Science*, *21*(3), 329–336.

Fernbach, P. M., & Erb, C. D. (2013). A quantitative causal model theory of conditional reasoning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *39*(5), 1327–1343.

Fugard, J. B., Pfeifer, N., Mayerhofer, B., & Kleiter, G. D. (2011). How people interpret conditionals: Shifts toward conditional event. *Journal of Experimental Psychology-Learning Memory and Cognition, 37,* 635–648.

Gilio, A. (2002). Probabilistic reasoning under coherence in System P. *Annals of Mathematics and Artificial Intelligence*, *34*, 5–34.

Gilio, A., & Over, D. E. (2012). The psychology of inferring conditionals from disjunctions: A probabilistic study. *Journal of Mathematical Psychology, 56*, 118–131.

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review, 116*(4), 661–716.

Hadjichristidis, C., Sloman, S. A., & Over, D. E. (2014). Categorical induction from uncertain premises: Jeffrey's doesn't completely rule. *Thinking & Reasoning, 20*(4), 405–431.

Hayes, B. K., Hawkins, G. E., Newell, B. R., Pasqualino, M., & Rehder, B. (2014). The role of causal models in multiple judgments under uncertainty. *Cognition, 133,* 611–620.

Holyoak, K. J., & Cheng, P. W. (2011). Causal learning and inference as a rational process: The new synthesis. *Annual Review of Psychology, 62,* 135–163.

Jeffrey, R. C. (1991). Matter of fact conditionals. *Aristotelian Society Supplementary Volume*, *65*, 161–183.

Johnson-Laird, P. N., & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, *109*, 646–678.

Kaufmann, S. (2005). Conditional predictions: A probabilistic account. *Linguistics and Philosophy*, *28*, 181–231.

Krzyżanowska, K., Collins, P., & Hahn, U. (2017). Between a conditional's antecedent and its consequent: Discourse coherence vs. probabilistic relevance. *Cognition*, *164*, 199–205.

Krzyżanowska, K., Wenmackers, S., & Douven, I. (2014). Rethinking Gibbard's riverboat argument, *Studia Logica*, *102*, 771–792.

Lassiter, D. (2022). Decomposing relevance in conditionals. *Mind & Language*. Available at: https://web.stanford.edu/~danlass/

Lewis, D. (1973). *Counterfactuals*. Oxford: Basil Blackwell.

Lewis, D. (1976). Probabilities of conditionals and conditional probabilities. *Philosophical Review, 85*, 297–315.

Lucas, C. G., & Kemp C. (2015). An improved probabilistic account of counterfactual reasoning. *Psychological Review, 122*(4), 700–734.

Madsen, J. K., Hahn, U., & Pilditch, T. (2020). The impact of partial source dependence on belief and reliability revision. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 46*(9), 1795–1805.

Mirabile, P., & Douven, I. (2020). Abductive conditionals as a test case for inferentialism. *Cognition*, *200*, 104232.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford: Oxford University Press.

Oaksford M., & Chater, N. (2013). Dynamic inference and everyday conditional reasoning in the new paradigm. *Thinking & Reasoning, 19*, 346–379.

Oaksford, M., & Chater, N. (2017). Causal models and conditional reasoning. In M. Waldmann (Ed.), *The Oxford handbook of causal reasoning*. Oxford: Oxford University Press.

Oaksford, M., & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology, 71*, 305–330.

Oberauer, K., & Wilhelm, O. (2003). The meaning(s) of conditionals: Conditional probabilities, mental models and personal utilities. *Journal of Experimental Psychology: Learning Memory and Cognition, 29,* 680–693.

Over, D. E. (2009). New paradigm psychology of reasoning. *Thinking & Reasoning, 15*, 431–438.

Over, D. E. (2017). Causation and the probability of causal conditionals. In M. Waldmann (Ed.), *The Oxford handbook of causal reasoning* (pp. 307–325). Oxford: Oxford University Press.

Over, D. E. (2020). The development of the new paradigm in the psychology of reasoning. In S. Elqayam, I. Douven, J. St. B. T. Evans, & N. Cruz (Eds.), *Logic and uncertainty in the human mind* (pp. 243–263). Abingdon: Routledge.

Over, D. E., & Baratgin, J. (2017). The "defective" truth table: Its past, present, and future. In N. Galbraith, D. E. Over, & E. Lucas, (Eds.), *The thinking mind: The use of thinking in everyday life* (pp. 15–28). Hove: Psychology Press.

Over, D. E., & Cruz, N. (2018). Probabilistic accounts of conditional reasoning. In L. J. Ball, & V. A. Thompson (Eds.), *International handbook of thinking and reasoning* (pp. 434–450). Hove: Psychology Press.

Over, D. E., Hadjichristidis, C., Evans, J. St. B. T., Handley, S. J., & Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology, 54,* 62–97.

Park, J., & Sloman, S. A. (2013). Mechanistic beliefs determine adherence to the Markov property in causal reasoning. *Cognitive Psychology, 67*(4), 186–216.

Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Francisco: Morgan-Kaufmann.

Pearl, J. (2000). *Causality: Models, reasoning, and inference.* New York: Cambridge University Press.

Pearl J. (2013). Structural counterfactuals: A brief introduction. *Cognitive Science*, *37*, 977–985.

Pfeifer, N. (2012). Experiments on Aristotle's Thesis: Towards an experimental philosophy of conditionals. *The Monist, 95*(2), 223–240.

Pfeifer, N., & Kleiter, G. D. (2009). Framing human inference by coherence based probability logic. *Journal of Applied Logic, 7*, 206–217.
Pilditch, T. D., Fenton, N., & Lagnado, D. (2019). The zero-sum fallacy in evidence evaluation. *Psychological Science, 30*(2), 250–260.
Politzer, G., Over, D. E., & Baratgin, J. (2010). Betting on conditionals. *Thinking & Reasoning, 16,* 172–197.
Ramsey, F. P. (1926/1990). Truth and probability. In D. H. Mellor (Ed.), *Philosophical Papers* (pp. 52–94). Cambridge: Cambridge University Press.
Ramsey, F. P. (1929/1990). General propositions and causality. In D. H. Mellor (Ed.), *Philosophical Papers* (pp. 145–163). Cambridge: Cambridge University Press.
Rehder, B., & Burnett, R. C. (2005). Feature inference and the causal structure of categories. *Cognitive Psychology, 50,* 264–314.
Rips, L. J. (2010). Two causal theories of counterfactual conditionals. *Cognitive Science, 34,* 175–221.
Rips, L. J., & Edwards, B. J. (2013). Inference and explanation in counterfactual reasoning. *Cognitive Science, 37,* 1107–1135.
Rott, H. (2019). Difference-making conditionals and the relevant Ramsey test. *The Review of Symbolic Logic,* 1–32.
Sanfilippo, G., Gilio, A., Over, D. E., & Pfeifer, N. (2020). Probabilities of conditionals and previsions of iterated conditionals. *International Journal of Approximate Reasoning*, *121*, 150–173.
Sanfilippo, G., Pfeifer, N., Over, D. E., & Gilio, A. (2018). Probabilistic inferences from conjoined to iterated conditionals. *International Journal of Approximate Reasoning*, *93*, 103–118.
Skovgaard-Olsen, N. (2020). Relevance and conditionals: A synopsis of open questions and semantic issues. In S. Elqayam, I. Douven, J. St. B. T. Evans, & N. Cruz (Eds.), *Logic and uncertainty in the human mind* (pp. 192–206). Abingdon: Routledge.
Skovgaard-Olsen, N., Collins, P, Krzyżanowska, K., Hahn, U., & Klauer, C. (2019). Cancellation, negation, and rejection. *Cognitive Psychology*, *108*, 42–71.
Skovgaard-Olsen, N., Kellen, D., Krahl, H., & Klauer, C. (2017). Relevance differently affects the truth, acceptability, and probability evaluations of 'and', 'but', 'therefore', and 'if then'. *Thinking & Reasoning, 23*, 449–482.
Skovgaard-Olsen, N., Singmann, H., & Klauer, K. C. (2016a). The relevance effect and conditionals. *Cognition*, *150*, 26–36.
Skovgaard-Olsen, N., Singmann, H., & Klauer, K. C. (2016b). Relevance and reason relations. *Cognitive Science*, *41*, 1202–1215.

Sloman, S. A., & Lagnado, D. A. (2005). Do we "do"? *Cognitive Science, 29,* 5–39.

Sloman, S. A., & Lagnado, D. A. (2015). Causality in thought. *Annual Review of Psychology, 66,* 223–247.

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causation, prediction, and search.* New York: Springer-Verlag.

Stalnaker, R. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory* (pp. 98–112). Oxford: Blackwell.

Stalnaker, R. (2020). Counterfactuals and probability. In L. Walters, & J. Hawthorne (Eds.), *Conditionals, paradox, and probability* (pp. 100–121). Oxford: Oxford University Press.

Thompson, V. A., & Byrne, R. M. J. (2002). Reasoning counterfactually: Making inferences about things that didn't happen. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 1154–1170.

Waldmann, M. R. (2017). Causal reasoning: An introduction. In M. R. Waldmann (Ed.), *The Oxford handbook of causal reasoning* (pp. 1–9). New York: Oxford University Press.

Wang, M., Over, D., & Liang, L. (2022). What is required for the truth of a general conditional? *Quarterly Journal of Experimental Psychology, 75*(11), 2105–2117. https://doi.org/10.1177/17470218221089331

Wang, M., & Yao, X. (2018). The dual reading of general conditionals: The influence of abstract versus concrete contexts. *Quarterly Journal of Experimental Psychology*, *71*, 859–869.

Zhao, J., & Osherson, D. (2010). Updating beliefs in light of uncertain evidence: Descriptive assessment of Jeffrey's rule. *Thinking & Reasoning, 16*, 288–307.

Inferentialism: A Manifesto

Igor Douven, Shira Elqayam, and Karolina Krzyżanowska

At various junctures in the history of thinking about indicative conditionals,[1] the idea has surfaced that their truth requires the presence of a *connection* between their antecedent and consequent. The current authors have been working on a new version of this view since 2013.

[1] Roughly, indicative conditionals are sentences of the form "If A, B" whose main auxiliary is in the indicative mood. From here on, we mostly refer to them simply as "conditionals."

The authors contributed equally to the paper and are listed alphabetically.

I. Douven (✉)
IHPST / CNRS / Panthéon–Sorbonne University, Paris, France
e-mail: igor.douven@univ-paris1.fr

S. Elqayam
School of Applied Social Sciences, De Montfort University, Leicester, UK
e-mail: selqayam@dmu.ac.uk

K. Krzyżanowska
ILLC, University of Amsterdam, Amsterdam, Netherlands
e-mail: k.h.krzyzanowska@uva.nl

Arché, University of St Andrews, Scotland, UK

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_7

In this paper, we summarize the progress that has been made over the past years and respond to objections that have been leveled at the new position. But first we recapitulate the main tenets of our position.

1 Inferentialism

Most work in the psychology of reasoning is concerned with how normal people respond to normal cases, most notably, normal instances of common inference patterns, or normal conditionals.[2] By contrast, logicians, and also philosophers interested in non-deductive forms of inference, have tended to focus on edge cases: sentences or inferences of which, at least initially, we do not know how to make sense. Paradigm cases are Gödel (1931) "This sentence is unprovable" and Tarski (1936) "This sentence is false," or the inference to the conclusion that our lottery ticket will lose, given its low probability of winning, or that there is bound to be some error in our recently finished monograph, given human fallibility—inferences which lead to highly probable conclusions but which, according to mainstream thinking, are to be resisted if we want to avoid ending up with contradictory beliefs (see Kyburg 1961; Makinson 1965).

Both approaches—the psychologists' and the philosophers'—are valuable. We need the former to become clearer about the descriptive adequacy of our theoretical commitments when applied to central cases, while the latter's focus on edge cases helps us develop a feeling for where the limits of our theories may lie.

Recently, both philosophers and psychologists working on conditionals have started paying attention to a type of conditionals that, although not typically encountered in quotidian speech, are not baffling in the way the Gödel or Tarski sentence is. To the contrary, there is no reason why they should be outside the scope of any of the main semantics for conditionals. However, when applied to the said type of conditionals, these semantics yield verdicts which sit badly with pre-theoretic judgments.

[2] Below, we are more specific about what we mean by "normal conditionals."

The conditionals we have in mind have been dubbed "missing-link conditionals" (Douven 2016a, 2017a), their characterizing feature being that—as the name suggests—there appears to be no link between their antecedent and consequent. The recipe for constructing such conditionals could not be more straightforward: take any propositions A and B that, for all you know, are entirely unrelated—being informed about the truth value of one of them would not tell you anything about the truth value of the other—and then construct the conditional "If A, B" (or, just as well, "If B, A"). For concreteness, consider

(1) If Eisenhower was the 34th American president, Newton is the author of the *Principia Mathematica*.

This conditional is odd. Given that the truth value of its consequent came to be settled long before that of the antecedent, it appears about as absurd as the notion of backwards causation. Nonetheless, supposing the material conditional account, according to which "If A, B" is equivalent to "not-A or B," (1) is true. Similarly for Stalnaker's possible world semantics, according to which a conditional is true if its consequent is true in the nearest possible world in which its antecedent is true. In this case, that nearest world is the actual world, and (1)'s consequent holds in the actual world.

To delve already a bit deeper into why conditionals like (1) strike us as being odd, note that there is a *reason* why sentences of the form "If A, B" are called "conditionals": we sense that the consequent *is conditional on*, *depends on*, the antecedent. What (1) and similar conditionals help to bring out is that this sense of conditionality is not adequately captured by the main semantics for conditionals. For the disjunction "not-A or B" to be true, there need not be any kind of connection between A and B. Nor need there be any connection between them for B to be true in the nearest A-world.

In psychology, the advent of the New Paradigm (Over 2009; Elqayam and Over 2013) has occasioned a shift of attention from the above semantics of conditionals to the probabilistic semantics as advocated by Adams (1975) and others. Central to this semantics is the probability conditional, which is any conditional whose probability equals that of its consequent on the supposition of its antecedent.[3]

Unfortunately, Adams' proposal does not fare any better than the material conditional account or Stalnaker's semantics when it comes to accounting for the oddness of (1) and other missing-link conditionals. On Adams' (1975) original proposal, conditionals do not have truth conditions and so also, seemingly paradoxically, no probabilities. There is no real contradiction here, however, because for Adams the probability operator, when applied to conditionals, is to be interpreted as measuring those conditionals' degree of acceptability and assertability. On this view, a conditional "If A, B" is highly acceptable/assertable precisely if Pr(B|A) is high.

But (1) is neither highly assertable nor highly acceptable, despite the fact that most people will assign its consequent unit probability, and will assign unit probability to that consequent also on the supposition of the conditional's antecedent, which is completely unrelated to the truth of the consequent.

More generally, note that Pr(B|A) will be high whenever B is highly probable and A and B are probabilistically independent. In fact, it can be high even if the probability of B is (slightly) diminished on the supposition of A. But although probabilistic independence and negative dependence suggest that the truth of B does not depend on that of A, respectively, that B is somewhat undermined, or disconfirmed, by the truth of A, according to Adams' proposal, "If A, B" can still be highly acceptable/assertable.

That on Adams' account conditionals lack truth conditions has repercussions unrelated to the present topic. Most significantly, Adams' account makes it hard to see how conditionals can occur in compound expressions, for instance, how we can account for conjunctions of conditionals, or for nested conditionals. The recognition of this problem

[3] It has been said that the claim that natural language conditionals are probability conditionals is intuitively supported by the fact that we sometimes report conditional probabilities—especially, objective conditional probabilities—using "if" instead of "given that" or "on the supposition that" (see van Fraassen 1976). Naturally, it does not follow from this (nor has anyone claimed otherwise) that conditional probabilities generally equal the probabilities of the corresponding conditionals.

inspired work aimed at combining Adams' proposal with de Finetti's three-valued semantics for conditionals. Leaving details aside here, the important observation to make is that, on the resulting account, any conditional with a true antecedent and consequent—such as (1)—still comes out true.

At this point, proponents of the above accounts tend to reach for pragmatics band aid, their point being that any seeming support for the thought that true conditionals require a connection between their component parts is, in actuality, only support for the idea that we have little *use* for conditionals whose components are unconnected—such conditionals can still be true or acceptable, but by *asserting* them we may mislead our audience, given that we normally only assert conditionals whose components *are* connected. Here is, for instance, how Over et al. (2007, p. 92) invoke pragmatics to account for their finding of an effect of probabilistic relevance of antecedent to consequent in their data:

> An Adams conditional [i.e., a probability conditional] is not equivalent to an explicit statement that A raises the probability of B, . . .nor that A causes B. . .A conditional probability Pr(B|A) can be high when A does not raise the probability of B and when A does not cause B. For example, Pr(B|A) can be high simply because Pr(B) is high. Does this mean that supporters of the view that these conditionals are Adams conditionals cannot account for the weak negative effect of Pr(B|A) in the current studies? [This is the just-mentioned relevance effect.] Not necessarily, for they can argue that the use of a conditional pragmatically suggests, in certain ordinary contexts, that A raises the probability of B or that A causes B.[4]

There is nothing wrong per se with the appeal to pragmatics here. As Grice (1989) convincingly argued, what we convey by our assertions is not just the semantic content of those assertions but also what our audience can reasonably be expected to infer from the fact *that* we asserted *what* we asserted *when* we asserted it *in the circumstances in which* we asserted it. To give a well-worn example, when you are attending a party and are being asked what time it is, then by responding "The guests are

[4] Notation slightly altered for uniformity of reading; comments in square brackets are ours.

already leaving" you will be interpreted as suggesting that it is already late and that, moreover, you are not in a position to give any more precise indication of the time. That is not because that is what your response *means*, but because it is the best explanation of why you gave that response to the given question under the given circumstances (Bach and Harnish 1979; Dascal 1979; Hobbs 2004; Douven 2012a, 2022, Ch. 1). Similarly, the best explanation of why someone asserts a conditional could be that there is some kind of link between its component parts, or at least that the person believes such a link to be present.

Still, there are reasons to be wary of the move Over and colleagues make in the above passage, and that others have made to account for linguistic phenomena apparently going against their favored semantics of conditionals. While our understanding of pragmatics is not nearly at the level of that of logic—we have been studying logic for over 2000 years, the study of pragmatics only seriously took off with the work of Austin, Grice, and Searle in the second half of the past century—by now a number of principles have been identified that guide us in working out the content conveyed by an assertion that goes beyond the semantic content of that assertion. The first candidates for such principles were provided by Grice and termed "maxims" by him, but they have been refined and supplemented by later authors (see, e.g., Levinson 2000). However, Over et al. (2007) make no reference to any specific pragmatic principles and, more generally, make no attempt to explain *how* the implicature ("the pragmatic suggestion") of a link (whether causal, probabilistic, or inferential) between a conditional's antecedent and consequent is brought about. May we request anyone wanting to rescue their semantics of conditionals by invoking pragmatics to at least *sketch* how the pragmatic explanation of whatever exactly it is that they are trying to explain pragmatically is supposed to go?

To forestall misunderstanding, there is a maxim of relevance in Gricean pragmatics, and the name might suggest that that is precisely what is needed here. But that is not so. According to this maxim, we should make our contributions to an ongoing conversation *relevant*. What needs explaining, however, is why a conditional whose antecedent is not relevant (in some sense) to its consequent does not relevantly contribute to an ongoing conversation (Douven 2008, 2016a).

One could conjecture that the relevance requirement applies also to the clauses of complex or compound sentences. It is in fact an assumption of discourse coherence theory that the hearer will always attempt to conjure up connections between any two consecutive elements of discourse, so that it can be interpreted as coherent (Kehler 2002; Asher and Lascarides 2003). Any discourse coherence violations might then seem as odd as missing-link conditionals, and if that is the case, the oddness of missing-link conditionals might be explained away in terms of discourse coherence violations.[5] This hypothesis turned out to be false, however. Krzyżanowska et al. (2017) report a study in which they compared the assertability of conditionals with the assertability of their consequents in the contexts in which the antecedents have already been asserted. They introduced two manipulations: the presence or absence of a common topic of the clauses, understood in discourse-coherence-theoretic terms, and the presence or absence of a stronger, inferential connection.[6] For instance, one of the vignettes used in the experiment introduced a protagonist, Patrick, who plans to take his girlfriend, Sophie, for holidays and discusses his ideas with Matt. Since they have enjoyed hiking in the Alps before, Patrick considers a trip to the Pyrenees. In the conversational exchange condition, Patrick tells Matt "Sophie likes the Alps," and Matt responds with a statement that is either both relevant (i.e., there is an inferential connection between the two statements) and on the same topic: "She will enjoy hiking in the Pyrenees," irrelevant (no inferential connection) and on a different topic: "More and more people in Western Europe care about animal welfare," or on the same topic but irrelevant: "Mountaineering can be dangerous." In the conditionals condition, in the same context, Matt asserts a conditional consisting of Patrick's statement in the antecedent and Matt's response in the consequent, for instance, "If Sophie likes the Alps, then she will enjoy hiking in the Pyrenees" in the same topic relevant condition. The participants are asked to evaluate to what extent Matt's assertion is natural or makes sense in the context. What Krzyżanowska

[5] Cruz et al. (2016) may be interpreted as hinting at such a possibility.

[6] Krzyżanowska and colleagues operationalized the presence or absence of an inferential connection as probabilistic relevance. In general, this is problematic; see below.

and colleagues found is that while the same topic is sufficient to make B assertable after A has been asserted (e.g., "Mountaineering can be dangerous" in response to "Sophie likes the Alps"), it is not enough to make "If A, B" (e.g., "If Sophie likes the Alps, then mountaineering can be dangerous") assertable in the same context. Conditionals turned out to require a stronger kind of relationship than the presence of a common topic of discourse.

One could also try to argue that the oddity of missing-link conditionals is not due to the lack of a connection between their antecedent and consequent but rather to a violation of Grice's maxim of quantity: "Make your contributions as informative as required (for the current purposes of the exchange)" (Grice 1989, p. 26). After all, it is typically when the antecedent and consequent of a missing-link conditional are both known to be true, or at least warrantedly assertable, that we are facing a discrepancy between what an account of conditionals renders acceptable, or true, and what we may be compelled to accept as such. A missing-link conditional whose antecedent and consequent are known to be true would not be odd because of its missing link, but because the speaker asserting such a conditional is violating the maxim of quantity—they assert a weaker statement when they are justified in asserting a stronger one, for instance, the consequent on its own, or the conjunction of both clauses.[7] This conjecture has been tested, too. Using different dependent variables in a series of four studies, Krzyżanowska et al. (2021) compared conditionals, whose antecedents and consequents were known to be true, to the corresponding conjunctions and, in Experiment 2, to materially equivalent disjunctions. When A and B were inferentially connected, participants did not prefer the supposedly more informative "A and B" over "If A, (then) B," and both types of sentences received relatively high ratings. But when there was no connection between A and B, the participants rated conditionals as significantly less assertable, less acceptable, and as making less sense to say. Thus, it is not the case that a true antecedent and a true consequent are sufficient to make a

[7] This argument was put forward by Grice (1989) in the context of a defense of the material account of conditionals, but see Krzyżanowska (2019) for a discussion of how it can be adapted to serve any theory of conditionals that validates and-to-if inferences and thus needs to deal with their counterintuitive consequences.

conditional unassertable. The presence or absence of the connection, however, did not affect the corresponding conjunctions, and neither did it affect the disjunctions. Indeed, all disjunctions of the form "not-A or B," regardless of whether A and B were connected, received very low ratings. These results do not only falsify the conjecture derived from Grice's maxim of quality, but they also show that the requirement that the clauses are inferentially connected is specific to conditionals.

The connection between the antecedent and consequent might be conversationally implicated nevertheless. It is possible, after all, that even if we have not yet established how it is supposed to be calculated and which of the Gricean or neo-Gricean maxims missing-link conditionals violate, we might do so in the future. Such a possibility notwithstanding, one could, in principle, test if a certain aspect of meaning is a conversational implicature by testing for its necessary characteristics. Grice (1989) proposed a number of such tests, none of which seems to be conclusive,[8] yet there seems to be a consensus that the least controversial and the most practical one is cancelability (Sadock 1978; Levinson 2000).[9] A proposition conveyed by an utterance is cancelable when it can be *canceled* by the speaker of that utterance or by the context in which it was uttered. In particular, when a speaker's utterance S conveys a proposition P in some context, but the speaker does not intend to convey P in that context, they can follow up S with ". . . but I didn't mean to say / suggest / imply that P." When P is a conversational implicature, such cancelations are felicitous. Take, for instance, a well-worn example of a scalar implicature: when Sally says "Some of my students passed the exam," she conversationally implicates that not all of them did, but if she adds "Oh, I didn't mean to suggest that not all of them passed, I've marked only a couple" that initial implicature is canceled. By contrast, an attempt to cancel the meaning which is semantically entailed rather than implicated is infelicitous. In fact, the speaker who makes such

---

8 See Sadock (1978) for a critical discussion of all these tests.

9 Already Sadock (1978) argues that, while cancelability (and, also, reinforceabilty which Grice does not discuss) are necessary characteristics of implicatures, they are not sufficient to distinguish them from other pragmatic phenomena. Others have suggested that there are conversational implicatures that are not cancelable (e.g., Lauer 2013), though Zakkou (2018) argues that the cancelability test is reliable when restricted to non-figurative use of language.

an attempt appears to be contradicting themselves, as for instance in: "Susan and Steve passed the exam . . . oh, I didn't mean to suggest that Steve passed the exam." If it then turned out that the connection between the antecedent and consequent of a conditional could be canceled, it would be a strong indication that it may be conversationally implicated, while if it is not cancelable, then the conversational implicature hypothesis is strongly undermined. In fact, Skovgaard-Olsen et al. (2019) tested the cancelability of the connection and demonstrated that a speaker who asserts "If A, then B" and then attempts to cancel the relation between A and B is perceived by participants as saying something contradictory. By contrast, the connection between the conjuncts in a conjunction turned out to be cancelable.[10] Taking all these considerations together, we can be rather confident that the connection or relevance relation between a conditional's antecedent and consequent is not conversationally implicated.

The oddity of missing-link conditionals does not seem to arise due to presupposition failure either. One of the defining features of presupposition is that it projects under embeddings, and under negation in particular. For instance, the sentence "John quit smoking" presupposes that John smoked in the past, and so does its negation, "It is not the case that John quit smoking." If we know that John never smoked in his life, both sentences sound inappropriate. If the inferential connection between a conditional's antecedent and consequent were a presupposition, then a missing-link conditional, "If A, C," and its negation, "It is not the case that if A, C," should receive similarly low ratings. However, Skovgaard-Olsen and colleagues (2019) tested this hypothesis in their Experiment 2 and did not find any evidence supporting it.

[10] Implicatures cannot only be canceled, but also reinforced. For instance, "Some of my students passed the exam" can be followed up with "Not all of them did" and, even though the former utterance conversationally implicates the latter, that extra bit of information is not perceived as redundant or unnecessary, unlike attempts to reinforce semantic entailments. A recently published study by Rostworowski et al. (2021) shows that when a conditional "If A, B" is followed by a statement emphasizing that there is a causal, deductive, or abductive connection between A and B (e.g., "A will result in B" or "A entails B"), the latter is perceived as redundant. Moreover, Krzyżanowska (2019) argues that the connection does not pass any other test for conversational implicature put forward in the literature.

Then how about stipulating "if" to generate the *conventional* implicature that the antecedent is relevant to its consequent? In that case, we would not have to rely on any maxims to work out the said suggestion but would simply infer this from the use of "if," which generates the suggestion by convention. There are again several problems to be faced. First, Bach (1999) makes a case against the existence of conventional implicatures generally. According to him, they are theoretical artefacts.[11]

Second, even granting there are conventional implicatures, the number of words that have been said to carry such implicatures is modest at best. So it requires an argument that a given word carries a conventional implicature and does not contribute the supposed implicature to what is *said* (as opposed to what is indicated or suggested).

In the case of the aforementioned examples, the arguments have all pointed at utterances which can be rephrased without the word at issue while retaining their truth value, although not their information content. For instance, it has been said that

(2) She is poor but honest.

is true precisely if

(3) She is poor and honest.

The supposed difference is that (2), but not (3), suggests a contrast between the conjuncts. Indeed, according to Grice (1989) *detachability* is a litmus test for whether a word generates a conventional implicature, which we can conduct by asking: can we say the same thing while avoiding the implicature?[12]

So, is the suggestion of a connection detachable in the case of a conditional? It is difficult to see how we could capture the semantic content of

[11] Potts (2015), who does hold that there are conventional implicatures, points out that it is not entirely obvious which side of the semantics–pragmatics divide they belong to.

[12] Admittedly, the usefulness of this test has been questioned by Sadock (1978, pp. 287–290), who has argued that it requires presupposing what it is supposed to be a test for.

(4) If global warming continues, London will be flooded.

while somehow abstaining from using “if” and not giving any impression that we see a connection between the continuation of global warming and London being flooded in the future. Here, advocates of the material conditional account might seem to have an advantage. They could claim that the following does the trick:

(5) Global warming does not continue or London will be flooded.

Note, though, that while the intuition that (2) and (3) have the same truth conditions is broadly shared, the claim that (4) and (5) have the same truth conditions is highly contested. Besides, the material conditional account faces a barrage of other problems, not least that it is inconsistent with virtually all known data about how people use conditionals—which is why no one in the psychology of reasoning community takes it seriously anymore.[13]

Setting the detachability question aside, one may wonder whether the connection between antecedent and consequent belongs to the *at-issue* content of a conditional, or to its *not-at-issue* content, where the latter term covers both conversational and conventional implicatures, as well as presuppositions.[14] And this is, indeed, what Skovgaard-Olsen and colleagues (2019) investigated in their Experiment 3, following the diagnostic tests proposed by Tonhauser (2012), who observed that the at-issue content is what can be accepted or denied directly, whereas the not-at-issue content can only be denied in a way that interrupts the flow of a conversation (e.g., “Hey, wait a minute . . .”). The participants were asked to react to an assertion by an English language learner asserting “If A, then C” (or “A therefore C”), choosing the most appropriate justification for their choice. For instance, in a scenario whose protagonists learn

[13] Proponents of the suppositional theory might claim that (4) can be paraphrased by, “London will be flooded supposing that/assuming that/provided that global warming continues.” But note that these paraphrases would be equally infelicitous when there is no connection between the component clauses.

[14] Note that whether the at-issue versus not-at-issue distinction is determined by semantic or pragmatic considerations itself depends on how implicatures and presuppositions are defined (Potts 2015).

that A and C are both true, but A is irrelevant for C, the participants could choose between "Yes, A and C" and "No, A is not a reason that C." If the participants preferred the former choice, that would indicate that the reason relation is not-at-issue. The study showed, however, that the relevance relation is, in fact, perceived as content at-issue. Since the same pattern of responses was obtained for both "If A, then B" and "A, therefore B"—the paradigmatic example of Gricean conventional implicature—Skovgaard-Olsen et al. maintain that their results are compatible with the possibility that the connection is a conventional implicature after all. However, this interpretation of their results is not compatible with treating conventional implicatures as strictly pragmatic phenomena. They are, after all, conventional *and* at-issue.

As an independent reason to be wary of the sort of appeal to pragmatics that Over et al. (2007), and also Over and Cruz (2021), make, note that pragmatics is about *assertion*, while the problem of dealing with missing-link conditionals also concerns their *acceptability*: we are not just disinclined to assert missing-link conditionals, we are also disinclined to incorporate such conditionals into our system of beliefs.

In fairness, we note that probably no party to the present debate (so including inferentialists) can, at this point, make very definite statements about the proper interpretation, qua semantic or pragmatic, of their experimental results. For one, that is because, at least currently, we are far from having a consensus view on where to draw the line between semantics and pragmatics, if such line can be drawn at all. See for instance the schema in Levinson (2000, p. 195), showing the wide variety of views on what should count as semantics and what as pragmatics, also illustrating his claim that "[the Gricean] program ...renders problematic and 'up for grabs' the correct division of labor between semantics and pragmatics in the explanation of many aspects of meaning" (Levinson 2000, p. 165). Furthermore, in recent years, many philosophers of language and linguists have been busy identifying aspects of meaning that cannot be easily classified as either semantic or pragmatic, that rather seem to lie *in between* the two. It is not only the aforementioned conventional implicatures that can be seen as having both semantic and pragmatic characteristics. There are also pragmatic processes or contextual contributions that affect the truth-conditional content, including ones that are necessary for the utterance to express a truth-evaluable proposition in the first place (see, e.g., Carston 2002, or Recanati 2003).

For another, the experimental results reported in Krzyżanowska and Douven (2018) should discourage anyone from making strong claims about whether they have shown a phenomenon to be semantic, or to be pragmatic. Krzyżanowska and Douven sought to determine whether people distinguish between the truth and the assertability of a sentence, and whether they distinguish between the assertability and the acceptability of a sentence. In their paper, they report the results from two experiments strongly supporting a negative answer to both of their research questions. They used as materials sentences that, according to standard semantic theorizing, were all true but that, according to standard pragmatic theorizing, also all carried false implicatures. Krzyżanowska and Douven found no reliable differences among assessments of the *truth* of the items in their materials, assessments of those items' *acceptability*, and assessments of their *assertability*. Distinctions among those concepts, theoretically important as they may be, appear to have little significance in the minds of laypeople.

However, as we and other inferentialists have noted, there are strictly theoretical reasons (including simplicity) to favor an account of conditionals that does not require explaining away data by reference to pragmatics, but can explain the oddness of conditionals such as (1) semantically—an account on which such conditionals come out as not being *true*. Any semantics that requires the presence of a link between a conditional's antecedent and consequent will be able to do so. And indeed, semantics of this type go back to the ancient Greek philosophers (Kneale and Kneale 1962), with later proponents including Mill (1843/1872), Ryle (1950), Mackie (1973), and, in psychology, Braine and O'Brien (1991). In particular, it has been proposed that the consequent must *follow*, in some sense, from the antecedent for the conditional to be true.

With the possible exception of Mill (Skorupski 1989, p. 73 f.), the aforementioned authors meant the sense in which the consequent ought to follow from the antecedent to be deductive. However, as pointed out in Krzyżanowska et al. (2014) and elsewhere, this insistence on a deductive–inferential link between antecedent and consequent makes the proposal open to immediate counterexamples. There are many conditionals we regard as true even though the truth of the antecedent does not guarantee the truth of the consequent. To give an example from Douven et al. (2018), we have no difficulty imagining a context in which we would deem true the statement "If Betty misses her bus, she will be late for the movies," even if, in that context, we are unable to rule out completely that Betty is transported from her present location to the cinema after missing the bus but still before the beginning of the movie.

As we have argued in a number of publications (Krzyżanowska et al. 2013, Krzyżanowska et al. 2014; Krzyżanowska 2015; Douven 2016a; Douven et al. 2018, 2020), it would be a mistake to insist on "inference" as meaning *deductive* inference. Rather, we should adopt a broader notion of inference which encompasses, besides deduction, also induction and abduction and possibly other forms of inference as well (such as analogical inference, if that is different from inductive inference; see Douven et al. 2021). To illustrate, while

(6) If $x + 1 = 7$, then $x = 6$.

embodies a clear deductive link, as does

(7) If the marble is green all over, it is not red all over.

supposing plausible meaning postulates (Carnap 1952), the conditional

(8) If John lives in Chelsea, he is rich.

rather features an inductive–inferential connection, on the (to our knowledge, true) supposition that virtually everyone living in Chelsea is rich. Or consider

(9) If Patricia and Peter are jogging together, they have patched up their friendship.

in which antecedent and consequent would appear to be connected via an abductive–inferential link. Supposing—as the conditional suggests—that they had ended their friendship, that they are jogging together is best explained by their having patched up their friendship. As said, there may be other relevant forms of inference still. For instance, Douven et al. (2021) look at conditionals such as

(10) If Jim's son likes ice skating, he will like ice hockey.

where the consequent follows from the antecedent by analogy, in the manner of Carnap (1980) and Paris and Vencovská (2017). It is to be noted that, unlike deductive inference, the forms of inference mentioned are not 100% safe, in that the truth of their premise or premises does not guarantee the truth of their conclusion; for instance, a couple of people living in Chelsea may be poor nonetheless, and John may happen to be one of them. Nevertheless, as Schurz and Hertwig (2019) point out, in people's actual reasoning, these non-deductive forms of inference probably play a much bigger role than deduction, for which, in our daily lives, we have limited use.

Specifically, our proposal has been that a conditional "If A, B" is true if there is a compelling argument from A plus contextually determined background premises to B, with A being pivotal to that argument (i.e., with A removed, the argument would cease to be compelling), false if there is a compelling argument from A plus contextually determined background premises to the *negation* of B, and indeterminate otherwise. "Compelling," as we pointed out, does not mean "conclusive." While the steps in a conclusive argument would all have to be deductively valid, an argument can be compelling even if it contains inductive or abductive steps, or other steps (e.g., ones involving analogical reasoning) that we take to transmit justification. An intuitive way to put the broad idea underlying the proposal is that anyone justified in believing A should become justified in believing B upon becoming justified in believing "If A, B," supposing the receipt of the conditional information that if A, B, does not undermine the person's justification for A.[15] That is what we

[15] That the latter can happen is demonstrated by the drivers license example from Douven (2012b).

take compelling arguments to do: to transmit whatever justification one may have for their premises to their conclusion.[16]

Thereby, inferentialism gives clear content to the idea of conditionality that was mentioned previously, the idea that the consequent is conditional on the antecedent.[17]

In general, being informed of a conditional "If A, B," we can move to accepting B on condition that we are in a position to accept A. We had already a first stab at explaining why missing-link conditionals like (1) strike us as odd. Inferentialism allows us to expand on this. According to inferentialism, the oddness of those conditionals is not explained simply by those conditionals' perceived lack of truth. Nor is it explained by the fact that we might be unable to *reconstruct* the argument connecting their constituent parts. After all, we have no issue accepting some conditionals as true despite not having effectively identified the connecting argument: being informed that if A, B by someone we trust will normally suffice to convince us of the presence of a compelling argument from A plus background premises to B even if the speaker does not provide that argument or we are unable to see it ourselves. Rather, the problem missing-link conditionals present us with is that it is so exceedingly clear that there is no compelling argument starting from their antecedent and ending with their consequent.

To this philosophical, or computational-level theory (to use Marr's 1982, terminology), we later added a psychological theory, Hypothetical Inferential Theory, or HIT, to provide an algorithmic-level explanation of how inferentialism is represented in the mind. HIT tops up inferentialism with a couple of psychological principles. According to

[16] For related ideas, see Oaksford and Chater (2010, 2013, 2014, 2017, 2020), Vidal and Baratgin (2017), and van Rooij (2019).

[17] Note that inferentialism is not the *only* way to cash out the idea of conditionality. As said in Douven (2016a, p. 36), to claim that there must be some kind of connection between a conditional's antecedent and consequent leaves the nature of that connection wide open: "[I]t could be logical, statistical, causal, explanatory, metaphysical, epistemic; or the 'connector' could be a second-order functional property, notably, the property that there is some first-order property or other that links antecedent and consequent, much in the way in which some have argued that truth is a second-order functional property, instantiated by correspondence to the facts in some domains of discourse, by assertability or verifiability in other domains, and by yet some other first-order property in yet other domains." Inferentialism is the substantive thesis that the nature of the "connector" is inferential.

the *principle of relevant inference*, the relevant mental representation of the conditional is by default the one in which there is an inferential relation between antecedent and consequent; and according to the *principle of bounded inference*, this inferential link between antecedent and consequent need only be strong enough, in the sense of being subjectively supported. Thus, the strength of the connection is bounded by Simon-style satisficing (Simon 1982).

Note that when we say that, in the case of missing-link conditionals, it is immediately obvious that there is no argument that could reasonably connect their component parts, we mean that this is so given a set of contextually determined background premises. In particular, we do not want to suggest that some conditionals are *intrinsically* or *objectively* missing-link conditionals. Indeed, it was already emphasized in Krzyżanowska et al. (2014) that whether a conditional embodies a deductive, abductive, inductive inferential connection, or no connection at all, is a question that can only be answered relative to a given body of background knowledge. What for one person is a deductive inferential conditional may be an abductive or inductive inferential conditional, or even a missing-link conditional, for another person, or for the same person at a different moment in time, when the person had or will have a different set of background beliefs. That also means that one and the same conditional can be true for one person and false for another, or true and false for the same person at different points in time. Thus, there is a clear perspectivalist aspect to inferentialism.[18]

In Sect. 3, we respond to Over and Cruz (2021)'s criticisms of inferentialism. We already mention here that they have clearly misunderstood the notion of a missing-link conditional. In particular, they appear to believe that such conditionals can be characterized probabilistically. According to them (p. 16), when the probability of the consequent of a conditional given its antecedent minus the probability of the same consequent given the negation of the antecedent is 0, the conditional is a

[18] Confusion could arise on this point given that, for obvious reasons, we have always chosen examples of missing-link conditionals whose status as such is likely to be preserved under all reasonable changes of our background knowledge.

missing-link conditional. Not so.[19] Consider a coin with unknown bias; the bias could be anything. Then Pr (The coin will land heads | The coin is fair) = 0.5 but also Pr (The coin will land heads | It is not the case that the coin is fair) = 0.5. (If the latter is not clear, integrate the probability of heads over the unit interval—which yields 0.5—and subtract the integral of heads over the single point 0.5, which equals 0.) Nevertheless, "If the coin is fair, it will land heads" is *not* a missing-link conditional. (To forestall further misunderstanding, that does not mean it is *true*. There is an inferential connection between antecedent and consequent alright, but it is too weak to afford a compelling argument from the former to the latter, making the truth value of the conditional indeterminate.)

2 Evidence for Inferentialism

Inferentialism has clear empirical content, and over the past years we have been concerned to test that content. Here, we focus on the main experimental work done on inferentialism.

2.1 Soritical Series and Inference Strength

The first evidence for inferentialism in the form we advocate came from an experiment reported in Douven et al. (2018), which concerned the soritical color series shown in Fig. 1. In this series, colored patches gradually shift from clearly green to clearly blue, through various shades of blue and green, including borderline blue-green shades. The participants in Douven and colleagues' experiment were asked to evaluate several conditionals pertaining to this series, all having the schematic form

If patch number i is X, then patch number j is X,

with $i \in \{2, 7, 8, 9, 10, 13\}$, and with X standing for either "blue" or "green," depending on whether the participant had been assigned to

[19] We cannot think of a missing-link conditional whose component parts are not probabilistically independent of each other. That does not mean that whenever a conditional's component parts are probabilistically independent of each other, that conditional is a missing-link conditional.

Fig. 1 The soritical color series from the materials of Douven et al. (2018)

the blue condition or to the green condition.[20] What values *j* could take depended on whether the participant had been assigned to the small or to the large condition: if the first, then the patch referred to in the consequent was either one or *two* steps away from the patch referred to in the antecedent; if the second, the distance between the patches was either one or *three* steps (Fig. 1).

As Douven et al. explain, one can naturally associate an argument with each of the resulting conditionals. For example, an argument backing

(11) If patch number 6 is green, then so is patch number 7,

would look something like this: Patches become greener as we move to the right in the color series; on the supposition that patch number 6 is green, and given that patch number 7 is to the right of patch number 6, patch number 7 must be green. Similarly, we can with

(12) If patch number 6 is green, then so is patch number 5,

associate an argument to the effect that because adjacent patches are very similar in color, and because patch number 5 is adjacent to patch number 6, patch number 5 must be green on the supposition that patch number 6 is green.

For the experiment, it was crucial that the arguments that can be associated with the conditionals in Douven et al.'s materials can vary in strength. For instance, while (11) and (12) both refer to adjacent pairs of patches, in the former the consequent patch is to the "greener" side of the antecedent patch, in the latter it is to the "bluer" side of the antecedent patch. The argument associated with (12) is certainly not weak, but it is not as strong as the argument associated with (11), given that, for the

[20] This split was made strictly for control purposes.

former argument but not for the latter, there is a consideration that at least somewhat weakens the conclusion.

Douven and coauthors were specific about what the main determinants for argument strength in the context of their materials were: direction—is the consequent patch to the left or to the right of the antecedent patch?—and distance: how close is the consequent patch to the antecedent patch? Comparing again (11) and (12) above should be enough to see why they thought direction mattered to argument strength. As for distance, compare (12) with

(13) If patch number 6 is green, then so is patch number 4.

With (13), we can associate an argument that is more or less identical to the one we associated with (12). However, because patches that are two steps away from each other are not quite as similar as patches that are only one step away, the argument associated with (12) is a bit weaker.

In their analysis, Douven et al. found that these factors indeed predicted with great accuracy the rates at which their participants had judged the conditionals to be true, in support of inferentialism.

Another noteworthy finding (replicated across four experiments) was a belief-bias analogue in truth judgments. Belief bias (Evans et al. 1983) is the effect of belief on inference, regardless of the inference's validity (for deductive arguments) or strength (for informal arguments). Classic belief bias has two components: a main effect of belief, in which arguments with believable conclusions are endorsed more than arguments with unbelievable conclusions; and an interaction of belief bias and argument validity, in which the difference between arguments with believable and unbelievable conclusions is larger for invalid arguments. Douven et al. drew an analogue between the conditional's antecedent and an argument's premise, and between a conditional's consequent and an argument's conclusion, an analogy not just in line with inferentialism, but necessitated by it. If the analogy is correct, then we should expect to see both belief bias effects in the truth evaluation of conditionals as well. First, we should expect a main effect of consequent, in which conditionals with a true consequent will be more often evaluated as true. Second, and importantly, we would expect an interaction: this effect should be stronger for "invalid" conditionals, that is, conditionals whose direction was "wrong." This is exactly what was found. We call this the "quacks like a duck" principle: if it quacks like a duck, it *is* a duck; if it behaves like an inference—bias and all—it *is* an inference.

In their (2020), the same authors went a step further and re-analyzed the data from their earlier paper to explicitly compare inferentialism with the main rival semantics of conditionals, including the material conditional account and Stalnaker's possible worlds semantics, finding that inferentialism predicted those data much more accurately than did any of the rivals.

While providing strong support for inferentialism, it is to be admitted that, as Mirabile and Douven (2020) note, the data from Douven et al. (2018) concerned a somewhat artificial setting. While their materials are not entirely abstract, they are not entirely realistic either. Naturally, it is more important to know how well a semantics of conditionals is able to handle realistic conditionals than it is to know how the semantics handles conditionals of a sort we rarely if ever encounter in everyday life.

Therefore, Mirabile and Douven devoted two of their experiments to testing the same hypothesis that had been the focus of Douven et al. (2018)—whether the strength of the argument from a conditional's antecedent to its consequent predicts the likelihood with which it will be endorsed—but now using realistic materials. More specifically, their materials for those experiments consisted of abductive conditionals, that is, conditionals, like (9), in which the connection between a conditional's component parts consists of an explanatory link between those parts: the consequent explains, to a higher or lower degree, the antecedent. For such conditionals, the strength of the argument they embody is a function of *how* well the consequent explains the antecedent (Douven et al. 2018). Using these materials, Mirabile and Douven tested the said hypothesis both between subjects and within subjects, both tests yielding strongly favoring evidence.

In their final experiment, Mirabile and Douven also looked at the endorsement rates of the conclusions of Modus Ponens arguments. Their aim was to determine whether the strength of the argument embodied by the major premise of a Modus Ponens argument would predict the likelihood with which the conclusion would be endorsed. Not only that:

they wanted to know whether argument strength was a better predictor of that likelihood than the probability of the major premise's consequent given its antecedent. To that end, they conducted an experiment in three phases, spaced one week apart. One phase sought to determine conditional probabilities, another phase sought to determine argument strength, and the third sought to determine endorsement rates. For instance, in the phase in which endorsement rates were determined, one of the items was

(14) Dennis tells you that John did well on his exam. Now suppose that if John did well on his exam, then he studied hard.

Participants were then asked to indicate how strongly they agreed that John studied hard. Corresponding to this argument, the participants were, in the phase that sought to determine explanation quality, presented with the following:

(15) Suppose we observe that John did well on his exam. We propose to explain this by the fact that he studied hard.

They were then asked to rate the quality of this explanation. Again corresponding to the same example, the remaining part asked to assign probabilities to the four rows in the truth table of conjunction with "John did well on his exam" and "John studied hard" as atomic propositions; from those probabilities Mirabile and Douven derived the conditional probability that John studied hard on the supposition that he did well on his exam. In their analysis, they found, again in support of inferentialism, that while conditional probability was a good predictor of conclusion endorsement, argument strength was a significantly better predictor.

2.2 Modus Ponens and Inference Strength

Psychologists have looked not only at how accurately various semantics of conditionals are able to predict truth judgments of conditionals but also at whether the inferences people are willing to make are in line with the commitments of those semantics. The inference rule studied more than any other is, of course, Modus Ponens (MP). In all experiments concerning this rule, it came out as being highly endorsed.

While typically *highly* endorsed, MP was, equally typically, not *universally* endorsed in those experiments. Absence of universal endorsement might be partially just noise. But proponents of the New Paradigm have also pointed out that, when experimenters request their participants to suppose the premises of an argument, they cannot expect those participants to follow suit exactly. Participants may bring their own beliefs about those premises to the experiment, and those beliefs may affect their judgment of whether the argument's conclusion follows from its premises. Most notably, uncertainty about the major premise in an MP argument might diminish a participant's willingness to endorse the conclusion.

Mirabile and Douven proposed an inferentialism-based explanation of the fact that endorsement rates of MP arguments tend not to be entirely at ceiling. As they note, from an inferentialist perspective, one can think of conditionals as *conduits* or *pipes* which, if accepted, allow one to transfer whatever grounds one has for believing the antecedent to the consequent. That, after all, is what compelling arguments do: transferring grounds of belief in the premises to grounds for belief in the conclusion (see Sect. 1). But precisely because compelling arguments need not be conclusive, we should conceive of conditionals as pipes that can, to varying extents, be *leaky*, in that the argument they embody may fail to carry over *all* the support we have or may have for the antecedent to the consequent.

Mirabile and Douven hypothesized that, if the inferentialism-based explanation were true, then endorsement rates of MP arguments with as major premise one of the abductive conditionals from the materials used in their experiments described earlier should be predicted by the strength of the explanatory argument connecting the antecedent and consequent of the given conditional. Experiment 3 in their paper tested this prediction and found again strong support for it. Not only that: Mirabile and Douven compared their hypothesis with the rival hypothesis that endorsements rates would be predicted by the conditional probability corresponding to the major premise—so the probability of the consequent of the premise given its antecedent—finding that the inferentialist predictor was much more reliable than the probabilistic one. That conclusion strongly favored inferentialism over the suppositional account.

2.3 Probabilities of Conditionals

No semantics of conditionals is complete if it does not account for the probabilities that people assign to conditionals. In our early work on inferentialism, we had been silent on the matter of probabilities. It was only recently addressed in Douven et al. (2021). To work out the implications of inferentialism for the probabilities of conditionals, these authors start by unpacking the truth conditions that inferentialism assigns to conditionals, noting that probabilities are probabilities of *truth*, and thus in particular that the probability of "If A, B" is the probability that "If A, B" is *true*, which is the probability that the truth conditions of "If A, B" are realized. As a result, inferentialists must hold that the probability of a conditional is the probability that there is a compelling argument from the conditional's antecedent (plus background knowledge) to the conditional's consequent, in the sense explained above.

As the authors also note, however, it is in general not a priori (in the colloquial sense of this expression) whether we can make a compelling case for a proposition on the basis of another proposition together with background knowledge. For example, we are somewhat confident that we can make a compelling case for the claim that the economy will speedily recover on the supposition that we get the COVID-19 outbreak under control, but we are, at the moment, not entirely convinced of this. We would have to think more carefully about whether other conditions for a quick economic recovery are in place (e.g., whether the pandemic has not done long-term damage to consumer confidence), whether other factors (e.g., Brexit) will not start to have a negative impact on European economies, what the effects of a growing Chinese economy will be in the coming years, and so on. If asked now for the probability we assign to

(16) If we can control the COVID-19 outbreak, the economy will quickly recover.

we will estimate the likelihood that we can make a compelling case for the consequent, starting from the antecedent plus background knowledge, and give that as our answer. Importantly, in making that estimate we use the heuristic of gauging the inferential strength between antecedent and consequent, that is to say, of gauging how strongly the consequent follows from the antecedent.

Douven et al. (2021) test this "inference heuristic" (as they call it) in two experiments, both presenting participants with three tasks, all of which used the same set of 50 conditionals. In one task, participants were asked to judge the probability of each of those conditionals; in a second task, they were asked to judge the strength of the inferential connection between antecedent and consequent for each of the conditionals; and the third task was meant to determine their conditional probabilities corresponding to the conditionals, where these conditional probabilities were measured via a probabilistic truth-table task in one experiment and by asking participants to engage in suppositional thinking in the other experiment.

In both experiments, inference strength judgments were strongly predictive of probability ratings, in support of inferentialism. Also in both experiments, and so independently of how conditional probabilities were measured, inference strength judgments predicted probability ratings much more accurately than conditional probability ratings did, an outcome strongly favoring inferentialism over the suppositional account.

It is particularly worth noting, also in connection with Over and Cruz's objections to be discussed below, that the materials used by Douven et al. (2021) included ten missing-link conditionals.[21] Inferentialism and the suppositional account make very different predictions about such conditionals, given that they are characterized by the absence of an inferential connection between antecedent and consequent but

[21] That is, in this case, conditionals that relative to any reasonable background premises will be perceived as lacking a connection between their component parts.

can nonetheless have any corresponding conditional probability. The ten missing-link conditionals in Douven and colleagues' materials had been chosen in the hope that the corresponding conditional probabilities would be more or less evenly distributed across the [0, 1] scale, which indeed turned out to be the case. Also, entirely as expected, ratings of inference strength were invariably low. As the authors noted, on the suppositional account, according to which inferential considerations do not matter in the interpretation of conditionals, one expects their probability ratings to be simply correlated with their conditional probabilities. By contrast, from an inferentialist viewpoint, one would expect inferential strength rather than conditional probability to be more strongly correlated with judgments of the probability of a missing-link conditional. The results were again clearly in favor of the inferentialist proposal, revealing a strong correlation between inference strength and probability ratings and a very weak one between conditional probabilities and those same probability ratings.

2.4 Similarity-Based Arguments

Douven et al. (2021) report experimental work primarily concerned with the study of a specific type of reasoning in the context of the conceptual spaces framework as developed in Gärdenfors (2000). Conceptual spaces are (typically) built on top of similarity spaces of the kind studied by Shepard (1964), Nosofsky (1988, 1989), and Petitot (1989), among many others. A similarity space is a one- or multidimensional metric space that is meant to represent people's judgments of how similar given items are in a specific respect. For instance, color similarity space is a three-dimensional space such that Euclidean distances in that space represent "dissimilarities" among color shades: the further apart two shades are—as represented in that space—the more dissimilar they are in people's perception; conversely, the closer they are in the space, the more similar we perceive them to be (Fairchild 2013; Jraissati and Douven 2018). The new proposal in Gärdenfors (2000) was that concepts (i.e., the mental correlates of words) can be represented as regions in similarity spaces. For example, the concept RED is a region in color similarity space, and the concept SOUR is a region in taste space. This opened up the possibility of studying concepts by geometric and topological means, which has led to a research program in its own right. Much of the theoretical modeling undertaken in this program was also recognized to have clear empirical content. This recognition in turn led to a significant amount of experimental work being devoted to the conceptual spaces framework, virtually all of it yielding supporting evidence.

Douven et al. (2021) used the conceptual spaces framework to study a particular type of non-deductive arguments, to wit, those which project a property from one object onto another, based on the similarity between the objects. Taking their cue from a theoretical proposal in Osta-Vélez and Gärdenfors (2020), they hypothesized that the strength of such arguments would depend on the degree of similarity between the object designated in the premise and the object designated in the conclusion. Their materials involved objects that could be precisely located within a conceptual space whose geometry and topology had been established in Douven (2016b). They found that distances in that space between premise-object and conclusion-object were indeed strongly predictive of how compelling participants deemed the corresponding argument to be.

Given the connection that inferentialism postulates between the truth value of a conditional and the strength of the argument for its consequent based on its antecedent, Douven et al. (2021) realized that they should also be able to predict endorsement rates of the conditionals corresponding to the arguments in their materials on the basis of distances in the conceptual space they had used for testing their hypothesis about argument strength. To clarify, if we can predict the strength of the argument for the conclusion that Jim's son will like ice hockey starting from the premise that the son likes ice skating based on how similar the two sports are, then, supposing inferentialism, that same similarity should allow us to predict the likelihood with which (10) would be endorsed.

This observation inspired Douven et al. (2021) to present participants with a set of conditionals matching the similarity-based arguments in their materials, asking the participants how strongly they agreed that those conditionals were true. In their analysis, they regressed the responses on the argument strength judgments as well as, separately, on the distances in the relevant conceptual space between the antecedent-object and the consequent-objects. Both turned out to accurately predict endorsement rates for conditionals, leading to the overall conclusion that the conceptual spaces framework can be fruitfully mustered for explaining certain non-deductive inferences as well as for the evaluation of conditionals embodying such inferences.

3 Objections and Replies

In their contribution to this volume, Over and Cruz (2021) criticize inferentialism, as part of a defense of their own preferred position, which combines Adams' and de Finetti's work on conditionals, yielding a version of the probability conditional coupled with a (non-classical) truth-conditional semantics. The concerns about inferentialism that Over and Cruz raise appear reasonable and might be shared by others. We thus believe it to be worth responding to them in some detail. Generally put, the objections are that (i) inferentialism is incomplete; (ii) it is too narrow; and (iii) it is implausible, for theoretical as well as empirical reasons. We discuss these objections in turn.

3.1 Inferentialism Is Incomplete

Inferentialism, in the version at issue, is a new semantics for conditionals. As mentioned, we started working on it in 2013. We are the first to admit that there remains important work to be done (see Sect. 4). By contrast, the position Over and Cruz are advocating—a version of the so-called suppositional account—has been in the making for almost a century, starting with Ramsey's and de Finetti's important work from the 1920s, and further building on equally important contributions from Adams made in the 1960s. So it should surprise no one that more is known about their favorite account than is known about inferentialism. That being said, Over and Cruz are right when they point out that, whereas there is a logic of the probability conditional, there is no logic of the inferential conditional. How damaging is this situation?

First, there is recent work by Crupi and Iacona (2020, 2021a, 2021b) and Raidl et al. (2021), which is aimed at formalizing the idea that a conditional's antecedent should *support* its consequent for that conditional to be true. In this work, the notion of support is mostly taken as primitive. In the end, one might want to have a logic of conditionals more closely tied to the finer mechanics of the notion of support, for instance, one that is sensitive to the different types of inference that *realize* the support. But of course, one could equally hope for a logic of the probability conditional that is sensitive to the reasons people have for assigning the probabilities they do, which one day we may be able to model formally as well. We do not think it is incumbent on the advocates of the probability conditional to do that work, but similarly, inferentialists might be happy to embrace one of the logics developed by the aforementioned authors.[22]

We are not here committing to any of these logics and indeed could imagine a very different approach to developing the logic of the inferential conditional. From the start, we have taken a rather detailed look at the sort of support the antecedent needs to provide to the consequent for the conditional to be true. First and foremost, this was to highlight the difference between our position and earlier attempts to give content to the idea that the truth of a conditional requires the presence of an inferential link between its component parts. As already pointed out, almost invariably, previous authors assumed that the link had to be deductive, which for the reasons indicated current proponents of inferentialism believe to be a bad idea. Hence, our emphasis on inductive and abductive inference in relation to inferentialism.

Recently, much has been done to clarify both inductive and abductive reasoning; on inductive reasoning, see Schurz (2019), on abductive reasoning, Douven (2017b, 2021, 2022). Neither of these authors offers anything deserving of the name "logic," but suppose we had logics of induction and abduction, and possibly of other forms of non-deductive inference as well. Then it would make a lot of sense to try to build a logic of the inferential conditional on those.

[22] The somewhat different approach to developing the logic of an inferentialist type of conditional taken by Berto and Özgün (2021) also appears promising to us.

But second, in the absence of such logics, and of a logic of the inferential conditional specifically, we point out that, for all anyone has shown, it is not true that only notions that can be properly formalized can play a role in human psychology. Indeed, we should at least reckon with the possibility that human psychology is inherently messy, and that neatly formalized notions of inference or support can at best play a marginal role in understanding how the mind works.[23,24]

To underpin this, we mention Douven and Williamson (2006) proof to the effect that there can be no purely logico-mathematical definition of the notion of (categorical) belief.[25] No one would suggest that the notion of belief cannot do any substantive work in the psychology of reasoning.

Even more to the point, Carnap spent a large part of his career trying to develop an inductive logic. Several published attempts were severely criticized by Goodman, Putnam, Quine, and others, which did not keep Carnap from trying to "get it right." Only in posthumously published work (Carnap 1980), he gave up, arguing that an inductive *logic* could not be had. In that same work, he presented a theory of inductive reasoning by introducing a precursor of the conceptual spaces

---

[23] From the perspective of the classical computational theory of mind, the idea of a conditional logic makes a lot of sense. If the mind is, at bottom, a Turing machine, then there must be rules for manipulating expressions involving the conditional symbol. Uncovering those rules would yield the logic of conditionals. But in particular in light of the successes of connectionist approaches to the mind, the computational theory has lost much of its erstwhile appeal.

[24] In this connection, we would also like to refer to a remark specifically about counterfactuals that Over and Cruz make (p. x), to wit, that we can profitably study such conditionals "for some time" even if we cannot precisely define what counts as a counterfactual and what does not. One could go one step further and omit the "for some time": even if we will *never* have a definition of the said kind, no one can deny that we know much more about counterfactuals now than we did fifty years back, and there is no reason to believe that any further progress can only be made by first finding a precise definition of counterfactuals. The decisive point is that we can identify clear instances of counterfactuals and also clear instances of conditionals that are *not* counterfactuals. If the class of counterfactuals remains vague around the edges, then that might hamper progress somewhat, but probably no more than vagueness does in many other areas of science that have nevertheless managed to report important successes. (Think of color science, which Clark 1993, p. vii, calls "*the* success story of scientific psychology so far," but in which vagueness is rampant; see, e.g., Douven et al. 2017).

[25] There have been attempts to escape the proof but these have serious drawbacks; see Douven et al. (2018) and Douven and Elqayam (2021).

framework. Suppose Carnap is right and there will never be an inductive logic (his arguments seem pretty compelling to us). Would that mean all the work that has been done on inductive reasoning (e.g., on category-based induction) was for naught? It seems to us that, to the contrary, that furthered our understanding of how people reason inductively considerably, whether or not we will ever have an inductive logic.

In short: True, there currently is no logic of the inferential conditional, but help may well be on its way. Even if, in the end, it turned out that the logics of conditionals now being developed cannot be brought fully in line with inferentialist commitments, or even if no logic could, it is hard to see why that would be bad for inferentialism. It is simply not a priori that our usage of everyday conditionals is governed by a logic. This is not to say that there are no inferential principles concerning conditionals that people tend to rely on. But those need not amount to anything worthy of the name "logic" (e.g., it might turn out to be impossible to gather them into an axiomatic system). And absent *any* logic of conditionals, we should still be able to make progress on understanding the role conditionals play in people's reasoning.

3.2 Inferentialism Is Too Narrow

We have made it clear from when we started working on inferentialism that our aim was to develop a semantics for *indicative* conditionals only, and then only for *standard* or *normal* ones.[26] Among the types of conditionals we excluded were so-called non-interference conditionals, such as,

(17) If hell freezes over, Betty will leave the department.

It is important to be clear about the claim we made about such conditionals. According to Over and Cruz (p. x), we "have gone so far as to deny that non-interference conditionals are conditionals." We have

[26] And really only a *semantics*. At this point, we have nothing to say about conditional threats or conditional promises, which are not the kind of things that can be true or false.

done no such thing. What we have done is draw attention to a distinction that linguists have been making for decades between standard and non-standard conditionals, the latter sometimes also being referred to as "nonconditional conditionals" (Geis and Lycan 1993; Lycan 2001) or "unconditionals" (Merin 2007).[27]

Among the non-standard conditionals are, next to non-interference conditionals, so-called relevance or speech act conditionals or biscuit conditionals, such as

(18) If you're hungry, there are biscuits on the table.

and Dutchman conditionals, such as

(19) If Harry passes the exam, I'm a Dutchman.

As said, we have, from the start, limited our proposal to standard conditionals, but not because we thought it would be impossible to account, in inferentialist terms, for conditionals of the aforementioned types, but rather because, in view of how little progress has been made on the semantics of conditionals, it would seem prudent for anyone wanting to develop a semantics for conditionals to start modestly and focus on standard conditionals first. One step at a time!

Even supposing inferentialism will not be able to account for non-interference conditionals, why would that be so bad? We cannot find a real argument in Over and Cruz (2021), except that they appear to think *any* semantics should be able to account for (at least) non-interference conditionals because (17) "looks like an acceptable conditional to us" (p. 20).

Not knowing who Betty is, how are we to tell? Well, the idea is of course to imagine a context in which a fictional colleague is firmly decided to leave our department and that nothing can change her mind. And yes, in such a context (17) may well be acceptable. What follows?

[27] The point is also missed in Mellor and Bradley (2021).

Consider that it is easy to imagine a context in which (18) is perfectly acceptable, but that there is nothing conditional about it: it asserts *unconditionally* that there are cookies on the table (the antecedent mentions the type of circumstance under which that information is relevant). Similarly, (19) may be acceptable in a context, but again there is nothing conditional about it: it expresses that the speaker deems it highly unlikely that Harry will pass his exam. And the same once more with respect to (17). There is nothing conditional about it: the consequent is asserted unconditionally. In the context we sketched, we understand the conditional as asserting that Betty has decided to leave the department and that nothing is going to change her mind.[28,29]

We have two further comments on this. First, while we have wanted to focus on *indicative* conditionals, it was already pointed out in Douven (2016a, p. 38 f.) that it would take little effort to extend the semantics to cover subjunctive conditionals. Moreover, it would seem equally easy to give an inferentialist account of concessives, by defining "[Even] if A, B" to be true if, and only if, there is a compelling argument for B from background premises alone and also from those premises revised (in the sense of Alchourrón et al. 1985) with A (i.e., given one's current background knowledge, there is a compelling argument from A to B, but A would not be essential to that argument). Finally, an inferentialist account of non-interference conditionals could plausibly look as follows: "If A, B" is true if, and only if, there is a compelling argument from background knowledge alone to B, also from background knowledge revised by A to B, *and* from background knowledge revised by not-A to B. Right now, these are just hypotheses, lacking any empirical support. We leave a full investigation for later, as there is still enough empirical work to be done on the original proposal pertaining to standard indicative conditionals.

Second, suppose inferentialism were forever limited to normal indicative conditionals, so non-interference conditionals never being in its

[28] We should also note that the first author has referred to the standard vs. non-standard distinction in publications long predating the time that we started working on inferentialism (see, e.g., Douven 2008). So the suggestion that the appeal to the distinction was ad hoc—not made in print but often in discussions—is demonstrably unfair.

[29] For an interesting discussion of what defines a conditional see also Elder and Jaszczolt (2016), whose starting point is an observation (based on the International Corpus of English-GB) of the disparity between the syntactic category of a conditional and the conditional meaning.

scope. It is a mystery to us why Over and Cruz (2021, p. x) believe that, in that case, inferentialism would be unfalsifiable. Consider, again, the experiments described in the previous section. Which of those was guaranteed to confirm inferentialism, or to favor it over the account Over and Cruz prefer, just because non-interference conditionals were excluded from the materials? For instance, there was no way in advance to tell that inferential strength would come out as being a much stronger predictor of the probabilities of conditionals than conditional probability, as was found in Douven et al. (2021). Similarly for the results reported in Mirabile and Douven (2020). Could Over and Cruz have predicted that the results would favor inferentialism over their account, just because the materials consisted of abductive inferential conditionals? If so, we would like to see their argument. As far as we can see, the account preferred by Over and Cruz could have prevailed in *all* experiments on inferentialism carried out so far. If it had, that would have been bad news for inferentialism.[30]

3.3 Inferentialism Is Implausible

As Over and Cruz rightly point out, Modus Ponens (MP) is not valid, given inferentialism. There can be a compelling argument from A to B, and B can still be false even if A is true. That is a consequence of the fact that "compelling" does not imply "conclusive." Over and Cruz appear to find this quite damning for inferentialism. And indeed, is MP not a rule of inference one would, pre-theoretically, want any semantics of conditionals to validate? Both from our own experience and looking at experimental data from cognitive psychology, it is obvious that we all tend to rely on this rule almost routinely in our reasoning. In light of this, inferentialism would appear implausible.

As explained in previous publications (e.g., Krzyżanowska et al. 2014), however, that inferentialism invalidates MP is not really a problem,

[30] Would it have *falsified* inferentialism? We are talking statistics here, so the old Popperian terminology is not very helpful. But it would have *disconfirmed* inferentialism, to an extent depending on *how* badly inferential strength would have failed to yield accurate predictions.

given that it will be *typically* the case that if there is a compelling argument from A to B, and A is true, then B is true as well. Because in daily practice we tend to rely much more on compelling-but-inconclusive arguments than on deductively valid ones (Schurz and Hertwig 2019), we would be in big trouble if the arguments we judge to be compelling were not highly truth-conducive. But then MP is, from an inferentialist perspective, highly truth-conducive as well. And why should that not suffice to account for people's reliance on that rule of inference? As was already noted by McGee (1985), who argued on independent grounds that MP is invalid for natural language conditionals, we should not expect our intuitions about validity to be sensitive to the difference between a rule of inference that is guaranteed to preserve truth and one that preserves truth in close to 100% of its applications.

Another objection involves the fact that, as Over and Cruz rightly remark, there are experimental data showing that people judging A true and B false tend to judge "If A, B" false, even if there is an inferential connection between A and B. For instance, (8) is, as said, generally considered to be an inductive–inferential conditional. Given that the vast majority of people living in Chelsea are rich, we are inclined to conclude that John is rich from the assumption that he lives in Chelsea. Although the inferential connection would appear quite strong, we tend to regard the conditional as false if we know that John is poor, or at least not rich, even if he lives in Chelsea. Over and Cruz appear to think that inferentialism is in tension with these data.

Over and Cruz are conflating two things here. According to inferentialism, a conditional is true if, relative to contextual background premises, there is a compelling argument from antecedent to consequent, where the antecedent is essential to the argument (without the antecedent, it loses its compellingness). Now consider that the kind of case Over and Cruz consider concerns conditionals whose consequent is known, or at least judged, to be false. And we simply do not deem anything a compelling argument for something we know, or judge, to be false. If we are convinced that John is poor, then nothing will strike us as a compelling argument for the claim that he is rich. Even if 99% of the people who live in Chelsea are rich, that will not convince us that John is rich, given that we know, or are independently convinced, that he is poor.[31]

To be sure, someone may point out to us that our grounds for believing John to be poor are faulty. Maybe we have been informed of that by an otherwise reliable witness, who, however, in this case had a self-serving reason to lie about John's financial status. Then we may abandon our belief that John is poor and we may come to consider his living in Chelsea—if that is where he lives—as being excellent grounds for believing that he is rich. If we are informed that John lives in Chelsea indeed, we may go from suspension of judgment about John's wealth to believing that he is rich. But already in the situation in which we have suspended our judgment on John's wealth, one might regard (8) to be true. (Whether we will may depend on whether we believe purely inductive support can be enough for a compelling argument, which is debatable; see Nelkin 2000; Douven 2003, for discussion.) To emphasize, in the experiments whose outcomes Over and Cruz deem problematic for inferentialism, the participants were *not* in this situation. They believed certain conditionals to have a false consequent, and thus they judged the conditionals to be false, entirely consistent with inferentialism.

To put Over and Cruz's misunderstanding more succinctly, according to inferentialism A can be true, B can be false, and still "If A, B" can be true (from someone's perspective) because there is a compelling argument from A to B (relative to that person's background knowledge). Over and Cruz seem to misread this as: A person can judge A to be true, judge B to be false, and yet judge "If A, B" to be true. The crucial difference is between the consequent *being* false, and the person judging the conditional's truth value *believing* (rightly or wrongly) the consequent to be false.[32]

---

31 This should also answer Over and Cruz's question of why inferentialists have not produced an intuitive example of a true conditional with a true antecedent and a false consequent (Over and Cruz 2021, p. 18). It is a bit as if Over and Cruz were challenging someone who holds that there are things she was once firmly convinced of that are no longer among her beliefs simply because they slipped from her memory to give an example of such a thing.

32 Skovgaard-Olsen et al. (2017) also miss this point.

3.4 Inferentialism Is Unfalsifiable

Above, we briefly touched upon Over and Cruz's claim that inferentialism is unfalsifiable because it focuses on standard conditionals. They make the same claim in connection with the belief bias analogue that, as mentioned earlier, we found (and replicated several times over). Specifically, we found that conditionals with believable consequent tended to be evaluated as true more often than conditionals with unbelievable consequent. We compared this to the well-documented effect of belief bias, in which inferences with believable conclusions tend to be evaluated as valid more often than inferences with unbelievable conclusion. Over and Cruz argue that this constitutes evidence against, rather than in favor of, inferentialism, and that the conditional probability hypothesis is directly supported by this pattern "without auxiliary hypotheses."

Alas, we think that Over and Cruz fell prey here to a normativist fallacy (Elqayam and Evans 2011): the idea that a theory of thinking must be backed by a normative system. This inevitably leads researchers to narrow their focus to where a normative system can be found. This is also the source of Over and Cruz's subsequent argument, that inferentialism cannot be empirically tested because it does not specify a logic. These are related arguments, and they lead Over and Cruz to argue that inferentialism is not falsifiable, or has limited falsifiability.

We beg to differ.[33] As argued, formal logic is neither necessary nor sufficient for an effective theory of conditionals, or for its empirical testing. Inferentialism has a very simple basic tenet: for agent A, the truth of a conditional, C, is a function of A's estimate of her ability to draw a compelling argument from C's antecedent to its consequent, given background knowledge. This tenet can be simply and directly tested by asking participants to draw an inference from antecedent to consequent, and measure the predictive power of this inference to the truth value of the conditional. Evidence against inferentialism can be entirely straightforward, if this predictive power fails. So far we found no such evidence.

---

[33] We find the term "falsifiable" rather puzzlingly Popperian and will instead refer to testability, and to evidence for or against the theory.

It is not, by any stretch of the imagination, an auxiliary hypothesis to expect inference to behave like inference. After more than half a century of empirical science of reasoning, we know a fair few things about how people draw inferences. Belief bias is a prominent feature of this knowledge. Moreover, this is not an isolated one-off, but rather a persistent pattern found across studies, and, more to the point, across other psychological patterns of inference. The similarity-based study in Douven et al. (2021) is a case in point. In both cases, understanding how people draw inference generates unique, testable predictions.

One final observation: Over and Cruz only refer to inferentialism, but as early on as Douven et al. (2018), we presented a rounded theory of conditionals, Hypothetical Inferential Theory (HIT), with inferentialism as the computational-level theory, supported by a suite of psychological principles constituting the algorithmic-level theory. What we say about inferentialism and the psychological patterns of inference is even more relevant to HIT as a psychological theory. This makes the argument that our belief bias hypothesis serves as an auxiliary even more unsustainable.

4 Conclusion

We hope to have shown that inferentialism, in the version we have been advocating for a number of years now, has already much going for it. Not in the least, it holds the promise of accounting for what is probably our most fundamental intuition about conditionals, to wit, that there is a dependency of the consequent of a conditional on its antecedent. Supposing inferentialism, there is no need to explain away this intuition as being somehow mistaken, nor for waving our hands in the direction of pragmatics, hoping that someday someone will come up with a detailed explanation of how the suggestion of a connection between a conditional's component parts is brought about pragmatically (or just hoping that readers will be happy enough with the hand-waving and not ask for details). In addition to this, inferentialism is backed by the outcomes of several experiments that have been undertaken in the past years, using a wide variety of materials and methodologies.

The position has not remained without criticism, as we saw. As argued, however, these criticisms have largely sprouted either from misunderstandings of inferentialism—concerning its scope, for instance, or concerning the role background knowledge plays in the semantics—or from (implicitly) making unreasonable demands, like that there are still open questions about the logic of the inferential conditional (as if the critics could honestly claim that their own position saw the light of day with all details fully worked out). We welcome the research on conditional logics that take seriously the idea of their being a connection between a conditional's component parts and find much of this research to be promising. At the same time, we noted that it is not a priori that there *must* be a logic of the conditional as used in everyday language. In our view, there is a real possibility that this usage is not governed by principles that can be regimented into anything worthy of the name "logic." We also noted that inferentialism was first presented as a semantics of normal indicative conditionals. We may well be able to extend it beyond those, but—to repeat—we prefer to take one step at a time.

Naturally, this is already to acknowledge that there is still work to be done. This is so even if, at least for a while, we keep confining ourselves to normal indicative conditionals. For example, Mirabile and Douven (2020) based predictions about endorsement rates of MP arguments on inferentialist tenets, but that work should be expanded to cover other argument forms involving one or more conditional premises as well—and there are many more than are usually considered in the psychology of reasoning literature (see Douven 2016a, Ch. 5). We also need to develop our processing account more fully. Although we made some progress by postulating an inference heuristic, we still lack a model of how this inference works.

Work also remains to be done on compounds of conditionals and embedded conditionals, such as negated conditionals (some of which are of the form "If A then not B"), which—inferentialism predicts—*deny* the presence of an inferential connection between antecedent and consequent. Besides, we have so far only looked at simple conditionals, that is, conditionals whose antecedent and consequent are not themselves conditional in form. There are many conditionals not of that sort—so-called nested conditionals—which make perfect intuitive sense yet which have proven a stumbling block for some of the main semantics (e.g., they pose well-known problems for the suppositional account). At least theoretically, inferentialism has no difficulty accounting for nested conditionals. Here is, for instance, one of our favorite examples of a both left- and right-nested conditional (i.e., a conditional whose antecedent and consequent are both conditional in form):

(20) If your mother gets angry if you come home with a B, then she'll get furious if you come home with a C.

According to inferentialism, (20) is true precisely if there is a compelling argument for the claim that your mother will get furious if you come home with a C from the premise that your mother gets angry if you come home with a B, which can be further analyzed as: there is a compelling argument for the conclusion that {you can compellingly argue that your mother gets furious from the premise that you come home with a C} from the premise that {you can compellingly argue that your mother gets angry from the premise that you come home with a B}. Whether this analysis would stand experimental scrutiny remains to be seen, however.[34]

References

Adams, E. W. (1975). *The logic of conditionals*. Dordrecht: Reidel.
Alchourrón, C. E., Gärdenfors, P., & Makinson, D. (1985). On the logic of theory change: Partial meet contraction and revision functions. *Journal of Symbolic Logic*, *50*, 510–530.
Asher, N. & Lascarides, A. (2003). *Logics of conversation*. Cambridge: Cambridge University Press.
Bach, K. (1999). The myth of conventional implicature. *Linguistics and Philosophy*, *22*, 327–366.

[34] We are greatly indebted to Paul Égré and Lance Rips for valuable comments on an earlier version of this paper.

Bach, K. & Harnish, R. (1979). *Linguistic communication and speech acts*. Cambridge MA: MIT Press.

Berto, F. & Özgün, A. (2021). Indicative conditionals: Probabilities and relevance. *Philosophical Studies*, in press.

Braine, M. D. S. & O'Brien, D. P. (1991). A theory of *if*: Lexical entry, reasoning program, and pragmatic principles. *Psychological Review*, *98*, 182–203.

Carnap, R. (1952). Meaning postulates. *Philosophical Studies*, *3*, 65–73.

Carnap, R. (1980). A basic system of inductive logic II. In R. C. Jeffrey (ed.), *Studies in Inductive Logic and Probability* (pp. 7–155). Berkeley CA: University of California Press.

Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication*. Oxford: Blackwell.

Clark, A. (1993). *Sensory qualities*. Oxford: Oxford University Press.

Crupi, V. & Iacona, A. (2020). The evidential conditional. *Erkenntnis*. https://doi.org/10.1007/s10670-020-00332-2.

Crupi, V. & Iacona, A. (2021a). Probability, evidential support, and the logic of conditionals. *Argumenta*, *6*, 211–222.

Crupi, V. & Iacona, A. (2021b). Three ways of being non-material. *Studia Logica*, https://doi.org/10.1007/s11225-021-09949-y.

Cruz, N., Over, D., Oaksford, M., & Baratgin, J. (2016). Centering and the meaning of conditionals. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 1104–1109). Austin TX: Cognitive Science Society.

Dascal, M. (1979). Conversational relevance. In A. Margalit (ed.), *Meaning and use* (pp. 153–174). Dordrecht: Reidel.

Douven, I. (2003). Nelkin on the lottery paradox. *Philosophical Review*, *112*, 395–404.

Douven, I. (2008). The evidential support theory of conditionals. *Synthese*, *164*, 19–44.

Douven, I. (2012a). The lottery paradox and the pragmatics of belief. *Dialectica*, *66*, 351–373.

Douven, I. (2012b). Learning conditional information. *Mind & Language*, *27*, 239–263.

Douven, I. (2016a). *The epistemology of indicative conditionals*. Cambridge: Cambridge University Press.

Douven, I. (2016b). Vagueness, graded membership, and conceptual spaces. *Cognition*, *151*, 80–95.

Douven, I. (2017a). How to account for the oddness of missing-link conditionals. *Synthese*, *194*, 1541–1554.

Douven, I. (2017b). Inference to the best explanation: What is it? And why should we care? In K. McCain & T. Poston (eds.), *Best explanations: New essays on inference to the best explanation* (pp. 4–22). Oxford: Oxford University Press.

Douven, I. (2021). How explanation guides belief change. In *Trends in cognitive sciences*, in press.

Douven, I. (2022). *The art of abduction*. Cambridge MA: MIT Press, in press.

Douven, I., & Elqayam, S. (2021). Rethinking the lottery paradox: A dual processing perspective. In I. Douven (ed.), *Lotteries, knowledge, and rational belief* (pp. 110–127). Cambridge: Cambridge University Press.

Douven, I., Elqayam, S., Gärdenfors, P., & Mirabile, P. (2021). *Conceptual spaces and the strength of similarity-based arguments*. Manuscript.

Douven, I., Elqayam, S., & Mirabile, P. (2021). *Inference strength predicts the probability of conditionals better than conditional probability*. Manuscript.

Douven, I., Elqayam, S., Singmann, H., & van Wijnbergen-Huitink, J. (2018). Conditionals and inferential connections: A hypothetical inferential theory. *Cognitive Psychology*, *101*, 50–81.

Douven, I., Elqayam, S., Singmann, H., & van Wijnbergen-Huitink, J. (2020). Conditionals and inferential connections: Toward a new semantics. *Thinking & Reasoning*, *26*, 311–351.

Douven, I., & Krżyzanowska, K. H. (2019). The semantics–pragmatics interface: An empirical investigation. In A. Capone, M. Carapezza, & F. Lo Piparo (eds.), *Further advances in pragmatics and philosophy: Theories and applications* (pp. 81–100). Dordrecht: Springer.

Douven, I. & Mirabile, P. (2018). Best, second-best, and good-enough explanations: How they matter to reasoning. *Journal of Experimental Psychology: Language, Memory, and Cognition*, *44*, 1792–1813.

Douven, I. & Rott, H. (2018). From probabilities to categorical beliefs: Going beyond toy models. *Journal of Logic and Computation*, *28*, 1099–1124.

Douven, I., Wenmackers, S., Jraissati, Y., & Decock, L. (2017). Measuring graded membership: The case of color. *Cognitive Science*, *41*, 686–722.

Douven, I., & Williamson, T. (2006). Generalizing the lottery paradox. *British Journal for the Philosophy of Science*, *57*, 755–779.

Elder, C.-H., & Jaszczolt, K. M. (2016) Towards a pragmatic category of conditionals. *Journal of Pragmatics*, *98*, 36-53.

Elqayam, S., & Evans, J. St. B. T. (2011). Subtracting 'ought' from 'is': Descriptivism versus normativism in the study of human thinking. *Behavioral and Brain Sciences*, *345*, 233-248.

Elqayam, S., & Over, D. E. (2013). New paradigm psychology of reasoning. *Thinking & Reasoning*, *19*, 249–265.

Evans, J. St. B. T., Barston, J. L., & Pollard, P. (1983). On the conflict between logic and belief in syllogistic reasoning. *Memory & Cognition*, *11*, 295–306.

Fairchild, M. D. (2013). *Color appearance models*. Hoboken, NJ: Wiley.

Gärdenfors, P. (2000). *Conceptual spaces*. Cambridge, MA: MIT Press.

Gärdenfors, P. (2014). *The geometry of meaning*. Cambridge, MA: MIT Press.

Geis, M. L., & Lycan, W. G. (1993). Nonconditional conditionals. *Philosophical Topics*, *21*, 35–56.

Gödel, K. (1931). Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme, I. *Monatshefte für Mathematik und Physik*, *38*, 173–198.

Grice, H. P. (1989). *Studies in the way of words*. Cambridge, MA: Harvard University Press.

Hobbs, J. R. (2004). Abduction in natural language understanding. In L. Horn & G. Ward (eds.), *Handbook of pragmatics* (pp. 724–741). Oxford: Blackwell.

Jraissati, Y., & Douven, I. (2018). Delving deeper into color space. *i-Perception*, *9*, 1–27, https://doi.org/10.1177/2041669518792062.

Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI Publications.

Kneale, W., & Kneale, M. (1962). *The development of logic*. Oxford: Oxford University Press.

Krzyżanowska, K., Wenmackers, S., & Douven, I. (2013). Inferential conditionals and evidentiality. *Journal of Logic, Language and Information, 22*, 315–334.

Krzyżanowska, K. H. (2015). *Between "If" and "Then."* Doctoral dissertation. University of Groningen.

Krzyżanowska, K. H. (2019). What is wrong with false-link conditionals? *Linguistics Vanguard*, *5*, s3, https://doi.org/10.1515/lingvan-2019-0006.

Krzyżanowska, K. H., Collins, P. J., & Hahn, U. (2017). Between a conditional's antecedent and its consequent: Discourse coherence vs. probabilistic relevance. *Cognition*, *164*, 199–205.

Krzyżanowska, K. H., Collins, P. J., & Hahn, U. (2021). True clauses and false connections. *Journal of Memory and Language*, in press.

Krzyżanowska, K. H., & Douven, I. (2018). Missing-link conditionals: Pragmatically infelicitous or semantically defective? *Intercultural Pragmatics*, *15*, 191–211.

Krzyżanowska, K. H., Wenmackers, S., & Douven, I. (2014). Rethinking Gibbard's river-boat argument. *Studia Logica*, *102*, 771–792.

Kyburg, H. (1961). *Probability and the Logic of Rational Belief*. Middletown, CT: Wesleyan University Press.

Lauer, S. (2013). *Towards a Dynamic Pragmatics*. Doctoral dissertation. Stanford University.

Levinson, S. C. (2000). *Presumptive meanings*. Cambridge, MA: MIT Press.

Lycan, W. G. (2001). *Real conditionals*. Oxford: Oxford University Press.

Mackie, J. L. (1973). *Truth, probability and paradox: Studies in philosophical logic*. Oxford: Oxford University Press.

Makinson, D. (1965). The paradox of the preface. *Analysis*, *25*, 205–207.

Marr, D. (1982). *Vision*. Cambridge, MA: MIT Press.

McGee, V. (1985). A counterexample to modus ponens. *Journal of Philosophy*, *82*, 462–471.

Mellor, D. H., & Bradley, R. (2021). Conditionals: Truth, safety, and success. *Mind & Language*, in press.

Merin, A. (2007). *Unconditionals*. Unpublished manuscript, available at http://semanticsarchive.net/Archive/WUwZTk5M/unconditionals.pdf.

Mill, J. S. (1843/1872). *A system of logic: Ratiocinative and inductive* (8th ed.). London: Longmans, Green, Reader, & Dyer.

Mirabile, P., & Douven, I. (2020). Abductive conditionals as a test case for inferentialism. *Cognition*, *200*, 104232, https://doi.org/10.1016/j.cognition.2020.104232.

Nelkin, D. K. (2000). The lottery paradox, knowledge, and rationality. *Philosophical Review*, *109*, 373–409.

Nosofsky, R. M. (1988). Exemplar-based accounts of relations between classification, recognition, and typicality. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*, 54–65.

Nosofsky, R. M. (1989). Further tests of an exemplar-similarity approach to relating identification and categorization. *Perception and Psychophysics*, *45*, 279–290.

Oaksford, M., & Chater, N. (2010). Causation and conditionals in the cognitive science of human reasoning. *The Open Psychology Journal*, *3*, 105–118.

Oaksford, M., & Chater, N. (2013). Dynamic inference and everyday conditional reasoning in the new paradigm. *Thinking & Reasoning*, *19*, 346–379.

Oaksford, M., & Chater, N. (2014). Probabilistic single function dual process theory and logic programming as approaches to non-monotonicity in human vs. artificial reasoning. *Thinking & Reasoning*, *20*, 269–295.

Oaksford, M., & Chater, N. (2017). Causal models and conditional reasoning. In M. R. Waldmann (ed.), *The Oxford handbook of causal reasoning* (pp. 327–346). Oxford: Oxford University Press.

Oaksford, M. & Chater, N. (2020). Integrating causal Bayes nets and inferentialism in conditional inference. In S. Elqayam, I. Douven, J. St. B. T. Evans, & N. Cruz (eds.), *Logic and uncertainty in the human mind* (pp. 116–132). London: Routledge.

Osta-Vélez, M., & Gärdenfors, P. (2020). Category-based induction in conceptual spaces. *Journal of Mathematical Psychology*, *96*, 102357, https://doi.org/10.1016/j.jmp.2020.102357.

Over, D. E. (2009). New paradigm psychology of reasoning. *Thinking & Reasoning*, *15*, 431–438.

Over, D. E., & Cruz, N. (2021). *Indicative and counterfactual conditionals in the psychology of reasoning*. Manuscript.

Over, D. E., Hadjichristidis, C., Evans, J. St. B. T., Handley, S. J., & Sloman, S. A. (2007). The probability of causal conditionals. *Cognitive Psychology*, *54*, 62–97.

Paris, J. B., & Vencovská, A. (2017). Combining analogical support in pure inductive logic. *Erkenntnis*, *82*, 401–419.

Petitot, J. (1989). Morphodynamics and the categorical perception of phonological units. *Theoretical Linguistics*, *15*, 25–71.

Potts, C. (2015). Presupposition and implicature. In S. Lappin & C. Fox (eds.), *The handbook of contemporary semantic theory* (2nd ed, pp. 168–202). Oxford: Wiley-Blackwell.

Raidl, E., Iacona, A., & Crupi, V. (2021). The logic of the evidential conditional. *Review of Symbolic Logic*, in press.

Ramsey, F. P. (1929/1990). General propositions and causality. In his *Philosophical Papers*, edited by D. H. Mellor (pp. 145–163). Cambridge: Cambridge University Press.

Recanati, F. (2003). *Literal meaning*. Cambridge: Cambridge University Press.

Rostworowski, W., Pietrulewicz, N., & Będkowski, M. (2021). Conditionals and specific links: An experimental study. *Synthese*, in press.

Ryle, G. (1950). “If”, “so”, and “because”. In M. Black (ed.), *Philosophical Analysis* (pp. 323–340). Ithaca, NY: Cornell University Press.

Sadock, J. M. (1978). On testing for conversational implicature. In P. Cole (Ed.), *Syntax and semantics Vol. 9: Pragmatics* (pp. 281–297). New York: Academic Press.

Schurz, G. (2019). *Hume's problem solved: The optimality of meta-induction*. Cambridge, MA: MIT Press.

Schurz, G., & Hertwig, R. (2019). Cognitive success: A consequentialist account of rationality and cognition. *Topics in Cognitive Science*, *11*, 7–36.

Shepard, R. N. (1964). Attention and the metric structure of the stimulus space. *Journal of Mathematical Psychology*, *1*, 54–87.

Simon, H. A. (1982). *Models of bounded rationality*. Cambridge, MA: MIT Press.

Skorupski, J. (1989). *John Stuart Mill*. London: Routledge.

Skovgaard-Olsen, N., Kellen, D., Krahl, H., & Klauer, C. (2017). Relevance differently affects the truth, acceptability, and probability evaluations of "and," "but," "therefore," and "if then." *Thinking & Reasoning*, *23*, 449–482.

Skovgaard-Olsen, N., Singmann, H., & Klauer, K. C. (2017). Relevance and reason relations. *Cognitive Science*, *41*, 1202–1215.

Skovgaard-Olsen, N., Collins, P., Krzyżanowska, K., Hahn, U., & Klauer, K. C. (2019). Cancellation, negation, and rejection. *Cognitive Psychology*, *108*, 42–71.

Stalnaker, R. (1975). Indicative conditionals. *Philosophia*, *5*, 269–286.

Tarski, A. (1936). Der Wahrheitsbegriff in den formalisierten Sprachen. *Studia Philosophica*, *1*, 261–405.

Tonhauser, J. (2012). Diagnosing (not-)at-issue content. *Proceedings of Semantics of Under-represented Languages of the Americas VI*, (pp. 239–254). Amherst: GLSA.

van Fraassen, B. C. (1976). Probabilities of conditionals. In W. L. Harper & C. A. Hooker (eds.), *Foundations of probability theory, statistical inference, and statistical theories of science* (pp. 261–301). Dordrecht: Reidel.

van Rooij, R., & Schulz, K. (2019). Conditionals, causality and conditional probability. *Journal of Logic, Language and Information*, *28*, 55–71.

Vidal, M., & Baratgin, J. (2017). A psychological study of unconnected conditionals. *Journal of Cognitive Psychology*, *29*, 769–781.

Zakkou, J. (2018). The cancellability test for conversational implicatures. *Philosophy Compass*, *13*, e12552.

Independence Conditionals

Nicole Cruz and David E. Over

Douven et al. (this volume) have replied to our critique (Over and Cruz this volume) of their version of *truth condition inferentialism* (TCI). They have generously allowed us to have the last word (in this volume) in this debate. Their TCI theory is that there must be a *compelling argument* from *p*, plus background information and not from this information alone, to *q* for a "standard" conditional, *if p then q*, to be true. This argument can be deductive, inductive, abductive, or possibly contain some other kind of step (such as an analogical inference), and it does not have to be conclusive (see Douven et al. this volume, for their truth conditions).

N. Cruz
Institute for Psychology, University of Innsbruck, Innsbruck, Austria
e-mail: Nicole.Cruz@uibk.ac.at

D. E. Over (✉)
Department of Psychology, Durham University, Durham, UK
e-mail: david.over@durham.ac.uk

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_8

Our primary aim, however, is not a point-by-point reply to their counter critique of our arguments, but the positive one of stressing the importance of what we will call *independence conditionals*. We define this term to refer to pragmatic uses of conditionals to convey information about the independence of the antecedents and consequents. Thus, a use of *if p then q* is an "independence conditional" when this use is intended to convey, along with background information, that p and q are *independent* of each other, which implies that the probability of q given p is the probability of q, $P(q|p) = P(q)$, or equivalently that $P(q|p) = P(q|not\text{-}p)$. We will argue that inferentialists should not, as they usually do, set aside these conditionals as somehow "nonconditionals", "nonstandard", or even "unconditionals". They often do this in very brief asides in their papers, or in footnotes to them. Consider the following footnote in Douven et al. (2020), which states that the TCI account introduced by Krzyżanowska et al. (2014) only proposes a semantics for "standard" conditionals and not "unconditionals":

> To avoid spurious debate, it is to be noted that linguists and philosophers have long recognized that there are special classes of conditionals - sometimes called 'nonconditional conditionals' (Lycan 2001) or 'unconditionals' (Merin 2007; Spohn 2013) - which do not require the existence of a connection between their antecedent and consequent. These include Dutchman conditionals (Jackson 1979, 1987), non-interference conditionals (Bennett 2003; Burgess 2004), and relevance conditionals (Bennett 2003). Krzyżanowska and coauthors explicitly propose their brand of inferentialism as a semantics for standard conditionals, not for unconditionals.

It is far from the case that most linguists and philosophers have thought of conditionals with no compelling arguments making a connection between the antecedents and the consequents as "unconditionals", with a different semantics from what Douven et al. call "standard" conditionals (see Csipak and Romero this volume, and Lassiter 2022, and the references given there). We wish to stress that uses of conditionals *if p then q* without a connection between p and q can be perfectly "standard", in an ordinary sense, in natural language and can have an important role in human reasoning, as what we are calling independence conditionals.

Suppose a class of students are to take an especially easy test, and we remark:

(1) If the students do not attend the lectures, they will pass that test.

In using (1) in this context, we would intend to convey the information that passing this test is independent of attending the lectures: the latter is unconnected with the former. Some theorists (referred to in the quote above) would call (1) a non-interference conditional, but we would also term it an independence conditional. We see nothing non-standard in it, or its use here to convey what could be, for the students, important information about independence. Inferentialists might claim that (1) should be taken as a *concessive conditional*, with an implicit "even if" in it:

(2) Even if the students do not attend the lectures, they will pass that test.

With limited space, we will not examine this claim here, but in any event, both non-interference and concessive conditionals, like (1) or (2), are standardly accompanied, usually implicitly but sometimes explicitly, by other indicative conditionals, which must also be true for independence to hold. In our example, this other indicative would be:

(3) If the students attend the lectures, they will pass that test.

In most contexts, the use of (3) would not be an independence conditional, as it would convey, pragmatically, that there is a causal connection between attending the lectures and passing the test, but in the context that we have described, it would be an independence conditional. Background knowledge tells us that the test is so easy to pass that attending the lectures would be redundant. By the TCI truth conditions given in Douven et al. (2020) and Douven et al. (this volume), (3) in this use, in this context, should be "neither true nor false". But we see no justification for coming to this conclusion. Independence in this case depends on the truth of both (1) and (3), and these truths could be established by a statistical analysis after the test, showing no correlation between attending the lectures and passing the test. Notice that (3) would be false when attending the lectures was correlated with failing the test, because the lectures contained misinformation. But that is not what we have in the example we have given.

If participants in an experiment responded that (1) and (3) had a high probability in the kind of context we have described, we would argue that a *probabilistic account* of conditionals (Over and Cruz 2018, this volume; Sanfilippo et al. 2020) had been supported and not TCI (see also Pfeifer this volume). For in this account, the probability of a conditional, *P(if p then q)*, is the conditional probability of *q* given *p*, *P(q|p)*, and *P(q|p)* is high when *P(q)* is high and *q* is independent of *p*. However, matters are not so quite straightforward with some supporters of TCI. They claim, as we have seen, that their theory does not have to apply to a wide range of uses of conditionals. They would have to include (1) and (3), in the context of our example, among the conditionals that their TCI theory does not have to apply to. Douven and his collaborators also claim that there is a *belief bias* effect in the results of Douven et al. (2018), which makes people endorse *if p then q* when they have a high degree of belief in *q*, i.e., *P(q)* is high, although there is not a compelling argument from *p* to *q*. Conditionals *if p then q* will simply be more often evaluated as "true" when *q* is evaluated as "true" (Douven et al. this volume). We have suggested (Over and Cruz this volume) that the falsifiability of TCI as a psychological hypothesis is called into question by adding these two claims to it: that critics cause a "spurious debate" by pointing out cases that go against TCI, and that people have a "belief bias" when their responses conflict with this theory.

Douven and his collaborators go much farther and make the even stronger claim that the supposed "belief bias" results support their TCI theory. They argue that they closely compare judging a conditional to making an inference, and any bias that affects the latter should affect the former. To analyze this claim more deeply, we must first say something about the psychological study of belief bias. As Thompson and Evans (2012) point out in their work on informal belief bias, this response was originally discovered in experiments on syllogisms and logical validity. To focus on one aspect of these experiments, there was some tendency for participants to judge an invalid syllogism as "valid" when its conclusion was believable. For example, there was a tendency for participants to respond that the following invalid syllogism with a believable conclusion was "valid":

No addictive things are inexpensive.
Some cigarettes are inexpensive.
Therefore, some addictive things are not cigarettes.

As Thompson and Evans (2012) outline, various explanations have been given of this belief bias response. To simplify, one proposal is that the participants with belief bias were trying to avoid difficulties. They were asked whether the above syllogistic conclusion necessarily followed from the premises. If the conclusion had been unbelievable, they might have examined the inference more closely, but finding questions about examples like the above difficult, they perhaps responded "yes" merely because they believed the conclusion. Clearly, there is at least a prima facie case, in this kind of experiment on deductive reasoning, with its instructions, for classifying this response as a "bias".

However, the inferences in Douven et al. (2018) that supposedly display "belief bias" are totally unlike syllogistic inferences that are difficult for people who have never studied logic. The former are elementary inferences about blue or green patches (Douven et al. this volume), which require no formal training, and it certainly should not be at all difficult for ordinary people to make unbiased judgments about conditionals referring to them. Consider:

(4) If grass is green, then the sky is blue.

According to everything Douven and his collaborators have said about the TCI truth conditions, (4) should be a prime example of a conditional that is "neither true nor false" by those conditions. On the other hand, according to what they have said about "belief bias", quite a few participants in an experiment would judge that (4) is "true" because of that "bias". But it is obvious that there is no compelling argument from the antecedent of (4) to its consequent, and there is no explanation of the "true" response as a supposed "bias".

Moreover, as Thompson and Evans (2012) also point out, it is more problematic to claim that it is a bias to make use of one's beliefs in wider inferences than logical deductions from assumptions. As they explain, Bayesian principles specify how we should use our prior beliefs in such reasoning. To take account of this fact, there is a new Bayesian approach in the psychology of reasoning (Oaksford and Chater 2007, 2020; Over and Cruz this volume). Suppose some students have a cold, and because they hope to recover from it quickly for an upcoming test, they have an interest in the probability of this conditional:

(5) If we take extra vitamin C, we will recover from the cold within a week.

They consider the argument to the conclusion that they will recover from the cold within a week, *w*, from the premise that they take extra vitamin C, *e*. A little thought about the colds they have had in the past could lead them to a judgment that the probability of *w*, $P(w)$, is high. That might lead on, after further reflection about their past colds, to the inference that *w* is highly probable given *e*, with the result that $P(w|e)$ is high for them. In our account, making this judgment would give them a high degree of confidence in (5). They could then notice that this high probability equals $P(w)$, i.e., $P(w|e) = P(w)$, and they could finally infer the conclusion that *w* is independent of *e*, with (5) having the same probability as:

(6) If we do not take extra vitamin C, we will recover from the cold within a week.

The arguments of Douven and his collaborators imply that there is a "bias" in the reasoning we have just described, and that (5) and (6) must be "neither true nor false". But it is unjustified, from a Bayesian point of view, to condemn an inference as necessarily "biased" when its conclusion is highly probable, and its premise, or premises, are independent of it. The students' reasoning, as we have described it, could be absolutely consistent with Bayesian principles and Bayes' theorem (Oaksford and Chater 2007, 2020). People would, of course, commit a fallacy if they inferred from (5), in this context, that taking extra vitamin C causes recovery from a cold within a week. But this is not what happens in our example, in which the students correctly conclude that recovering from a cold within a week is independent of taking extra vitamin C.

Douven et al. say that we have misunderstood them, and they do not themselves consider what we are calling independence conditionals "unconditionals". We are sorry for any misunderstanding, but we are less convinced than ever that it is a "bias" for the participants in Douven et al. (2018) to endorse a conditional *if p then q* when *q* holds independently of *p*. In the TCI account of Douven et al., as they now explain it, *if p then q* has more than one meaning and can legitimately be used, in a "non-standard" sense, when there is no compelling argument from *p* to *q*. The participants in Douven et al. (2018) may be thinking of this sense when they endorse a use of *if p then q* without a compelling argument from *p* to *q*, and so we still do not see why they are said to have a "bias". The problem for Douven et al. is to explain precisely how one can tell which sense, "standard" or "non-standard", is being assigned to *if p then q* in a given context, by the participants in an experiment or anyone else.

We have a parallel problem of specifying how we can distinguish one pragmatic use of a conditional from another. How can we tell whether (5), for instance, is being used to support an argument that taking extra vitamin C causes quick recovery from colds, or is part of an argument that the recovery is independent of taking the vitamin supplement? We cannot write more about this problem here, but we would point to Lassiter (2022) as an important source which is relevant to this problem in pragmatics.

Douven et al. (this volume) also respond to our questions about the falsifiability of their theory. They refer to a number of experimental results that, in their view, support their position. We do not have the space here to discuss further problems, from our point of view, with the interpretation of their experiments. But we are happy to agree that a straightforward version of TCI can be disconfirmed (which as Bayesians is what we mean by "falsifiable"): the psychological hypothesis that people will only endorse *if p then q* when they believe there is a compelling argument from *p* to *q*. Some experiments do disconfirm this version of TCI: Cruz et al. (2016), Pfeifer (this volume), and Skovgaard-Olsen et al. (2017). This version is also disconfirmed by the supposed "belief bias" results in Douven et al. (2018).

As we have seen, Douven et al. (2018, 2020, this volume) argue that those "belief bias" results support their less straightforward version of TCI. "Belief bias" is supposed to give people a tendency to endorse *if p then q* as true when they believe *q*, even if there is not a compelling argument from *p* to *q*. But consider their own example, in this volume, of a conditional that is supposed to be absurd on semantic grounds:

(7) If Eisenhower was the 34th American president, then Newton is the author of *Principia Mathematica*.

Douven et al. can certainly predict that some participants in an experiment will not judge (7) to be true. But according to their interpretation of their "belief bias" results, other participants, who believe that Newton is the author of *Principia Mathematica*, will endorse (7) as true. Yet another possibility, by what Douven et al. say, is that a further group of participants will interpret (7) as true as a "non-standard" conditional. We could, of course, make the same points about example (4). Therefore, we would still question the falsifiability of the Douven et al. version of TCI.

We also remain puzzled by what Douven and his collaborators are claiming about *modus ponens* (MP): inferring *q* from *if p then q* and *p*. Douven et al. (2020), and Mirabile and Douven (2020), argued that MP does not always preserve truth, and so is logically invalid, according to TCI. Their argument was that there can be a compelling, but non-conclusive, argument from *p* to *q*, making *if p then q* true by the TCI truth conditions, although *p* is true and *q* is false. Mirabile and Douven suggested that this fact could explain, from a TCI point of view, why fewer than 100% of participants in experiments on reasoning endorse MP as a valid inference. The endorsement rate of MP is extremely high, but not quite 100%. Douven and his collaborators know that the slightly lower rate of just under 100% might be the noise that is found in all experiments on reasoning, but they also hypothesize that this lower rate could result from some participant awareness that MP does not always preserve truth. In our chapter (this volume), we asked for an example of *if p then q* that was "true" when *p* is true and *q* false, and we referred to experimental evidence that, when participants are presented with a true *p* and a false *q*, they respond that *if p then q* is false, whether or not there is a general correlation between *p* and *q* (Skovgaard-Olsen et al. 2017).

With questionable consistency, Douven et al. (this volume) respond that our request is impossible to satisfy, because, when *p* is true and *q* is false, there is not really a compelling argument from *p* to *q* to make *if p then q* true, and that participants in experiments are aware of this fact. To begin with, this claim undercuts the Mirabile and Douven proposed TCI explanation of why MP is not endorsed at a rate of 100% in experiments. If participants in an experiment cannot conceive of a true *if p then q* with a true *p* and a false *q*, then they will not hesitate to endorse the validity of MP, and the slightly less than 100% endorsement rate will only be the noise that the vast majority of psychologists of reasoning take it to be.

Douven et al. now seem to be saying no more than that some people can believe that *if p then q* is "true" when they are unaware that *p* is true and *q* is false. Such people can mistakenly believe that *q* is true when they come to believe *p*, but with the major premise false, this is not a case in which MP fails to preserve truth. Curiously for us as well, Douven and his collaborators refer to the claimed counterexample to MP in McGee (1985) as if it could be a case in which *if p then (if not-q then r)* is true when *p* is true and *(if not-q then r)* is false. But if a debatable example of the supposed failure of MP to preserve truth, as implied by TCI, can be given in this compound case, then we do not see why a debatable example cannot be given, as we requested, for a simple *if p then q* and *p* and *q*. Basically, if MP does sometimes fail to preserve the truth for such an *if p then q*, as held by Douven and his collaborators, then there should be an example of this.

We would like, however, to end on a positive note, stressing the importance of what we call independence conditionals. As we have recalled above, we are arguing for a probabilistic account of conditionals, in which $P(\text{if } p \text{ then } q) = P(q|p)$. This fundamental relation allows conditionals to be interpreted pragmatically as epistemic or causal conditionals in some contexts, when *p* raises the probability of *q*, $P(q|p) > P(q|\text{not-}p)$, and also as independence conditionals in other contexts, when *q* is independent of *p*, *P(q|p)* = *P(q|not-p)*. But we would observe further that these two general uses of conditionals are closely related in our reasoning. For example, we could not successfully reason about causes and effects unless we could sometimes infer, and convey in our use of conditionals, that *q* is independent of *p*. Reasoning about causes and effects goes hand in hand with reasoning about independence, and conditionals are central to all of this reasoning in our probabilistic account.

References

Bennett, J. (2003). *A philosophical guide to conditionals*. Oxford: Oxford University Press.

Burgess, J. P. (2004). Review of J. Bennett. A philosophical guide to conditionals. *Bulletin of Symbolic Logic, 10(4)*, 565–570. https://doi.org/10.1017/S1079898600003632

Cruz, N., Over, D., Oaksford, M., & Baratgin, J. (2016). Centering and the meaning of conditionals. In A. Papafragou, D. Grodner, D. Mirman, & J. C. Trueswell (Eds.), *Proceedings of the 38th Annual Conference of the Cognitive Science Society* (pp. 1104–1109). Austin, US: Cognitive Science Society.

Douven, I., Elqayam, S., Singmann, H., & van Wijnbergen-Huitink, J. (2018). Conditionals and inferential connections: A hypothetical inferential theory. *Cognitive Psychology*, *101*, 50–81.

Douven, I., Elqayam, S., Singmann, H., & van Wijnbergen-Huitink, J. (2020). Conditionals and inferential connections: Toward a new semantics. *Thinking & Reasoning*, *26(3)*, 311–351.

Jackson, F. (1979). On assertion and indicative conditionals. *Philosophical Review*, *88(4)*, 565–589. https://doi.org/10.2307/2184845

Jackson, F. (1987). *Conditionals*. Oxford: Blackwell.

Krzyżanowska, K., Wenmackers, S., & Douven, I. (2014). Rethinking Gibbard's riverboat argument. *Studia Logica*, *102*, 771–792.

Lassiter, D. (2022). Decomposing relevance in conditionals. *Mind & Language*. Available at: https://web.stanford.edu/~danlass/

Lycan, W. G. (2001). *Real conditionals*. Oxford: Oxford University Press.

McGee, V. (1985). A counterexample to modus ponens. *Journal of Philosophy*, *82(9)*, 462–471.

Merin, A. (2007). Unconditionals (Unpublished manuscript). Retrieved from http://semanticsarchive.net/Archive/WUwZTk5M/unconditionals.pdf

Mirabile, P., & Douven, I. (2020). Abductive conditionals as a test case for inferentialism. *Cognition*, *200*, 104232.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford: Oxford University Press.

Oaksford, M., & Chater, N. (2020). New paradigms in the psychology of reasoning. *Annual Review of Psychology*, *71*, 305–330.

Over, D. E., & Cruz, N. (2018). Probabilistic accounts of conditional reasoning. In Linden J. Ball & Valerie A. Thompson (Eds.), *International handbook of thinking and reasoning* (pp. 434–450). Hove: Psychology Press.

Sanfilippo, G., Gilio, A., Over, D. E., & Pfeifer, N. (2020). Probabilities of conditionals and previsions of iterated conditionals. *International Journal of Approximate Reasoning*, *121*, 150–173.

Skovgaard-Olsen, N., Kellen, D., Krahl, H., Klauer, C. (2017). Relevance differently affects the truth, acceptability, and probability evaluations of ‘and’, ‘but’, ‘therefore’, and ‘if then’. *Thinking & Reasoning*, *23*, 449–482.

Spohn, W. (2013). A ranking-theoretic approach to conditionals. *Cognitive Science*, *37*(6), 1074–1106. https://doi.org/10.1111/cogs.12057

Thompson, V. A., & Evans, J. St. B. T. (2012). Belief bias in informal reasoning. *Thinking & Reasoning*, *18*, 278–310.

Experimenting with (Conditional) Perfection: Tests of the Exhaustivity Theory

Fabrizio Cariani and Lance J. Rips

1 Introduction

Conditional claims like (1) can, in the right context, convey additional information—the sort of information carried by (2) or (3):

(1) If she turns in the final paper, she will pass the class.
(2) Only if she turns in the final paper, will she pass the class.
(3) If she does not turn in the final paper, she will not pass the class.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/978-3-031-05682-6_9.

F. Cariani
University of Maryland, College Park, MD, USA
e-mail: fabrizio@umd.edu

L. J. Rips (✉)
Psychology Department, Northwestern University, Evanston, IL, USA
e-mail: rips@northwestern.edu

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_9

More generally, in the appropriate contexts, from an utterance of *If A, B* hearers can infer the biconditional *A if and only if B*. Following Geis and Zwicky (1971), we will call this inference *conditional perfection* (sometimes simply *perfection*, since it is the only kind of perfection we will be concerned with).

There is strong reason to think that perfection is not a purely logical inference—that is, it's not to be explained exclusively as a semantic entailment of conditional sentences.[1] After all, theories of conditionals universally agree that *If A, B* does not entail *If B, A* (and that it does not entail *A if and only if B*). Moreover, there are many conditional statements and contexts that do not trigger perfection.

(4) If this cactus grows native to Idaho, then it's not an Astrophytum (Lilje 1972).
(5) If you look at this Canaletto painting, you'll get a good idea of what the Canal Grande looks like (Herburger 2015).

It is natural, then, to explore the idea that perfection arises, when it does, from some kind of pragmatic reasoning. Perhaps this sort of pragmatic reasoning is systematic enough to be derivable from general pragmatic principles (that is without leveraging very specific assumptions about context).

In this paper, we present a series of experiments designed to test one of the most promising pragmatic accounts of perfection. This is von Fintel's (2001) idea that whether perfection arises depends on what kinds of questions are, implicitly or explicitly, driving the inquiry of the participants to the conversation (we present the essentials of this account in the next section). Because our findings are mostly negative, we emphasize at the outset that von Fintel's proposal is avowedly speculative. We are interested in testing it experimentally because it is a plausible working hypothesis. Even if it were not the whole story, it is worthwhile exploring which elements of it are supported by experimental scrutiny and which aren't.

[1] We will later consider a view on which conditional perfection does arise as a semantic entailment, but only when additional covert material is present. In other words, even at that point, we remain committed to the view that it is not a semantic entailment *of the conditional*.

Before proceeding, note that, while perfection is interesting in its own right as a problem for pragmatics, it has a wider significance. For instance, we believe that an account of conditional perfection might illuminate why participants in reasoning experiments are sometimes willing to endorse the fallacious inferences of *Affirming the Consequent* and *Denying the Antecedent*—the invalid cousins of *Modus Ponens* and *Modus Tollens*.[2] Indeed, the link between perfection and these argument forms is at the center of our experiments.

2 From Implicatures to Exhaustivity

Conditional perfection, we said, is not a purely logical inference. More generally, it seems wrong to maintain that it is the sort of inference that arises solely on the basis of the linguistic material that is explicitly made available by a conditional sentence. Geis and Zwicky (1971) illustrated the phenomenon of perfection with:

(6) If you mow the lawn, I will give you five dollars.

No doubt, they had in mind a context in which (6) conveys the information that mowing the lawn is the only way for the hearer to get the five dollars. But it's easy to come up with contexts in which this is not true. For instance, imagine (6) being uttered immediately after:

(7) If you tidy your room, I will give you five dollars.

In this case, the surrounding discourse seems to block perfection. In general, linguistic context can determine whether the inference is licensed or blocked.

[2] The connection between conditional perfection and the fallacies is already made in Geis and Zwicky (1971, p. 562). For reviews of the experimental findings concerning the conditional fallacies, see Evans and Over (2004, Chapter 3) and Oaksford and Chater (2007, Chapter 5). For a proposal about how to account for the findings, given a story about perfection, see Cariani and Rips (2017).

Another important aspect of conditional perfection is that there are multiple paths to generate a biconditional interpretation from a conditional (Van Canegem-Ardijns and Van Belle 2008). In particular, for a given conditional *If A, B,* we may note at least four distinct paths to perfection, depending on whether we strengthen with:

- converse: *if B, A*
- obverse: *if* $\sim A$, $\sim B$
- exhaustified conditional: *Only If A, B*
- exhaustified contrapositive: *Only If* $\sim B$, $\sim A$.

As Van Canegem-Ardijns and Van Belle (2008) note, these paths are not equivalent (see also Bonnefon and Polizer 2010; Franke 2009, p. 235). For instance, they claim that (8) invites the inference to (9) but not to (10).

(8) If you empty a bucket of oil on the street, the street will get slippery.
(9) If you do not empty a bucket of oil on the street, the street won't get slippery.
(10) Only if you empty a bucket of oil on the street, will the street get slippery.

The focus in our experiments is on perfection inferences that go via converse and obverse.

So, what features of context license perfection? There is a long history of controversial pragmatic analyses.[3] Fast-forwarding to one of the later stages of this debate, Horn (2000) proposes a view that (with interpretational help from von Fintel 2001) we understand as follows: conditional perfection is licensed when *If A, B* is in Gricean competition with *B, no matter what*. By uttering *If A, B*, a cooperative speaker conveys her inability to make the stronger statement that *B* holds unconditionally (that is, to assert: *B, no matter what*). That does license a hearer to infer something that goes beyond *If A, B*. But, as von Fintel (2001) points out, that license is well short of perfection. Even if one's assertive utterance of

[3] In addition to Geis and Zwicky (1971), see van der Auwera (1997), Horn (2000), and von Fintel (2001).

If A, B signals reluctance to assert *B, no matter what*, it does not rule out another antecedent D, not equivalent to A, such that *If D, B*. So it does not imply that A is the only such antecedent (i.e., that A is necessary as well as sufficient for B).

The exact upshot of Horn's proposal depends on the background theory of conditionals. Let us adopt one theory as a starting point. Suppose that conditionals are universal quantifiers over a contextually set domain of possible worlds.

> **Strict:** *If A, B* is true in context C at world w iff for all $v \in C(w)$, either A is false at v or B is true at v .

Here, $C(\cdot)$ denotes a function from worlds to sets of worlds. Intuitively, this is interpreted as the set of worlds that are relevant to the evaluation of the conditional, given that the world of evaluation is w.

Given this semantics, the most likely explanation for one's not being in a position to assert *B, no matter what* would be that one believes that B might be false in some of the relevant worlds. But if the additional information conveyed by *If A, B* is just that B might be false in some salient possibility, we are far from perfection. Perfection would require not just that B fails at some of the relevant worlds at which A fails; it requires the truth of the converse conditional (i.e., *If B, A*), and so it requires that every relevant B-world is an A-world.[4] Taking stock: Gricean competition between *If A, B* and *B, no matter what* seems to get at a real phenomenon. But that phenomenon is not perfection.

Why, then, is it sometimes possible to get the full perfection inference? Von Fintel (2001) sketches a different sort of pragmatic account, following an insight in Cornulier (1983). Commenting on the example,

(11) One can take this seat if one is disabled or one is older than 70.

Cornulier remarks:

[4] Or at least, minimally, that both A and B fail at the actual world. The Horn-inspired proposal fails this more minimal standard as well.

> For we can suppose, very roughly, that in [*One can take this seat if one is disabled or if one is older than 70*] the word *if* keeps its merely sufficient condition meaning, and that the utterance situation suggests that if other sufficient conditions (allowing one to sit there) did exist, they would have been mentioned, so that the only mentioned property (to be disabled or to be older than 70) is the only property which gives one the right to sit there (presumption of exhaustivity). (Cornulier 1983: 248)

Incidentally, Cornulier's example is especially interesting because it is a conditional permission, which might actually trigger a particularly strong form of perfection inference (see Sect. 3 for discussion; however, our experiments did not target conditional permissions).

Von Fintel's (2001) central move is to connect Cornulier's talk about exhaustivity with off-the-shelf work on exhaustivity in response to questions (specifically Groenendijk and Stokhof 1984; for a survey of work on exhaustivity see van Rooij and Schulz 2003; for another development of Cornulier's insight, see Franke 2009). The key assumption is that we generally, though not always, interpret simple answers to questions as exhaustive. If the doctor asks, "What did you drink last night?" and Lucy replies "Two glasses of wine," we take Lucy to provide an exhaustive list of what she drank.

This idea can be extended to conditionals. Consider a question Q such that (i) *If A, B* is a possible answer to Q and (ii) when *if A, B* is provided as an answer to Q, the conditional is naturally given an exhaustive interpretation. There are many different questions conditionals might be used to answer. It will be useful to isolate two categories of questions: the first category consists of questions about how the consequent might come about. We call them *consequent-directed* (abbreviated [CONS?]). Here are some examples of [CONS?] questions:

- What are all the ways in which *B*?
- How might *B* happen?
- Is *B* true?

The second category consists of questions about what follows from the antecedent.

- What follows from *A*?
- What happens if *A*?

To make this concrete, consider an example like (6). This might be an answer to [CONS?] questions like: "What are all the ways in which I might get five dollars?", "Will I get five dollars?"; it might also be an answer to [ANT?] questions like "What happens if I mow the lawn?". Moreover, we should not suppose that these are the only questions: [CONS?] and [ANT?] questions are not exhaustive categories.

Von Fintel's claim is that [CONS?] questions, but not [ANT?] ones, might help yield something like perfection. His argument starts, like ours, with the assumption that the truth conditions for *If A, B* are captured by **Strict**.

Here, then is one possible way of deriving perfection given the assumptions we have on the table. Suppose that a [CONS?] question Q_{cons}, was asked and that it was answered by *If A, B*. Then:

> *Step 1:* provided that *If A, B* is understood as an exhaustive answer to Q_{cons}, the speaker is not in a position to assert *If D, B* for any *D* that competes with *A*.
>
> *Step 2:* provided that the speaker is informed about the truth-values of these conditionals, she must believe all conditionals of the form *If D, B* are false.
>
> *Step 3:* if all conditionals of the form *If D, B* are false for every antecedent that competes with *A*, then *if not A, not B* must be true.

To justify the reasoning from Step 2 to Step 3 we need another assumption. One approach might stipulate that the space of competitors to *A* is particularly rich:

> **Competitors**: the antecedents that compete with *A* are all those antecedents *D* that do not entail *A*.

This entails that there are *lots* of alternatives. In particular, it entails that for each *not A* world w, there is an alternative conditional (*If* Sw, B) where Sw is a sentence that is only true in w.[5]

Another way of justifying this step, a more plausible one in our view, is to assume (i) that conditionals satisfy the principle of Conditional Excluded Middle (CEM) so that from *~(If D, B)* one might infer *If D, ~B* and (ii) that every relevant possibility is included in some competitor or other. We think that this second approach fits best with the rest of von Fintel's theoretical framework—as von Fintel's (1997, Section 7) argues that **Strict** truth conditions for *If* can be made compatible with CEM by hypothesizing that conditionals come with a homogeneity presupposition.[6]

Either way, von Fintel's account predicts that conditional perfection should arise precisely when the other assumptions in the above reasoning are satisfied. We take **Strict** to be a valuable working hypothesis, so the key assumptions to focus on, as we move to the experimental part of our paper are: (i) that there is some principled way in which the relevant competitors to A are generated; (ii) that *If A, B* is understood as an exhaustive answer to [CONS?] questions; and (iii) that the speaker is relevantly informed.

[5] Concluding the reasoning: asserting *If A, B* would implicate ~(*If* Sw, B), which entails that there is a world that verifies Sw & ~B. But since only w verifies Sw, then w must verify ~B. Since w was an arbitrary world that does not verify A, it follows that every ~A-world is a ~B-world.

[6] Of course, the canonical way of validating CEM is to adopt Stalnaker's semantics for conditionals (Stalnaker 1968, 1981): *If A, B* is true at w iff B is true at the selected A-world. Yet another option, one that we have some stake in, starts with the argument in Cariani and Santorio (2018) that *will* is a "selectional" modal (that its contribution is to select a world out of a modal base). They note that combining a restrictor semantics in the style of Kratzer (1991, 2012) with selectional modals gives something roughly like Stalnaker's semantics for conditionals of the form *If A, will B*—specifically, a semantics that validates CEM. One might extend this insight to a broader class of conditional sentences by postulating that conditionals can sometimes restrict covert selectional modals (Cariani 2021).

3 Designing Experimental Tests of Perfection Inferences

Before going through the battery of experiments we ran, it's useful to describe informally how we set about experimenting with perfection. There are many decisions to be made in setting up such an experiment, and small variants might result in significant changes.

One guiding idea that motivated us is that if perfection arises, we should see unusually high endorsement rates for the conditional fallacies of *Affirming the Consequent* (AC) and *Denying the Antecedent* (DA) (see Geis and Zwicky 1971).

(AC) *B, If A, B.* Therefore: *A*
(DA) *not A*, *If A, B*. Therefore: *not B*.

The idea here is that if, in context, *If A, B* conveys *If B, A* then, barring interference, the endorsement rate for AC should approach the endorsement rate for Modus Ponens. Similarly, if *If A, B* conveys *If not A, not B*, then DA should also approach Modus Ponens.

We have chosen to test endorsement rates for conditional inferences partly because we could rely on a wealth of established and very robust data concerning people's endorsement of these patterns. Figure 1 summarizes endorsement rates for MP, MT, AC, and DA from earlier experiments. Incidentally, it reveals one of the important discoveries in the psychology of reasoning: The endorsement rate of Modus Ponens is higher than the endorsement rate of Modus Tollens (even for bare, non-modalized conditionals).[7]

[7] See the references mentioned in Footnote 3 for reviews of this evidence. The data in Fig. 1 come from Evans et al.'s (1993) review of earlier studies of conditional inference (their Table 2.4). As Evans et al. state, the participants were adults (typically college students). They were "normally given the premises and conclusions and asked whether it follows, or else given a list of conclusions including the normal one and 'nothing follows' to choose from. All these studies involve either so called 'abstract' problem material or ones which are concrete but arbitrary, so that prior beliefs and pragmatic associations are not likely to be cued" (pp. 35–36). Recall that our approach is to test the endorsement rates of AC and DA. When perfection is triggered, these inferences aren't fallacies, but actually valid applications of modus ponens or modus

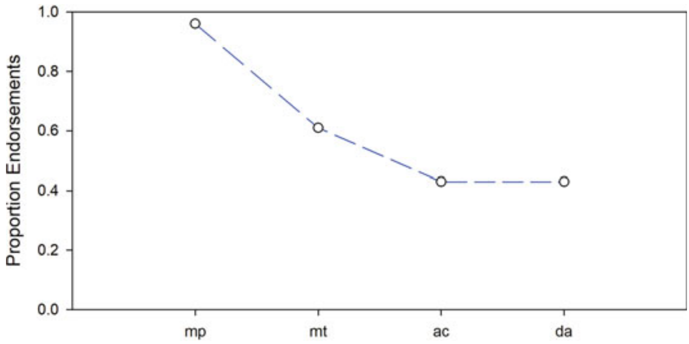


Fig. 1 Standard endorsement rates for the four inference patterns modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da) across previous experiments (data from Evans et al. 1993, Table 2.4 and are weighted averages from seven earlier experiments)

We could have chosen a different approach had we decided to test whether people endorse the inference from *If A, B* to *If not A, not B*. As we just noted, our choice of two-premise arguments (e.g., *If A, B* and *Not A* to *Not B*) allowed us to use the large set of earlier results, summarized in Fig. 1, as a baseline for our manipulations. One could argue that the one-premise approach is a more direct test of whether people perfect conditionals. However, we know of no evidence to suggest that people find the one-premise arguments easier to think about than their two-premise counterparts. For example, the rate of endorsement for the inference from *If A, B* to *If not A, not B* (for "contingent inferences," such as "If the mushroom is red, it is edible," in Fillenbaum 1975) is within the range of endorsements for DA arguments in the studies summarized in Fig. 1. While our data are primarily of the two-premise kind, we recognize both kinds of data as significant to an account of conditional perfection. We report evidence of the one-premise type in Experiment 7.

tollens. Thus, when we think about, say, AC, there are two possibilities that might lead to different endorsement rates:

i. people infer from *If A, B* to *If B, A* and then apply MP to *If B, A* and *B*.
ii. people infer from *If A, B* to *If not A, not B* and then apply MT to *If not A, not B* and *B*.

This difference might matter to our project, and we should be mindful of it in interpreting our results.

We note that psychological research provides many examples in which college-aged participants endorse AC and DA at rates greater than those that appear in Fig. 1. The Fig. 1 experiments used conditionals for which people do not have strong beliefs in the necessity of the antecedent for the consequent. But if prior beliefs *do* suggest (causal) necessity, participants find AC and DA more congenial (e.g., Cummins et al. 1991; Marcus and Rips 1979; Staudenmayer 1975; Thompson 1994). For example, a conditional like *If the butter is heated, then it melts* accords with people's belief that heating is both necessary and sufficient for the butter melting. It's difficult to think of other ways to melt butter aside from heating it. So an AC or DA argument with such a conditional appears correct to many participants (e.g., *If the butter is heated, then it melts; the butter melts; therefore, the butter was heated*). Effects of this sort are similar to those of conditional perfection in that they encourage a reading like that of *A iff B*, but they arise for reasons other than the pragmatic considerations that we have discussed so far. Because our intent in these experiments is to test the pragmatic account of Sect. 2, we chose conditionals with antecedents that don't already suggest necessity.[8]

A more pertinent set of studies have examined conditional promises [e.g., (6)–(7)] and threats (e.g., *If you continue to disrupt the class, you'll have to leave the room*). These studies have shown greater acceptance of AC and DA for promises and threats than for ordinary indicatives (e.g., *If there is an electrical failure, school will be closed*, Markovits and Lesage 1990; *If the student is doing economics then he is a socialist*, Newstead

[8] For related reasons, the research we describe here is neutral with respect to probabilistic theories of conditionals. A number of theorists have proposed that whether a conditional sentence is appropriate depends on the conditional probability of its consequent given its antecedent (e.g., Adams 1965; Evans and Over 2004). If these theories are correct, then we should expect conditional perfection when both the conditional probability of the consequent given the antecedent and that of the antecedent given the consequent are sufficiently high (as they are in the butter heating example). In addition to constraining the semantics to a lesser degree, the pragmatic theory we explore here maintains that whatever the success of these probabilistic accounts, a further source of conditional perfection is a conversational demand for exhaustivity, which can perfect conditionals even if the conditional probability of the antecedent given the consequent is not initially high.

et al. 1997). Similarly, participants are more willing to accept the inference from *If A, B* to *If not A, not B* under the same circumstances (Fillenbaum 1975). The goal of a promise or threat is usually to get the addressee to perform some action (e.g., mow the lawn) or to refrain from one (disrupting the class). For these speech acts to be effective, speakers presumably intend not to provide the promised reward if the action is not taken and not to carry out the threatened punishment if the infraction is not committed. These presumptions can be overridden in the right circumstances, as we've already noted in Sect. 2. But by default, promises and threats convey an exhaustive interpretation, as the experiments we've just cited suggest. Our aim in the present experiments is to manipulate participants' impression of exhaustivity by means of [CONS?] questions, so we picked conditionals for our experiments that (unlike promises and threats) don't by themselves convey perfection.

The second guiding idea in shaping our tasks was that we needed to create matched items for purposes of experimental control. One version of each item used [CONS?] questions; the other either used no question at all or [ANT?] questions—depending on the experiment. In the following, we call this variable *question type*. Here is an example of one of our [CONS?] items (*modulo* some differences in exact wording across experiments):

> John has taken a test on Chapters 4–6 that has not been graded yet.
> [You ask Mary, "Did John do well on the test?"]
> Mary says, "If John understood Chapter 5, then John did well on the test."
> Assume that Mary's response is true and that John did well on the test. Given this information, then, does Mary's statement imply that John understood Chapter 5?

When the [CONS?] question was a polar question (i.e., a yes/no question), the matching item did not have a question at all (so it is simply the result of removing the bit in square brackets).

Other experiments involved more complicated [CONS?] questions, such as:

> You ask Mary, "What are all the ways John could manage to do well on the test?"

Here, the matching item replaced the line in which we asked the [CONS?] question with a line in which we asked the [ANT?] question:

> You ask Mary, "What are all the things that could happen in case John understood Chapter 5?"

Note also that the sample item above is testing for Affirming the Consequent. For each vignette (there were 16 of them), we had items that tested Modus Ponens, Tollens, Affirming the Consequent, and Denying the Antecedent.

In total, this means that we associated each vignette with eight possible items (four inferences for each of the two possible questions). Participants saw the vignettes presented one-at-a-time on a computer screen, in a new random order for each participant. They responded by clicking on one of two options (e.g., "implies" vs. "does not imply").

Participants were college students enrolled in an introductory psychology course, and they completed the experiment as part of a course requirement. No participant took part in more than one experiment.

Note that [CONS?] questions like "What are all the ways John could manage to do well on the test?" appear more complex, at least syntactically, than other possible [CONS?] questions, such as "Did John do well on the test?". In most of our experiments, we have privileged the lengthier question because it reduces the permissibility of *mention-some* answers, which are partial answers. Consider the question:

(12) Q: Where can I buy Stephen King novels?
A: At Powell's Books.

In (12), we do not expect the answer to be an exhaustive catalog of the places where the questioner can buy Stephen King novels. It is important to avoid this interpretation, for the account of perfection we sketched above breaks down if the answer is understood to be partial.

To have a better chance of ruling out *mention-some* answers (that is: for the question to set up the presumption that any answer would be exhaustive), we might ask a different sort of question, such as[9]:

(13) Q: What are all the places where I can buy Stephen King novels?
A: At Powell's Books.

This is why most of our experiments use "What are all the ways in which *B*?" as the [CONS?] question. The Online Appendix lists all 16 vignettes with the full [CONS?] and matched [ANT?] questions for the problems testing Denying the Antecedent. These are the items as they appeared in Experiment 2. We mention experimental variations in wording as they come up in the relevant parts of Sect. 4.

While this is a good reason to run experimental tests with the more complex phrasing, it is not a good reason *not to* test the more natural [CONS?] questions. It is for this reason that our leading experiment involves a simple polar question, as in our initial example above (in which "you" ask Mary, "Did John do well on the test?").[10]

4 The Sequence of Experiments

A quest for perfection drives the experiments we report here. In each of them, we present participants with a series of problems, variations on the standard set that we described earlier. These problems vary the inference type (Modus Ponens, Tollens, Affirming the Consequent, and Denying the Antecedent) and question type ([CONS?] and [ANT?]) in the hope that a demand for an exhaustive set of reasons for the consequent—[CONS?] but not [ANT?]—will lead participants to perfect the conditional. According to the exhaustivity hypothesis, perfection should

---

[9] Perhaps, even the question in (13) sometimes permits a *mention-some* answer. We discuss the implications of this point in Sect. 5.

[10] The reason why we did not use an [ANT?] question as a comparison in this case is that it is implausible to target the antecedent with one such question (say, "Did John understand Chapter 5?"). Except for some recherché contexts, the conditional "If John understood Chapter 5, then John did well on the test" is not an acceptable answer to the question: "Did John understand Chapter 5?".

be manifested by increased endorsement of the “fallacies,” Denying the Antecedent and Affirming the Consequent. As mentioned, these arguments switch from invalid to valid under a perfected conditional, which should increase the likelihood that participants will accept them. The same manipulation, however, should have only a weak effect, if any, on the endorsement of Modus Ponens and Modus Tollens, since these inferences are already valid under the “unperfected” reading of the conditional and remain so under the perfected reading. This, then, is the perfection pattern of our quest: Greater endorsement of Affirming the Consequent and Denying the Antecedent under [CONS?] questions but not [ANT?] questions, but little change in endorsement of Modus Ponens and Modus Tollens.

Experiment 1: Polar Questions

We start with our barest experiment.[11] As described, this experiment straightforwardly contrasted a polar [CONS?] question with the same item with the question removed.

> John has taken a test on Chapters 4–6 that has not been graded yet.
> [You ask Mary, “Did John do well on the test?”]
> Mary says, “If John understood Chapter 5, then John did well on the test.”
> Assume that John did not do well on the test.
> Given this information, does Mary’s statement imply that John did not understand Chapter 5?

The design of the experiment included the standard set of items that we described earlier. The Online Appendix lists the full set of vignettes in their Denying the Antecedent versions. Although a given participant saw a specific vignette only once (i.e., in only one of its eight versions), a participant saw each combination of inference type and question type, instantiated in two different vignettes. So a participant received 16 trials in all. Across participants, each vignette version appeared equally often. We tested 32 participants from the population mentioned earlier,

[11] Although we lead with this experiment, it is not temporally the first experiment we ran. It was suggested to us by Kai von Fintel as a simplification of some later experiments.

choosing this number based on the earlier experiments summarized in Fig. 1 (see Evans et al. 1993, Table 2.4). The experiments we report later share this same structure.

At the beginning of the session, a computer presented written instructions informing participants that they would see on each trial several sentences about a particular conversation. They were asked to read the problem carefully and then to ask themselves "whether the last of these statements is implied within the conversation." They clicked on a button labeled "implied" or on one labeled "not implied" to record their decision. We did not provide an explication of "implied," since we are interested in participants' natural understanding of whether the conclusion of the inferences followed from the information in the vignette. "Implied" has an ordinary language sense—roughly the dictionary sense of conveying something indirectly—that includes both the semantic and pragmatic components that are at stake in theories of perfection.

The proportion of "yes" responses showed no difference between the [CONS?] question and no question at all. As Fig. 2 illustrates, we did find the typical difference in endorsements as a function of inference type, $F(3, 24) = 11.35$, $p < 0.001$. As the figure suggests, however, we found no reliable overall difference due to the [CONS?] question ($F(1, 15) < 1$), and no differential effect of the [CONS?] question on the rate of endorsement for the individual inference types ($F(3, 24) < 1$).[12]

We also checked whether any of the individual vignettes showed an effect of the [CONS?] question. The analysis used the data from just Affirming the Consequent and Denying the Antecedent, since the theory under consideration predicts a difference only for these. However, we found no overall effect of vignettes and no interaction of the vignettes and the [CONS?] versus no question variation ($F(15, 224) < 1$ in both cases). Planned comparisons found no effect of question for any of the

[12] The statistical tests in this and the following experiments are based on a generalized linear mixed model for binomial data. The models were "maximal" (in the sense of Barr et al. 2013) in including as random effects: (a) the main effects of participants and vignettes, and (b) all interactions of participants and vignettes with the fixed effects of interest (question and inference type).

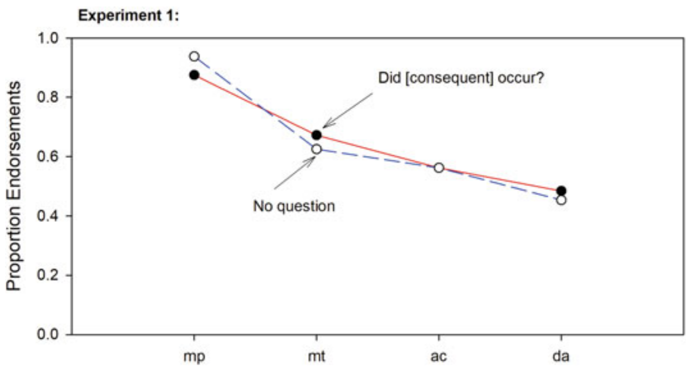


Fig. 2 The effect of polar question vs. no question on endorsement of modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da), Experiment 1

individual vignettes (for the largest of these, $F(1, 224) = 2.16$, $p = 0.14$).[13]

Experiment 2: Explicit Demands for Exhaustive Answers

Experiment 2 involved the more complicated [CONS?] questions, ones that attempt to emphasize the demand for exhaustivity in the question to Mary:

> John has taken a test on Chapters 4–6 that has not been graded yet.
> You ask Mary, "What are all the ways John could manage to do well on the test?"
> Mary responds, "If John understood Chapter 5, then John did well on the test."
> Assume that John did well on the test.
> Given this information, does Mary's statement then imply that John understood Chapter 5?

As mentioned, the [CONS?] question was intended to make *mention-some* answers as impermissible as possible. The [ANT?] question was

---

[13] This analysis was similar to the one just reported but treated vignettes and their interaction with question as fixed effects rather than as random effects. Denominator degrees of freedom were estimated using the Satterthwaite method.

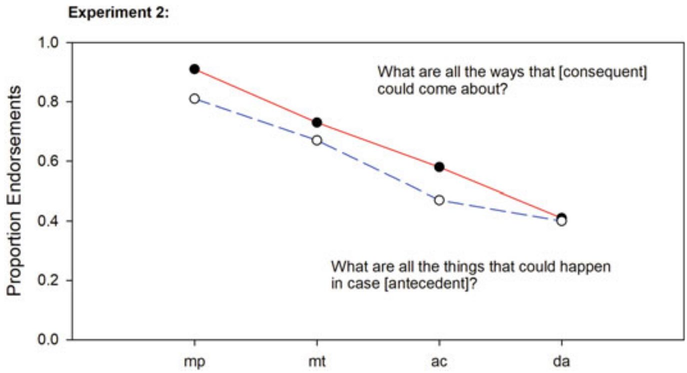


Fig. 3 The effect of [CONS?] versus [ANT?] questions on endorsement of modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da), Experiment 2

also introduced to make [ANT?] items parallel to the [CONS?] items, controlling for the length and complexity of the vignettes. Specifically, in the "Chapter 5" case, we used the [ANT?] question: "What are all the things that could happen in case John understood Chapter 5?".

None of these changes, however, led to greater endorsement of Affirming the Consequent and Denying the Antecedent for [CONS?] questions relative to [ANT?] questions. Figure 3 graphs the data from 32 participants and shows the usual decrease in endorsement rates from Modus Ponens to Tollens to Affirming the Consequent to Denying the Antecedent, $F(3, 24) = 12.46$, $p < 0.001$. However, we found no significant difference favoring problems with [CONS?] over those with [ANS?] ($F(1, 15) = 2.73$, $p = 0.120$), and no significant difference appeared in the shape of these two functions, $F(3, 24) < 1$. Explicit demands for an exhaustive answer to how the consequent could occur do not seem to elevate Affirming the Consequent and Denying the Antecedent to a greater extent than they did Modus Ponens or Modus Tollens.

Experiment 3: Checks for Memory for the Questions

Could it be that our participants were ignoring the question? After all, one could perform the task while skipping the question "you" are asking Mary and reading just the conditional and the minor premise. That might explain our difficulty in obtaining the perfection pattern.

To explore this hypothesis, we ran a variant of Experiment 2. After answering each item, participants viewed a new screen that prompted them to recall which question, [CONS?] or [ANT?], they had seen earlier, and they picked one of them by clicking on it. For the sample vignette, the choice was between "What are all the ways John could manage to do well on the test?" and "What are all the things that could happen in case John understood Chapter 5?" If participants could not remember what question had been asked, we inferred that the question did not play a role in their reasoning. Although they could correctly recognize the question without using it in reasoning, a correct answer would at least suggest that the question was available for them to reason with. In addition, since participants saw multiple items, this forced choice encouraged them to attend to the question as they were working through the later items in their allotted sequence.

Alas, the results did not fit the perfection pattern. Figure 4 plots the proportion of endorsements from 33 participants after we removed all trials on which the participants made a memory error (i.e., selected the [CONS?] question when they had actually seen [ANT?] or the reverse error). These errors ranged from 9 to 18% across the eight conditions shown in the figure. As the figure suggests, we found the typical effect of inference types, $F(3, 23) = 10.42$, $p < 0.001$. But neither the effect of the question nor the interaction between the question and the inference types was statistically reliable (both F's < 1).

Given these results, one might become quite skeptical. Either there is a problem in the theoretical proposal or in the particular way we had sought to test it.

Experiment 4: Explicit Specification of the Antecedent as the Only Condition

Are there cases where we *do* get the perfection pattern using a similar experimental setup? The next idea, then, was to get very close to forcing a biconditional reading by semantic means. In the items for the new experiment, Mary does not just answer the question with the conditional: she overtly asserts information corresponding to one of the paths to perfection (to make things easier, we underlined the new bit below: it was not underlined in the stimuli presented to participants).

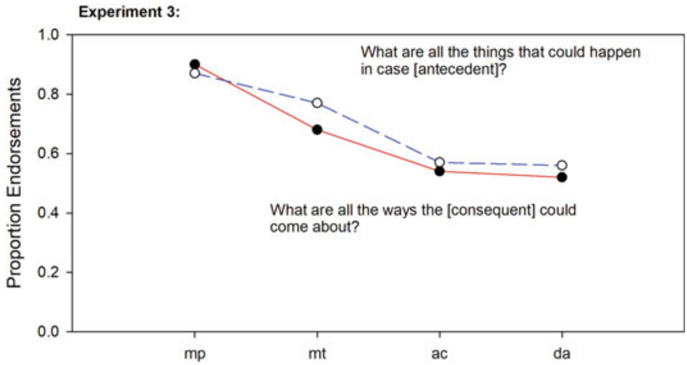


Fig. 4 The effect of [CONS?] versus [ANT?] questions on endorsement of modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da). Experiment 3, with memory checks

> John has taken a test on Chapters 4–6 that has not been graded yet.
> You ask Mary, "What are all the ways John could manage to do well on the test?
> Mary responds, "If John understood Chapter 5, then John did well on the test.
> That is the ONLY way John could have done well on the test."
> Assume that John did well on the test.
> Given this information, does Mary's statement imply that John understood Chapter 5?

In the [ANT?] variant, in addition to the usual difference in question (i.e., "What are all the things that could happen in case John understood Chapter 5?"), Mary's response changes to:

> Mary responds, "If John understood Chapter 5, then John did well on the test.
> That is the ONLY thing that could have happened if John understood Chapter 5."

Note that the continuation in these [ANT?] variants does not go any distance toward conveying biconditional information (for it does not rule out that John's doing well on the test might have come about by some other means).

At last, these explicit changes did yield a perfection pattern, as shown in Fig. 5 (based on 32 participants). Not only did we find a significant effect of inference type ($F(3, 24) = 4.70$, $p = 0.010$), we also found a significant effect of question type ($F(1, 15) = 13.38$, $p = 0.002$), and crucially an interaction of the two ($F(3, 24) = 10.56$, $p < 0.001$). This last effect is the perfection result, apparent in the difference in the shape of the curves in the figure.

Essentially, if we explicitly provide the strengthening that is supposed to be conveyed by pragmatic means, we get precisely the pattern we would expect. This suggests that our procedure is sensitive to (at least some kinds of) information that can get participants to interpret a conditional as a biconditional.

Experiment 5: Speaker's Knowledge of the Answers and Willingness to Relate Them

The biconditional interpretation (predictably) arises if we have an explicit continuation with *only*, as in Experiment 4. So why were we unable to find a similar pattern in Experiments 1–3? We have already attempted to rule out that the questions were not sufficiently clear in their demand for

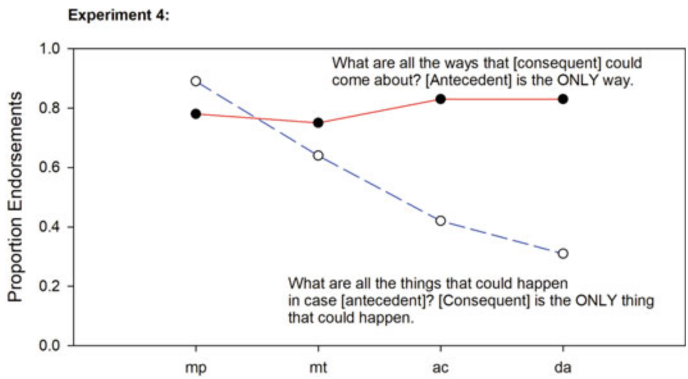


Fig. 5 The effect of [CONS?] versus [ANT?] questions on endorsement of modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da), Experiment 4, with explicit indication of exhaustivity

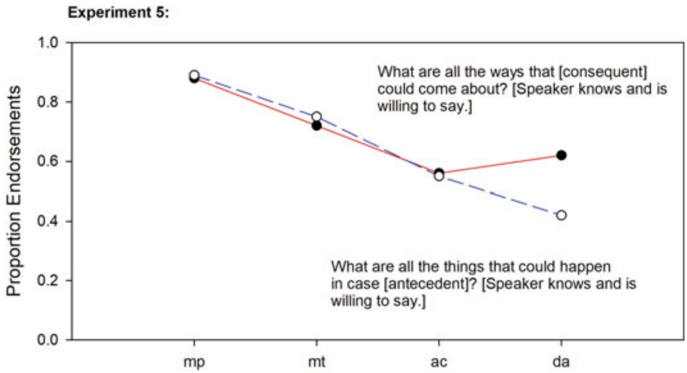


Fig. 6 The effect of [CONS?] versus [ANT?] questions on endorsement of modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da), Experiment 5, with information about speaker's knowledge and willingness

an exhaustive answer and that participants were inattentive to the questions. Another possibility is that we have not completely eliminated a "mention some" answer. Although our [CONS?] question asks Mary for all the ways the consequent could come about, her response may reflect only her partial knowledge of these ways or her limited willingness to produce them. Participants may have interpreted Mary's statement in a way that allows for these constraints. If so, they may have doubted whether her answer does indeed convey all the ways the consequent could happen.

To encourage participants to think that Mary's response was exhaustive, we tried "loading up" the context. The items in Experiment 5 looked like this (we underlined the new bit; it was not underlined in the text that participants saw):

> John has taken a test on Chapters 4–6 that has not been graded yet.
> You ask Mary, "What are all the ways John could manage to do well on the test?"
> <u>In fact, Mary knows all the ways and is willing to relate them.</u>
> Mary responds, "If John understood Chapter 5, then John did well on the test."
> Assume that John did well on the test.

> Given this information, does Mary's statement imply that John understood Chapter 5?

The same underlined information was also inserted in the [ANT?] version of the problem.

We expected that stipulating that Mary knows all the ways in which John could do well on the test and is willing to relate them would have a similar effect to saying explicitly that [Antecedent] is the only way that [Consequent] could happen.

However, this expectation was not met. Endorsement rates for the four inference types again differed significantly ($F(3, 24) = 8.50$, $p < 0.001$), based on data from 32 participants. But neither the overall difference between questions ($F(1, 15) < 1$) nor the interaction between question and inference type ($F(3, 24) = 1.52$, $p = 0.235$) are significant. Figure 5 shows a trend toward higher endorsement of Denying the Antecedent for the [CONS?] questions, but this difference was not fully significant, $F(1, 24) = 3.93$, $p = 0.059$. Moreover, the [CONS?] question did not affect endorsement rates for Affirming the Consequent ($F(1, 24) < 1$). Note, too, that even for Denying the Antecedent, the boost in endorsement rates is smaller than what we have seen in the full-blown perfection of Experiment 4.

Experiment 6: Explicit Specification of Exhaustiveness

Even when participants know that the speaker of a conditional is (a) under explicit pressure to produce an exhaustive answer to the question of how the consequent could come about, and (b) knows all the ways it could come about and is willing to relate them, they do not produce the full perfection pattern. This failure may be due to a residual unwillingness on the participants' part to believe that the speaker really has produced all the ways. Although Mary may know all the ways and is willing to tell you about them, she may nevertheless give you just a sample, perhaps because the list is too long, too complex, or too unrelated to present concerns. These considerations suggest that we might be able to reinstate full perfection if Mary explicitly states that she is in fact giving all the ways when she asserts the conditional.

To check this prediction, we used the vignettes from Experiment 5, but added Mary's assertion that she was giving all the ways the consequent could come about:

> John has taken a test on Chapters 4–6 that has not been graded yet.
> You ask Mary, "What are all the ways John could manage to do well on the test?"
> In fact, Mary knows all the ways and is willing to relate them.
> Mary responds, "Here are ALL of them: If John understood Chapter 5, then John did well on the test."
> Assume that John did well on the test.
> Given this information, does Mary's statement imply that John understood Chapter 5?

The corresponding [ANT?] version likewise included the "Here are ALL of them" prefix in Mary's answer.

The results from 32 participants appear in Fig. 7 and show that the new "ALL of them" clause was enough to produce the perfected interpretation.

When the speaker was asked to give all the ways the consequent could occur, and the speaker then made it clear that she was providing all the

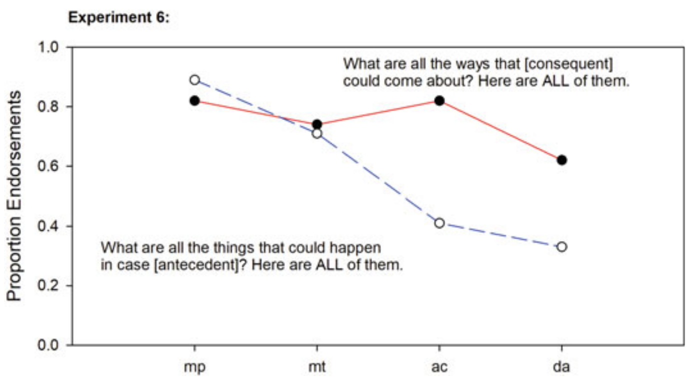


Fig. 7 The effect of [CONS?] versus [ANT?] questions on endorsement of modus ponens (mp), modus tollens (mt), affirming the consequent (ac), and denying the antecedent (da), Experiment 6, with explicit exhaustivity

ways, participants were likely to endorse each of the inference types. When the speaker was asked about all the things that could happen if the antecedent occurred, however, we find the typical decreasing pattern across inference types. This produced significant effects of inference type ($F(3, 24) = 8.29$, $p < 0.001$), question type ($F(1, 15) = 9.67$, $p = 0.007$), and an interaction between the two ($F(3, 24) = 6.80$, $p = 0.002$). Unlike the results of Experiment 5, the difference due to the question is significant for both Affirming the Consequent ($F(1, 24) = 22.60$, $p < 0.001$) and Denying the Antecedent ($F(1, 24) = 9.81$, $p = 0.004$).

Experiment 7: Conditional Transformations

The studies we've reported so far suggest that demands for exhaustive answers fail to produce perfection. We get the expected pattern only when the speaker explicitly states that her answer is exhaustive. A gap seems to exist between requesting an exhaustive response and accepting the response as exhaustive. This gap may have been widened, though, by some aspects of our procedure. For one thing, the question we put to participants was whether the conditional statement (e.g., "If John understood Chapter 5, then John did well on the test"), together with the given assumption (e.g., "John did well on the test") implied the conclusion ("John understood Chapter 5"). But one might argue that pragmatic effects of the kind we're seeking depend on people's understanding of what the speaker implied rather than what her statement implied. Emphasis on the statement might have led participants to think that only what explicitly appears in that statement matters for the inference. This would help explain the difference between Experiments 4 and 6, where the speaker claims her answer is exhaustive, and Experiments 1–3 and 5, where she doesn't.

A second reason why the latter experiments may have failed to produce perfection has to do with the type of inference we asked participants to assess. As in traditional experiments on conditional reasoning, participants had to integrate information from the speaker's conditional statement with the minor premise that we asked them to assume. Combining these statements may have required cognitive effort that suppressed the effect of interest. Of course, this complexity would not explain the obtained differences between endorsement rates among the four inference types, since all used the same format. For the same reason, complexity would not explain the differences between Experiments 1–3 and 5, on one hand, and Experiments 4 and 6, on the other. However, it is possible that complexity masked a weak effect of exhaustivity in the former experiments.

A third possibility we entertained is that participants may feel that the problems are too abstract. Our items involve judgments about a character's (Mary's) conditional claims which in turn are about decisions by third parties (e.g., John's teacher). We sought to simplify this dynamic by turning the participant into *both* the decision-maker and the interpreter of the conditional sentence.

To examine these factors, we revised the vignettes. First, instead of asking whether the conditional and the further assumption implied a conclusion, we asked whether the speaker intended to convey by her conditional a second conditional: either *If A, B* (the identical statement), *If not B, not A* (the contrapositive), *If B, A* (the converse), or *If not A, not B* (the obverse). These correspond to Modus Ponens, Modus Tollens, Affirming the Consequent, and Denying the Antecedent when we conditionalize the conclusion of the latter arguments on their minor premise. So we can usefully compare the acceptance rates for the new problems to their counterparts in the earlier experiments. Additionally, we changed the setup so that the participant would also be the decision-maker. For example, the converse inference for the test-taking vignette appeared like this (in its [CONS?] version):

> John is a high school student, and he is taking *Introduction to Calculus.* John pays attention in class, and he studies with his friend Sarah for all of the exams. John has taken a test on Chapters 4–6 that has not been graded yet.
> As a high school advisor, you need to give advice to John on whether or not he should take Calculus II based on his current performance. **You want to decide how well John has performed in the class so far.**
> You ask Mary, "What are all the ways John could manage to do well on the test?"
> In fact, Mary knows all the ways and is willing to relate them. Mary responds, "If
> John understood Chapter 5, then John did well on the test."

> Does Mary intend to convey, among other things, that if John did well on the test then John understood Chapter 5?

Notice that the problem asks whether Mary intended to convey the conclusion rather than whether her statement implies it.

However, with these changes, we found that [CONS?] questions produced an increase in the endorsement rates for the converse, which was not fully significant ($F(1, 23) = 4.04$, $p = 0.056$), and, this time, for the contrapositive ($F(1, 23) = 5.80$, $p = 0.024$). But there was no increase for the obverse ($F(1, 23) < 1$). Figure 8 shows this trend in data from 32 participants. An analysis parallel to that of the earlier experiments found the expected difference due to the type of inference ($F(3, 23) = 12.44$, $p < 0.001$) and also an effect of question type ($F(1, 15) = 5.84$, $p = 0.029$), but the interaction between question type and inference was not significant ($F(3, 23) < 1$).

Another way to assess these findings is to compare them to the results of Experiment 4, which showed the full perfection pattern. To do this, we considered just the data from the [CONS?] questions from these two studies, treating experiment and inference type as the factors of interest. As in the earlier analyses, inference type had four levels, aligned across experiments according to the correspondence that we mentioned earlier (e.g., Affirming the Consequent in Experiment 4 was paired with

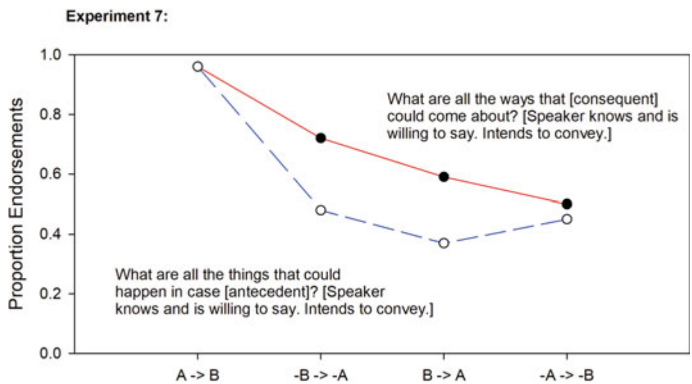


Fig. 8 The effect of [CONS?] versus [ANT?] questions on endorsement of a conditional, its contrapositive, its converse, and its obverse, Experiment 7

the inference to the converse in Experiment 7).[14] This analysis found a significant effect of inference type ($F(3, 24) = 4.10$, $p = 0.018$) and a significant interaction between experiment and inference type ($F(3, 179) = 7.42$, $p < 0.001$). These effects can be seen by comparing the red solid lines in Figs. 5 and 8. The main effect of experiment was not significant, $F(1, 8) = 2.59$, $p = 0.146$. Planned comparisons showed that participants endorsed affirming the consequent in Experiment 4 more often than they endorsed the converse inference in Experiment 7 ($F(1, 179) = 8.46$, $p = 0.004$) and denying the antecedent more often than the obverse inference ($F(1, 179) = 13.68$, $p < 0.001$). If an inference from one conditional to another (e.g., from a conditional to its converse) is simpler than a conditional syllogism (e.g., Affirming the Consequent), then these differences are not due to simplicity.

Despite our efforts to clarify the structure of the problems, Experiment 7 failed to get participants to perfect conditionals to the same extent as Experiment 4. Moreover, it produced a near-significant difference in the acceptance of the converse but not the obverse, which is not predicted by any extant model. A clear statement that the antecedent represents the only way of bringing about the consequent produces the biconditional reading. But merely stating a conditional in response to a demand for an exhaustive answer does not fully do so. One could maintain that the positive results for Experiments 4 and 6 and the null results for the remaining experiments simply show that the former experiments provide the right ingredients for obtaining the exhaustivity effect. But although Experiments 4 and 6 do provide evidence for a biconditional reading, they don't provide evidence that the reading was due to the pragmatics of exhaustivity. To obtain the biconditional effect in these two experiments (i.e., increased endorsement of DA and AC), the vignettes added semantic material ("This is the ONLY way [the consequent could come about]," "Here are ALL the ways [the consequent could come about]") that signaled the necessity of the antecedent. What's missing is support

[14] We again used a generalized linear mixed model for binomial data, including as random effects: (a) the main effects of participants (nested within experiments) and vignettes (crossed with experiments), and (b) the interactions of participants with inference type, vignettes with inference type, and vignettes with experiments.

for the idea that people will perfect a conditional when it is simply a response to a question about how the consequent could happen.

It is possible, of course, that further tinkering with the vignettes or with the instructions could succeed where Experiments 1, 2, 3, 5, and 7 did not. At this point, though, we think it worthwhile to consider some possible reasons for their lack of perfection.

5 Theoretical Discussion

The positive results in Experiments 4 and 6 suggest that true conditional perfection can, in fact, be linked to exhaustivity. However, the negative results in Experiments 1–3 and (perhaps to a lesser extent) Experiments 5 and 7 require us to put this finding in perspective.

Specifically, a common assumption is that a strong relationship exists between questions and exhaustivity. This relationship includes at least the idea that some overt questions trigger exhaustive readings in conditionals. The experimental evidence we have considered suggests that this is not quite right. In order to trigger exhaustive readings of conditionals, we need more than just questions.

In closing, we briefly consider how our experiments qualify the relation among questions, exhaustivity, and perfection. But first, we consider a similar experiment that raises some of the same issues.

5.1 Relation to a Prior Experiment

There is some similarity between the question-variations we used in our experiments and an earlier experiment by Farr (2011). Farr gave her participants vignettes like this:

> Monika sells seafood on the market. She gets 1 euro for a crab, 2.50 euros for an eel, 15 euros for a lobster, and 2.50 euros for a pike. Kerstin, an employee of Monika, cannot remember the prices. Since she does not want to ask Monika again, she asks Sahra, who also works for Monika. Sahra knows the prices exactly.

At this point in Farr's experiment, there is a dialogue between Kerstin and Sahra. This dialogue starts with one of two questions:

what-if-p Kerstin: What happens if I sell an eel?
when-q Kerstin: When do I get 2.50 euros?
Sahra: If you sell an eel, you get 2.50 euros.

Participants are then asked:

Did Sahra answer Kerstin's question sufficiently? [Yes] [No]

Farr (2011) found that when the *what-if-p?* question preceded the conditional, participants more often responded that Sahra's answer was sufficient than when the *when-q?* question preceded it. Farr's interpretation is that *when-q?* demands an exhaustive answer with respect to the conditional's consequent (What are all the cases in which I get 2.50 euros?) and triggers a perfected reading of the conditional (if and only if you sell an eel do you get 2.5 euros). Because the conditional mentions only one of the two ways to get 2.50 euros as given in the background story, participants see it as insufficient. The *what-if-p?* question also demands an exhaustive answer, but to a different question (What are all the things that happen if I sell an eel?) and does not perfect the conditional. So participants see Sahra's answer to this question as sufficient.

Despite the similarities in the setup to the present studies, Farr's experiment did not end up testing whether Sahra's conditional gets perfected. As noted, participants learn from the story that there are two ways to get 2.50 euros but only a single price for an eel. So even a simple (nonconditional) answer like "An eel costs 2.50 euros" is complete with respect to questions like "How much is an eel?" or "What happens if I sell an eel?" But the same simple answer is incomplete with respect to "What costs 2.50 euros?" or "When do I get 2.50 euros?" It seems possible, then, that participants' dissatisfaction with the answer after the *when-q?* question does not depend on their perfecting the conditional but instead on their sense of Sahra's lack of full disclosure. For this reason, we think it important to have a more direct indicator of perfection, such as participants' willingness to accept the conditional's converse or to accept an inference like Affirming the Consequent.

5.2 How the Experiments Constrain Explanations of Perfection

At first sight, our experiments appear to refute the idea that questions trigger perfection. Even explicit questions that demand exhaustive answers about the consequent (e.g., *What are all the ways [the consequent] could come about?*) don't always yield a perfected conditional (*antecedent iff consequent*).

According to the von Fintel/Cornulier account that we have been pursuing, questions about some event *B* set up the expectation that a conditional answer of the form *If A, B* implies that *A* is the only way *B* could come about. Thus, *if ~A, ~B*. Together, *If A, B* and *If ~A, ~B*, yield the perfected interpretation, *A iff B*. For example, questions like *Did John do well on the test?* or *What are all the ways John could do well on the test?* imply that the answer *If John understood Chapter 5, then he did well on the test* supplies all ways he could do well. So John did well if and only if John understood Chapter 5. Our experiments, however, failed to produce this pattern of reasoning. In principle, then, this failure could come about either (a) because participants failed to infer a perfected conditional from (what they perceived as) an exhaustive answer, or (b) because they failed to interpret the conditional as exhaustive, in the first place. Let's consider these two possibilities in turn.

5.2.1 Do People Infer Perfected Conditionals from Exhaustive Answers?

Perhaps our negative results are partial evidence that people do not reason from (i) to (ii):

(i) For each alternative *D* to the antecedent *A*, *~(If D, B)*
(ii) *If ~A, ~B*

In the case of our experiments, failure to infer (ii) from (i) amounts to the idea that participants understood the antecedent (e.g., *John understood Chapter 5*) as the only way that the consequent (*John did well on the test*) could occur, and yet did not infer that if John did not understand Chapter 5, he did not do well.

But on the contrary, Experiments 4 and 6 show that given discourse that basically entails claims of the form of (i), people will reason their way to a claim of the form of (ii).[15] This suggests that the failure to find the perfection pattern in the remaining experiments is due to participants failing to interpret the conditional (e.g., *If John understood Chapter 5, he did well on the test*) as an exhaustive answer to the question (*What are all the ways John could do well on the test?*).

To back up this possibility, we asked participants in a further experiment to decide whether the conditionals mentioned all the ways the consequent could come about.[16] The experiment was very similar to Experiment 5 (in which Mary is said to know all the ways and is willing to relate them), but in addition to asking whether participants agreed with the inference, we also asked them, "Did Mary's response mention all the ways?" (Half the participants answered the inference question first and half answered the "all the ways" question first, though the order had no statistically reliable effect on the results.) For [CONS?] questions (e.g., *What are all the ways John could do well on the test?*), participants believed that the conditional response (*If John understood Chapter 5, he did well on the test*) mentioned all the ways on only 14.5% of trials.

Of course, our results do not mean that questions *never* produce the presumption that a conditional answer is exhaustive. Rather, the results suggest that even very explicit questions of the proper sort don't always trigger an exhaustive reading. Something more is needed to ensure it.

[15] Note, incidentally, that if one rejects the Competitor assumption we sketched in Sect. 2, this would be indirect, and admittedly very defeasible, experimental evidence that people reason with something like Conditional Excluded Middle.

[16] We thank Robert Stalnaker for suggesting this experiment.

5.2.2 Why don't People Believe Conditional Answers to Questions Are Exhaustive?

Questions don't always yield exhaustive answers. Still, we might be able to recover the spirit of the Cornulier/von Fintel proposal on conditional perfection by supposing that questions yield exhaustive answers by default and giving an independent explanation for why this default pattern did not emerge in Experiments 1–3, 5, and 7.

Perhaps the explanation is that some property of the vignettes in these experiments encouraged a *mention-some* reading. Consider this dialogue:

(14) Q: Who are all the people who came to the party?
A: John came.

Even though the question is phrased as demanding a complete list, a possible *mention-some* interpretation of the answer seems available. The respondent rejects the burden of providing a complete answer to the question and volunteers instead whatever information she is able to provide (perhaps expecting that other conversational participants will be able to fill out the rest of the party-goers list). In the specific case of (14), this sort of interpretation might even be invited by the background knowledge that only extremely unusual parties have only one attendee. Similarly, it is possible that, when we ask, "What are all the ways in which [consequent] could come about?" the respondent's answer is given a *mention-some* reading.[17]

Though we do not think that this assessment is without merit, it has too many surprising consequences to be plausible. What we found is that, unless the respondent explicitly avows providing a complete answer, there is no significant pattern of perfection. Though there are slight

[17] Demands for exhaustive answers often call for memory searches that exceed people's abilities, especially in the context of an on-going conversation. In such situations, speakers may be thrown back on a satisfactory answer that is informative, but partial. Hearers may likewise make allowances for this kind of satisficing by leaving open the possibility that the speaker's answer is all she can come up with at the moment—that it is temporarily exhaustive, rather than exhaustive period. In the case of (14), for example, this interpretation is enhanced if the speaker indicates some hesitancy: "Well…John came." Viewed in this way, what our results suggest is that this temporarily exhaustive reading is more available than one might have expected.

increases in endorsement rates for the fallacies Affirming the Consequent and Denying the Antecedent in some experiments, they are typically not enough to meet standard significance thresholds. One might have expected that in the cases where a *mention-all* interpretation is possible but not mandated, we should have seen some participants reach for it. In Experiments 5 and 7, we even tried as much as possible to "load" up the context so as to invite a *mention-all* answer, but without much success. This is especially striking because *mention-all* interpretations of answers to questions are not unusual. According to the received wisdom (see, for example, van Rooij and Schulz 2003), they should be preferred unless they are contrary to expectations (as in (14)). More generally, much recent work on implicatures argues that the computation of implicatures happens by default (see e.g., Chierchia 2013).

Of course, it is still possible that participants read the [CONS?] questions as having open-ended answers that respondents were unlikely to answer exhaustively. In the case of our running example, participants may have taken the question, *What are all the ways John could manage to do well on the test?* as placing an impossible demand on the addressee, given the many ways John could do well (e.g., cheating, bribing the instructor, divine intervention, lucky guessing,…). However, some of the vignettes in our experiments were explicit in listing alternatives for the antecedent. For example, one of the vignettes began with the sentence *Someone has put a fertilizer, either Easy-gro or Bloom-builder, on the plants.* This was then followed by the [CONS?] question *What are all the ways the plants could manage to grow quickly?* and the conditional *If Easy-gro was put on the plants, then the plants grew quickly*. Although items like these are not completely immune to *mention-some* interpretations, they seem at least less open to these interpretations than those in which the alternatives are unspecified. The initial sentence seems to limit the relevant alternatives to just a few (in this case, Easy-gro or Bloom-builder); so the speaker of the conditional should find it less of a burden to provide an exhaustive list. However, a re-analysis of Experiment 1 shows that participants were not more likely to endorse the inferences for the items with explicitly-provided alternatives than for the remaining open items (Vignettes 1, 4, 12, and 14 in the Online Appendix have unspecified alternatives, and the remainder specified ones). Considering just the items with [CONS?] questions, we find that participants endorsed Affirming the Consequent on 54% of trials for the items with alternatives and 62% for the open items. Similarly, they endorsed Denying the Antecedent on 48% of trials for items with alternatives and 50% for the open items. These differences are in the wrong direction to explain the results. We should be cautious here about drawing strong conclusions, since we did not design the experiment with this difference in mind. Still, it provides some presumptive evidence against an explanation based on response burden.

It appears then that pragmatic reasoning based on background questions is not enough to trigger the relevant exhaustive readings.[18] What turns out to be necessary—in our experiments, at least—is linguistic material that explicitly directs hearers toward an exhaustive interpretation.

A very austere development of this idea would be to claim that the biconditional interpretation requires that such material always be explicitly represented. This amounts to the claim that there is no distinctive pragmatic phenomenon of conditional perfection. Biconditional interpretations arise as entailments when an utterance of *If A, B* is conjoined with additional claims that are strong enough to entail their converses.

The immediate problem with this explanation is that several previous experiments have found evidence of perfection stemming from the conditional's status as a promise or threat (e.g., *If you disrupt the class, you'll have to leave the room*). (See the sources cited in Sect. 3.) These conditionals don't entail their obverses—they merely suggest them—so perfection is obtainable without explicit entailment of *If ~A, ~B*. What

[18] It is important to acknowledge in this context that there are theories like Franke (2009, Section 5.2.2) which maintain that there are two, quite distinct kinds of perfection inferences and that alternative-based strengthening is only one of them. According to Franke some perfection inferences arise on account of expectations of normality in context. We think this is quite possible, and if so, there may be ways to construct vignettes (other than those we used here) that would produce perfection. However, we still require an explanation for why alternative-based strengthening failed to yield this pattern in our studies.

is the case is that questions, in particular, are not always enough to yield the obverse.[19]

On the positive side, then, we strongly suspect that there's more than one path to perfection. One way to obtain it is through threats, promises, permissions, and obligations, which convey via practical reasoning that if the addressee does not perform the key action the reward, punishment, or some other kind of normative status will not be forthcoming. A second path to perfection is through background information that the converse of the conditional is true. Section 3 reviewed some evidence for these possibilities. A third route is through explicit exhaustification devices like the ones we used in Experiments 4 and 6. Pluralism about perfection may be due to the absence in natural language of expressions as simple as *if* that mean *iff*, leaving *if* to cover for *iff* in any of the variety of situations (outside math classes) in which *iff* would be more exact. Provided that there is some contextual reason to think that *If B, A* is true, then an utterance of *If A, B* may suggest *A iff B*. Because there are

[19] Herburger's (2015) "Whole Truth" account of conditional perfection provides one possible route to deriving perfection without the need for explicit exhaustification devices. Her preliminary statement of the theory is:

> Conditional Perfection and upper-bounding inferences arise as logical entailments when a sentence *S* is silently conjoined with ~~only *S*~~, resulting in the conjunction *S* ~~and only *S*~~. *S* ~~and only *S*~~ is then taken to express 'the truth and the whole truth'. (Herburger 2015, p. 6)

Note that Herburger strikes through "*and only S*" to indicate the fact that it may not be pronounced. If it is possible to interpret utterances of *If A, B* as utterances of *If A, B* ~~and only If A, B~~, then that's when we should expect perfection inferences to arise. According to this intermediate take, perfection inferences can arise due to overt or covert exhaustification.

This sort of approach is difficult to evaluate, experimentally at least, absent some systematic ideas about when we are allowed to supplement *S* with the silent *and only S*. Since it is part of the proposal that there are no systematic principles that connect the questions made salient by the discourse to the availability of the strengthened interpretation, it is hard to see what such principles might look like.

But perhaps a non-experimental argument is available. If perfection is achievable with overt exhaustification devices (as our experiments demonstrate), and if there is precedent for covert exhaustification, we should expect that nothing prevents an exhaustified interpretation of conditionals. The remaining question, once again, is why we did not find much trace of these exhaustified interpretations in our first experiments.

many reasons why *If B, A* could be true, there are many ways to achieve perfection.

Here's where we've got to: Questions that explicitly request exhaustive answers don't seem sufficient to produce perfection. That's the experimental finding. This seems to be because the answers aren't read as exhaustive rather than because an exhaustive answer doesn't yield perfection. Although it's possible that the question-to-exhaustive-answer link is the normal case and something about our experimental materials militated against it, the usual suspects—the respondent's lack of knowledge, uncooperativeness, and response burden—seem to be ruled out by the results. There may be some unusual suspects responsible for blocking the normal route to interpreting the answer as exhaustive, but what could these be? Instead, it seems more likely that an exhaustive interpretation of an answer requires more than just a question demanding one. This something more could be an explicit avowal that the response is exhaustive, but it seems likely that people could settle on exhaustive readings from weaker evidence. Perhaps what listeners require is some reason to think that an exhaustive conditional answer is in the respondent's interest. Otherwise, the listeners' experience with their own communicative foibles may make them hesitant to think they've gotten the full story.

Acknowledgements We thank Jonathan Cohen, Paul Égré, Michael Franke, Joshua Knobe, Paolo Santorio, Katrin Schulz, Eva Wittenberg, and Yimei Xiang for comments on this project. We also thank audiences at the 2016 Pacific APA, a conference on causal reasoning at the University of Amsterdam, the UIC Cognitive group, Northwestern Phling, Sembabble at UCSD, and the Gothenburg Logic Seminar. Kai von Fintel delivered comments on the paper at the 2016 APA that significantly shaped its later incarnations. For their assistance with the experiments, we thank John Glines, Gabrielle McCarthy, Jane Ko, Sabrina Kohanzad, Michael Yi, Kathy Tian, Camdyn Parrott, Joshua Kennedy, Alyssa Pierce, and Nathaniel Yoon. The final version is also indebted to three anonymous reviewers from a prior journal submission.

References

Adams, E. (1965). The logic of conditionals. *Inquiry, 8,* 166–197.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effect structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68,* 255–278.

Bonnefon, J.-F., & Politzer, G. (2010). Pragmatic conditionals, conditional pragmatics, and the pragmatic component of conditional reasoning. In M. Oaksford, & N. Chater (Eds.), *Cognition and conditionals: Probability and logic in human thinking* (pp. 233–249). Oxford: Oxford University Press.

Cariani, F. (2021). *The modal future.* Cambridge: Cambridge University Press.

Cariani, F., & Rips, L. J. (2017). Conditionals, context, and the suppression effect. *Cognitive Science, 41,* 540–589.

Cariani, F., & Santorio, P. (2018). *Will* done better: Selection semantics, future credence, and indeterminacy. *Mind, 127, 129–165.*

Chierchia, G. (2013). *Logic in grammar*. Oxford: Oxford University Press.

Cummins, D. D., Lubart, T., Alksnis, O., & Rist, R. (1991). Conditional reasoning and causation. *Memory & Cognition, 19,* 274–282.

de Cornulier, B. (1983). If and the presumption of exhaustivity. *Journal of Pragmatics, 7*, 247–249.

Evans, J. St. B. T., Newstead, S. E., & Byrne, R. M. J. (1993). *Human reasoning*. Hillsdale, NJ: Erlbaum.

Evans, J. St. B. T., & Over, D. E. (2004). *If.* Oxford, UK: Oxford University Press.

Farr, M. C. (2011). Focus influences the presence of conditional perfection: Experimental evidence. In I. Reich, E. Horch, & D. Pauly (Eds.), *Proceedings of Sinn und Bedeutung, 15,* 225–239.

Fillenbaum, S. (1975). If: Some uses. *Psychological Research, 37,* 245–260.

Franke, M. (2009). *Signal to act: Game theory in pragmatics*. [Doctoral dissertation, Amsterdam University]. https://www.illc.uva.nl/Research/Publications/Dissertations/DS-2009-11.text.pdf

Geis, M., & Zwicky A. (1971). On invited inferences. *Linguistic Inquiry, 2,* 561–566.

Groenendijk, J., & Stokhof, M. (1984). *Studies in the semantics of questions and the pragmatics of answers*. CiteSeerX. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.867.7509&rep=rep1&type=pdf

Herburger, H. (2015). Conditional perfection: The truth and the whole truth. In S. D'Antonio, M. Moroney, & C. R. Little (Eds.), *Proceedings of SALT 25* (pp. 615–635).

Horn, L. (2000). From IF to IFF: Conditional perfection as pragmatic strengthening. *Journal of Pragmatics, 32,* 289–326.

Kratzer, A. (1991). Conditionals. In A. von Stechow, & D. Wunderlich (Eds.), *Semantics: An international handbook of contemporary research* (pp. 651–656). Berlin: de Gruyter.

Kratzer, A. (2012). *Modals and conditionals*. Oxford: Oxford University Press.

Lilje, G. W. (1972). Uninvited inference. *Linguistic Inquiry, 3,* 540–542.

Marcus, S. L., & Rips, L. J. (1979). Conditional reasoning. *Journal of Verbal Learning and Verbal Behavior, 18*, 199–223.

Markovits, H., & Lesage, C. (1990). Pragmatic reasoning schemas for conditional promises: Context and representation. In J.-P. Caverni, J.-M. Fabre, & M. Gonzalez (Eds.), *Cognitive biases* (pp. 183–192). Amsterdam: North-Holland.

Newstead, S. E., Ellis, M. C., Evans, J. St. B. T., & Dennis, I. (1997). Conditional reasoning with realistic material. *Thinking and Reasoning, 3,* 49–76.

Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford: Oxford University Press.

Stalnaker, R. (1968). A theory of conditionals. In N. Rescher (Ed.), *Studies in logical theory* (pp. 98–112). Oxford: Blackwell.

Stalnaker, R. (1981). A defense of conditional excluded middle. In W. Harper, R. Stalnaker, & G. Pearce (Eds.), *Ifs: Conditionals, belief, decision, chance, and time* (pp. 87–104). Dordrecht: Reidel.

Staudenmayer, H. (1975). Understanding conditional reasoning with meaningful propositions. In R. J. Falmagne (Ed.), *Reasoning: Representation and process in children and adults* (pp. 55–79). Hillsdale, NJ: Erlbaum.

Thompson, V. A. (1994). Interpretational factors in conditional reasoning. *Memory & Cognition, 22,* 742–758.

Van Canegem-Ardijns, I., & Van Belle, W. (2008). Conditionals and types of conditional perfection. *Journal of Pragmatics, 40,* 349–376.

van der Auwera, J. (1997). Pragmatics in the last quarter century: The case of conditional perfection. *Journal of Pragmatics, 27,* 261–274.

van Rooij, R., & Schulz, K. (2003). *Exhaustification.* CiteSeerX. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.13.1744&rep=rep1&type=pdf

von Fintel, K. (1997). Bare plurals, bare conditionals, and *only*. *Journal of Semantics, 14*, 1–56.

von Fintel, K. (2001). *Conditional strengthening: A study in implicature*. MIT. http://mit.edu/fintel/fintel-2001-condstrength.pdf

The External Syntax of Conditional Clauses

Liliane Haegeman and Manuela Schönenberger

1 Introduction

1.1 Main Goals and Outline of the Chapter

The chapter looks at the typology of conditional clauses against the background of the wider typology of adverbial clauses, focusing on their external syntax. Clauses introduced by the conjunction *if* display (at least) three readings: (i) an event conditional (1a) expresses a condition

Manuela Schönenberger's research was funded by SNSF grant 188933.
This work was first presented by Liliane Haegeman at the CNRS summer school *Conditionals in Paris—Logic, Linguistics and Psychology* (ÉCOLE THÉMATIQUE—CNRS) on June 4, 2019. We thank the organizers of the summer school for their invitation and the audience for their comments. Special thanks are due to Andrew Radford and Andrew Weir for help with the English data, to Anne Breitbarth for her comments on a prefinal version of the paper and to two reviewers for this volume who gave extensive comments on an earlier version. Of course, none of them can be held responsible for the way these comments have been interpreted and used.

L. Haegeman (✉)
Research Groups Ugent, DiaLing, University of Geneva, Geneva, Switzerland
e-mail: Liliane.Haegeman@UGent.be

M. Schönenberger
University of Geneva, Geneva, Switzerland
e-mail: manuela.schoenenberger@unige.ch

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_10

on the realization of the eventuality encoded in the associated clause, (ii) a factual conditional (Iatridou 1991: 58–96) (1b) introduces a background assumption which serves as the basis for the contextualization of the proposition encoded in the associated clause,[1] (iii) a speech-event conditional (1c) encodes a condition on the realization of the speech event.

(1) a. If your back-supporting muscles tire, you will be at increased risk of lower-back pain. (*Independent on Sunday*, *Sports*, 14.10.2001, page 29, col 3)
   b. When Simenon was asked how the Maigret novels differed from his other books – his *romans durs* – he described them as 'sketches' […] If the books are sketches, they are the sketches of an old master.
   (*Observer*, 05.01.2020, page 42, cols 1 + 5)
   c. David Davis even said he had been a personal friend of Mr. Clarke for 30 years. He had just faced his greatest test. If I may say so, I commend how he has personally responded to, and risen to that occasion.
   (*Guardian*, 21.07.2005, page 2, col 8)

The three types differ with respect to their external syntax, i.e. how closely they are integrated with the associated clause: the event conditional in (1a) is more closely integrated than the factual conditional in (1b), which is more closely integrated than the speech-event conditional in (1c). See also Van der Auwera (1986) for an early discussion.

The chapter is mainly based on English data focusing on the question whether and if so how the three types of clauses can/must be given different representations in terms of their formal relation with the

[1] Various terms are used to refer to the three types of conditionals. What we call factual conditionals have also been labelled pragmatic conditionals (Haegeman 1984b), premise-conditionals (Haegeman 2003), relevance conditionals (Iatridou 1991), factual P-conditionals (Declerck 2000), conditional assertions (Kearns 2006). Speech-event conditionals are also referred to as *biscuit* conditionals (Austin 1961), for discussion of *biscuit* conditionals see a.o. Ebert et al. (2008).

associated clause.[2] While some properties discussed here pertain only to conditional clauses, most are shared across all adverbial clauses, reflecting the general line of reasoning adopted which considers conditional clauses as a subcase of adverbial clauses.

The point of departure is the hypothesis in Haegeman (1984b) that, depending on their degree of integration with the associated clauses, we can distinguish 'central' adverbial clauses from 'peripheral' adverbial clauses. The event conditional in (1a) is a central adverbial clause, the factual conditional in (1b) is a peripheral adverbial clause. Originally, Haegeman and Wekker (1984) and Haegeman [1984b]) grouped speech-event conditionals (1c) with factual conditionals (1b) as peripheral adverbial clauses. However, based on the properties of their external syntax, this initial binary distinction between 'central' adverbial clauses and 'peripheral' adverbial clauses has to be revised and replaced by a ternary distinction along the lines of that discussed in Takami (1988) and Iatridou (1991), a.o., and recently developed in work by Frey (2016, 2018, 2019, 2020).

The chapter is organized as follows: Section 1 is a first descriptive inventory of the typology of conditional clauses in line with that developed in earlier work by Haegeman (1984a, b, c, 2003). Section 2 focuses on the external syntax of conditional clauses, presenting the distinction between event conditionals (1a), which modify the state of affairs encoded in the associated clause, and factual conditionals (1b), which provide a background proposition for the contextualization of the associated proposition. While the contrasts outlined are relevant to adverbial clauses in general, this chapter focuses on their relevance for conditional clauses. Section 3 explores the external syntax of factual conditionals

---

[2] One subset of conditional clauses, those referred to as event conditionals, also function as clausal arguments. In (i) the bracketed conditional clause is the complement of the preposition *for*:

(i) exactly what it says it is perfect for [*if* $_1$ you need some cash in your blizzard wallet or you want to give a gift card to someone for holiday or bday].

(https://www.amazon.com/gp/customer-reviews/R1W5EO8HSSX2CW?ASIN=B012JMS4W2)

We will not go into this pattern here, though it is obviously of independent interest, in particular because to the best of our knowledge, factual *if*-conditionals cannot function as arguments.

(1b). It turns out that with respect to the distinctive properties identified in Sect 2, factual conditionals coincide with epistemic modal adverbials. An analysis is proposed which aligns the two in relation to the functional hierarchy of the clause as developed by Frey (2016, 2018, 2019, 2020). Section 4 discusses the third type of conditional (and adverbial) clauses, namely those like (1c) which modify the speech event as such, rather than its propositional content. While in earlier work (Haegeman and Wekker 1984; Haegeman 1984a, b, 1991/2009) speech-eventconditionals were treated on a par with factual conditionals, both being categorized as peripheral adverbial clauses, the two types must be kept distinct. In particular, evidence from the syntax of Verb Second patterns in Dutch is shown to support the need for this distinction. The original binary classification is replaced by a ternary classification that differentiates between central adverbial clauses (CACs), peripheral adverbial clauses (PACs) and non-integrated adverbial clauses (NiCs), but we speculate on a possible reinterpretation of this new classification. Section 5 summarizes the chapter.

1.2 Starting Point: Two Types of Conditional Clauses: Event Structuring vs. Discourse Structuring

It is well known that a conjunction introducing an adverbial clause can often be seen to introduce clauses with quite distinct readings. This is illustrated for the conjunction *while* in (2), in which we focus on two readings.[3] For ease of identification, we will identify the two uses of the conjunction *while* in (2) as *while*$_1$ and *while*$_2$. The *while*$_1$-clause in (2a) has a temporal reading, *while* is roughly equivalent to 'during the time that', and the clause modifies the event time of the associated clause. In (2b) *while* is similar to 'whereas' introducing a 'concessive'/'adversative'/'contrastive[4]' *while*$_2$-clause which serves to develop the

[3] We turn to the speech-event related reading corresponding to (1c) in Sect. 4.1.

[4] From now on we will use the term 'concessive', though closer study of the semantics of *while*$_2$-clauses might reveal the relative appropriateness of the three labels.

argumentation, introducing a background assumption which contextualizes the propositional content of the associated clause. Observe that in (2c) the content of the *while*$_2$-clause clearly picks up on, i.e. echoes, the preceding context, a property that can also be found in factual conditionals (see (1b)). Example (2d) illustrates both uses of *while* in one utterance.

(2) a. According to Smith, a group of Arkansas state troopers who worked for Clinton while$_1$ he was governor wanted to go public with tales of Clinton's womanizing. (*Guardian*, G2, 12.03.2002, page 3, col 2-3)

b. While$_2$ [Dr Williams'] support for women priests and gay partnerships might label him as liberal, this would be a misleading way of depicting his uncompromisingly orthodox espousal of Christian belief.
(*Guardian*, 02.03.2002, page 9, col 1-2)

c. Instead the patriotic duty was dismissing 'random acts of criminality'. While$_2$ criminal the rioting indubitably was, random it was not.
(*Guardian*, 06.09.2012, page 36, col 2)

d. While$_2$ [the lawsuit challenging the legitimacy of lethal injections] probably won't stop the use of lethal injection altogether, it will certainly delay its use while$_1$ the supreme court decides what to do.
(*Guardian*, G2, 12.12.2003, page 4, col 4)

As shown in (1a) and (1b), *if*-clauses provide evidence for the same dual use, though these two uses are perhaps slightly less salient because, informally speaking, both clause types arguably encode 'conditions'. We repeat examples (1a) and (1b) and add some additional illustrations. Examples (3a) and (3b) contain event conditionals: the conditional clauses specify the circumstances under which the eventuality in the associated proposition will or will not become true, or, put differently, the conditional clause encodes an eventuality whose realization has a causal connection with the realization of the eventuality in the associated clause. In such examples, the conjunction *if*$_1$ can sometimes be replaced by *in case*, and some event conditionals can be paraphrased by replacing the conjunction *if*$_1$ with *if and when*, as in (3c, d). Event conditionals differ from factual conditionals illustrated in (3e, f) which echo contextually accessible propositions providing a premise for the contextualization of the associated clause (Comrie 1982; Declerck and Reed 2001; Biasio and Castro 2019). Factual conditionals can often be paraphrased with 'if it is true that', 'now that', 'given that' and they carry the implication that someone believes that the content of the factual conditional is true. (For discussion see Iatridou 1991: 60, Declerck 2000). Though factual conditionals may echo contextually accessible propositions, as in (3e, f), they do not have to be literal echoes of actual utterances. 'They may also be echoes of an internal or mental proposition (thought) such as the interpretation of an experience, perception etc.' (Declerck and Reed 2001: 83). A paraphrase with *if and when* is not available for a factual conditional, as shown in (3e, f).

(3) a. If$_1$ your back-supporting muscles tire, you will be at increased risk of lower-back pain. (*Independent on Sunday, Sports*, 14.10.2001, page 29, col 3)

b. If$_1$ last week you had shown me the piece of pipe system that Laila and I built on Tuesday, I would never have believed it. (*Guardian*, G2, 27.02.2004, page 8, col 3)

c. [President Bush and Mr. Blair]will be taking even more [risks] if$_1$, and when, a land war starts.
(*Independent on Sunday*, Comment, 14.10.2001, page 25, col 2)

d. Air support for the marines will come from US navy fighter-bombers, some of which may be based at Kandahar airport if$_1$ and when it is considered secure. (*Guardian*, 27.11.2001, page 3, col 8)

e. When Simenon was asked how the Maigret novels differed from his other books – his *romans durs* – he described them as 'sketches' [...] If$_2$ (*and when) the books are sketches, they are the sketches of an old master.
(*Observer*, 05.01.2020, page 42, cols 1 + 5)

f. If$_2$ (*and when) I'm no longer going to be arrested for possessing cannabis for my own consumption ('Cannabis laws eased in drugs policy shake-up', October 24), shouldn't I be able to grow my own?
(Jason Cundy, Letter *Guardian*, 25.11.2001, page 9, col 8)

Iatridou (1991: 98) signals that while event conditionals are compatible with negative polarity items (NPIs), e.g. *lift a finger* in (4a), factual conditionals are not (4b).

(4) a. If$_1$ John lifts a finger to help, I'll be surprised. (Iatridou 1991: 98, her (101c))
b. *If$_2$ he (indeed) lifted a finger to help, you should pay him. (Iatridou 1991: 98, her (103a))

If, following Giannakidou (1998), we assume that NPIs are licensed by non-veridicality, defined as in (5), (4) shows that event conditionals are non-veridical, while factual conditionals are veridical.

(5) A propositional operator F is veridical iff F entails p: $Fp \rightarrow p$; otherwise F is non-veridical.

The contrast in veridicality correlates with some other differences. So-called 'putative' *should* expresses a 'potentialis' mood in event conditionals (6a). This mood is not compatible with factual conditionals (6b). Putative *should* may give rise to conditional inversion (Iatridou and Embick 1994, Biberauer and Roberts 2017), which thus also remains restricted to event conditionals (6c, d). The bracketed conditional clauses in (6e) and (6f) are necessarily read as event conditionals.

(6) a. If$_1$ you (should) see him, tell him I want to talk to him.
b. If$_2$ crime is/*should be falling, why are our prisons full?
c. Should you see him, tell him I want to talk to him.
d. *Should crime be falling, why are our prisons full?
e. It is to your credit that you are concerned with your husband's future [should you leave him]. (*Guardian*, G2, 16.01.2004, page 6, col 1)
f. [Should the court refuse to set the judgement aside, and the Orams fail to comply], then we will try, within the spirit of EU regulations, to have the judgement registered in the UK and have it enforced that way.
(*Observer*, 20.02.2005, page 7, col 1)

In factual conditionals, the modal *should* will receive either an epistemic reading or an obligation reading. The constructed exchanges in (7) illustrate this point. In these contexts, *should* inversion is not licit (cf. the B'-examples).

(7) a. A: John and Mary should be able to get here by 8 pm.
B: If$_2$ they should be able to get here by 8 pm, then we might just as well wait for them.
B': *Should they be able to get here by 8 pm, then we might just as well wait for them.
b. A: John is head of department. He really should be present at the final meeting.
B: If$_2$ John should be present at the final meeting, it cannot be scheduled on a Monday because that's his day off.
B': *Should John be present at the final meeting, it cannot be scheduled on a Monday because that's his day off.

Factual conditionals are also incompatible with irrealis mood and with the conditional inversion available for auxiliaries encoding irrealis (Iatridou and Embick 1994; Biberauer and Roberts 2017). In (8a) the conditional clause is an irrealis event conditional; in the factual conditional in (8b) an irrealis reading is unavailable. Conditional inversion of irrealis *had* is available in (8a), as shown in (8c), but, because in English conditional inversion is restricted to non-veridical auxiliaries (Biberauer and Roberts 2017), it is not available in the factual conditional in (8b), as shown in (8d). The bracketed conditional clauses in the attested examples (8e) and (8f), which illustrate conditional inversion with irrealis *had*, are necessarily read as event conditionals.

(8) a. If$_1$ I had seen him, I would have told him to come back.
b. If$_2$ crime had been falling, why were our prisons still full?
c. Had I seen him, I would have told him to come back.
d. *Had crime been falling, why were our prisons still full?
e. Back in 1991 I was a very bad actor and would have been out of work [had I not busked my way into a number of kitchens]. (*Guardian*, G2, 31.03.2004, page 14, col 2)
f. [Had the money not been returned], the evidence would have pointed strongly to a conclusion that the NRCC 'financed' the Forum.
(*Washington Post*, 29.04.2003, page A18, col 3)

2 Differentiating Between the Two Types of Conditionals: Diagnostics

The interpretive difference between event conditionals and factual conditionals is matched by a number of distributional distinctions, a subset of which will be illustrated in the present section. The differences shown jointly lead to the conclusion that event conditionals are more closely related to or integrated with the associated proposition than factual conditionals, hence the labels 'central' and 'peripheral', used originally in Haegeman and Wekker (1984) and in Haegeman (1984a), a.o., to distinguish the two. The contrasts discussed have been used as a basis for postulating a difference in external syntax between event conditionals and factual conditionals (see, for instance, Rutherford 1970 for early generative discussion, Haegeman and Wekker 1984; Haegeman 1984b, 1991/2009; Iatridou 1991; Haegeman 2003). Most patterns discussed here for conditional clauses can be replicated in other adverbial clauses. For a discussion of the internal syntax of conditionals see Haegeman (2010, 2012).

2.1 Coordination of Likes

As shown in (9), a coordination of two event-conditional clauses is acceptable (9a); an event-conditional clause can also coordinate with a central adverbial clause (9b). Similarly, two factual conditional clauses can be coordinated (9c).

(9) a. The party is also in danger of alienating older people above the poverty line, Mr. Cable argues. 'Both these groups will swing to the Conservatives if_1 the Tories are smart enough and if_1 we have nothing much to offer them.'
(*Guardian*, 11.02.2002, page 6, col 5)

b. When I was playing at fly half in 2001-02 and if_1 something went wrong behind the scrum, he'd turn and have a go at me.
(*Observer*, 15.05.2005, page 13, col 5)

c. Not only has Sir Richard failed to keep his warring department in check but he is claimed to have swerved from readiness to do a deal with Mr. Sixsmith to fury at a government 'complete cock-up', before finally throwing in his lot with Mr. Byers […] But if_2 Sir Richard has been tainted by the affair, and if_2 Mr. Sixsmith's role may not have been as entirely well-intentioned as he claims, the individual most damaged by the row remains Stephen Byers.
(*Guardian*, 25.02.2002, page 4, col 3)

However, a central adverbial clause and a peripheral adverbial clause cannot coordinate, even when introduced by the same conjunction. Example (10a) contains an event conditional and a factual conditional. Though (10b) is acceptable, both *if*-clauses must get a factual reading.[5] Example (10c) is based on (2d), which contains a central *while*$_1$-clause and a peripheral *while*$_2$-clause: though these *while*-clauses are associated with the same clause, they cannot coordinate.

[5] We added *indeed* in (10a) and (10b) to ensure the factual reading of the conditional clause.

(10) a. If$_2$ the head of department ought indeed to be present, then the meeting will be cancelled if$_1$ he's unable to travel.
b. If$_2$ the head of department ought indeed to be present and if$_{*1/2}$ he's unable to travel, then the meeting will be cancelled.
c. *While$_2$ [the lawsuit challenging the legitimacy of lethal injections] probably won't stop the use of lethal injection altogether and while$_1$ the supreme court decides what to do, it will certainly delay its use.

In the formal literature, it is standardly assumed that coordination is subject to a 'likeness' condition (Williams's 1978 *Law of coordination of the likes*). For discussion of the definitions and problems associated with this condition see Whitman (2004). We interpret the condition here also[6] in syntactic terms along the lines of Huddleston and Pullum (2005):

> A coordination of α and β is admissible at a given place in sentence structure if and only if each of α and β is individually admissible *at that place* with the same function. (Huddleston and Pullum 2005: 201, italics lh, ms)

Adopting this syntactic view on coordination, the proposal would then be that central adverbial clauses, such as event conditionals, and peripheral adverbial clauses such as factual conditionals, cannot coordinate because they do not occupy the same 'structural place', which is interpreted here in the sense that they are attached at different heights in the structural hierarchy, event conditionals being lower, i.e. more closely integrated with the associated clause, than factual conditionals. We return to the level of integration of conditional clauses in Sect. 3.

[6] Observe that in a cartographic view (cf. Cinque and Rizzi [2008] for an introduction), in which syntactic structure closely matches semantic interpretation, the height of attachment of the conditional clauses correlates with a semantic distinction, a point that will become clearer in Sect. 3.2. So the constraint on coordination of 'likes' is both semantic and syntactic.

2.2 Scope Phenomena

Various scopal properties distinguish central and peripheral adverbial clauses: in a nutshell, central adverbial clauses are in the scope of operators in the associated clause while peripheral adverbial clauses are outside the scope of the same operators. Below are some illustrations of such scope differences with a focus on conditional clauses (for more examples see Haegeman 2003, 2012). The scopal distinctions fall out naturally from an account which postulates different heights for attachment of the conditional clauses.

2.2.1 Temporal and Modal Subordination

Central adverbial clauses are temporally and modally subordinated to the associated clause. Temporal and modal subordination is reflected in the tense forms used in English event-conditional clauses. For instance, in the central *if*$_1$-clause (1/3a), repeated as (11a), the present tense verb *tire* inherits a futurity reading from the future time expression *will be* in the associated clause, a phenomenon sometimes referred to as '*will* deletion' (Jespersen 1909; Palmer 1965; McCawley 1971; Leech 1971; Palmer 1974; Zandvoort 1975; Wekker 1976, 1977; Haegeman and Robinson 1979; Close 1980; Comrie 1982; Declerck 1984; Niewint 1986; Declerck 1991; Iatridou 1991; Biasio and Castro 2019; a.o. for discussion and additional references). The central *if*$_1$-clause in (3b), repeated as (11b), illustrates modal subordination: the past tense form *had* in the conditional clause inherits the irrealis reading encoded by the irrealis modal *would* in the associated clause.

(11) a. If$_1$ your back-supporting muscles tire, you will be at increased risk of lower-back pain.
b. If$_1$ last week you had shown me the piece of pipe system that Laila and I built on Tuesday, I would never have believed it.

Declerck and Reed (2001: 131) summarize the result of the subordination as follows:

> [In (11) lh, ms], the speaker makes a *single* (but complex) prediction: she presents the contents of the two clauses as forming a unit. (Declerck and Reed 2001: 131, italics lh, ms)

In contrast with central adverbial clauses, peripheral adverbial clauses do not manifest temporal or modal subordination in relation to the associated clause (but see Sect. 2.4). In the factual conditional (3f), repeated as (12a), futurity is encoded independently (*I'm no longer going to*'). The examples (12b) and (12c) from the literature illustrate the same point. In the attested (12d), the present tense *is worried* is not temporally subordinated: present tense here 'means' present time.

(12) a. If$_2$ I'm no longer going to be arrested for possessing cannabis for my own consumption ('Cannabis laws eased in drugs policy shake-up', October 24), shouldn't I be able to grow my own? (Jason Cundy, Letter to the editor, *Guardian,* 25.11.2001, page. 9, col 8)

b. If$_2$ the lava will come down as far as this, all these houses must be evacuated at once. (Close 1980: 103)

c. If$_2$ he won't arrive before nine, there's no point in ordering for him.
(Comrie 1982: 148)

d. If$_2$ Tony Blair is worried about public confidence already, in this bright weather, he should think about what it's going to be like when we are huddled into the December winds. (*Independent*, Comment 01.11.2001, page 5, col 1)

Declerck and Reed (2001: 131) provide the following characterization:

> When the Future Perspective System [*will, be going to*, lh, ms] is used in both clauses [as in (12) lh, ms], the *speaker makes two independent predictions*: there are, as it were, *two illocutionary speech acts*. (Declerck and Reed 2001: 131, italics lh, ms)

Following Hornstein (1990: 43) and much later work, we take it that temporal and modal subordination are regulated by syntactic structure.

2.2.2 Negation

Sentential negation can scope over event conditionals (13a, b), but it cannot scope over factual conditionals (13c). To illustrate: (13a) means that it is not the case that if it rains the speaker will go to the park; (13b) means that it will never be the case that if you take the tram you'll get home on time. The negation in these examples bears on the causal relationship between the condition and the main clause event. On the other hand, in (13c) sentential negation as encoded in *not* does not bear on the causal relation between the state of affairs expressed in the conditional, i.e. the manufacturers being pleased, and the state of affairs in the associated clause, i.e. the manufacturers being satisfied; the fact that the manufacturers are pleased is not presented as leading to them being satisfied. Thus (13c) could be paraphrased as 'the manufacturers are pleased, but they are not satisfied'. For reasons of space, we refer the reader to earlier work by Haegeman (1984a; b, c, 1991/2009) for further illustrations.

(13) a. If$_1$ it rains, I won't go to the park.
b. You will never get home on time if$_1$ you take the tram.
c. Only the chocolate manufacturers could look with pleasure at these statistics. But if they are pleased, they are not satisfied. (*Guardian*, G2, 08.03.2002, page 6, col 1)

2.2.3 Focus

The two types of adverbial clauses pattern differently in relation to their focusing potential. While event conditionals can be focused by *only* (14a), with subject-auxiliary inversion as a concomitant, this is not possible for factual conditionals (14b) (cf. Quirk et al. 1985: 1070f.).

(14) a. Only if$_1$ you have the courage to follow your heart will you succeed on the path of love.
b. The chocolate manufacturers looked with pleasure at the statistics. # Only if$_{1/*2}$ the manufacturers are pleased are they not satisfied.

While event conditionals can be the focus of an *it*-cleft (15a),[7] this is not possible for factual conditionals (15b). The latter is only acceptable if the *if*-clause is interpreted as an event conditional: the manufacturers are not satisfied if they are pleased.

(15) a. If you download 2 billion links from one index, they will be unique. It is only if$_1$ you try to merge both the Fresh index and Historic index that you will get duplicate links. (https://blog.majestic.com/development/topicaltrustflow/)

b. The chocolate manufacturers looked with pleasure at the statistics. It is only if$_{1/*2}$ the manufacturers are pleased that they are not satisfied.

Along the same lines, event conditionals can be the focus of interrogatives: (16a, b) illustrate *yes/no* questions, (16c) illustrates a *wh*-question. Examples (16a) and (16b) enquire whether, in the event of being abroad, the interlocutor would read Belgian papers. In (16c), the *if* $_1$-clause is a reply to the *when*-question about the conditions under which the interlocutor would give up eating meat. As shown by the attested (16d), factual conditionals associated with an interrogative clause remain outside the scope of the interrogative operator. In (16d) it is taken as a given that 'crime is falling' and therefore this proposition is not within the scope of the interrogative operator.

(16) a. Would you also read Belgian papers if$_1$ you were abroad?

b. If$_1$ you were abroad, would you also read Belgian papers?

c. A: When would you decide to give up eating meat?

B: If$_1$ there were enough vegetarian restaurants in my hometown.

d. We are seeing a fall in the incidence of crime, particularly serious crime, and I think we're right to say, 'What's going on?' If$_2$ crime is falling, why are we seeing a continuing rise in the prison population.
(*Guardian*, 01.11.2001, page 2, col 6)

[7] In clefted *if* $_1$-clauses the addition of *only* is obligatory, a point which we won't go into here.

2.3 VP Ellipsis and VP Anaphora

The hypothesis that the two types of conditionals are distinguished by the height of their syntactic attachment leads us to expect that they are affected differently by syntactic processes which operate on a relatively low structural level such as, for instance, *so* VP anaphora. This expectation is borne out. *So* VP anaphora may subsume an event conditional, as shown in (17a), but it cannot affect a factual conditional, as shown in (17b). As a result, in (17a) both a strict (i) and a sloppy (ii) identity reading are available for the possessive pronoun *his* in the recovered event conditional 'if his paper is discussed'. In (17b) the factual conditional clause is not covered by anaphoric *so* and a sloppy identity reading is unavailable.

(17) a. John will leave if_1 his paper is discussed and so will Bill.
(i) 'Bill will also leave if John's paper is discussed.'
(ii) 'Bill will also leave if his (own) paper is discussed.'
b. A: John's paper is going to be discussed now.
B: John must leave, if_2 his paper is going to be discussed now, and so must Bill.
(i) 'Bill must also leave, if John's paper is going to be discussed now.'
(ii) *'Bill must also leave, if his (own) paper is going to be discussed now.'

2.4 Embedding of the Conditional Clause with the Associated Clause

Due to their closer integration with the associated clause, we may expect that event conditionals can be embedded with the associated clause. This is illustrated in the attested examples in (18), in which the bracketed conditionals embed with their associated clause:

(18) a. No, I think [that [if$_1$ he had known he would be president], he would have started dying his hair, like, 10 years ago]. (Michelle Obama https://quotefancy.com/quote/867878/, Michelle-Obama-No-I-think-that-if-he-had-known-he-would-be-president-he-would-have)

b. When asked by FRONTLINE in 2015 about the use of vaccines to combat vaccine-prev entable illnesses, she said: "if you ask 99.9 percent of parents who have children with autism if we'd rather have the measles versus autism, we'd sign up for the measles," adding [that [if$_1$ she had another child], she wouldn't vaccinate]. (https://www.businessinsider.nl/jenny-mccarthy-became-the-face-of-the-anti-vaxx-move ment-2019-4/?jwsource=cl)

The examples in (19) show that, despite their lesser integration with the associated clause, factual conditionals also embed with the associated clause, which means that factual conditionals are also syntactically integrated with the associated clause. Observe in particular that the past tense (*thought*) in the factual conditional in (19b) arises as the result of sequence of tenses due to embedding under a past tense verb (here *led*), which is evidence that the factual conditional is in the scope of the past tense in the main clause. The temporal dependency confirms that the factual conditional clause is integrated in the syntax of the domain of embedding.

(19) a. The party now has to prepare for a possible general election as early as next May and is already looking at appointing an advertising agency to help run the campaign. "I think [that [if$_2$ Theresa May is going to go [for an early election, lh, ms]], it will be May or October]," he says. (https://www.theguardian.com/politics/2016/sep/24/corbyn-pledge-on-grassroots-after-leadership-win)

b. He unilaterally announced the award of the George cross to the RUC, and was effusive about its history – which led Unionists to question [that, [if$_2$ he thought the RUC were so honourable], why did he later agree with Patten and strip the police of their 'Royal' Title?] (*Observer*, 23.04.2000, page 13, col 4)

2.5 Verb Second

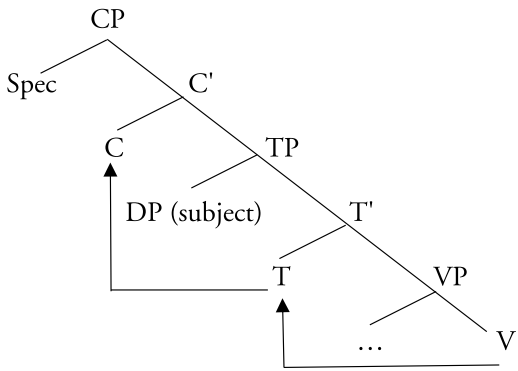
Evidence from Verb Second (V2) phenomena in Germanic languages confirms the conclusion in Sect. 2.4 that both event conditionals and factual conditionals are syntactically integrated with the associated clause (cf. for similar claims in a.o. Frey 2016, 2018, 2019, 2020; Frey and Meinunger 2019).

The label V2 describes the typical word-order restriction in German and Dutch, by which the finite verb in the root clause is preceded by one and only one constituent. In (20), for instance, the finite auxiliary *heeft* 'has' is preceded by a temporal adjunct (20a), by a direct object (20b), by the subject (20c), or by the VP (20d), but it cannot be preceded by more than one such constituent (20e–g).

(20)	a.	[Gisteren]	heeft	Hans	dat huis	gekocht.
		yesterday	has	Hans	that house	bought
	b.	[Dat huis] heeft Hans gisteren gekocht.				
	c.	[Hans] heeft gisteren dat huis gekocht.				
	d.	[Dat huis gekocht] heeft Hans gisteren.				
	e.	*[Gisteren][dat huis] heeft Hans gekocht.				
	f.	*[Hans] [dat huis]heeft gisteren gekocht.				
	g.	*[Gisteren] [Hans] heeft dat huis gekocht.				

V2 has received a range of syntactic analyses (see Woods and Wolfe 2020 for an overview), but there is a fair amount of consensus that the phenomenon is syntactic. According to the approach initiated by den Besten (1977/1983/1989) and since modified for theory-internal reasons, the V2 order is derived by movement of the finite verb to the position C, with an additional constituent moving from a TP-internal position to the specifier position, SpecCP (Holmberg 2020). The simplified representation in (21) ignores additional functional structure as that postulated in cartographic approaches (cf. Haegeman 1996; Holmberg 2020).

(21)



Like non-clausal adjuncts (20a), event conditionals fulfil the V2 requirement: when they are the first constituent of the root clause, they must be followed immediately by the finite verb (22a). Factual conditionals also serve to fulfil the V2 requirement: as the first constituent of the root clause they must also be followed immediately by the finite verb, as shown in (22b).

(22) a. Als$_1$ het regent, blijf ik/ *ik blijf thuis.
if it rains stay I/ I stay home
(but see Haegeman and Greco 2018, 2020 for West Flemish)

b. Als$_2$ het morgen waarschijnlijk gaat regenen,
if it tomorrow probably goes rain
kunnen we/ *we kunnen beter met de trein gaan.
can we/ we can better with the train go
'If it's probably going to rain tomorrow, we'd better go by train.'

Assuming with Holmberg (2020) that the initial constituent in a V2 configuration originates in a clause-internal position, we conclude that both types of conditional clauses originate in a position internal to the clause (labelled TP in (21), but to be modified in Sect. 3.2). Conversely, the fact that neither event conditionals nor factual conditionals can be the first constituent in a V2 transgression (Catasso 2015) means that neither the event conditional nor the factual conditional can function as clause-external constituents in the sense of Broekhuis and Corver (2016: 1679–1733).

3 The External Syntax of Factual Conditionals and Judgement Phrase

The empirical data discussed in Sect. 2 correlate with a structural analysis according to which both types of conditional clauses, event conditionals and factual conditionals, are inserted clause-internally and in which the differences highlighted in Sect. 2 correlate with the relative height of their insertion, i.e. in terms of the degree of syntactic integration.[8]

This section investigates the level of attachment of factual conditionals (and peripheral adverbial clauses in general).

3.1 Epistemic Modals

Importantly, the diagnostics that help to determine the syntactic attachment site of factual conditionals and to differentiate between factual conditionals and event conditionals, also single out a set of 'high' adverbial modifiers, in the sense of Cinque (1999), illustrated here using the epistemic modal *probably*. From (23a) it is clear that the epistemic modal *probably* must be located at some clause-internal level: it follows the canonical subject, the pronominal *he*, which is assumed to occupy the highest specifier position of the clausal domain (e.g. SpecTP in a standard Minimalist format), and the finite auxiliary *will*, which occupies the associated head position (T in the Minimalist format). Importantly for our purposes, the temporal interpretation of the epistemic modal is tied to the present speech time rather than to the future: (23a) means that 'it is now probable that he will take early retirement', and it does not mean that 'at some future time it will become probable that he takes early

[8] A radical alternative to this approach taken in Declerck and Reed (2001) is to deny that the distinction between central and peripheral adverbial clauses is syntactic:

> a subordinate clause is a syntactically dependent clause. Such questions as the scope of negation, focusing, modality, etc.; in the head clause are immaterial to this, as they pertain, not to syntactic, but to semantic integration (Declerck and Reed 2001: 37f.).

Given that we endorse a view according to which the various patterns discussed in Sect. 2—i.e. scope, temporal subordination, focus, etc.—are syntactically encoded, this viewpoint is not pursued.

retirement'. The assessment of the probability is in the present speech time. Similarly, in (23b), with a reading such as 'I consider it probable that he did not take early retirement', *probably* is not within the scope of sentential negation encoded by *not*. Like factual conditionals (and other peripheral adverbial clauses), the epistemic modal *probably* cannot be focused, as shown in (23c) and (23d), the latter intended with the cleft reading, and it cannot be the focus of *wh*-questions as in (23e). As shown in (23f), VP ellipsis does not subsume the epistemic modal *probably*: 'his wife has too' means that 'his wife has also gone home' and not that 'his wife has also probably gone home'. Like factual conditionals, the epistemic modal can occur in embedded domains, in which case it is interpretively tied to the subject of the embedding predicate. In (23g), for instance, the source of the probability assessment is the main clause subject *Jane*. The Dutch analogue of *probably*, *waarschijnlijk*, can constitute the first constituent in a V2 clause in Dutch (23h), and, conversely, it does not give rise to a V2 transgression (23i). The latter points again confirm strongly that the epistemic modal *probably* is syntactically integrated with the clause it modifies and that, following Holmberg (2020), it must originate in a clause-internal position.[9]

(23) a. He will probably take early retirement.
b. He probably did not take early retirement.
c. He has (*even) probably taken early retirement.
d. *It is PROBABLY that he has taken early retirement, not definitely.
e. A: How did he leave?
B: *Probably.
f. He has probably gone home, and his wife has too.
g. Jane thinks that he will probably take early retirement.
h. Waarschijnlijk komt hij morgen terug.
probably comes he tomorrow back
i. *Waarschijnlijk hij komt morgen terug.
probably he comes tomorrow back

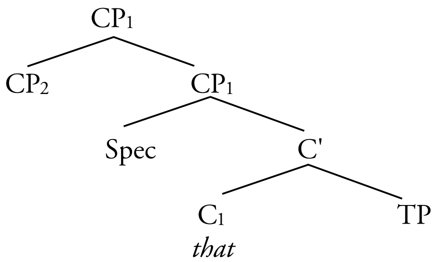
[9] Andrew Radford (p.c.) confirms that it is hard to get a *probably* reading for the deleted VP in (23f).

Given the overlap between the properties of the epistemic modal *probably* and those of factual conditionals discussed in Sect. 2, it becomes tempting to view the height of insertion of factual conditionals as related to the structural layer associated with that of the epistemic modal, as proposed in Frey (2016, 2018, 2019, 2020).

3.2 Layers of Structure: Krifka (2017; to Appear), Frey (2020)

In earlier work (Haegeman 1984b, 2003; Coniglio 2011; Frey 2011) it was proposed that peripheral adverbial clauses such as factual conditionals are adjoined to the topmost layer of the clause, 'CP' (or 'ForceP' in the cartographic tradition) and thus remain outside the scope of CP- or TP-internal operators.

(24)



However, if factual conditionals were located in this position, i.e. outside the associated CP, they should necessarily precede the complementizer *that*, which occupies the position C, as in (25a). But, like central adverbial clauses, embedded factual conditionals follow *that*, as shown in (19a), the crucial parts of which are repeated for convenience in (25b):

(25) a. *I think [[if$_2$ Theresa May is going to go for an early election], that it will be May or October].
b. I think [that [if$_2$ Theresa May is going to go for an early election], it will be May or October].

Based on Krifka (2017; to appear), Frey (2016, 2018, 2019, 2020) assumes the functional hierarchy in (26a), with the various layers characterized as in (26b), based on Frey (2020: 13).

(26) a. ActP > CmP > JP > TP
b. i. The Tense phrase (TP) encodes a proposition φ.
ii. The Judgement phrase (JP) encodes a judge and expresses an evaluation of the proposition φ by the judge.
iii. The Commitment phrase(CmP) encodes a committer and expresses public commitment of the committer.
iv. The Speech act phrase (ActP) encodes the speaker and expresses the occurrence of a specific speech act.

Crucially for our purposes, the original layer 'TP' as used in the Minimalist representation is here decomposed into an articulated structural layer in which TP labels a lower sub-component. The layer JudgementP (JP) dominates this lower TP-layer and these two layers have specialized functions: TP encodes the propositional/at-issue content of the clause while JP contains proposition-external, not-at-issue content. For further motivation for JP we refer the reader to Krifka's own work and to Frey (2018, 2019, 2020); for the concept 'at-issue' we refer the reader to Potts (2015) and to Charnavel (2020), the latter in relation to French *puisque* ('since').[10]

---

[10] Krifka's (2017, to appear) functional layer JP, adopted by Frey (2018, 2019, 2020), could be reinterpreted as a 'telescoped' variant of Cinque's (1999) topmost four high modal projections: $\text{MoodP}_{\text{speech act}}$, $\text{MoodP}_{\text{evaluative}}$, $\text{MoodP}_{\text{evidential}}$ and $\text{ModP}_{\text{epistemic}}$. Cinque's hierarchy is replicated in (i).

(i) $\text{MoodP}_{\text{speech act}}$ > $\text{MoodP}_{\text{evaluative}}$ > $\text{MoodP}_{\text{evidential}}$ > $\text{ModP}_{\text{epistemic}}$ > TP (Past > TP (Future) > $\text{MoodP}_{\text{irrealis}}$ > $\text{ModP}_{\text{alethic}}$ > $\text{AspP}_{\text{habitual}}$ > $\text{AspP}_{\text{repetitive}}$ > $\text{AspP}_{\text{frequentative}}$ > $\text{ModP}_{\text{volitional}}$ > $\text{AspP}_{\text{celerative}}$ > TP(Anterior) > $\text{AspP}_{\text{terminative}}$ > $\text{AspP}_{\text{continuative}}$ > $\text{AspP}_{\text{retrospective}}$ > $\text{AspP}_{\text{proximative}}$ > $\text{AspP}_{\text{durative}}$ > $\text{AspP}_{\text{generic/progressive}}$ > $\text{AspP}_{\text{prospective}}$ > $\text{ModP}_{\text{obligation}}$ > $\text{ModP}_{\text{permission/ability}}$ > $\text{AspP}_{\text{completive}}$ > VoiceP > $\text{AspP}_{\text{celerative}}$ > $\text{AspP}_{\text{repetitive}}$ > $\text{AspP}_{\text{frequentative}}$ (Cinque 2004: 133, his (3)).

Further research will have to shed light on the question as to what extent the four distinct levels postulated by Cinque (1999) could or should be correlated to specific peripheral adverbial clauses such as, for instance, factual conditionals, concessive *while*-clauses, rationale *since*$_2$/*as*$_2$/*vermits*-clauses, etc. For proposals the interested reader is referred to Endo and Haegeman (2019), who propose a general mechanism for the insertion of adverbial clauses in relation to their internal syntax, and Charnavel (2020) for a discussion of French rationale *puisque* ('since') clauses as modifiers of Cinque's $\text{MoodP}_{\text{evidential}}$.

Subjective epistemic modals such as *probably* are proposition-external and hence not-at-issue; they are associated with JP. Because they are not-at-issue, they cannot be focused (for more discussion see Krifka 2017; to appear; Frey 2018).

To capture the fact that peripheral adverbial clauses are syntactically integrated but remain outside the scope of TP operators, Frey (2016, 2018, 2019, 2020) postulates that they are associated with JP (27a). In this chapter Frey's analysis is applied to factual conditional *if* 2-clauses (27b).

(27) a. [$_{\text{CP1}}$ *that* [$_{\text{JP}}$ [$_{\text{CP2}}$ peripheral adverbial clause] [$_{\text{TP}}$ subject…
b. [$_{\text{CP1}}$ *that* [$_{\text{JP}}$ [$_{\text{CP2}}$ *if*$_2$-clause] [$_{\text{TP}}$ subject…

Representation (27) remains compatible with the observation that when embedded, factual conditional *if* 2-clauses follow the complementizer and that in V2 languages factual conditionals may be moved from their lower position to constitute the first constituent in a V2 clause.

Frey's original analysis, which correlates the syntactic location of peripheral adverbial clauses, here exemplified by factual *if* 2-clauses, with that of the epistemic modals is strengthened by the observations presented in Sect. 3.1 that several distributional properties of *if* 2-clauses overlap with the distributional properties of the English epistemic modal *probably* and its Dutch analogue *waarschijnlijk*.

3.3 Against an Orphan Account for Factual Conditionals

For completeness' sake, we briefly turn to an alternative hypothesis concerning the difference between central and peripheral adverbial clauses developed in Haegeman (1991/2009), based on Safir (1986), Fabb (1990), and also endorsed in Shaer and Frey (2004). The idea was that peripheral adverbial clauses are 'orphan constituents', i.e. they are merged with the associated clause outside the narrow syntax, as extra-sentential orphan constituents. The proposal is schematically represented in (28): in this view, the associated clause entertains a discourse-type relation with the peripheral adverbial clause. For concrete proposals in relation to the syntax of extra-sentential constituents the reader is referred to Safir (1986), Fabb (1990), Koster (2000), Shaer and Frey (2004), Cinque (2008), Axel and Wöllstein (2009), Haegeman et al. (2009), Giorgi (2014) and Haegeman and Greco (2018, 2020), among many others.

(28) [$_{CP2}$ peripheral adverbial clause] [$_{CP1}$ associated clause]

Upon closer inspection, the orphan analysis raises several issues. One concerns the diagnostics invoked in Sect. 2 as the basis for (28). As shown in Sect. 3.1, it turns out that, although peripheral adverbial clauses, such as factual conditionals, can be distinguished from central adverbial clauses, such as event conditionals, these diagnostics do not constitute conclusive evidence for the orphan analysis. On the contrary, the observed parallelisms with epistemic modals, which are definitely not orphan constituents, and the data from embedding and from V2 patterns, show that factual conditionals must be structurally integrated.

In the next section, though, the orphan analysis schematized in (28) will be shown to be relevant for a different set of conditional clauses, namely those like that in (1c), which are modifiers of the speech event.

4 Speech-Event Modifiers as Syntactic Orphans

4.1 A Third Type of Adverbial Clause

Central adverbial clauses were opposed to peripheral adverbial clauses, contrasting event conditionals, as in (29a), with factual conditionals, as in (29b). However, a third use of conditional clause was illustrated in (1c), in which what looks like a conditional clause, *if I may say so*, functions as a modifier not of the event encoded in the associated clause but of the speech event as such. Another example of this use of a conditional, based on the literature, is given in (29c). Henceforth we will annotate the conjunction which introduces a speech-event modifying adverbial clause with a subscripted diacritic '?' because the clause it introduces is distinct from the two categories established so far. Haegeman's own earlier work (see also Haegeman and Wekker 1984) grouped conditionals that modify the speech event (29c) with factual conditionals (29b) and considered both as peripheral conditionals. Following Takami (1988) and Iatridou (1991: 50–57), a.o., who argue in favour of a ternary distinction, we reconsider this binary classification here. This section expands on Haegeman (2012: 181f.), Schönenberger and Haegeman (to appear) and is heavily indebted to Frey (2016).

(29) a. If$_1$ you don't pass the final exam, you won't get your degree.
b. If$_2$ he won't arrive before nine, there is no point in ordering dinner for him. (based on Quirk et al.1972: 781; in Takami 1988: 271, his (28))
c. If$_?$ I may change the subject, I visited one of my friends in America last month. (based on Takami 1988: 271, his (29))

The need for a ternary classification extends to other adverbial clauses. The temporal *while*$_1$-clause in (30a) is classified as a central adverbial clause. The concessive *while*$_2$-clause in (30b) provides a background proposition for the processing of the associated clause, and the *while*$_?$-clause in (30c) provides a temporal modification of the speech event. In previous work Haegeman (1984b, 2003; 1991/2009) grouped concessive *while*$_2$-clauses (30b) with examples such as (30c), treating both as peripheral adverbial clauses.

(30) a. While$_1$ we were talking about Theresa May, the BBC announced her resignation.
b. While$_2$ the Prime Minister may be a conservative, her recent proposals are very innovative.
c. While$_?$ we are talking about Theresa May, her recent proposals are very innovative.

In (31) additional attestations of conditional *if*$_?$-clauses are provided, which modify the speech event rather than the event encoded in the associated clause. Crucially, interpretively these are not factual conditionals: they do not echo a contextually salient proposition.

(31) a. 'If$_?$ you don't mind me asking,' ventures the BAT press officer. 'Why are you particularly interested in Iran?' (*Guardian*, 02.09.2005, page 24, col 1)

b. Leasing out video-conferencing facilities was where it was at, if$_?$ I recall – that and being some kind of ceremonial bag carrier for the Lord Lieutenant of Devon. (*Guardian*, G2, 18.10.2005, page 14, cols 2-3)

c. In fact, poor old 't' is disappearing even in the middle of words – e.g. 'butters' is commonly said as 'buyers'. The most stable letters are 'm' and 'n', if$_?$ you're interested: they're very unlikely to disappear from spoken language.
(*Observer*, 14.08.2005, page 8, col 3)

d. If$_?$ Hughes and Oaten are anything to go by, it's a matter of time until Lib Dem Front-runner 'Ming' Campbell gets bogged down in Scandal.
(*Independent,* 02.01.2006, page 34, col 3)

While the adverbial clause modifying the speech event arguably has an event reading in (29c), in (30c) and in (31), and thus interpretively aligns with ‘central adverbial clauses’, this is not the only type used as a speech-event modifier: in (32a), the *since$_?$*-clause also functions as a speech-event modifier but it provides the rationale for the speech event. Interpretively it would thus correspond to the rationale *since$_2$*-clause in (32b), which has the properties of the peripheral adverbial clauses set out in Sect. 2 (see for French *puisque* (‘since’) Charnavel (2020), Schönenberger and Haegeman, to appear).

(32) a. I possibly have rose-tinged memories because I'd just attained my first girlfriend and earned and spent my first own money (on a small bottle of Brut for Men by Fabergé, [since$_?$ you ask], and, yes, it still astonishes me how the acquisition of the second did not more violently militate against the acquisition of the first). (*Observer*, 06.11.2005, page 18, col 1)

b. Dr Durieux says it is unrealistic to expect the provincial capital to move, [since$_2$ never in human history has a city evacuated before a natural catastrophe]. (*Guardian*, 22.11.2004, page 9, col 4)

4.2 Scope Effects and Speech-Event Modifiers

As far as scopal properties are concerned, speech-event modifiers align with peripheral adverbial clauses, which was the basis for grouping the two sets together in earlier work. We illustrate this point with conditional clauses.

Like factual conditionals, speech-event conditionals are not temporally subordinated to the tense in the associated clause: the *if*?-clause is related to the speech time and its tense form and its temporal interpretation are independent of the tense form or the temporal interpretation of the associated clause, as shown in (33). Moreover, (33c) shows that speech-event modifiers are not within the scope of the sentential negation of the associated clause.

(33) If? I may change the subject,
  a. I visited one of my friends in America last month.
  b. I am visiting one of my friends in America next week.
  c. I never got a reply to that email I sent to my supervisor.

Like factual conditionals, speech-event conditionals cannot be clefted and they cannot be focused by *only:*

(34) a. *It is if? I may change the subject that I visited one of my friends in America last month. (based on Takami 1988: 271, his (29))
  b. *Only if? I may change the subject did I visit one of my friends in America last month. (based on Takami 1988: 271)

Like factual conditionals, speech-event conditionals cannot be the focus of an interrogative: in (35a), the speech-event conditional is not an appropriate reply to the *wh*-question; in (35b), the speech-event conditional is not within the scope of the *yes/no* operator.

(35) a. When did you visit your American friend?
  ! If? I may change the subject.
  b. If? I may change the subject, did you visit your friend last month?

In terms of the diagnostics deployed in Sect. 2, speech-event conditionals do pattern with factual conditionals. The next section will show, however, that speech-event conditionals differ from factual conditionals in several respects. The conclusions reached here extend to other speech-event modifying adverbials.

4.3 Re-Evaluating the Analyses

4.3.1 Coordination

Section 1.2 invoked the fact that a central adverbial clause, such as an event conditional or a temporal *while*$_1$-clause, and a peripheral adverbial clause, such as a factual conditional or a concessive *while*$_2$-clause, cannot coordinate (see examples in (10)) in support of the need to differentiate between the two clause types in terms of their structural position. The gist of the argument was that the illicit coordination would violate the coordination of the *likes constraint* because in the structural perspective of Huddleston and Pullum (2005), central adverbial clauses and peripheral adverbial clauses do not occupy the same 'structural place', i.e. their level of attachment to the host clause is different.

Observe now that factual conditionals and conditional speech-event modifiers also cannot coordinate, a point made in Ros (2005: 94f.), shown in (36) (which corresponds to his (24b)). Although both *if*-clauses modify the same associated clause in (36a), the factual conditional *if*$_2$ *John's wife is French* cannot coordinate with the speech-event conditional *if*$_?$ *I might say so* in (36b).

(36) a. John should know about wines if$_2$ John's wife is French, if$_?$ I might say so.
b. *John should know about wines if$_2$ John's wife is French$_2$ and if$_?$ I might say so.

This argument extends to adverbial clauses introduced by other conjunctions. Both *while*-clauses modify the same associated clause in (37a), but the concessive *while*$_2$-clause and the *while*$_?$-clause modifying the speech time cannot coordinate in (37b).

(37) a. While$_?$ we are talking about Theresa May, while$_2$ the Prime Minister may be a conservative, her recent proposals are very innovative.

b. *While$_?$we are talking about Theresa May and while$_2$the Prime Minister may be a conservative, her recent proposals are very innovative.

As shown in (36b) above, speech-event conditionals do not coordinate with factual conditionals, and they do not coordinate with event conditionals either:

(38) a. Her behaviour will not improve if$_1$ you do not react soon, if$_?$ I might say so.

b. *Her behaviour will not improve if$_1$ you do not react soon and if$_?$ I might say so.

4.3.2 *Then* Resumption (Takami 1988: 271f.)

Takami (1988: 271f.) signals an additional contrast between event conditionals and factual conditionals, on the one hand, and speech-event conditionals, on the other: like event conditionals (39a), factual conditionals (39b) can be resumed by *then*. With speech-event conditionals (39c, d), *then* resumption is not licit. This point is also made in Ros (2005: 96), who cites Wakker (1996).

(39) a. If$_1$ you don't pass the final exam, (then) you won't get your degree.

b. If$_2$ he won't arrive before nine, (then) there is no point in ordering dinner for him. (based on Quirk et al. 1972: 781; in Takami 1988: 271, his (28))

c. If$_?$ I may change the subject, (*then) I visited one of my friends in America last month. (based on Takami 1988: 271, his (29))

d. If$_?$ you're interested, (*then) I visited one of my friends in America last month.

4.3.3 Non-Veridicality

Speech-event conditionals pattern differently from both central adverbial clauses and peripheral adverbial clauses. However, arguably, being event modifiers, they are to some extent similar to event conditionals, as they are also compatible with NPIs: (31d), repeated as (40a), contains the NPI *anything*. Moreover, speech-event conditionals, just like event conditionals, are compatible with putative *should* (40b) and they also allow conditional inversion with putative *should* (40c).

(40) a. If$_{?}$ Hughes and Oaten are anything to go by, it's a matter of time until Lib Dem Front-runner 'Ming' Campbell gets bogged down in Scandal.
(*Independent*, 02.01.2006, page 34, col 3)
b. If$_{?}$ you should be interested, the conference is going to be in Paris this year.
c. Should you be interested, the conference is going to be in Paris this year.

4.3.4 Embeddability

While event conditionals and factual conditionals can be embedded with the associated clause, as discussed in Sect. 2.4, speech-event conditionals do not embed, as shown in (41). See also Charnavel (2020) for a similar observation concerning French *puisque*-clauses. The conditional *if you are interested* in (41a) can be interpreted as a speech-event conditional. As a speech-event conditional it frames the speech event in relation to the interlocutor's interest. However, the conditional clause in (41a) can also spell out the circumstances in which the eventuality in the associated clause itself takes place: whether or not Tom is going to cook dinner is dependent on the interlocutor being interested. Hence (41a) is ambiguous, it can be interpreted as a speech-event conditional or as an event conditional. But (41b), which embeds (41a) as indirect speech, is not ambiguous. The embedded conditional in (41b) must now be interpreted as spelling out the circumstances in which the eventuality in the associated clause itself will be realized, i.e. whether or not Tom is going to cook dinner depends on 'me', the indirect object of the main clause and the addressee of the speech act verb *told*, being interested. The embedded conditional in (41b) cannot be interpreted as modifying the embedded speech event itself.

(41) a. If$_{1/?}$ you're interested, Tom is going to cook dinner next Saturday.
b. Harry told me [that [if$_{1}$/ $_{*?}$ I was interested], Tom was going to cook dinner next Saturday].

The embeddability restriction extends to other speech-event modifying adverbial clauses. For instance, the *before*-clause in (42a) is a temporal modifier of the speech event, whereas the embedded *before*-clause in (42b) cannot be interpreted as a temporal modification of the embedded speech event. However, (42b) would be acceptable with the reading in which the temporal *before*-clause modifies the eventuality encoded in the embedded clause.

(42) a. Before we start, five cabinet ministers will be voting with the opposition.
b. The invited speaker announced [that [before we started], five cabinet ministers would be voting with the opposition].

4.3.5 Verb Second

Based on evidence from Dutch, Sect. 2.5 revealed that both event conditionals and factual conditionals fulfil the V2 requirement, providing evidence for their syntactic integration with the associated clause. In contrast, as shown in (43a) and (43b), speech-event conditionals do not satisfy the V2 requirement. Conversely, they can pattern as a first constituent in a V2 transgression (Catasso 2015) and thus behave like extra-sentential constituents (Broekhuis and Corver 2016: 1679–1733). This is also true of the *before*-clause in (43c).

(43) a. Als? je het echt moet weten, ik was/*was ik in Rome
if you it really must know I was / was I in Rome
'If you really want to know, I was in Rome.'
b. Als? u het zich herinnert,
if you it yourself remember
het boek verscheen/*verscheen het boek voor het eerst in 1982.
the book appeared / appeared the book for the first in 1982
'If you remember, the book first appeared in 1982.'
c. Voor? we beginnen, de vergadering gaat/ *gaat
before we start the meeting goes/ goes
de vergadering morgen niet door.
the meeting tomorrow not through
'Before we start, tomorrow's meeting has been cancelled.'

4.3.6 Summary

Table 1 summarizes the similarities and differences between the various conditional clauses.

Table 1 Three types of conditional clauses

	Event conditional	Factual conditional	Speech-event conditional
Temporal/modal subordination	+	–	– (anchored to present)
In scope of host clause negation	+	–	–
In scope of host clause interrogative	+	–	–
Host clause cleft/host clause focus	+	–	–
Then resumption	+	+	–
Negative polarity items/putative *should*/conditional inversion	+	–	+
Embeddable	+	+	–
Dutch: V2 first constituent	+	+	–
Dutch: V2 transgression	–	–	+

4.4 A Ternary Typology

4.4.1 Peripheral Adverbial Clauses vs. Non-Integrated Adverbial Clauses

If the fact that factual conditionals—and peripheral adverbial clauses in general—are embeddable and interact with V2 is taken as evidence that they are part of the narrow syntax then the fact that speech-event conditionals pattern differently on these two scores provides support for the hypothesis that speech-event conditionals are orphan constituents or extra-sentential constituents in the sense of Broekhuis and Corver (2016: 1679–1733): they are not part of the narrow syntax, i.e. that they are not syntactically integrated with the associated clause. Let us, therefore, adopt Frey's (2016, 2018, 2019, 2020) ternary typology of adverbial clauses for the classification of conditional clauses, as shown in (44).

(44) Adopting Frey's typology of adverbial clauses (2016) to conditional clauses:
  (i) Event conditional:
    a type of central adverbial clause (Frey's CAC)
  (ii) Factual conditional:
    a type of peripheral adverbial clause (Frey's PAC)
  (iii) Speech-event modifier:
    a type of non-integrated adverbial clause (Frey's NonIC)

Both CACs (event conditionals) and PACs (factual conditionals) are syntactically integrated, but they differ in terms of their level of attachment: CACs are attached lower in the tree than PACs, because CACs modify VP or TP while PACs modify JP.

Elaborating on Frey's (2016, 2018, 2019, 2020) proposals speech-event conditionals are considered to be non-integrated adverbial clauses, which he labels as 'NonICs'. For a discussion of the discourse syntax of these non-integrated adverbial clauses, see also Greco and Haegeman (2018, 2020) and Schönenberger and Haegeman (to appear), in which the label 'NiC' rather than 'NonIC' is used because of certain differences between Frey's NonICs and their NiCs. For simplicity's sake, we shall adopt the label 'NiC' here.

4.4.2 PAC vs. NiC: Some Problematic Cases

The empirical data and the diagnostics discussed so far have led to a ternary distinction between CAC, PAC and NiC. It has been noted in the literature, however, that some adverbial modifiers, including those that are clausal, may behave either like PACs or like NiCs (see Meinunger 2004 and Frey 2016 for German non-integrated adverbial clauses, see also Sweetser 1990 for English conditionals). The patterns will briefly be illustrated here for Flemish.[11]

In Flemish (45a), the conditional clause *als ge het mij vraagt* ('if you ask me') is extra-sentential: like NiCs it precedes a full-fledged V2 root clause. This would correlate with its status as a speech-event modifier: the conditional clause modifies the speech event itself (which corresponds to the V2 root clause). In (45b), however, what looks like the same conditional clause is the first constituent of a V2 clause, hence, by our reasoning (*pace* Axel and Wöllstein 2009) it is not extra-sentential and it cannot be classified as a NiC. Rather, it must be an integral part of the syntax of the root clause (*pace* Axel and Wöllstein 2009), and thus it would be either a PAC or a CAC. Given its interpretation and the fact that sentential negation cannot scope over it, the PAC label seems the better option. The two co-existing patterns—V2 transgression and V2—suggest that the same conditional clause may behave either as a NiC or as a PAC, and with Frey (2016: 172), we would like to suggest that a subtle difference in interpretation obtains: while the conditional clause in (45a) represents a genuine speech-event modifier, that in (45b) should rather be interpreted as encoding an adverbial modification similar to that encoded in evidential expressions such as 'in my opinion', 'according to me'. Such modifiers may constitute the first constituent in a V2 clause (45c), but they can also be found in middle-field position in Flemish (45d). Crucially, in the latter configuration, they do not require a special parenthetical intonation.

[11] For examples from English in relation to resumptive *then* with conditional clauses (cf. Sect. 4.3.2), we refer the reader to Dancygier and Sweetser (1997: 128f.).

(45) a. Als? ge het mij vraagt,
if you it me ask
het is vandaag niet zo koud als gisteren.
it is today not as cold as yesterday
b. Als2 ge het mij vraagt,
if you it me ask
is het vandaag niet zo koud als gisteren
is it today not as cold as yesterday
'If you ask me, today it is not as cold as yesterday.'
c. Volgens mij is het vandaag niet zo koud als gisteren.
according me is it today not as cold as yesterday
d. Vandaag is het volgens mij niet zo koud als gisteren.
today is it according me not as cold as yesterday
'According to me, today it is not as cold as yesterday.'

Similarly, in (46a) the conditional *als ik het me goed herinner* ('if I remember correctly') precedes a full-fledged V2 clause, it is extra-sentential and patterns with NiCs. It is interpreted as a speech-event modifier. In (46b) the same conditional clause is the first constituent of a V2 clause and we assume that here it will be interpreted as a PAC: the PAC encodes evidential modality, in particular defining the evidential basis on which the propositional content of the clause is founded, i.e. the speaker's own memory.[12]

(46) a. Als? ik het me goed herinner,
if I it me well remember
het boek verscheen voor het eerst in 1982.
the book appeared for the first in 1982
b. Als2 ik het me goed herinner,
if I it me well remember
verscheen het boek voor het eerst in 1982.
appeared the book for the first in 1982
'If my memory serves me right, the book first appeared in 1982.'

[12] It might be objected that the *if*2-clause in (47b) is not echoic in any obvious way. One could perhaps argue that (47b) constitutes a reply to an implicit or explicit question asking whether the speaker remembers the date of publication.

Not all NiCs can double up as PACs: in (47a), based on Te Velde (2013: 22, his (IIg)), the conditional clause *als u het zich herinnert* ('if you remember') patterns as an NiC modifying the speech event, and it cannot also pattern as a PAC, as shown by the fact that the same clause cannot be the first constituent in a V2 pattern (47b). This is to be expected: appealing to the interlocutor's memory ('if you remember') cannot constitute the speaker's own evidential basis for a propositional content, since the interlocutor's memory is inaccessible to him or her.

(47) a. Als? u het zich herinnert,
if you it yourself remember
het boek verscheen voor het eerst in 1982.
the book appeared for the first in 1982
'If you recall, the book appeared for the first time in 1982.'
(cf. Te Velde 2013: 22, his (IIg))

b. *Als2 u het zich herinnert,
if you it yourself remember
verscheen het boek voor het eerst in 1982.
appeared the book for the first in 1982

4.4.3 Speculations on the Ternary Typology

Frey (2018, 2019, 2020) replaces a binary classification of adverbial clauses by a ternary one (see (44)). Many general issues are still to be addressed regarding the classification of adverbial clauses in which non-integration plays a crucial role. Ultimately, however, we would like to envisage that the ternary model can be reframed in terms of a binary one in which integrated clausal modifiers—and indeed other modifiers—are opposed to non-integrated ones.

Integrated clauses vary in relation to the point of integration (see also Endo and Haegeman 2019): for instance, CACs are TP-internal, hence part of the proposition conveyed by the associated clause, and hence potentially at-issue, and PACs are TP-external, thus outside of the at-issue content of the clause. A revised typology of adverbial clauses is suggested in (48), which includes a tentative non-exhaustive list of adverbial modifiers that would be non-integrated and which could be analysed along the lines of recent proposals in Haegeman and Greco (2018, 2020). Further work must clarify the possible internal organization of such non-integrated clauses and shed light on the extent to which these are to be treated in terms of a (discourse) syntactic analysis.

(48) A revised typology of adverbial clauses
  (i) Integrated clauses, differentiated by the height of attachment (Endo and Haegeman 2019)
    CAC TP-internal e.g. temporal *while*, event conditional
    PAC TP-external e.g. concessive *while*, factual conditional
  (ii) Non-integrated clauses (merged in discourse syntax, cf. Greco and Haegeman 2018, 2020)
    - CAC-like speech-event modifiers;
    - PAC-like speech-event modifiers;
    - CAC modifiers of root proposition (Greco and Haegeman 2020) (49a);
    - PAC modifiers of root proposition (not discussed in Greco and Haegeman 2018, 2020) (49b).

(49) a. Oa-me tuskwamen, de deure stond open en de lucht was an.
  when we home came the door was open and the light was on
  'When we arrived home, the door was open and the light was on.'
  (Greco and Haegeman 2020: 78, their (25))
 b. Oat et programma toch goa veranderen,
  if the schedule PART goes change
  ge zou beter ofwachten.
  you should better off-wait
  'If the programme is going to change anyway, then you'd better wait.'

5 Summary

This chapter revisited the typology of conditional clauses as part of the wider typology of adverbial clauses developed in Haegeman (1984a, 1984b, 1984c, 1991/2009, 2003, 2012). Like other adverbial clauses, clauses introduced by the conjunction *if* display (at least) three readings: (i) an event conditional encodes a condition on the event expressed in the main clause; (ii) a factual conditional encodes a background assumption which serves as the basis for the processing of the root proposition; (iii) a speech-event conditional encodes a condition on the speech event. In earlier work by Haegeman (Haegeman and Wekker 1984; Haegeman 1984b, 1991/2009, 2003), two classes of adverbial clauses were distinguished: peripheral adverbial clauses and central adverbial clauses. Both factual conditionals and speech-event conditionals were treated as belonging to the class of peripheral adverbial clauses and event conditionals as belonging to the class of central adverbial clauses.

Based on additional properties of the external syntax of adverbial clauses including the relation to the syntax of V2, the binary distinction between 'central' adverbial clauses and 'peripheral' adverbial clauses was replaced by the ternary distinction argued for in Frey (2016). For the syntactic analysis, both variation in height of insertion as well as the opposition between syntactic integration and non-integration were seen to be relevant, the former pertaining to the difference between event conditionals and factual conditionals, and the latter serving to distinguish between event conditionals and factual conditionals on one hand, and speech-event conditionals on the other. An important area of study for future work is the issue of the formalization of the (discourse-)syntax of non-integrated modifiers.

References

Austin, J.L. 1961/1970. Ifs and cans. In J.O. Urmson and G.J. Warnock (eds.), *Philosophical Papers*, 153–180. Oxford: Oxford University Press. Second edition (1970), 205–232.

Auwera, Johan van der. 1986. Conditionals and speech acts. In Elisabeth A. Traugott, Alice ter Meulen, Judy Snitzer Reilly and Charles A. Ferguson (eds.), *On Conditionals*, 197–214. Cambridge: CUP.

Axel, Katrin and Angelika Wöllstein. 2009. German verb-first conditionals as unintegrated clauses. A case study in converging synchronic and diachronic evidence. In Susanne Winkler and Sam Featherston (eds.), *The Fruits of Empirical Linguistics*, II, 1–37. Berlin, New York: Mouton de Gruyter.

Besten, Hans den. 1977/1983/1989. On the interaction of root transformations and lexical deletive rules. In Werner Abraham (ed.), *On the Formal Syntax of Westgermania*, 47–131. Benjamin: Amsterdam. Reprinted in Hans den Besten 1989, *Studies in Westgermanic syntax*. Doctoral dissertation, Katholieke Universiteit Brabant. Amsterdam and Atlanta, GA: Rodopi.

Biasio, Marco and Enrico Castro. 2019. *Will* in the English protasis: a proposal for motivating the constraint. *Paper presented at the Incontro di Grammatica Generativa*. Università degli Studi di Padova.

Biberauer, Theresa and Ian Roberts. 2017. Conditional inversion and types of parametric change. In Bettelou Los and Pieter de Haan (eds.), *Word Order Change in Acquisition and Language Contact: Essays in Honour of Ans van Kemenade*, 57–77. Amsterdam and Philadelphia: John Benjamins. [Linguistik Aktuell/Linguistics Today, 243].

Broekhuis, Hans and Norbert Corver. 2016. *Syntax of Dutch. Verbs and Verb Phrases. Volume 3: Chapter 14: Main Clause-External Elements*. Amsterdam: Amsterdam University Press.

Catasso, Nicholas. 2015. On post initial *Aber* and other syntactic transgressions: some considerations on the nature of V2 in German. *Journal of Germanic Linguistics* 27(04): 317–365.

Charnavel, Isabelle. 2020. French causal *puisque*-clauses in the light of (not)-at-issueness. In *Romance Languages and Linguistic Theory - Selected Papers from the 47th Linguistic Symposium on Romance Languages* (LSRL47), Newark, Delaware: John Benjamins Publishing Company.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads*. New York and Oxford: Oxford University Press.

Cinque, Guglielmo. 2004. Issues in adverbial syntax. *Lingua* 114: 683–710.

Cinque, Guglielmo. 2008. Two types of non-restrictive relatives. In Olivier Bonami and Patrizia Cabredo Hofherr (eds.), *Empirical Issues in Syntax and Semantics* 7, 99–137. Paris: CNRS.

Cinque, Guglielmo and Luigi Rizzi. 2008. The cartography of syntactic structures. *Studies in Linguistics*, *CISCL working papers*. Vol. 2.

Close, Reginald A. 1980. Will in if-clauses. In Sidney Greenbaum, Geoffrey Leech and Jan Svartvik (eds), *Studies in English Linguistics*, 100–109. London: Longmans.

Comrie, Bernard. 1982. Future time reference in the conditional protasis. *Australian Journal of Linguistics* 22: 143–152.

Coniglio, Marco. 2011. *Die Syntax der deutschen Modalpartikeln: ihre Distribution und Lizenzierung in Haupt- und Nebensätzen*. Berlin: Akademie.

Dancygier, Barbara and Eve Sweetser. 1997. *Then* in conditional constructions. *Cognitive Linguistics* 8(2): 109–136.

Declerck, Renaat. 1984. Pure future will in if clauses. *Lingua* 63: 279–312.

Declerck, Renaat. 1991. *Tense in English*. London and New York: Routledge.

Declerck, Renaat. 2000. On how to interpret canonical conditional clauses in discourse. *Journal of Literary Semantics* 29(3): 149–169.

Declerck, Renaat and Susan Reed. 2001. *Conditionals: A Comprehensive Empirical Analysis*. Berlin: Mouton de Gruyter.

Ebert, Christian, Cornelia Endriss and Stefan Hinterwimmer. 2008. A unified analysis of indicative and biscuit conditionals as topics. In Tova Friedman and Satoshi Ito (eds.), *SALT* XVIII, Ithaca, NY: Cornell University.

Endo, Yokio and Liliane Haegeman. 2019. Adverbial clauses and adverbial concord. *Glossa-A Journal of General Linguistics* 4: 1.

Fabb, Nigel. 1990. The difference between English restrictive and non-restrictive relative clauses. *Journal of Linguistics* 26: 57–79.

Frey, Werner. 2011. Peripheral adverbial clauses, their licensing and the prefield in German. In Eva Breindl, Gisella Ferraresi and Anna Volodina (eds.), *Satzverknüpfung – Zur Interaktion von Form, Bedeutung und Diskursfunktion*, 41–77. Berlin: de Gruyter.

Frey, Werner. 2016. On some correlations between formal and interpretative properties of causal clauses. In Ingo Reich and Augustin Speyer (eds.), *Co- and Subordination in German and Other Languages. Special Issue of Linguistische Berichte* 21: 153–179. Hamburg: Buske.

Frey, Werner. 2018. On the syntax discourse interface with different kinds of not-at-issue expressions. CGSW 33, 27–28 August 2018, Georg-August-Universität Göttingen.

Frey, Werner. 2019. Linking some syntactic and semantic features of concessives and adversatives. *Paper presented at DGfS, AG 5*, Bremen.

Frey, Werner. 2020. On the categorical status of different dependent clauses. Ms.

Frey, Werner and Andre Meinunger. 2019. Topic marking and illocutionary force. In Valéria Molnár, Verner Egerland and Susanne Winkler (eds.), *Architecture of Topic*, 95–138. Berlin: de Gruyter.

Giannakidou, Anastasia. 1998. *Polarity Sensitivity as (Non) Veridical Dependency*. (Linguistics today 23). Amsterdam and Philadelphia: John Benjamins.

Giorgi, Alexandra. 2014. Prosodic signals as syntactic formatives in the left periphery. In Anna Cardinaletti, Guglielmo Cinque and Yoshio Endo (eds.), *On Peripheries, Exploring Clause Initial and Clause Final Positions*, 161–188. Tokyo: Hituzi Syobe Publishing.

Haegeman, Liliane. 1984a. Parasitic gaps and adverbial clauses. *Journal of Linguistics* 20: 229–232.

Haegeman, Liliane. 1984b. Pragmatic conditionals in English. *Folia Linguistica* 18: 485–502.

Haegeman, Liliane. 1984c. Remarks on adverbial clauses and definite anaphora. *Linguistic Inquiry* 15: 712–715.

Haegeman, Liliane. 1991. Parenthetical adverbials: the radical orphanage approach. In Shuki Chiba, Akira Ogawa, Yasuaki Fuiwara, Norio Yamada, Osamu Koma and Takao Yagi (eds.), *Aspects of Modern English Linguistics: Papers presented to Masatomo Ukaji on His 60th Birthday*, 232–254. Tokyo: Kaitakushi.

Haegeman, Liliane. 1996. Verb second, the split CP and null subjects in early Dutch finite clauses. *GenGenP*. Available on: http://ling.auf.net/lingBuzz/001059.

Haegeman, Liliane. 2003. Conditional clauses: external and internal syntax. *Mind and Language* 18: 317–339.

Haegeman, Liliane. 2009. Parenthetical adverbials: the radical orphanage approach. In Benjamin Shaer, Philippa Cook, Werner Frey and Claudia Maienborn (eds.), *Dislocated Elements in Discourse: Syntactic, Semantic and Pragmatic Perspectives*, 331–347. London: Routledge. Reprinted from Haegeman (1991).

Haegeman, Liliane. 2010. The movement derivation of conditional clauses. *Linguistic Inquiry* 41: 595–621.

Haegeman, Liliane. 2012. *Adverbial Clauses, Main Clause Phenomena and the Composition of the Left Periphery*. Oxford: OUP.

Haegeman, Liliane and Brian Robinson. 1979. A note on some aspects of the use of will in affirmative declarative sentences. *Journal of Linguistics* 15: 109–110.

Haegeman, Liliane and Herman Wekker. 1984. The syntax and interpretation of futurate adverbials in English. *Journal of Linguistics* 20: 263–281.

Haegeman, Liliane, Benjamin Shaer and Werner Frey. 2009. Postscript: problems and solutions for orphan analyses. In Philippa Cook, Werner Frey and Claudia Maienborn (eds.), *Dislocated Elements in Discourse: Syntactic, Semantic, and Pragmatic Perspectives*, 348–365. New York: Routledge.

Haegeman, Liliane and Ciro Greco. 2018. West flemish V3 and the interaction of syntax and discourse. *Journal of Comparative Germanic Linguistics* 21(1): 1–56.

Haegeman, Liliane and Ciro Greco. 2020. Frame setters and microvariation of subject-initial verb second. In Rebecca Woods and Sam Wolfe (eds.), *Rethinking Verb Second*, 61–89. New York and Oxford: Oxford University Press.

Holmberg Anders. 2020. On the bottleneck hypothesis of verb second in Swedish. In Rebecca Woods and Sam Wolfe (eds.), *Rethinking Verb Second*, 40–60. Oxford and New York: Oxford University Press.

Hornstein, Norbert. 1990. *As Time Goes By*. Cambridge, MA: The MIT Press.

The Handbook of English, 159–219. Oxford: Blackwell.

Iatridou, Sabine. 1991. *Topics in Conditionals*. Ph.D dissertation. Boston: MIT.

Iatridou, Sabine and David Embick. 1994. Conditional inversion. In Mercè Gonzàlez (ed.), *Proceedings of the North Easter Linguistic Society* 24, 189–203. Amherst: Graduate Linguistics Student Association.

Jespersen, Otto. 1909–49. *A Modern English Grammar*. Heidelberg/Copenhagen/London: Karl Winter/Einar Munksgaard/Allen and Unwin.

Kearns, John T. 2006. Conditional assertion, denial, and supposition as illocutionary acts. *Linguistics and Philosophy* 29(4): 455–485.

Koster, Jan. 2000. Extraposition as parallel construal. Ms. University of Groningen.

Krifka, Manfred. 2017. Assertions and judgements, epistemics and evidentials. *Paper presented at the Workshop Speech Acts: Meanings, Uses, Syntactic and Prosodic Realizations*, Leibniz-ZAS Berlin, May 29–31, 2017. https://amor.cms.hu-berlin.de/~h2816i3x/Talks/CommitmentEpistemicsHandout.pdf.

Krifka, Manfred. To appear. Layers of assertive causes: propositions, judgements, commitments, acts. In Jutta M. Hartmann and Angelika Wöllstein (eds.), *Propositionale Argumente im Sprachvergleich: Theorie und Empirie. Propositional Arguments in Cross-Linguistic Research: Theoretical and Empirical Issues.* [Studien zur Deutschen Sprache]. Tuebingen: Gunter Narr Verlag.

Leech, Geoffrey N. 1971. *Meaning and the English Verb*, xiv + 132. London: Longman (2nd and 3rd editions: 1987, 2004).

McCawley, James D. 1971. Tense and time reference in English. In Charles J. Fillmore and D. Terence Langėndoen (eds.), *Studies in Linguistic Semantics*, 96–113. Irvington: Holt, Rinehart & Winston.

Meinunger, Andre. 2004. Interface restrictions on Verb Second. *Linguistics in Potsdam* 22: 51–81.

Nieuwint, Pieter. 1986. Present and future in conditional protases. *Linguistics* 24: 371–92.

Palmer, Frank. 1965. *A Linguistic Study of the English Verb*. London: Longman.
Palmer, Frank. 1974. *The English Verb*. London: Longman.
Potts, Christopher. 2015. Presupposition and implicature. In Shalom Lappin and Chris Fox (eds.), *Handbook of Contemporary Semantic Theory* (2nd edition), 168–202. Oxford: Wiley/Blackwell.
Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik. 1985. *A Comprehensive Grammar of the English Language*. London: Longman.
Ros, Hilke. 2005. Functional phrase structure: integrating functional ideas in generative adverbial syntax. Ph.D. dissertation, Ghent University.
Rutherford, William. 1970. Some observations concerning subordinate clauses in English. *Language* 46: 97–115.
Safir, Kenneth. 1986. Relative clauses in a theory of binding and levels. *Linguistic Inquiry* 17: 663–690.
Schönenberger, Manuela and Liliane Haegeman. To appear. English rationale *since* and a reassessment of the typology of adverbial clauses. Ms. Ghent University/University of Geneva.
Shaer, Benjamin and Werner Frey 2004. 'Integrated' and 'non-integrated' left peripheral elements in German and English. In Benjamin Shaer, Werner Frey and Claudia Maienborn (eds.), *Proceedings of the Dislocated Elements Workshop*, ZAS Berlin, November 2003. *ZAS Papers in Linguistics* 35(2): 465–502.
Sweetser, Eve E. 1990. *From Etymology to Pragmatics*. Cambridge: Cambridge University Press.
Takami, Ken-ichi. 1988. The syntax of if-clauses: three types of if-clauses and X'-theory. *Lingua* 1988: 263–281.
Velde, John te. 2013. West Germanic left-dislocated adverbial clauses: the role of the semantic interface. *Interdisciplinary Journal for Germanic Linguistics and Semiotic Analysis* 18: 163–206.
Wakker, Gerry. 1996. Conditionals at different levels of the clause. In Betty Devriend, Louis Goossens and Johan van der Auwera (eds.), *Complex Structures*, 177–199. Berlin: Mouton de Gruyter.
Wekker, Herman C. 1976. *The Expression of Future Time in Contemporary British English*. Amsterdam: North Holland.
Wekker, Herman C. 1977. Future reference in adverbial clauses. *Interlanguage Studies Bulletin* 2: 64–77.
Whitman, Neal. 2004. Semantics and pragmatics of English verbal dependent coordination. *Language* 80(3): 403–434. Retrieved August 20, 2020. www.jstor.org/stable/4489719.

Williams, Edwin.1978. Across the board rule application. *Linguistic Inquiry* 9: 31–43.
Woods, Rebecca and Sam Wolfe. 2020. *Rethinking Verb Second*. New York and Oxford: Oxford University Press.
Zandvoort, Reinart Willem. 1975. *A Handbook of English Grammar*. London: Longman.

Toward a Unified Linguistic Approach to Conditionals—Some Empirical Evidence

Ghanshyam Sharma

Abbreviations

ACC	accusative case
ADV	adverb
COND	conditional
CPM	complementizer
DAT	dative case
ERG	ergative case
F	feminine gender
FUT	future tense
GEN	genitive case
HAB	habitual aspect
IMPFV	imperfective aspect

G. Sharma (✉)
Department of South Asian Studies, National Institute of Oriental Languages and Civilisations (INALCO), Paris, France
e-mail: paris.conditionals@gmail.com

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_11

LOC	locative case
M	masculine gender
NOM	nominative case
PART	particle
PFV	perfective aspect, perfect tense
PL	plural number
POSS	possessive case
PROG	progressive aspect
PRON	pronoun
PRS	present tense
SUFF	suffix

1 Introduction

The purpose of this article is to shed light on the characterization of conditionals by looking into their typological linguistic structure. As is well known, a proper understanding of conditional constructions in human languages is essential for a number of disciplines, including logic, linguistics, psychology, cognitive science, artificial intelligence, and others. However, despite their fundamental importance to a range of disciplines and despite being a primarily linguistic phenomenon, conditionals have not attracted as much scholarly attention from linguists as they demand. While there have been some noteworthy typological as well as theoretical studies in the last few decades (Comrie 1986; Wierzbicka 1997; Haegeman 2003; Xrakovskij 2005; Bhatt and Pancheva 2006; Thompson et al. 2007), more linguistic research into conditional constructions is needed. This paper attempts to make a small contribution to filling this gap in the literature. Since the subject itself is extremely large, and the debates surrounding it have been wide-ranging, it will not be possible here to deal thoroughly with every argument or to address all aspects of conditionals. Instead, this paper will limit its scope to pursuing two principal objectives: revisiting Greenberg's Universal 14 and, in the light of recent research, reviewing some of the proposals made by Comrie (1986).

2 The Conditional Construct and Its Syntax

Giving a concise definition of a conditional construct has always been simpler in logic than in linguistics. In extensional logic, the 'material conditional' relation between two propositions that make up a conditional construction, namely $P \supset Q$, is considered the sole criterion for determining a conditional construction and giving an account of its truth conditions. As a result, in extensionally oriented linguistic studies, the concept of the material conditional has been found applicable to the study of natural language conditionals (Williamson 2020), and is considered to be synonymous with an *if–then* construction.[1] As we shall see later, however, there are languages in which the two clauses are not always marked with any specific morphological devices comparable to English *if* and *then* markers. Furthermore, since most of the foundational research into conditional constructions has been carried out using examples drawn from the English language, our understanding of the complex phenomenon of conditionals is largely dependent on the way conditionals are structured in this particular language. Obviously, any generalizations based solely on the conditional constructs of English will inevitably be limited.

Acknowledging the difficulty of describing a conditional construction in languages other than English, Wierzbicka (1997) comes up with a seemingly workable solution. She claims that the construct introduced by *if* is a primitive lexico-grammatical universal and 'is one of those relatively simple and clear concepts which cannot be made clearer by decomposing them into simpler concepts' (1997: 15). Furthermore,

[1] On the viability of using the term 'material conditional' in natural language, Barwise (1986: 21) has the following to say: "For those of us involved in the attempt to spell out the relation between statements and those aspects of reality they are about, conditionals are a thorny issue. Within this semantic tradition, common wisdom can be summarized rather contentiously as follows: classical model theory gives us the semantics of the material conditional. It works fine for mathematical conditionals, but is a disaster if applied to ordinary language conditionals, especially counterfactual conditionals. Within the possible worlds framework, there are various treatments, some of which are quite successful for certain types of natural language conditionals, including counterfactuals, but they are all a disaster when applied to mathematical conditionals."

defining the concept is of no great importance, according to Wierzbicka, as 'there is little point in trying to define simple concepts (such as, for example, WANT, THINK, KNOW or SEE) in terms of more complex ones (such as, for example, "volition", "deontic modality", "cognition", "epistemic modality", "information", "vision" and so on). Similarly, there is little point in trying to define *if* in terms of more complex concepts such as "hypothetical", "inference", or "possible worlds" (1997: 17). Although Wierzbicka's claim might seem to pinpoint the difficulty in defining a conditional construct, it needs to be substantiated by more empirical data and further research. Following Wierzbicka's line of argument, one is confronted with a rather obvious question: assuming that *if* belongs to a class of lexico-grammatical universal primitives, why do many languages not exhibit conditionality through a specific marker similar to *if*? Undoubtedly, conditionality as an abstract concept is a part of human reasoning and can thus be considered a universal phenomenon. However, more empirical research is needed to establish how primitive lexico-grammatical universal concepts are displayed cross-linguistically. In the context of Wierzbicka's project—namely, paraphrasing the object language in the object language itself without any recourse to a metalanguage or truth conditions—it makes sense to say that IF cannot be reduced to anything simpler. Regrettably, as things currently stand, her proposal to abandon any sophisticated terminology, such as volition, epistemic modality, deontic modality, etc., and adopt instead her rather embryonic universal primitive lexico-grammatical category IF, does not lead to any significant insight into the complexity of conditional constructions.

With respect to the syntax of conditionals, Bhatt and Pancheva (2006) present a detailed summary of the different proposals for describing the conditional clause found in modern syntactic theory. These include regarding the conditional clause as an adverbial, an interrogative or a correlative. Bhatt and Pancheva (2006: 640) seem to hold to the first of these proposals and maintain that a conditional clause (i.e. an *if*-clause) is similar to an adverbial clause:

> Conditional structures involve an adverbial clause, often referred to as the CONDITIONAL CLAUSE, ANTECEDENT or PROTASIS (the underlined constituent in (1)), and a main clause, known as the CONSEQUENT or APODOSIS. Conditional structures are interpreted, in general terms, with the proposition expressed by the antecedent clause specifying the (modal) circumstances in which the proposition expressed by the main clause is true. Thus, (1) states that the possible worlds/situations in which Andrea arrives late (the denotation of the conditional clause) are possible worlds/situations in which Clara gets upset (the denotation of the main clause).
>
> (1) If Andrea arrives late, Clara will get upset.
>
> ...
>
> Conditionals are not unique in their overall structure, rather conditional clauses belong to a class of adverbial clauses that includes, among others, clausal adverbials of time, cause, and concession, as illustrated in (3).
>
> (3) a. If Andrea arrived late, Clara must have gotten upset.
> b. When Andrea arrived late, Clara got upset.
> c. Because Andrea arrived late, Clara got upset.
> d. Although Andrea arrived on time, Clara got upset.
>
> Like the other clausal adverbials, conditional clauses are typically introduced by a CP-related element, a complementizer or an operator in Spec, CP (cf. *if*, *when*, *because*, *although* in (3) above). And like the other adverbial clauses, conditional clauses may precede or follow the main clause.

The idea that the *if*-clause is an 'adverbial' is fully developed and forcefully argued in Haegeman and Schönenberger (this volume), who present the typology of conditional clauses as part of the wider typology of adverbial clauses. They maintain that: "like other adverbial clauses, clauses introduced by the conjunction *if* display (at least) three readings: (i) an event conditional encodes a condition on the event expressed in the main clause; (ii) a factual conditional encodes a background assumption which serves as the basis for the processing of the root proposition; (iii) a speech-event conditional encodes a condition on the speech event."

While the conditional clause may appear to be similar to adverbial clauses syntactically—and thus thought to be comparable with adverbials of time, cause or concession in presenting a uniform syntactic analysis—a closer examination reveals that from a semantic point of view these different types of clauses do not carry the same epistemic stance that a speaker adopts when making an utterance using them. Thus, an *if*-clause cannot be equated semantically with adverbials of time, cause or concession, as the semantic contribution it makes differs from the contribution made by these adverbials.

To highlight the difference between an *if*-clause and other types of adverbials, let us consider the epistemic stance adopted by the speaker in uttering each of the propositions listed above in the quotation from Bhatt and Pancheva, along with an additional counterfactual clause:

Table 1 Speaker's epistemic stance

Subordinate proposition P	Clause type	Speaker's epistemic stance
a. When Andrea arrived late,	Time adverbial	$K_S \Box P$
b. Because Andrea arrived late,	Cause adverbial	$K_S \Box P$
c. Although Andrea arrived on time,	Concession adverbial	$K_S \Box P$
d. If Andrea arrived late,	*If*-clause	$\neg K_S\ P \vee \neg B_S P$
e. If Andrea had arrived late	*If*-clause (counterfactual)	$K_S \neg P \vee B_S \neg P$

P stands for the proposition "Andrea arrived late"

As Table 1 clearly shows, in uttering *When Andrea arrived late* …, *Because Andrea arrived late* … or *Although Andrea arrived late* …, the speaker adopts an epistemic stance of certainty (i.e. for all the speaker knows it is necessarily P), thus indicating that the speaker knows Andrea arrived late, whereas in uttering *If Andrea arrived late* …, the speaker neither knows nor believes that Andrea arrived late. Thus, the truth-condition of Andrea arriving late in *If Andrea arrived late* has to be ascertained in an alternative world, which is introduced by the protasis marker *if*. When it comes to counterfactuals, the speaker's epistemic stance is quite the opposite. In uttering *If Andrea had arrived late*, both the speaker and the addressee either know or believe for certain that Andrea did not arrive late. Whether this distinctiveness of the *if*-clauses can have repercussions for the syntactic analysis of a conditional statement is a complex subject worthy of further discussion. Nevertheless, we believe that the difference between the factual nature of time, cause or concession adverbials and the non-factual nature of *if*-clause cannot be considered solely as a semantic issue, because it is also of relevance to syntactic theory.

Another view on the conditional clause worth mentioning—and the one we consider the most promising for a comprehensive cross-linguistic survey of conditional clauses—is that advanced by Bhatt and Pancheva (2006: 639). This sees conditional clauses as: "… essentially free relatives of possible worlds. Similarly to the more familiar instances of free relatives of individuals, (i) conditional clauses likely involve clause-internal operator-movement to Spec, CP; (ii) they receive the interpretation of definite descriptions; and (iii) they may participate in correlative structures, as happens in the subcase of conditionals with the proform *then*."

According to this line of research, just like correlative constructions, conditional constructions involve a free relative clause adjoined to the matrix clause and co-indexed with a proform inside it: [free relative]$_i$ [... proform$_i$...]. In fact, in many languages, particularly in many South Asian languages, conditional constructions are to some extent comparable to correlative constructions. It has also been claimed that in some languages, the *if*-clauses are historically derived from correlative constructions:

> Our proposal that if-clauses are free relatives, i.e., definite descriptions of possible worlds, naturally predicts that they should be able to appear in the correlative construction. Geis (1985), von Fintel (1994), Izvorski (1997) among others have suggested that conditional constructions are related to correlatives. Geis was perhaps the first to note that conditional

> constructions in English are the remnants of a strategy of correlativization that was once more productive in the language. Treating some conditionals as correlatives helps us to understand several aspects of the behavior of conditionals crosslinguistically. In languages where correlativization is a productive strategy, it is apparent that conditionals are correlatives (e.g. Marathi). (Bhatt and Pancheva 2006: 661)

The literature has made it abundantly clear that providing an exhaustive definition of conditionals which would encompass all the divergent syntactic characteristics that are attested cross-linguistically remains a challenging task. Moreover, this is a task that requires research based on empirical evidence, rather than pure theoretical discussion. Hoping that further research into the syntax of conditional constructions and their detailed cross-linguistic survey will eventually reveal some hitherto unknown characteristics of conditionals, for the moment, we have no option but to accept the general description of conditionals given by Traugott et al., who write as follows: "Conditional (*if*–then) constructions directly reflect the characteristically human ability to reason about alternative situations, to make inferences based on incomplete information, to imagine possible correlations between situations, and to understand how the world would change if certain correlations were different" (1986: 3).

3 Clause Order and Clause Marker

A conditional statement is made up of two clauses, namely the protasis (also known as the conditional clause, subordinate clause, P-clause, *if*-clause or antecedent) and the apodosis (also known as the conclusion, principal or main clause, Q-clause, *then*-clause or consequent).[2]

[2] In line with the linguistics tradition, I have chosen to use the terms 'protasis' and 'apodosis' here, although I believe the terms 'antecedent' and 'consequent' are a more appropriate way of denoting the two clauses in a conditional construct.

Concerning the ordering of these two clauses, Greenberg's Universal of Word Order 14 states the following:

> In conditional statements, the conditional clause precedes the conclusion as the normal order in all languages. (1963: 84)

Greenberg's Universal 14 is a well-established and widely accepted syntactic principle. The idea that the protasis-apodosis clause order is the universal order was examined in detail by Lehman (1974), who reported that no empirical data could be found to call into question the normal protasis-apodosis clause order. He concluded that whenever apodosis-protasis ordering does occur, it is to be considered either non-normal or highlighted for some reason.

Protasis-apodosis order is congruent with the speaker's communicative strategy, whereby the speaker both engages the addressee in contemplating a potential disjunction and also uses this potential disjunction as the grounds for developing an argument. Furthermore, protasis-apodosis ordering resembles the order of human reasoning and shows 'parallels between order of elements in language and order of elements in experience' (Traugott et al. 1986: 9). The protasis-apodosis order is also supported by Haiman (1978), who claims that a conditional clause (i.e. protasis) shares the typical properties of topic elements found in many languages and is thus placed first.

Granted the foregoing statement, broadening our knowledge of conditional clauses requires a cross-linguistic typological study in which this widely accepted generalization can be scrutinized. Greenberg's Universal establishes that the protasis-apodosis is the normal order; however, this suggests that apodosis-first and protasis-second is possible. In the following paragraphs, I will argue that, in a conditional statement, protasis-first and apodosis-second is the only possible order. I will further maintain that the seeming apodosis-protasis order purportedly attested in various languages is due to the phenomenon of 'fronting' the proposition contained in the apodosis. In other words, it is the proposition contained in the apodosis which is pre-posed or fronted, not the entire apodosis, since no language to our knowledge exhibits a "marked apodosis-marked protasis" order. Furthermore, there are languages such as Hindi in which the protasis is optionally marked and the apodosis is obligatorily marked, but the so-called apodosis marker does not move with the apodosis when it is fronted. This peculiarity reveals the need for further investigation into the Hindi conditional constructions.

Our subsidiary claim is that the placing of the protasis in the second position is not due to an afterthought on the part of speaker, as suggested by Comrie (1986), who says, 'Given that it seems to be commoner cross-linguistically for the protasis to be marked overtly as non-factual than for the apodosis to be so marked …, placing the overtly marked protasis in front of the unmarked apodosis avoids the apodosis being interpreted as a factual statement' (Comrie 1986: 84). As I maintain elsewhere (Sharma 2011), contrary to widely held belief, no language actually seems to exhibit a marked apodosis in the sentence-initial position. In other words, there is no evidence of any sort to demonstrate the apodosis-protasis ordering in conditionals. This finding may have repercussions for syntactic theories that classify the protasis under the blanket term *adverbials*.

3.1 Markers of Protasis and Apodosis

The conventional wisdom is that one or both of the two clauses that make up a conditional construction are marked either overtly (through a separate morphological device) or covertly (through special verb forms). Furthermore, it is also believed that the marking of the conditional clauses may be obligatory either for both clauses or for one clause only, as is the case in English. To have cross-linguistic validity, however, a unified theory of conditionals has to develop tools to analyze data from as many divergent languages as possible, and necessarily from languages that are structurally different from well-studied languages such as English. Keeping in mind this necessity, we can postulate different types of clause combinations to cover some, if not all, of the possible scenarios (Sharma 2010, 2011):

Table 2 Types of marked clauses in conditionals

1. Non-overtly marked P + Non-overtly marked Q	–
2. Obligatorily marked P + Obligatorily marked Q	such as Ngiyambā
3. Obligatorily marked P + Non-obligatorily marked Q	such as English, French and Italian.
4. Non-obligatorily marked P + Obligatorily marked Q	such as Hindi
5. Obligatorily marked P + Non-overtly marked Q	such as Chinese
6. Non-obligatorily marked P + Non-overtly marked Q	–
7. Morphologically marked P + Unmarked Q	such as Tamil, Telugu and Kannada.
8. Non-overtly marked P + Obligatorily marked Q	–

As can be seen from Table 2, different types of possible clause markings can be hypothesized. A unified approach to conditional clause markings has to take into account this cross-linguistic diversity—a task which requires a detailed survey of languages belonging to divergent families. However, to offer a glimpse of the diversity involved here, let us consider five cases that have attracted major attention already: CLASS I: Ngiyambā; CLASS II: English, French, Italian, etc.; CLASS III: Hindi, among others; CLASS IV: Chinese, etc.; and CLASS V: Tamil, Telugu, Kannada, Malayalam, etc.

3.1.1 Class I: Overtly and Obligatorily Marked P + Overtly and Obligatorily Marked Q

Ngiyampā (or Ngiyambā)—an Australian aboriginal language—is reported to belong to this peculiar class in which there is an overt marking of both clauses. Furthermore, both clauses in this language are said to be marked by the same clitic, *-ma*. Given this phenomenon, it is not clear how a protasis and an apodosis can be identified independently. From the literature, all we know is that a conditional sentence has a rigid clause order with no possibility of clause inversion, and that the first clause of a conditional sentence is considered the protasis. Comrie (1986: 84) cites the following example from Ngiyambā (referencing Donaldson 1980: 251–252), observing that 'in Ngiyambā, with past tense counterfactuals, both clauses have the same overt marking (with the clitic *-ma*), and the first must be interpreted as protasis …':

(1) Nginuu-*ma*-ni burray giyi, ngindu-*ma*-ni yada gurawiyi
Lit. 'your-counterfactual-this child was, you-counterfactual-this well looked-after'
'If this child had been yours, you would have looked after it well.'

I believe further research is needed to establish whether the so-called past tense counterfactual morpheme attested in the protasis and the apodosis in (1) does, in fact, mark both the protasis and the apodosis, rather than perform different functions in different contexts. A plausible hypothesis is that the *-ma* particle is a marker of counterfactuality rather than a marker of both protasis and apodosis. In fact, marking counterfactuality through the same morphological device is a widespread phenomenon. It also remains to be seen whether this phenomenon is limited solely to counterfactuals, or whether it is found as well in other types of conditionals in this language. However, the essential point is that there are languages such as Ngiyampā which are said to exhibit overt clause marking of both clauses in counterfactuals but which do not allow a clause inversion. This supports the claim that Q-P order is not possible.

3.1.2 Class II: Overtly and Obligatorily Marked P + Overtly But Not Obligatorily Marked Q (The dubious nature of the English 'then')

English, French and Italian, among many other European languages, belong to this class. It is the most investigated class of languages in which only the protasis is believed to be obligatorily marked. The apodosis is thought to be marked by an optional marker. Since our understanding of conditionals in human language is shaped mainly by the results obtained from analyses of conditional constructions attested in this class of languages, particularly English, it is important to look closely into the structural properties of conditionals in languages belonging to this class. The data from English, French and Italian, for example, clearly suggest an obligatory marking of the protasis, which is obtained through *if*, *si* and *se*, respectively:

(2) a. *If* Mary invites John, he will go to her party.
b. *Si* Mary invite John, il ira à sa fête.
c. *Se* Mary invita John, lui andrà alla sua festa.

As can be observed in (2), the presence of the protasis marker *If*, *Si* and *Se* in the three examples from English, French and Italian is obligatory. All three examples are grammatically correct without their respective apodosis markers *then*, *alors* and *allora*, but not without their respective protasis markers.

With regard to the former, the optional marking has been widely discussed. In particular, there has been a long debate on the real contribution of the so-called English apodosis marker *then* in a conditional sentence (Geis and Zwicky 1971; Iatridou 1994; Dancygier and Sweetser 1997; van der Auwera 1997; Horn 2000; Cariani and Rips this volume). It has been argued that the English marker *then* carries a bi-conditionality meaning which is derived from the pragmatic scalarity in the protasis. For example:

(3) a. If you mow my lawn, I'll pay you ten dollars.
b. If you mow my lawn, *then* I'll pay you ten dollars.

The example in (3b), according to this line of research, means that ten dollars will be paid *if and only if* the lawn is mowed. Given that the purported English apodosis marker *then* carries a bi-conditionality meaning derived from the pragmatic scalarity in the protasis, a plausible explanation of *then* is that it is associated with the protasis rather than with the apodosis. According to this proposal, the so-called apodosis marker *then* is, in fact, a pragmatic marker which induces implicatures, giving rise to bi-conditional readings, as discussed by the abovementioned scholars. To show the pragmatic affiliation of *then* with the protasis, we can roughly present (3a) and (3b) in the following manner:

(4) a. [If you mow my lawn], I'll pay you five dollars.
b. [If you mow my lawn then], I'll pay you five dollars.

As (4b) shows, "*If*" and "*then*" belong together, which supports the proposal that *then* is a pragmatic marker, as it seems to belong to the protasis not to the apodosis. In order to investigate this phenomenon further, let us consider some other characteristics of the English *then*.

First, as Bhatt and Pancheva (2006) have noted, *then* has to be adjacent to the protasis, as in (5a):

(5) a. If it rains, then I think that we should stay at home.
b. *If it rains, I think that then we should stay at home.[3]

In syntactic terms, it has been argued that the surface location of *then* marks a predicate that combines with the *if*-clause, and therefore that *then* must be structurally adjacent to that clause. However, the fact that *then* in such circumstances has to be adjacent to the protasis also proves that it is semantically and pragmatically associated with the protasis rather than with the apodosis, because it has to pick out the pragmatic scalarity meaning from the protasis.

Secondly, there is a restriction on the use of *then* in those situations in which the protasis contains pragmatic elements expressing other pragmatic scales. In fact, itself being a pragmatic scalarity marker, *then* conflicts with other scalarity markers such as *even if* and *only if*, as examples (6) and (7) clearly show:

(6) Even if it rains, (*then) the football game will happen.

(7) Only if it is sunny, (*then) will I visit you.

Thirdly, Iatridou (1994) and Dancygier & Sweetser (1997) have variedly argued that there is a restriction on the use of *then* when the protasis includes a reference to a generic time or event, as in (8):

(8) If Mary bakes a cake, (*then) she gives some slices of it to John.

The ungrammaticality of (8) derives from the fact that the protasis does not exhibit a definite pronominal anaphora which *then* could pick up. Instead, it has a generic pronominal reference *at all times* or *whenever*, which is not compatible with *then.*

Yet another restriction on *then* is ascribed to von Fintel's observation concerning its incompatibility with *unless* (reported by Bhatt and Pancheva 2006), as in (9):

(9) Unless it rains tomorrow, (# then) I won't leave.

The ongoing discussion makes it clear that, in English, *then* is a pragmatic marker (or discourse marker) rather than a logical connective or marker of the apodosis. Its presence in a conditional statement seems to highlight the speaker's attitude to the protasis and the fact that the speaker is prepared to assert what is coming next on the basis of whatever it is that *then* is referring back to earlier in the sentence. Moreover, whenever *then* occurs alone, its role is to invoke the protasis in discourse:

(10) A: Did you know that Oswald would be coming to the party tomorrow.
B: *Then*, I won't be.
= If Oswald is coming to the party tomorrow, *then* I won't be coming.

In sum, any cross-linguisticresearch into conditional constructions must keep in mind this peculiarity of English conditionals and not immediately go in search of equivalent apodosis markers in other languages, since the presence and contribution of the so-called marker of apodosis *then* in English—and probably in all languages belonging to this class of languages—remains highly dubious. As we will see in Sect. 3.1.3, the same is true of the supposed apodosis marker *to* in Hindi which is an integral part of the protasis (or antecedent). For this reason, we presume that an apodosis (or consequent) is never marked in any language, even though an apodosis is commonly called "*then*-clause".

3.1.3 Class III: Overtly but Not Obligatorily Marked P + Overtly and Obligatorily Marked Q (The Supposed Hindi Apodosis Marker *to*)

Now, let us focus on the class of languages in which the protasis can be marked (optional marking), but the apodosis has to be marked (obligatory marking). Hindi belongs to this class of languages. The use of the so-called Hindi apodosis marker *to* (then) is obligatory in a conditional construction, regardless of the presence or absence of the protasis marker *agar/yadi* (*if*), as can be observed in Table 3:

Table 3 Grammaticality test of presence or absence of P and Q markers

		If Ram comes/came, I will/would ask him						
		agar	*Rām*	*āyā,*	*to*	$m\tilde{\varepsilon}$	*us* = *se*	*pūchū̃gā*
		if	Ram	come.pfv.m.sg	then	I	he-abl	ask-fut.m.sg
(a)	√	P-marker			Q-marker			
(b)	√	Ø			Q-marker			
(c)	?	P-marker			Ø			
(d)	*	Ø			Ø			

Ø stands for a null marker

It is notable that, in the case of clause inversion, i.e. apodosis-protasis order, the marker *to* (then) follows the protasis even when the apodosis has to dislocate due to fronting of the proposition, as can be observed in Table 4. Consequently, absence or displacement of *to* (then) renders a Hindi conditional either ungrammatical or semantically odd, as can be observed in (b), (c) and (d) in Table 4:

In a nutshell, as Tables 3 and 4 show, the conditional constructions in Hindi exhibit a peculiarity with respect to the use of the so-called apodosis marker, in that it is required even when the proposition contained in the apodosis, for whatever reason, has to be pre-posed or fronted. Table 5 presents the grammaticality test of Hindi conditional constructions, as outlined above. Notice that the Hindi Q-marker—whose presence is obligatory—does not dislocate with Q when it is pre-posed, as (f) in Table 5 clearly shows:

The unique linguistic characteristic of Hindi conditional sentences requires further inquiry. Let us consider the so-called 'biscuit' conditionals (11), imperative conditionals (12) and interrogative conditionals (13) in Hindi. Notice that, in the case of clause inversion, the tendency to place the particle *to* (then) at the end of the conditional sentence is prevalent across the language:

(11) a. Biscuit (or relevance) conditionals P-Q

(agar) āp=ko bhūkh lagī ho ***to,*** *biskuṭ ālmārī=mẽ*
if you=ACC hunger felt be-SUB then biscuits sideboard=in
rakhe hẽ
placed are
'If you are hungry, there are biscuits in the sideboard.'

b. Biscuit (or relevance) conditionals Q-P

biskuṭ ālmārī=mẽ rakhe hẽ agar
biscuits sideboard=in placed are if
āp=ko bhūkh lagī ho ***to***
you-ACC hunger felt be-SUB then
'There are biscuits in the sideboard if you are hungry.'

Table 4 Markers in an inverted ordering of clauses

Q-P				I will/would ask Ram if he comes/came					
(a)	√		*mɛ̃*	*Rām = se*	*pūchū̃gā*	*agar*	*vo*	*āyā*	*to*
			I	Ram-abl	ask-fut.m.sg	P-marker	he	come-pfv.m.sg	Q-marker
(b)	*	Q-marker				P-marker			
(c)	*	Q-marker				Ø			
(d)	*	Ø				Ø			

* stands for ungrammatical sentence
Ø is a null marker

Table 5 Grammaticality test of Hindi conditional constructions

	P-clause		Q-clause	Marker	Acceptability
(a)	Marked P	-	Marked Q		√
(b)	Unmarked P	-	Marked Q		√
(c)	Marked P	-	Unmarked Q		*
(d)	Unmarked P	-	Unmarked Q		*
	Q-clause		P-clause		
(e)	Unmarked Q	-	Marked P		?
(f)	Unmarked Q	-	Marked P	**Q-marker**	√
(g)	Marked Q	-	Marked P		*
(h)	Marked Q	-	Unmarked P		*
(i)	Unmarked Q	-	Unmarked P		*

(12) a. Imperative conditionals: P-Q
agar tum he apnī jān bacānī ho ***to*** *bhāgo yahā̃=se*
if to you self's life save be-SUBJ then run away here =from
'If you want to save your life, run away from here!'
b. Imperative conditionals: Q-P
bhāgo yahā̃=se agar tumhe apnī jān bacānī ho ***to***
run away here from if to you self's life save be-SUBJ then
'Run away from here if you want to save your life!'

(13) a. Interrogative conditionals: P-Q
agar mɛrī=ne bulāyā ***to*** *kyā tum uskī pārṭī=me jāoge*
if Mary=ERG invited then WH you her party=in will go
'If Mary invited you, would you go to her party?'
b. Interrogative conditionals: Q-P
kyā tum mɛrī=kī pārṭī=me jāoge agar us=ne bulāyā ***to***
WH you her party=in will go if she- ERG invited then
'Would you go to Mary's party if she invited you?'

It is interesting that this tendency seems to hold also in those Hindi conditionals containing an element of pragmatic scalarity, although in Q-P ordering there are issues concerning the acceptability of the conditional sentences. For example, in protasis-apodosis order, (14a) and (15a), the apodosis marker *to* 'then' behaves normally, whereas in apodosis-protasis order, (14b) and (15b), the presence of the marker *to* 'then' seems to be in conflict with the element of pragmatic scalarity, leaving the Hindi sentences grammatically unacceptable.

(14) a. *Only if* Conditionals: P-Q

(agar)	*mɛrī*	*john=ko*	*bulāe*	*to*	*hī*
if	Mary	John=ACC	invites	then	only
vo	*uskī*	*pārṭī=mẽ*	*jāegā*		
he	her	party=in	will go		

'Only if Mary invites John, will he go to the party.'

b. *Only if* Conditionals: Q-P

?john	*pārṭī=mẽ*	*jāegā*	agar	*mɛrī*	*use*
John	party=in	will go	if	Mary	him
bulāye	*to*	*hī*			
invites	then	only			

'John will go to the party only if Mary invites him.'

(15) a. *Even if* Conditionals: P-Q

(agar)	*mɛrī*	*john=ko*	*bulāe*	*to*	*bhī*
if	Mary	John= ACC	invite-SUBJ	then	even
john	*pārṭī =mẽ*	*nahī̃*	*jāegā*		
John	party=in	not	will/would go		

'Even if Mary invites/invited John, he will/would not go to the party.'

b. *Even if* Conditionals: Q-P

?john	*pārṭī =mẽ*	*nahī̃*	*jāegā*	(agar)	*mɛrī*	*use*
John	party=in	not	will go	(if)	Mary	him
bulāye	*to*	*bhī*				
invite-SUBJ	then	even				

'John will not go to the party even if Mary invites him.'

Examples (11) through (15) suggest that the so-called Hindi apodosis marker *to* (then) is a peculiar case. Even though it looks like a marker of apodosis, it does not move leftward together with the apodosis when the latter is fronted or pre-posed. This strongly suggests the hypothesis that it is only the proposition contained in the apodosis which is fronted, not the entire apodosis. Another hypothesis worth examining is that the Hindi apodosis marker *to* has a dual role to play in the language: at times, as a PROFORM, it plays the role of a connective, but it can also be

Table 6 Clause marking in Chinese

If Zangsan drinks wine, (then) I will scold him									
	acceptability	(*rúguŏ*) if	*Zhangsān* Zhangsan	*hē* drinks	*jǐu* wine	*wŏ* I	(*jìu*) will	*mà* scold	*tā* him
a	√	*rúguŏ*	*Zhangsān*	*hē*	*jǐu*	*wŏ*	*jìu*	*mà*	*tā*
b	√	*rúguŏ*	*Zhangsān*	*hē*	*jǐu*	*wŏ*	Ø	*mà*	*tā*
c	√	Ø	*Zhangsān*	*hē*	*jǐu*	*wŏ*	*jìu*	*mà*	*tā*
d	√	Ø	*Zhangsān*	*hē*	*jǐu*	*wŏ*	Ø	*mà*	*tā*

Ø stands for a null marker

employed as a marker of pragmatic scalarity. Obviously, there is a strong need for more in-depth research into the role played by the so-called Hindi apodosis marker *to* ‘then’.

3.1.4 Class IV: Overtly but Not Obligatorily Marked P + Overtly but Not Obligatorily Marked Q

Mandarin is said to overtly, but not obligatorily, mark both protasis and apodosis. Although the marking of two clauses in a conditional statement in Mandarin is not obligatory, their ordering is obligatorily fixed. In other words, the first clause is always protasis whether it is marked or not. Keeping in mind this peculiarity, Comrie (1986: 85) asserts that ‘the protasis necessarily precedes the apodosis, whether the protasis alone is marked for non-factuality (by a conjunction such as *rúguŏ* ‘if’), whether the apodosis alone is marked (for instance by *nà* and/or *jìu* ‘then, in that case’), whether both are marked, or whether neither is marked’.

This distinctiveness of Mandarin conditional constructions is illustrated by the sentence in Table 6, a widely cited example of a construction that is correct in all circumstances.

However, according to our informant,[3] the so-called apodosis marker in Mandarin, namely *jìu*, might mean different things according to context. This is because *jìu* can play different grammatical or semantic roles in the Chinese language. To put it simply, it is questionable

[3] Luo Yujia, a native speaker of Chinese and a doctoral candidate at INALCO, Paris.

to call it an apodosis marker. Having said that, we believe that the case of Mandarin lends even stronger support to our hypothesis that protasis-apodosis ordering is the only ordering in conditional statements, regardless of the presence or absence of any overt clause markers.

3.1.5 Class V: Protasis Marked through a Special Morphological Form of the Verb + Apodosis Remaining Unmarked

Languages belonging to this class express conditionality not through any distinct marker of protasis or apodosis, but rather through a special verb form contained in the protasis. Many Dravidian languages spoken in India, such as Tamil, Malayalam, Telugu and Kannada, apparently belong to this class, as can be seen from the examples (16)–(17) below (Bhatt 1999). The apodosis seems to remain unmarked except for the case of counterfactuals in Telugu, where the apodosis also exhibits a morphological mark of conditionality in the verb, as can be seen in (17c). This is a highly complex issue worthy of a detailed survey.

(16) If Mary invites him, John will go to her party.

a. Tamil

Mary	*avan-ai*	*azhai-t-āl*	*John*
Mary.3F	he-ACC	invite-COND	John-3M

aval	*parti-kku*	*po-v-ān*
her	party-LOC	go-FUT.3M

b. Malayalam

Mary	*jōṇi-ṉe*	*kṣaṇicc-āl*	*avaṉ*	*avaḷ-uṭe*	*pāṟṭṭi-kkə*	*pō-(k)um*
Mary	John-ACC	invite-COND	he	she-GEN	party-DAT	go-FUT

c. Telugu

Mary	*pilus-te*	*John*	*parti*	*ki*	*veL-tā-Du*
Mary	call-COND	John	party	DAT	go-FUT-PRON.SUFF

d. Kannada

Mary	avan-annu	kareda-are,	*John-n-u*
Mary.NOM	he-ACC	invite-PRS-COND	John-Mas-NOM

Ava-L a	*parti-ge*	*hoog-utt-āne*
she-f-POSS	party-DAT	go-PRS.3SM

(17) If Mary had invited him, John would have gone to her party.

a. Tamil

*mēri-v-in avan-ai azhai-tt-iru-**nt-āl,***
Mary.3SF.GEN he-ACC invite-ADV.PART-BE-PAST-COND
John aval parti-kku po-y-iru-nth-iru-pp-ān
John her party-LOC go-BE.PAST-BE.FUT.3M

b. Malayalam

*Maṛy joṇi-ṉe kṣaṇicc-iru-**nneŋkil** avaṉ*
Mary John-ACC invite-be-COND he
avaḷ-uṭe pāṛṭṭi-kkə pō-(k)um-ā(y)-irunnu
she-GEN party-DAT go-FUT-be-PFV

c. Telugu

mēri pilic- i un-Tee John parti ki
Mary call-CPM be-COND John party DAT
veLL-I un-De vāDu
go- CPM be-COND PRON Copy

d. Kannada

mēri avan-annu kared-u-idda-are,
Mary.NOM he-ACC invite-PTCPL-was-COND
John-n-u ava-L-a parti-ge hoog-utt-idd-anu
John-Mas-NOM she-F-POSS party-DAT go-PROG-was-3 SM

3.2 Summary

Table 7 sums up our discussion of the language classes that exhibit different possible orderings of the two clauses in a conditional statement.

Table 7 The five classes of conditionals discussed above

	P-Q	
Class 1	morphologically marked P and morphologically marked Q	Ngiyampā
Class 2	obligatory P-marker and not obligatory Q-marker	English
Class 3	not obligatory P-marker and obligatory Q-marker	Hindi
Class 4	not obligatory P-marker and not obligatory Q-marker	Chinese
Class 5	morphologically marked P and no overt marking in Q	Tamil

As Table 7 shows, languages belonging to classes 2 and 3 seem to exhibit the so-called apodosis-protasis (Q-P) ordering. In fact, data from these languages have led linguists to believe that a conditional statement may exhibit both clause orderings (i.e. P-Q and Q-P). However, as we have argued above, these cases in no way violate the universal of conditional clause ordering (namely P-Q), since in cases of Q-P ordering, only the proposition contained in the apodosis is pre-posed. A null marker at the end of the conditional string in class 2 effectively indicates that in these languages, whenever the Q-proposition is pre-posed, it occurs as a pragmatic strategy of the speaker. Furthermore, data from the languages belonging to class 3, such as Hindi, indubitably demonstrate that, in these languages, the so-called apodosis marker, for example, *to* 'then', cannot be omitted and, except for a few cases, must remain at the end of the conditional string, even in those cases where the apodosis-proposition has to be pre-posed or fronted. We believe that the pre-posing of the apodosis-proposition attested to in languages belonging to classes 2 and 3 (see Table 7) requires further empirical research to be better understood.

One of the reasons for Q-proposition pre-posing seems to be that conditional statements are always discourse-bound (see e.g. Akatsuka 1986). In fact, the phenomena of Q-proposition pre-posing, reduced conditionals (i.e. deletion of either P or Q), nonconditional conditionals (see Lycan 2001 for details) and pseudo-conditionals can only be understood in the light of the discourse-bound nature of conditional statements. Thus, we believe that, depending on the elements of knowledge shared by the speaker and hearer (available from the previous part of the discourse), the speaker may consider it necessary to prepose the apodosis in order to highlight the information contained in it, thus violating the normal clause ordering. Another reason for apodosis-proposition pre-posing has to do with the type of modality it contains. When the speaker expresses deontic modality in making requests or orders, he/she invariably begins the conditional statement by pre-posing the apodosis-proposition, as in,

(18) a. Sit down, if you want!
b.? If you want, sit down!

We believe that similar observations can be made about other conditional statements where the apodosis carries different kinds of non-assertive illocutionary force (as is the case in uttering exclamations, interrogatives, etc.) and thus expresses a non-epistemic modality. Subject to further language-specific research, we can expect apodoses carrying non-epistemic modal meanings to be always pre-posed—irrespective of syntactic differences in languages. In sum, our analysis does not support the widely held belief that conditionals can have both orderings: P-Q and Q-P.

3.3 The Link between P and Q

In logic, the material conditional allows any two unrelated but true propositions to be linked together (e.g. 'If Paris is the capital of France, two is an even number'). However, as we know, not all combinations of two propositions result in conditional constructs in a natural language. In the last few decades, there has been a tremendous amount of research involving various types of psychological experiments aiming to discover the type of relation that P and Q may hold in a conditional statement. Although it would be off-topic and rather presumptuous to try to contribute to this debate in this paper, we deem it appropriate to mention that there are roughly two schools of thought with opposing views on this issue. The first group of scholars maintains that not all P and Q are linked together by a clear relation in a conditional construction. According to their line of reasoning, there are perfectly "standard" conditionals in natural language called "independence conditionals", which do not necessitate P and Q being joined together through any particular relation (Over 2017; Cruz and Over this volume; Over and Cruz this volume). Scholars belonging to the other group, instead, claim that the relation between P and Q is essentially inferential, involving all three types of reasoning, namely induction, deduction and abduction (Douven et al. this volume). As mentioned earlier, for the purpose of typological linguistics research, it will be helpful to adopt a view that envisages a kind of relation between P and Q, be it causal, inferential or whatever. In linguistic typological studies, the relation that holds the two clauses together is a prerequisite for them being called conditionals, and is often labeled 'causal', although it is not always easy to establish the cause–effect relationship between P and Q.

The causal relation that is thought to exist between an antecedent and its consequent has also been the topic of considerable debate in linguistics, and has been studied from different angles, including: mental spaces (Sweetser 1990), semantic consistency (Wierzbicka 1997; Athanasiadou and Driven 1997; Declerck and Reed 2001) and syntactic parameters (Haegeman 2003). The idea of the causal relation as a requirement for a cross-linguistic study, as proposed by Comrie (1986), has been examined by Wierzbicka (1997: 19), who writes (example lettering mine):

> It is true that "if" implies some sort of connection between two propositions, and also that a causal link is often involved, too; I claim, however, that the "if" connection is sui generis, and cannot be reduced to anything else; and that a link with "because" is not always present. For example, the sentence:
>
> (a) *If he insults me, I will forgive him.*
>
> does not imply that I will forgive him BECAUSE he has insulted me: it is true that I can forgive him only if he has done something bad to me (e.g., if he has insulted me), but it is not true that the insult will be the "cause" of my forgiveness. Similarly, the sentence:
>
> (b) *If he invites me to dinner, I will not go.*
>
> does not mean that I will not go because he has invited me: if he doesn't invite me I will not go either; and the sentence:
>
> (c) *If he is asleep, I will not wake him up.*
>
> does not mean that I will not wake him up because he is asleep: on the contrary, I could wake him up only if he were in fact asleep.
> Consider also the following *if*-sentences, of a different kind from those cited above:
>
> (d) *If you do this, people will know about it.*
> (e) *If you do this, this will be bad.*
>
> Clearly, here, too, there is no causal connection between the two propositions.

Despite this observation, we believe that Comrie's idea concerning a causal link between the two propositions withstands Wierzbicka's criticism, because (a) and (b) in the citation are not representative examples of pure conditional constructs, since they require an 'even if' reading which is a special case of conditionals.[4] As far as (c) is concerned, there is no anomaly, since X's being asleep is in fact the cause for Y's decision to not wake up X. Likewise, we can say that "X's doing something causes Y's knowing it" in (d) and "X's doing Y will be the cause of its being judged bad" in (e). At any rate, the causal link between protasis and apodosis is to be viewed from the point of view of speaker and hearer, and not judged on the basis of common knowledge of how things exist or ought to be. We maintain that there is a sort of link, causal or otherwise, between the two propositions which can be established through the epistemic stance of the speaker and the hearer.

Furthermore, for the purpose of developing the criteria for a cross-linguistic typological study, we must also bear in mind that not all combinations of two propositions that are labeled 'conditionals' in English are so in other languages. For instance, the so-called zero conditionals in English (e.g. 'If you freeze water, it becomes solid.') are not conditionals in many languages. Nor are those combinations of two propositions which have a covert adverbial *whenever* or *at all times* in the protasis (e.g. 'If I drink too much coffee, I can't sleep at night.'). Similarly, all those conditional statements whose protasis is part of the shared knowledge of speaker and addressee (e.g. "If you are going there, I will come with you") are not conditionals in many languages, as these sentences have other forms similar to 'Since you are going there, I will come with you'.

4 Time Reference in Conditionals

It has been convincingly argued that the role of grammatical tenses, aspects and moods—particularly of grammatical tenses—in non-factual

[4] I am grateful to David E. Over for reminding me that even in "even if" conditionals, a link, albeit of "topic", can be established.

conditional constructions is markedly different from their use in other constructions in a language. The past time reference in counterfactuals in particular has been a dominant topic of debate in linguistics for many decades. There are significant but conflicting, linguistic theories that address the use of tenses in conditionals, particularly counterfactuals (Dahl 1997; Iatridou 2000; Ippolito 2013; Kaufmann 2005, this volume; Arregui 2007, 2009; Karawani 2014; von Fintel and Iatridou 2020; Mackay 2015; amongst many others), the detailed explication of which is beyond the scope of this paper. Historically, the oldest and the accepted view concerning the hypothetical use of the English past tense is that a type of "back-shifting" in hypothetical conditionals that allows the speaker to convey a belief concerning the fulfillment of the condition of the precise formulation, which, in turn, is predicated on the time reference of the conditional clause (Quirk et al. 1985). Modifying this back-shifting theory, Dahl (1997) considered Tedeschi's Branching-futures model (1981) as a suitable tool to elucidate the past tense in counterfactuals. According to this model, "at any point in time, there is one past and an infinite set of (possible) futures. A counterfactual situation, with respect to a point in time t, is located at a branch of the tree that can be found by going backwards in time from t and then forwards along an alternative path." (Dahl 1997: 101). Tedeschi (1981) used example (19) to illustrate the Branching-futures model, as shown in Fig. 1:

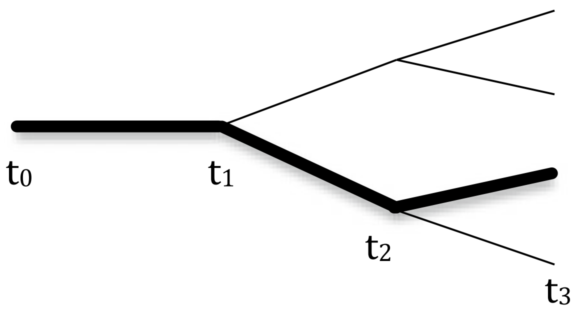


Fig. 1 Branching-futures model (adapted from Dahl 1997)

(19) If Germany had invaded England, they would have won the war.

According to this line of research, assume that we are at t_3, and that t_1 is the crucial point at which Hitler decided not to invade England. In the alternative branch, he invaded England at this point (or a little later), and he won the war at t_2. This suggests an interpretation of (19) as in (20):

(20) It was the case: if Germany invades England, it will be the case: Germany wins the war.

> "In Tedeschi's words, 'we evaluate counterfactual conditional sentences as if we returned to the past and looked at possible futures with respect to that past.' (Tedeschi 1981)" (Dahl 1997: 102). Therefore, this account of the semantics of counterfactuals would be able to precisely portend a combination of past and future marking in such sentences.

Various versions of the aforementioned Branching-futures model have been expounded upon in linguistic research, notably in formal syntax and semantics. Against this backdrop, Iatridou (2000) develops a new line of research and postulates that the past tense semantics can be applied both temporally and modally. She contends that the past tense morphology has an "exclusion feature" and, consequently, the "topic set" excludes the "speaker set".[5] The modal use of the past in counterfactuals, therefore, signals "remoteness". When used temporally, the past tense marks the times talked about as distinct from the speaker's time. On the other hand, when the past involves a modal use, it marks the worlds talked about as distinct from the actual world of the speaker. Another line of research is developed by Ippolito (2003, 2006, 2013) and Arregui (2005, 2009) among others, who both posit that the additional pasts in subjunctive conditionals do retain their usual temporal meaning after all. Using a possible-world semantics technique, Kaufmann (this volume) offers perhaps the most advanced theory, including a unified account of

[5] It seems that Iatridou (2000) intends this to be either sets of times or sets of worlds; in the latter case, the "speaker set" is akin to "the world according to the speaker".

indicative and subjunctive conditionals. That said, we cannot delve into the strengths and weaknesses of all these theories in this paper.

Obviously, much of the data discussed in the aforementioned formulation of the theories of tenses in conditionals comes either from English or from a handful of other thoroughly investigated languages. Hence, it is important to examine the possible contributions through the observations of the different uses of tenses and aspects in conditional constructions in less-commonly investigated languages. In this section, to highlight the importance of these phenomena from a cross-linguistic perspective, I discuss the use of verbal tenses and aspects in counterfactuals in Hindi, a subject that has already come under scrutiny in some important works (Bhatt 1997; Karawani 2014; von Fintel and Iatridou 2020). In doing so, I would like to advance a new proposal concerning Hindi counterfactuals that has not hitherto been considered in detail. A closer examination of the structure of Hindi counterfactuals reveals that it is perhaps the verbal aspect, not the tense, which plays the major role in expressing counterfactuality in this language. To understand this, the distinction between 'normal' and counterfactual conditionals should be considered from the perspective of verbal aspects rather than merely from that of tenses. According to our analysis of Hindi counterfactual conditionals, it is the imperfective aspect—not the past tense marker—that exhibits counterfactuality. If this idea is tenable, there is no need to coin new terms such as "fake habitual aspect" (von Fintel and Iatridou 2020) for an objective analysis of Hindi counterfactual morphology.

As discussed in the literature (van Olphen 1975; Shapiro 2003; McGregor 1995; Sharma 2002), the Hindi verbal system is structured around the tense, aspect and mood elements, as set out in Table 8. Believed to have been derived from the Old Indo-Aryan morphology (Masica 1991), the perfective/imperfective opposition has flourished in all the major New Indo-Aryan languages and is one of the most important characteristics of the Hindi verbal system, in which it is the aspect that plays a major role in structuring all the conditional constructions, particularly counterfactuals. No tense markers exist in Hindi counterfactuals—neither in the protasis nor in the apodosis. Both clauses exhibit the same imperfective morphology. This unique characteristic—not only

Table 8 A complete inflectional inventory of Hindi verb *calnā* 'to walk' or 'to move' in combination with the Hindi TAM elements (grammatical accord: 3rd person, masculine and singular)

	Moods									
Non-aspectual	Root Subjunctive	1. *cale*								
	Future Indicative	2. *cale-gā*								
	Imperative	3. *cal, calo, calie*								
		Imperfective						Perfective		
		Habitual			Progressive					
		ROOT	HAB	AUX	ROOT	PROG	AUX	ROOT	PVF	AUX
	Indicative	4. *cal-*	*tā*	Ø	–			5. *cal-*	*ā*	Ø
Aspectual		6. *cal-*	*tā*	*hai*	7. *cal*	*rahā*	*hai*	8. *cal-*	*ā*	*hai*
		9. *cal-*	*tā*	*thā*	10. *cal*	*rahā*	*thā*	11. *cal-*	*ā*	*thā*
	Subjunctive	13. *cal-*	*tā*	*ho*	13. *cal*	*rahā*	*ho*	14. *cal-*	*ā*	*ho*
	Presumptive	15. *cal-*	*tā*	*ho-gā*	16. *cal*	*rahā*	*ho-gā*	17. *cal-*	*ā*	*ho-gā*
	Counterfactual	18. *cal-*	*tā*	Ø	–			–		
		19. *cal-*	*tā*	*ho-tā*	20. *cal*	*rahā*	*ho-tā*	21. *cal-*	*ā*	*ho-tā*

prog = progressive aspect
pfv = perfective aspect
aux = auxiliary verb

of the Hindi counterfactuals but also of those in most New Indo-Aryan languages—has been discussed by Bhatt (1997). Bhatt makes some important generalizations regarding the marking of counterfactuals through imperfective morphology in New Indo-Aryan languages such as Hindi. First, the imperfective participle alone is the most important ingredient in the counterfactual morphology. Second, Hindi counterfactuals do not exhibit any periphrastic tense marking. Third, both the protasis and the apodosis exhibit the same imperfective morphology. These unique characteristics of Hindi counterfactual morphology have led some scholars to propose new concepts for explaining counterfactuals, such as "fake habitual aspect" (von Fintel and Iatridou 2020) and "aspect stacking" (Karawani 2014). Iatridou (2009) maintains that the imperfective aspect is fake in Hindi counterfactuals, as there is nothing in the semantics of the imperfective that would make it a necessary ingredient for rendering a counterfactual reading. Hence, according to Iatridou, it makes no semantic contribution to counterfactuality. Her conclusion is based on examples such as (21), as she maintains that there is a slot for the fake imperfective (i.e. the habitual morpheme) and a slot for the real imperfective (i.e. the progressive morpheme):

(21) a. *vo gaa rahaa hotaa
he sing PROG be-HAB
b. agar vo gaa rahaa hotaa to log wah wah kar
if he sing PROG be-HAB then people wow wow do
rahe hote
PROG.M.PL be.HAB
'If he were singing, people would be going wow wow.'
(example from Bhatt 1997)

Although Iatridou's observation that there is a slot for the real imperfective (the progressive marker *rahaa* above) and another slot for the fake imperfective (the so-called habitual suffix *-taa* above) in Hindi appears to be correct, the entire concept of "fake aspect" is based on the flawed assumption that languages can only exhibit counterfactuality through tense. We maintain that an aspect-based model of interpretation may turn out to be useful in analyzing counterfactuality in languages such as Hindi.

Let us consider the Hindi verbal system as depicted in Table 8. As can be seen, Hindi only has three markers for tenses, namely *hai* 'is', *thā* 'was' and a suffix, *-gā*, which marks the future tense. The remainder of the verbal system is structured around the aspectual opposition: that is, perfectivity versus imperfectivity. The imperfective aspect marker, namely *tā*, is the marker of imperfectivity throughout the language, although it also marks the habitual aspect by default in certain contexts.

Moreover, the Hindi imperfective marker *tā* does not signal any trace of the habitual aspect in the other contexts in which it is frequently employed, as can be seen in (22a), (22b), (22c) and (22d). Thus, it is incorrect to claim that *tā* is a marker of habitual aspect.

(22) a. rotā huā laṛkā
cry-IMPFV be-PFV boy
'The boy who was crying ...' ≠ 'The boy who cries habitually...'

b. bhāgtā huā cor ...
flee-IMPFV be-PFV.M.SG thief
'The fleeing thief ...' ≠ 'The thief who flees habitually ...'

c. Ram=ne cor=ko bhāgte hue dekhā
Ram=ERG thief=ACC flee-IMPFV.OBL be-PFV.OBL see-PVF
'Ram saw the thief fleeing' ≠ 'Ram saw the thief who flees regularly.'

d. Ram=ne bhāgte hue cor=ko dekhā
Ram=ERG flee-IMPFV.OBL be-PFV.OBL thief=ACC see-PVF
'Ram saw a thief fleeing' ≠ 'Ram saw a thief who flees regularly.'

Let us now consider the question of aspect stacking as discussed by Karawani (2014). Unfortunately, the account that she provides is only a partial account. Contrary to her claim that it is limited solely to counterfactuals (2014: 24), aspect stacking is a widespread phenomenon outside of the realm of conditionality, and is also abundantly attested in factual expressions. Consider the following examples:

(23) a.

yah	baccā	hamesā	rotā	rahtā	hai
this	child	always	cry-IMPFV	stay-IMPFV	AUX.PRS

‘This child keeps on crying all the time.’

b.

Ram	bāzār	jātā	rahtā	hai
Ram	market	go-IMPFV	stay-IMPFV	AUX.PRS

‘Ram keeps on going to the market.’

c.

pichle sal		Ram aksar mandir		jāyā	kartā	thā
last	year	Ram often	temple	go-PFV	do-IMPFV	AUX-PST

‘Last year Ram used to go to the temple (very often).’

d.

Ram	kulfī	khātā	jā	rahā	hai
Ram	ice cream	eat-IMPFV	go	stay-PFV	aux-PRS

‘Ram continues (keeps on) eating the ice cream.’

As can be seen from examples (23a)–(23d), contrary to Karawani’s claim, aspect stacking is the only tool for obtaining iterativity in Hindi. Nonetheless, I believe that this paper is not the right place to discuss all the characteristics of aspect stacking in Hindi, nor to illustrate the perfective-imperfective dichotomy in the language. Nevertheless, in summary, it can safely be affirmed that the imperfective marking suffix *-tā* is employed in at least four different contexts in Hindi: (1) It is suffixed to the verbal root, where it expresses the imperfective aspect and then by default the habitual aspect; (2) it is exploited to obtain imperfective adjectival and adverbial participles; (3) it is employed throughout the language to express different types of iterativity via aspect stacking; and (4) it is used as a modal to express counterfactuality. Thus, we maintain that the imperfective aspect, not the past tense, is the integral element in Hindi counterfactuals, and that the imperfective/perfective aspectual dichotomy is sufficient to provide an explanation for all the issues concerning counterfactuality in Hindi.

In order to determine what an aspect-based branching might look like, let us now consider a Hindi version of example (19) which was analyzed by Tedeschi (1981) in his account of his Branching-futures model:

(24) (agar) Germany=ne England=par hamlā kiyā hotā
(if) Germany=ERG England=on attack do-PFV.M be-IMPFV.M
to vo jīt gayī hotī
then she win go-PFV.F be-IMFV-F
'If Germany had invaded England, they (she) would have won the war.'

Let us use the same branching-futures tree to illustrate this (Fig. 2):

Assume that we are at t_3, and that at t_2, for whatever reason, the action—that could have given Germany a win—was not carried out or called off. In the alternative branchwhich is expressed via perfectivity, not tense, in Hindi—Hitler did indeed accomplish the task and win the war. Thus,

(25) It was the case: if Germany completes invasion of England, it will be the case: Germany accomplishes the task of winning the war.

At this point, tense-theorists might argue that, since the points in time in the diagram, namely t_1, t_2 and t_3, are expressed through tenses, how can an aspect-based model address the question of distancing one event from another—an issue that can only be dealt with appropriately by tenses? If we look at the English example closely, it is the perfective aspect (or perfect)—not the simple past tense—that provides the distance between the two past actions in example (19). Also note that

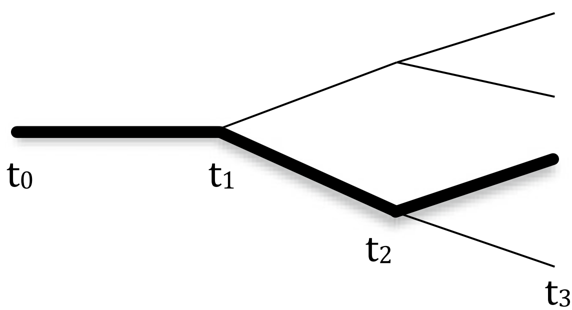


Fig. 2 Tree representing Branching-aspects model

only four of the thirteen English tenses listed by Reichenbach (1947: 297) can be considered "pure" (i.e. aspectless) tenses: *simple past* (E,R—S), *simple present* (S,R,E), and *simple future* (S—R,E and S,R—E). The remaining nine tenses are obtained in combination with aspect: *anterior past* or *past perfect* (E–R–S), *posterior past 1* (R–E–S), *posterior past 2* (R–S,E), *posterior past 3* (R–S–E), *anterior past* or *present perfect* (E–S,R), *anterior future* or *future perfect 1* (S–E–R), *anterior future* or *future perfect 2* (S,E–R), *anterior future* or *future perfect 3* (E–S–R) and *posterior future* (S–R–E). Thus, a closer look reveals that the role played by aspects in counterfactuals has not yet been fully explored. In an aspect-based model, the anteriority of an imperfect action can be determined by the speech time. Furthermore, in an aspect-based model, a perfective aspect may indicate the completion of action prior to the speech time or reference time (That is E > R > S or E, R > S) whereas an imperfect aspect indicates the noncompletion of the action either at the reference time or the speech time. Obviously, E can overlap with the speech time, or follow it, depending on the aspect.

With regard to the question of the crucial role played by the grammatical aspect in counterfactuals, Arregui (2005, 2007, 2009) develops a new line of research and discusses the following examples from English in support of her approach. According to her proposal, aspect plays a central role in counterfactuality, even in English, in which aspect is not exhibited systematically:

(26) You: Could you look after my plants next week while I am away?
Me: Of course, but I am rather nervous. If your plants died next week, I would be very upset.

(27) You: Don't worry about looking after my plants next week. They died.
Me: I am sorry, but also a bit relieved. If your plants had died next week, I would have been very upset.
Me: I am sorry, but also a bit relieved. #If your plants died next week, I would be very upset.

(28) Your plants do not have enough light. If they had enough light,they would be fine.

With regard to the question of the simple past subjunctive conditionals in examples (26) and (27), Arregui (2005, 2007, 2009) observes that the sentence *If your plants died next week, I would be very upset* is acceptable in (26), but not in the second option in (27). According to Arregui, the second option in (27) is anomalous due to aspect restrictions. The past perfect subjunctive conditionals in (28), however, can always be counterfactual, regardless of the type of predicate in their antecedent. Needless to say, we cannot go into details of this proposal in this paper, despite the fact that it is well worth considering.

5 Degrees of Hypotheticality

Recognizing the difficulty of using traditional terminologies such as open versus closed, real versus irreal or real versus hypothetical in cross-linguistic research, Comrie (1986: 88) claims that hypotheticality is a continuum and therefore cannot be classified according to any bipartite or tripartite divisions. As mentioned in the previous sections, it is difficult to find two languages with a correspondingly equal number of classes of conditionals. For example, not many Indian languages possess two separate categories for expressing the difference between the following two English conditional sentences:

(29) a. If you give me a kiss, I'll buy you a beer.
b. If you gave me a kiss, I'd buy you a beer. (from Comrie 1986)

This could be true of many languages that have not been studied thoroughly. In this scenario, a cross-linguistic investigation of conditionals becomes extremely difficult. We believe that, in order to gather data for a sound analysis of conditionals in a less-commonly investigated language, it is necessary to determine the exact number of types of conditional by examining the speaker and hearer's epistemic stock. In addition, one has to begin with only two basic categories of conditionals that are unanimously considered to be universal and leave the third category, namely "If X happened, Y would happen" for subsequent investigation:

> Cross-linguistic evidence suggests that "an awareness of the conditional", an ability to say "what if …?" or "if … then…", is indeed a human universal. I submit that "an awareness of the counterfactual", an ability to say "had this not happened, that wouldn't have happened", may also be a human universal, a vital path in the human mental process, an indispensable element of human language and cognition.
> By contrast, cross-linguistic evidence appears to suggest that the "hypothetical conditionals", situated half-way between the conditionals of real possibility and counterfactuals, (e.g. "if this happened, that would happen"), may be a language-specific phenomenon, and not a universal feature of human language and human thought. (Wierzbicka, 1997: 52)

Thus, for any cross-linguistic investigation, it is imperative to look beyond the use of tenses in the English language. Following the logic of English tenses in any linguistic survey will inevitably lead to only a partial understanding of the wider situation. For example, the use of the English future tense in protasis is limited to such rare cases as the following:

(30) A: The Universe won't come to an end for several million years yet.
B: If it won't come to an end for several million years yet, we'll still be able to go to Florida this winter. (example from Comrie 1986)

By contrast, in most of the languages spoken in South Asia, the use of the future tense in protasis is a common phenomenon. Similarly, most of the examples of English conditionals that exhibit the present indicative tense are rendered in many languages, including Hindi, via the subjunctive mood (Oranskaya 2005). Hence, we believe that the frequently used terms—indicative and subjunctive—are misleading for any serious cross-linguistic research. Following Wierzbicka, we believe that we should begin any cross-linguistic survey by identifying two basic categories of conditionals which could be called: contingent conditionals, and counterfactual conditionals.

5.1 Types of Speaker's Epistemic Stances in Factual Expressions

We believe that before investigating the controversial question of the precise number of conditional types to be investigated in a particular language and the various degrees of hypotheticality expressed thereby, one needs to thoroughly examine the tense-aspect-mood system of that language. In this regard, some generalizations concerning the speakers' epistemic stances are of the utmost importance. We propose that, like all sentences of a natural language, conditionals should be analyzed in light of the speaker's communicative stance. However, before we embark upon a general typology of conditionals (non-factual expressions), we need to provide a brief sketch of epistemic stances in factual expressions in a language. The speaker's communicative epistemic stances and their respective modal meanings can be described in the following manner (Sharma 2002):

(31) **K*s* □*p***: "In order for you to take notice of it and act accordingly, I would like to communicate to you that, for all I know, it is necessarily P." In other words, it is not possible not-P (i.e. Ks¬◊¬P). Almost all natural languages exhibit this epistemic stance of the speaker through "indicatives", which may be loaded with different aspectual morphologies marking habitual, progressive and perfective. Using Reichenbachian terminology, the aspectual characterisations can have various representations: past habitual or progressive E,R>S [such as 'It used to rain last month', 'It was raining this morning', and so forth]; past perfective: E>R>S ['It had rained a lot before I went out'], present habitual or progressive S,R,E ('It rains every day in Ireland' and 'It is raining in Dublin right now', respectively), present perfective ('It has rained a lot this morning in Dublin') and future S<E,R, ['It will rain tomorrow in Dublin'].

(32) **K*s* ◊*p***: "In order for you to take notice of it and act accordingly, I would like to communicate to you that, for all I know, it is possibly *p*." Thus, it is not necessarily *P* and it is not necessarily not-*P*: (that is, $Ks\ \neg\Box P\wedge Ks\neg\Box\neg P$). In general, natural languages do not need to mark this epistemic stance of the speaker in their grammar. This can be paraphrased in the following manner: The speaker knows that it is possible that it rains in Dublin at any time. We are not aware of any language that expresses this epistemic stance through grammatical means, although some languages, such as Hindi, express this epistemic stance through an iterative habitual aspect.

(33) **B*s* □*p***: In order for you to take notice of it and act accordingly, I would like to communicate to you that, although I don't know that *P*, I nonetheless believe that it is necessarily *P*. This means, the speaker believes that it is not possible not-P (that is, $Bs\ \neg\Diamond\neg P$). Many languages, including English, exhibit this epistemic stance through modals such as must, may, might and the like [for example, 'It must be raining every day in Dublin in September' and 'It must be raining in Dublin right now']. In many languages, including Hindi, this epistemic stance can be well supported by morphological devices that express various aspectual elements in line with its tense-aspect system. For example, a Hindi speaker can exhibit this epistemic stance with three distinct aspectual markings, which are habitual, progressive and perfective.

(34) **B*s* ◊*p***: In order for you to take notice of it and act accordingly, I would like to communicate to you that, although I don't know that *P*, I nonetheless believe that it is possibly *P*. Thus, the speaker believes that not necessarily *P* and not necessarily not-*P* (that is, $Bs\ \neg\Box P\wedge Bs\ \neg\Box\neg P$). Many languages have morphological ways of expressing this epistemic stance, which is often called the subjunctive (also called the optative in certain languages. For example, 'It may rain in Dublin tomorrow'), which may also be loaded with markers of various aspects such as habitual ('It is possible that it rains daily in Dublin this week'), progressive ('It is possible that it may be raining in Dublin right now') and perfective ('It may have rained in Dublin this morning') markers.

5.2 Types of Speaker's Epistemic Stance in Conditionals

Having illustrated the speaker's epistemic stance in factual statements (31)–(34), we now turn to conditionals to discuss the two major classes of conditionals, namely, contingent conditionals (also known as indicative conditionals) and counterfactuals (also known as "subjunctive conditionals"). Other classes of conditionals can be discussed in a cross-linguistic survey following the same line of research.

5.2.1 Conditional Type 1 (The Lowest Degree of Hypotheticality)

Obviously, one may gather different data-sets from a language that may appear to belong to the first type. However, we propose that, for an in-depth cross-linguistic survey, it is important to focus solely on a small specimen of a conditional. Accordingly, let us consider the following example in (35) in light of the epistemic stock of the speaker and the hearer, as presented in (36):

(35) If Mary invites John, he will go to her party.

(36) Speaker and Hearer's epistemic stock:
α = Mary's Party will take place tomorrow (i.e. S < E,R)
β = Mary hasn't extended an invitation to John yet.
γ = Mary will extend an invitation to John.

(37) (a) $K_s\, \alpha \wedge K_s\, K_h\, \alpha \vee B_s\, K_h\, \alpha$
(b) $K_s\, \beta \wedge K_s\, K_h\, \beta \vee B_s\, K_h\, \beta$
(c) $B_s\, \gamma \wedge B_s\, \neg K_h\, \gamma$

Thus, in view of the Speaker-Hearer's epistemic stock in (36), the conditional statement in (35) will have the speaker's epistemic stance as formulized in (37), which reads as follows: (a) The speaker knows that

Mary's party will take place tomorrow and also knows that hearer knows about it. Or, the speaker at least believes that the hearer knows that *Mary's party will take place tomorrow*; (b) the speaker knows that *Mary hasn't extended an invitation to John yet* and knows that the hearer also knows about it. Or, the speaker at least believes that the hearer knows that *Mary hasn't extended an invitation to John yet*; and (c) the speaker believes that Mary *will extend an invitation to John* and believes that the hearer does not know that Mary *will extend an invitation to John.*

Thus, in view of the Speaker-Hearer's epistemic stock in (36), the first type of conditional in (35) will have the following modal meaning: "In order for you to take notice of it and act accordingly, I would like to invite you to evaluate Q (that is 'John will go to Mary's party') in light of P (that is 'Mary invites John to her party'). In other words, the speaker does not believe that the hearer knows that Mary will invite John to the party and thus wants to inform him that John going to the party is contingent upon Mary's invitation to John.

5.2.2 Conditional Type 2 (The Highest Degree of Hypotheticality)

This class of conditionals is the opposite of type 1; these are counter-factuals because the speaker invites the hearer to evaluate the counter to the fact conditions.

(38) If Mary had invited John, he would have gone to her party.

(39) Speaker and Hearer's epistemic stock:
α = Mary's party took place yesterday [i.e. E,R > S]
β = Mary didn't extend an invitation to John.
γ = John didn't go Mary's party.

(40) (a) $K_s\,\alpha \wedge K_s\,K_h\,\alpha$
(b) $K_s\beta \vee B_s\,\beta \wedge K_s\,\neg K_h\,\beta \vee K_s\,\neg B_h\,\beta$
(c) $K_s\,\gamma \vee B_s\,\gamma \wedge K_s\,K_h\,\gamma \vee B_s\,K_h\,\gamma$

Thus, in view of the Speaker-Hearer's epistemic stock in (39), the conditional statement in (38) will have the speaker's epistemic stance as formulized in (40) which reads as follows: (a) The speaker knows that *Mary's party took place yesterday* and also knows that the hearer knows that *Mary's party took place yesterday*; (b) the speaker furthermore either knows or believes that *Mary didn't invite John to the party*. He also knows that the hearer either does not know or does not believe that *Mary didn't invite John to the party*; in addition, (c) the speaker either knows or believes that *John didn't go to Mary's party*, and the speaker knows or believes that the hearer knows that *John didn't go to Mary's party*. Equipped with this epistemic stock, the speaker invites the hearer to consider β as the sole reason for γ, and to imagine a world in which both β (antecedent) and γ (consequent) were contrary to the fact.

6 Conclusion

In the preceding sections, we have attempted to evaluate the tenability of Greenberg's Universal Word Order 14 in relation to different classes of languages, and argued that P-Q was the only ordering that was acceptable in conditional statements. It has been affirmed that the Q, P ordering attested in some languages is not due to any type of afterthought on the part of the speaker (such as placing P after Q in order to avoid the risk of making a factual statement, as suggested by Comrie, 1986), but is the result of a discourse-related (Akatsuka 1986) requirement (for example, an apodosis is fronted if that is the focus of the discourse), some pragmatic universals (fronting an apodosis in imperative conditionals is a universal pragmatic phenomenon) and other factors that are language specific. The classification of language-specific features across different classes of languages is a broad topic that requires further in-depth research and a detailed survey of less-commonly scrutinized languages. We have also attempted to show that, contrary to widely held belief, the so-called apodosis markers, such as *then* in English, are actully related to the protasis rather than to the apodosis. With regard to the types of conditionals found in different languages, there seem to be only two universal categories: the so-called indicative conditionals, which we would prefer to call contingent conditionals, and counterfactual conditionals. Languages that exhibit more than two types of conditionals, such as English, French and Italian, amongst many others, have developed sophisticated morphological tools through which the speaker can express his/her epistemic stance in the protasis. In addition, as far as the time reference in counterfactuals is concerned, we have argued that the past tense reference alone is not sufficient for the formation of a unified account of tenses, as it is the verbal aspect that is responsible for obtaining counterfactuals in many languages.

Acknowledgements I am deeply indebted to my mentors, Stefan Kaufmann and David E. Over, for their unwavering support and active encouragement (both academic and personal) without which I would not have been able to finish this article. As a linguist by trade, I eagerly and gladly accepted their advice on a variety of subjects. Obviously, any errors in the paper are my own. I would also like to thank friends and colleagues from different universities for their help with the data judgment: John Lowe (Oxford), Tanmoy Bhattacharya (Delhi), K.V.R. Subbarao (Delhi), Amba Kulkarni (Hyderabad), Anuradha Sudharsan (Hyderabad) and Prashant Pardeshi (Tokyo). Again, none of these individuals should be held accountable for any inaccuracies in the paper.

References

Akatsuka, N. (1986). Conditionals are discourse-bound. In *On Conditionals*, Traugott, E. C., Meulen, A. T., Reilly, J. S. & Ferguson, C. A., editors, 333–351. Cambridge: Cambridge University Press.

Arregui, A. (2005). *On the accessibility of possible worlds: The role of tense and aspect*. University of Massachusetts (Amherst) Doctoral Dissertation.

Arregui, A. (2007). When aspect matters: The case of would-conditionals. *Natural Language Semantics 15*(3): 221–264. https://doi.org/10.1007/s11050-007-9019-6

Arregui, A. (2009). On similarity in counterfactuals. *Linguistics and Philosophy 32*: 245–278.

Athanasiadou, A. & Driven, R., editors, (1997). *On Conditionals Again*. Amsterdam: John Benjamins Publishing Company.

Barwise, J. (1986). Conditionals and conditional information. In *On Conditionals*, Traugott, E. C., Meulen, A. T., Reilly, J. S. & Ferguson, C. A., editors, 21–54. Cambridge: Cambridge University Press.

Bhatt, D. N. S. (1999). *The Prominence of Tense, Aspect and Mood*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bhatt, R. (1997). *Counterfactuality in Indo-Aryan. Notes assembled for Sabine Iatridou's Counterfactuals Project*. Philadelphia, PA: University of Pennsylvania. https://people.umass.edu/bhatt/papers/sjv.pdf

Bhatt, R. & Pancheva, R. (2006). Conditionals. In *The Blackwell Companion to Syntax*, Everaert, M. & van Riemsdijk, H., editors, 638–687. Malden, MA: Blackwell.

Cariani, F. & Rips, L. J. (this volume). Experimenting with (Conditional) Perfection.

Comrie, B. (1986). Conditionals: A typology. In *On Conditionals*, Traugott, E. C., Meulen, A. T., Reilly, J. S. & Ferguson, C. A., editors, 77–99. Cambridge: Cambridge University Press.

Cruz, N. & Over, D. E. (this volume). Independence conditionals.

Csipak, E. & Romero, M. (this volume). Counterfactual hypothetical vs. biscuit conditionals: A semantic/pragmatic analysis of their morphological differences.

Dahl, Ö. (1997). The relation between past time reference and counterfactuality: A new look. In *On Conditionals Again*, Athanasiadou, A. & Driven, R., editor, 97–114. Amsterdam: John Benjamins Publishing Company.

Dancygier, B. (1998). *Conditionals and Prediction: Time, Knowledge and Causation in Conditional Constructions*. Cambridge: Cambridge University Press.

Dancygier, B. & Sweetser, E. (1997). *Then* in conditional constructions. *Cognitive Linguistics 8*(2): 109–136.

Declerck, R. & Reed, S. (2001) *Conditionals, A Comprehensive Empirical Analysis*. Berlin: Mouton de Gruyter.

Donaldson, T. (1980). *Ngiyambaa: The language of the Wangaaybuwan*. Cambridge: Cambridge University Press.

Douven, I., Elqayam, S. & Krzyżanowska, K. (this volume). Inferentialism: A manifesto.

Edgington, D. (1986). Do conditionals have truth conditions? *Crítica, Revista Hispanoamericana de Filosofía 18*(52): 3–39.

Edgington, D. (2007). On conditionals. In *Handbook of Philosophical Logic*, 127–221. Dordrecht: Springer.

Edgington, D. (2014). Conditionals. In *The Stanford Encyclopedia of Philosophy*, Zalta, E. N., editor. http://plato.stanford.edu/archives/win2014/entries/conditionals/

Edgington, D. (this volume). Conditionals, indeterminacy, probability and truth.

Geis, M. & Zwicky, A. M. (1971). On invited inferences. *Linguistic Inquiry 2*: 561–566.

Greenberg, J. H. (1963). Some universals of grammar with particular reference to the order of meaningful elements. In *Universals of Language*, Greenberg, J. H., editor, 73–113. (second edition 1966). Cambridge, MA: MIT Press.

Haegeman, L. (2003). Conditional clauses: External and internal syntax. *Mind & Language 18*(4): 317–339.

Haegeman, L. & Schönenberger, M. (this volume). The external syntax of conditional clauses.

Haiman, J. (1978). Conditionals are topics. *Language 54*: 565–589.

Haiman, J. (1986). Constraints on the form and meaning of the protasis. In *On Conditionals*, Traugott, E. C., Meulen, A. T., Reilly, J. S. & Ferguson, C. A., editors, 215–227. Cambridge: Cambridge University Press.

Horn, L. (2000). From IF to IFF: Conditional perfection as pragmatic strengthening. *Journal of Pragmatics 32*: 289–326.

Iatridou, S. (2000). The grammatical ingredients of counterfactuality. *Linguistic Inquiry 31*(2): 231–270.

Iatridou, S. 2009. (Collaborative thoughts) about the imperfective in counterfactuals. Handout of a talk given at the Imperfective Workshop at Yale University.

Iatridou, S. & Embick. D. (1994). Conditional inversion. In *Proceedings of NELS*, Vol 24, 189–203. University of Massachusetts at Amherst: GLSA.

Iatridou, S. (1994). On the contribution of conditional *then*. *Natural Language Semantics 2*: 171–199.

Ippolito, M. (2003). Implicatures and presuppositions in counterfactuals. *Natural Language Semantics 11*: 145–186.

Ippolito, M. (2006). Semantic composition and presupposition projection in subjunctive conditionals. *Linguistics and Philosophy 29*: 631–672.

Ippolito, M. (2013). *Subjunctive Conditionals: A Linguistic Analysis*. Cambridge, MA: MIT Press.

Karawani, H. (2014). *The Real, the Fake, and the Fake Fake: In Counterfactual Conditionals, Crosslinguistically*. LOT 357: Netherlands Graduate School of Linguistics.

Kaufmann, S. (2005). Conditional truth and future reference. *Journal of Semantics 22*: 231–280.
Kaufmann, S. (2013). Causal premise semantics. *Cognitive Science 37*: 1136–1170.
Kaufmann, S. (this volume). How fake is fake past?
Lehmann, C. (1974). Prinzipien fur ‘Universal 14’. In *Linguistic Workshop II*, Seiler, H., editor, 69–97. Paderborn: Wilhelm Fink Verlag.
Lycan, W. G. (2001). *Real Conditionals*. Oxford: Oxford University Press.
Mackay, J. (2015). Actuality and fake tense in conditionals. *Semantics and Pragmatics 8*(12): 1–12. https://doi.org/10.3765/sp.8.12
Masica, C. P. (1991). *The Indo-Aryan Languages*. Cambridge: Cambridge University Press.
McGregor, R. S. (1995). *Outline of Hindi Grammar*. Oxford: Oxford University Press.
Oranskaya, T. I. (2005). Conditional constructions in Hindi. In *Typology of Conditional Constructions*, Xrakovskij, V. S., editor, 218–245. Munich: Lincom Europa.
Over, D. E. (2017). Causation and the probability of causal conditionals. In *The Oxford Handbook of Causal Reasoning*, Waldmann, M., editor, 307–325. Oxford: Oxford University Press.
Over, D. & Cruz, N. (this volume). Indicative and counterfactual conditionals in the psychology of reasoning.
Quirk, R., Greenbaum, S., Leech, G. & Svartvik, J. (1985). *A Comprehensive Grammar of the English Language*. London & New York: Longman.
Reichenbach, H. (1947). *Elements of Symbolic Logic*. New York: Dover Publications [Originally published by Macmillan Company].
Shapiro, M. C. (2003). Hindi. In *The Indo-Aryan Languages*, George, C. & Jain, D., editors, 276–314. London and New York: Routledge.
Sharma, G. (2002). On the modal meanings of subjunctive in Hindi. In *The Yearbook of South Asian Languages and Linguistics*, 177–198. New Delhi/Thousand Oaks/London: Sage Publications.
Sharma, G. (2010). On Hindi conditionals. In *Annual Review of South Asian Languages and Linguistics*, 107–134. Berlin: Mouton de Gruyter.
Sharma, G. (2011). On the role of protases in conditional statements: Some evidence from Hindi. In *Annual Review of South Asian Languages and Linguistics*, 49–78. Berlin: Mouton de Gruyter.
Sweetser, E. (1990). *From Etymology to Pragmatics*. Cambridge: Cambridge University Press.

Tedeschi, P. (1981). Some evidence for a branching-futures semantic model. In *Tense and Aspect, Volume 14 of Syntax and Semantics*, Tedeschi, P. & Zaenen, A., editors, 239–270. New York: Academic Press.

Thompson, S. A., Longacre, R. E. & Hwang, S. J. J. (2007). Adverbial clauses. In *Language Typology – A Syntactic Description*, Shopen, T., editor, Vol 2, 237–299. Cambridge: Cambridge University Press.

Traugott, E. C., Meulen, A.T., Reilly, J. S. & Ferguson, C. A., editors, (1986). *On Conditionals*. Cambridge: Cambridge University Press.

van der Auwera, J. (1997). Conditional perfection. In *On Conditionals Again*, Athanasiadou, A. & Dirven, R., editors, 169–180. Amsterdam: John Benjamins Publishing.

van Olphen, H. (1975). Aspect, tense, and mood in the Hindi verb. *Indo-Iranian Jornal 16* (4): 284–301.

von Fintel, K. (2011). Conditionals. In *Semantics: An International Handbook of Natural Language Meaning*, von Heusinger, K., Maienborn, C. & Portner, P., editors, Vol 2, 1515–1538. Berlin: Mouton de Gruyter.

von Fintel, K. & Iatridou, S. (2020). Prolegomena to a theory of X-marking. Manuscript submitted to *Linguistics & Philosophy*.

Wierzbicka, A. (1997). Conditionals and counterfactuals: Conceptual primitives and linguistic universals. In *On Conditionals*, Athanasiadou, A. & Driven, R., editors, 15–59. Amsterdam: John Benjamins Publishing Company.

Williamson, T. (2020). *Suppose and Tell: The Semantics and Heuristics of Conditionals*. Oxford: Oxford University Press.

Xrakovskij, V. S., editor. (2005). *Typology of Conditional Constructions*. Munich: Lincom Europa.

Should Past-as-Modal Theorists Also Be Past-as-Past Theorists?

John Mackay

1 Introduction

In many languages, the conditionals traditionally labelled "subjunctive" or "counterfactual" are marked with the past tense. This past tense appears even where the events or states described by the conditional do not precede the time of evaluation.

Many accounts of this phenomenon agree that these conditionals feature an instance of past tense that takes wide scope outside the antecedent and consequent. Thus, consider example (1), taken from Romero (2014). Leaving aside all other internal elements of its composition, it has a structure like the following:

J. Mackay (✉)
Department of Philosophy,
University of Wisconsin-Madison, Madison, WI, USA
e-mail: john.mackay@wisc.edu

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_12

(1) If Hubert were here, Steffi would be happy.
PAST (If Hubert is here, Steffi will be happy).

If we further follow Kratzer (1981) in maintaining that the antecedent of a conditional restricts a modal operator, and follow Abusch (1997) in identifying "would" as the past tense version of the modal WOLL, we get a structure like the following:

PAST ([WOLL: if Hubert is here] [Steffi is happy]).

However, there is a divergence in view about the meaning of this wide-scope past tense. According to one type of view, variously known as "past-as-past" or "temporal distancing", this past tense receives its standard temporal interpretation. The *would*-conditional (1) is true if and only if the corresponding *will*-conditional was true at some past time. Implementations of this broad strategy include Ippolito (2013), Romero (2014), Arregui (2009), and Khoo (2015). In this approach, the conditional is standardly evaluated with a historical modal base according to which one world is historically accessible at a time from another if it shares its history with the other world up until that time.[1] Since a counterfactual antecedent is not historically accessible at the time of utterance, a present-tense conditional with that antecedent is vacuous at the time of utterance. In order to select worlds of the antecedent, we need to go back to a time at which the possibility of the antecedent had not yet been foreclosed. Suppose that (1) is true; in this case, it is so because the conditional "if Hubert is here, Steffi will be happy" was true at some time before Hubert decided not to come.

Another type of view is known as "past-as-modal" or "modal distancing". In this approach, the wide-scope past tense in these conditionals receives a modal interpretation conveying something about the relation between the selected worlds of the antecedent and the worlds of the context. It does not shift the time of evaluation earlier. Rather,

[1] These theorists differ over whether the modal base is always historical; Khoo allows that it is sometimes epistemic.

it expands the available worlds to be selected beyond those of the context set, though these views differ among themselves in the notion of context involved. Views of this kind include Iatridou (2000), Schulz (2014), Karawani et al. (2019), and Mackay (2019).[2] There are, one may grant, some broad methodological themes that favour a past-as-past view, in which the past tense has just the usual temporal interpretation: in semantics, one tries to assign uniform meanings where possible. But the issues with explaining the presence of past tense in subjunctive conditionals have remained sufficiently complex that both research programmes remain active.

Setting up the debate in this way, however, presents a puzzling asymmetry. The past-as-past theorist denies that the modal reading of the past exists at all. The past-as-modal theorist, by contrast, does not deny that the past tense receives a temporal interpretation in ordinary non-modal examples. It is not in dispute that the past tense receives a temporal interpretation in a basic non-modal sentence like "It was raining yesterday". Thus, the past-as-modal theorist postulates two readings for the past tense where the past-as-past theorist postulates just one. *Prima facie*, if the past tense has two readings, and the subjunctive conditional contains the past tense, then the subjunctive conditional itself should have two readings. The existing past-as-modal literature, however, generally just discusses the modal reading of the past in the subjunctive conditional and is not always explicit about whether there exists another reading of the same conditional in which the past tense is read temporally. This raises the question whether the intended past-as-modal view is that the subjunctive conditional itself has two readings, or instead just the one in which the wide-scope past tense receives the modal interpretation.

These considerations should lead us to divide the available views into three rather than just two, dividing the "past-as-modal" depending on whether the wide-scope past tense in the subjunctive conditional can also receive the temporal reading. This leads to the following options:

[2] von Prince (2019) has a view somewhat harder to classify, with elements of both views.

(i) **One reading past-as-past**. The past tense has just the temporal reading. All readings of the subjunctive conditional are obtained from the interaction of the past tense with the other elements of the conditional.
(ii) **One reading past-as-modal**. The past tense has a modal reading as well as a temporal reading. However, the subjunctive conditional does not appear without the modal interpretation; every subjunctive conditional contains at least one instance of the past tense with the modal interpretation in wide-scope position.
(iii) **Two readings**. The past tense has a modal reading as well as a temporal reading. Every subjunctive conditional has (at least) two readings: one where the wide-scope past tense receives a modal reading and one where it receives a temporal reading.

To be clear, the readings being counted here in the names of the views are the readings of the subjunctive conditional as a whole sentence. Both the one-reading past-as-modal view and the two-readings view agree that the past-tense morpheme has two readings. But the one-reading past-as-modal theory maintains that the subjunctive conditional does not have a reading where all instances of the past tense are read temporally.

I will set aside some foundational questions about the relationship between the two readings of the past tense. Past-as-modal theorists have generally suggested that the past tense's meaning involves an abstract structural property in common between modality and tense, rather than an accidental ambiguity. But this question will not concern me directly here.

The past-as-modal theorist owes us an account of whether they favour a view of type (ii) or of type (iii). At first glance, there is something puzzling about either option. On the one hand, the one-reading past-as-modal theory seems to propose that a reading is unavailable in some environments in which general considerations of compositionality would suggest that it could be interpreted. Why should the past tense be unable to receive its temporal interpretation in the wide-scope position of the subjunctive conditional? This theorist owes us an account of what blocks this reading. On the other hand, the two-readings theory seems to postulate two distinct readings for single conditionals where it is not pretheoretically clear that two readings of the sentence are available.

In this paper, I shall present some data that, I shall argue, suggest that these two readings do indeed exist. This favours a two-readings theory. Subjunctive conditionals, therefore, have two readings, one where the wide-scope past tense takes a modal reading and one where it takes a temporal reading. However, I shall also present an explanation of why the temporal reading is unavailable in most contexts; this provides some explanation of why the existence of the two readings is not obvious.

2 One-Past and Two-Past

It should be clarified that all sides agree that the past tense receives the temporal interpretation when it appears in narrow-scope position on the antecedent or the consequent itself.

Compare Romero's (1) above to (2):

(2) If Hubert had been here, Steffi would have been happy.
PAST ([WOLL: if PAST (Hubert is here)] [PAST(Steffi is happy)])

In (2) the antecedent and consequent clauses themselves have another layer of past morphology, appearing within the conditional as a past perfect form. Forms like (1) are known as "one-past" while forms like (2) are "two-past", in the sense that they contain two layers of past-oriented morphology. I will follow various theorists on both sides of the debate in assuming that the combination of wide-scope and narrow-scope past in (2) manifests itself in the English surface form as a past perfect; Ippolito (2006) did advocate an aspect-based theory in which the perfect rather than a second past was involved but came to reject it in Ippolito (2013).[3] The levels of past, meanwhile, can vary independently between the antecedent and consequent, as for example in (3), which has a two-past antecedent and a one-past consequent.

[3] See also Schulz (2017).

(3) If Hubert had been here, Steffi would be happy.
PAST ([WOLL: if PAST (Hubert is here)] [(Steffi is happy)])

In general—though with exceptions, to be discussed below—we use the two-past form when the event time of the clause is itself in the past. Thus, if the event from which Hubert is absent is ongoing at the time of speech, we use (1); if it precedes the speech time, we use either (2) or (3) depending on whether Steffi's counterfactual happiness would also be in the past. The past-as-modal theorists grant that this second layer of tense in narrow-scope position receives the standard temporal interpretation, and that the modal reading is unavailable in this position. And they have a principled explanation for doing so. If we follow Heim (1994) in positing that tenses express a presupposition about the value of a variable, a tense must have a variable of the relevant kind to operate on. The antecedent and consequent clauses can contain a time variable, but if they lack a modal themselves, there will be no modal base variable for a narrow-scope past to operate on, and it can only receive the temporal reading. This stands in contrast with the question about the wide-scope past tense, where our problem is that the modal theorist does not have a clear reason to exclude the availability of the temporal reading.

Thus, the debate concerns just the reading of the past tense in the wide-scope position of the subjunctive conditional. Both sides are in agreement that the narrow-scope tense on the antecedent and consequent is just interpreted temporally.

3 One-Past Forms in Past-Oriented Narratives

The data I will focus on involve the use of the one-past form when the events or states described by the antecedent and consequent are in the past relative to the speech time.

As noted above, the general tendency in subjunctive conditionals is that there is one more layer of past than the event time of the clause would intuitively suggest. Consider again the examples of Hubert and Steffi at the party. If the event from which Hubert is absent is ongoing at the speech time, we use (1). If the party is in the past, and we mean to convey that Steffi would have been happy in the past as well, we use (2). If the party is in the past but Steffi's possible happiness extends to the present, we use (3).

However, there are two types of exception to this pattern. Ippolito (2006) observes that in some cases, we can use two more past layers than the event times would suggest; the conditional (2) can be used for an ongoing party. This usage conveys some kind of additional distance between the actual world and the possibility introduced by the antecedent, for example if Hubert is dead or the party in question has been cancelled. My focus here, however, is on an exception in the opposite direction: in some cases, we can use the one-past form even when the event time is in the past.

Theorists on both sides of the debate have claimed that the one-past form for a past event is unacceptable; in introducing (1), Romero stated that it cannot be used for past events, and on the past-as-modal side I endorsed a similar claim in Mackay (2017). However, this claim is too strong. The form is indeed usually unacceptable for past events out of the blue, but it can be used for an event that precedes the time of utterance in the midst of a narrative that makes salient some reference time in the past which precedes that event or is simultaneous with it. Consider discourses like (4) and (5).

(4) Anna looked around the room to see which of her friends were present. She saw that Steffi was there, and looked further for Hubert. If Hubert were here, Steffi would be happy.
(5) The party was to begin in an hour. If Hubert came to the party, Steffi would be happy.

In these cases, the conditional is acceptable even though there is only one layer of past morphology and the antecedent and consequent pertain to the past relative to the speech time. In the framework of Reichenbach (1947), the discourse preceding the conditionals introduces a reference time which precedes both the speech time and the event times of the conditional's antecedent and consequent. Relative to that reference time, the event times of the conditional are not in the past. Thus the right generalization is not that the one-past form requires event times that are non-past relative to the speech time, but rather that it requires event times that are non-past relative to the contextually salient reference time, which is often the speech time but need not be.

This phenomenon raises two questions. First, why is this reading available at all, rather than unavailable? Second, why is the reading available only in rare cases, in which the discourse has made some past time salient, instead of being generally available out of the blue?[4]

4 Explaining the Availability of These Forms

Let us turn to the first question: how can there be an available reading where there is just one layer of past morphology even though the event times are in the past?

Repeating the structure from above, we have:

PAST ([WOLL: if Hubert is here] [Steffi is happy])

Now, the modal WOLL requires the event time of the prejacent to be either at or after the time of evaluation. WOLL is standardly given a semantics along the following lines, allowing for various accounts of what the relation R is:

$$[\![\text{WOLL } \phi]\!]^{w,t,g} = 1 \text{ iff } \forall w' R w : \exists t' \geq t : [\![\phi]\!]^{w',t',g} = 1.$$

---

[4] It is worth pointing out (as a reviewer suggests) a connection to the phenomenon of modal subordination discussed by Roberts (1989). However, pointing to modal subordination does not, on its own, explain the phenomenon. In Roberts' canonical examples of modal subordination, analogous distinctions in the levels of tense appear.

(i) A wolf might come in. It would eat you first.
(ii) A wolf might have come in. It would have eaten you first.

Ordinarily, we would use the first of these where the hypothetical entry of the wolf is in the future, while the second would be used where it is in the past, with exception analogous to the conditionals we have been considering. But the relationship between the times involved and the tense morphology here is a distinct topic from the resolution of anaphora in modal subordination.

There are a number of different accounts along these general lines, differing principally in which worlds are selected by what I call the relation *R*; these include Kaufmann (2005), Giannakidou and Mari (2018), and Copley (2009). (See also Kaufmann [this volume], which contains some independent discussion of some examples of the kind that interest me in this paper.) This schematic view is neutral with respect to whether the selected worlds always include the world of evaluation, and thus as to whether *will p* entails *p*. It is also consistent with there being just one world selected, as in the view of Cariani and Santorio (2017). But regardless of the details of which worlds are selected, the event time for the prejacent of WOLL cannot precede the time of evaluation for the higher clause including WOLL.

In the past-as-modal semantics of Iatridou, Schulz, and Mackay, the modal reading of the past does not manipulate the time of evaluation. Rather, it expresses a presupposition about the either the modal base or the worlds of evaluation. But without any effect on the time of the prejacent of WOLL, it does not remove the requirement that the event time be at or after the time of evaluation. Thus, this reading should be unavailable if the wide-scope past tense receives a modal reading.

This suggests that the past tense is interpreted temporally in these conditionals. The past tense therefore can receive a temporal interpretation even in the wide-scope position of the form of the conditional traditionally labelled subjunctive. This is inconsistent with the one-reading past-as-modal view.

In the abstract, this phenomenon is neutral between one-reading past-as-past theories and two-reading theories. However, some past-as-past views do not predict the availability of these views. This is due to differences among these views, internal to the overall past-as-past approach. In the view of Khoo (2015), the clauses under the scope of WOLL have a null tense when they do not have an additional (narrow-scope) layer of past tense. This gives (1) a structure like the following[5]:

[5] In these representations, I am eliding some differences among the formal frameworks of Khoo, Ippolito, and Romero, which are not directly relevant to the difference among the views of interest here. Ippolito, for example, has separate morphemes HIST and SIM that govern the selection of historically possible worlds.

PAST ([WOLL: if Steffi is here] [Steffi is happy])

This will have truth conditions like the following:

$[\![$PAST ([WOLL: if Hubert is here] [Steffi is happy])$]\!]^{w,t,g} = 1$ iff $\exists t' < t : \forall w' R w : \exists t'' \geq t' : [\![$ Hubert is here $]\!]^{w',t'',g} = 1 : [\![$ Steffi is happy $]\!]^{w',t'',g} = 1$.

Here the event time of Hubert's presence and Steffi's happiness must be non-past with respect to the past time at which Hubert's presence is historically possible (assuming R involves a historical accessibility relation), but it is neutral with respect to whether they are before or after the evaluation time of the whole sentence.

In the view of Ippolito (2013) and Romero (2014), by contrast, the clauses have a present tense which appears morphologically as the past due to the sequence of tense phenomenon.

PAST ([WOLL: if PRES (Steffi is here)] [PRES(Steffi is happy)])

This present tense is a non-past operator, anchored to the evaluation time for the whole sentence, that requires an event time that does not precede the evaluation time. This gives us truth conditions like the following:

$[\![$PAST ([WOLL: if PRES (Hubert is here)] [PRES(Steffi is happy)])$]\!]^{w,t,g} = 1$ iff $\exists t' < t : \forall w' R w : \exists t'' \geq t' : \exists t''' \geq t : [\![$ Hubert is here $]\!]^{w',t''',g} = 1 : \exists t''' \geq t : [\![$ Steffi is happy $]\!]^{w',t''',g} = 1$.

Here the event times of the antecedent and consequent cannot be before the evaluation time for the whole conditional. Therefore, both Romero and Ippolito predict that the one-past form should be unavailable when the event times of the antecedent and consequent are in the past. The one-past form is indeed usually unavailable, and these theorists cite this prediction as an advantage of their view over certain others. But as we have seen, the prediction is too strong. Thus, regardless of whether the past tense has an additional modal reading, as in (iii) above, or not, as in (i), the temporal interpretation should receive an account like Khoo's, which allows the event time in the one-past form to precede the speech time so long as it does not precede the reference time.

5 The Difficulty in Accommodating Reference Times

The second question is why this reading is only available inside certain narratives. While the thesis that the two-past form is mandatory for past events is too strong, the two-past form is still unavailable for past events in many typical contexts. The claim that it is unavailable was motivated by the fact that out of the blue, or otherwise outside a narrative of the kind we have been considering, sentence (2) can pertain to a past event but (1) cannot.

Given that the acceptability of the form under consideration depends on the surrounding discourse, I suggest that the unavailability of the reading in many contexts is due to the pragmatic difficulty of accommodating a past reference time. As observed by Partee (1973) and Enç (1987), a tense needs to be anchored to a reference time, and when the tense is widest in scope in a sentence, the reference time comes from the context of utterance. This basic idea does not depend only on a referential theory of tense; it can also be implemented within a quantificational theory, as for example by Altshuler (2016). Of course, such a reference time can often be accommodated even when it is not salient in the discourse prior to the utterance of the past-tense sentence.

A general pattern is that it is easy to accommodate a past reference time that matches the event time, while it is difficult to accommodate a past reference time that differs from the event time. Typically, the reference time matches the event time in a simple past sentence, but they can differ when the tense interacts in certain ways with modality or with aspect. Contrast the felicitous discourse in (6) with the problematic ones in (7) and (8).

(6) Steffi is here. She came by car, and it took her fifteen minutes to get here.
(7) Steffi is here. She would arrive by car fifteen minutes later.

(8) Steffi is here. She had left fifteen minutes earlier.

In (7) and (8), it is hard to interpret what times the second sentence in the discourse is talking about. Now, in all three of (6)–(8), the utterance of the initial present-tense sentence causes the time of utterance to be the most salient candidate reference time. In each case, a past tense sentence follows, meaning that a contextually salient past reference time needs to be accommodated. However, in a typical context, this past time can be accommodated in (6) but not in (7) or (8). In (6), where the reference time and event time of the simple past "she came by car" are the same, it is easy to accommodate the interval during which she came as a reference time. The truth conditions of the second sentence in (7) are roughly equivalent to those of (6) if the reference time for the past tense that takes scope over WOLL is the time at which Steffi departed. But (7) is infelicitous in the imagined context. It is not feasible for the speakers to accommodate Steffi's departure time as a reference time in the past when the past tense takes scope over a modal whose event time, her arrival, is later. Similarly, (8) is infelicitous, because we cannot accommodate Steffi's arrival time as a reference time for the past that takes scope over the perfect which involves an event time fifteen minutes before that.

By contrast, the second sentences of (7) and (8) can be made felicitous by the insertion of an intervening simple past sentence that shifts the reference time to the relevant earlier time.

(9) Steffi is here. She left half an hour ago. She would arrive by car fifteen minutes later.

(10) Steffi is here. She arrived fifteen minutes ago. She had left fifteen minutes earlier.

In these cases, the interpretation of the second sentence involves the accommodation of a past reference time, but it is now straightforward since it matches the event time for that sentence. Then after the acceptance of the second sentence, the reference time for the final sentence is now the contextually salient reference time and does not need to be accommodated.

The contrast between (7) and (9) is particularly relevant to our study here, since both involve the past-tense taking scope over WOLL, like the so-called subjunctive conditionals that have been considered in this paper. We see that the same pattern appears with past tense taking scope over WOLL whether or not the modal is restricted by an if-clause: it is not easy to accommodate a past reference time for a past tense that takes scope over a modal, and so such sentences are felicitous only when an intervening simple past sentence has already shifted the reference time for WOLL. This explains why even though the truth conditions of the one-past subjunctive conditional allow it to be used for past event times, it generally cannot be so used unless an intervening past sentence has already made a past reference time salient. Thus, this reading is generally not available out of the blue.

Since this rule has to do with pragmatic rules of accommodation rather than the truth conditions of the sentence, it is not without exceptions. An example is the opening sentence of Gabriel Garcia Marquez's *One Hundred Years of Solitude*:

(11) Many years later, as he faced the firing squad, Colonel Aureliano Buendía was to remember that distant afternoon when his father took him to discover ice.

The novel begins with a past-tense future-oriented claim out of the blue. But here, it is surely in part due to Garcia Marquez's flouting of this norm that the sentence is so celebrated. And even still, given our general expectation that a novel begins in the past relative to the time of narration, this sentence arguably violates the norm in question less severely than it would in the midst of a present-oriented discourse.

6 Towards a Two-Readings Theory

Thus, we have an explanation as to why the one-past form usually cannot be used for past events except where the discourse has already made a past reference time salient. A past reference time that differs from the event time is generally difficult to accommodate and requires an intervening expression to explicitly introduce the reference time.

However, this explanation poses a problem for one-reading past-as-past theories. The problem is that this explanation overgeneralizes. According to the past-as-past account, every subjunctive conditional, not merely the one-past form as used for past events, involves an instance of past tense, which requires a reference time, taking scope over WOLL. Given a historical modal base and an antecedent that is counterfactual, this reference time will need to be before the event time, at a time at which the antecedent was still historically possible. The pragmatic considerations discussed in the previous section suggest that this past reference time should be difficult to accommodate in contexts in which the salient time prior to the utterance is in the present. This would mean that these conditionals are unassertable in these contexts regardless of whether the event times are in the past. But in fact, they are easily assertable when the event times are in the present or future.

Consider again the one-past conditional (1). We have been exploring its use in contexts in which the event under consideration precedes the speech time; in those contexts, it is usually unacceptable but can be made acceptable if a prior narrative has made salient a past reference time. But now let us return to its more typical use, in a context in which the party from which Hubert is absent is ongoing at the speech time. If Hubert is absent from the party, there was some last time before which it was still historically possible that he should come to the party. Suppose that this time was an hour before the speech time, because it would have taken him an hour to arrive by any means of transportation available to him. The historical conditional "if Hubert is here, Steffi will be happy" is true relative to some times that extend up to that moment. In this context, the candidate reference times for the past tense are at least an hour prior to the event time. This suggests that interpreting the conditional should involve the accommodation of a past reference time an hour prior to the event time. As we saw in the previous section, this is generally difficult without any previous explicit reference to that time.

In fact, however, there is no difference in assertability between these two discourses.

(12) This is a nice party, but Steffi is upset. If Hubert were here, Steffi would be happy.
(13) This is a nice party, but Steffi is upset. An hour ago, Hubert decided not to come. If Hubert were here, Steffi would be happy.

In (13), there is an intervening simple past sentence before the conditional that shifts the contextually salient reference time to the past. In (12), there is no such intervening sentence. What we saw above where the party was in the past was that there was a difference in assertability between the analogous cases; only the analogue of (13) was acceptable. This was explained on the basis that we cannot accommodate a past reference time that differs from the event time of the sentence whose assertion is requiring the accommodation. But this explanation would also predict that only (13) is assertable while (12) is not.

We need an account of why the discourse in (12) is felicitous. I cannot claim to prove conclusively that no such account is available in a one-reading past-as-past framework, but I do not see one available. The solution, I propose, is that in (12), the wide-scope past tense in the conditional is interpreted modally rather than temporally. Given that it is a one-past form, the conditional does not involve any temporally interpreted instances of past tense at all. Thus, the interpretation of the conditional does not involve the accommodation of a past reference time that differs both from the event times of the conditional and from the reference times that were previously salient in the discourse.

Thus, we are pushed towards a two-readings theory according to which subjunctive conditionals have both the modal and the temporal reading. With the temporal reading, the one-past form can be used for any event time, even for events in the past. The bare truth conditions of the form allow event times that are either before or after the speech time, so long as they are not before the reference time for the past tense. But general pragmatic considerations about the difficulty of accommodating reference times that differ from event times mean that this temporal reading is generally available only where a past time already has been made salient in the discourse. With the modal reading, the one-past form is unavailable for past event times; with no temporally interpreted past tense in the sentence at all, the sentence cannot pertain to a past event. But the modal reading, unlike the temporal reading, is easily interpretable in contexts where no past time is salient, because it does not require any past reference time to be accommodated. These contexts comprise the usual contexts in which we use subjunctive conditionals for present or future events; thus, the modal reading is in a sense the default one for these contexts. So the resulting view is still closer in spirit to the overall past-as-modal framework than to the past-as-past one.

Given the methodological reasons not to multiply readings where possible, it might be seen to be undesirable to conclude that subjunctive conditionals have these two readings. But really the two-readings type of view is no more problematic from the point of view of this principle, and arguably less so, than a one-reading past-as-modal view. Given that a morpheme has two readings, it should be the default hypothesis that both readings are available in any given environment. A view in which both modal and temporal readings of the past tense in subjunctive conditionals are available is what one would expect to follow from the thesis that the past tense has both a modal and a temporal reading.

7 Conclusion

We are pushed to this view by the following considerations. The one-past form is unavailable for past events except where a past reference time, at least as far in the past as the event in question, has already been made salient. In that type of narrative context, the form is available. In a one-reading past-as-modal view, the one-past form should be unavailable for past events altogether, since the form contains no instances of temporally interpreted past tense. Thus, the availability of this reading suggests that the past tense is read temporally in these contexts. But the bare truth conditions of the past-as-past analysis would predict that this reading should be widely available, regardless of whether a past time is already salient in the context. To explain why it is only available when a past time has already been made salient in the discourse, we use the pragmatic principle—independently motivated by other data—that speakers generally do not accommodate new past reference times that differ from event times. But this principle also means that in a one-reading past-as-past theory, the same pattern of availability should apply to subjunctive conditionals more generally, not just to one-past forms for past events. This pattern, however, is not displayed: subjunctive conditionals are widely felicitous when no past reference time is salient. This suggests that subjunctive conditionals used when no past reference time is salient involve a modal reading of the wide-scope past tense.[6]

References

Abusch, D. (1997). Sequence of tense and temporal de re. *Linguistics and Philosophy*, 20(1):1–50.

Altshuler, D. (2016). *Events, states and times: An essay on narrative discourse in English*. Walter de Gruyter GmbH.

Arregui, A. (2009). On similarity in counterfactuals. *Linguistics and Philosophy*, 32(3):245–278.

Cariani, F., and Santorio, P. (2017). Will done better: Selection semantics, future credence, and indeterminacy. *Mind*, 127(505):129–165.

Copley, B. (2009). *The semantics of the future*. Routledge.

Enç, M. (1987). Anchoring conditions for tense. *Linguistic Inquiry*, 18(4):633–657.

Giannakidou, A., and Mari, A. (2018). A unified analysis of the future as epistemic modality. *Natural Language & Linguistic Theory*, 36(1):85–129.

Heim, I. (1994). Comments on Abusch's theory of tense. In Kamp, H., editor, *Ellipsis, tense and questions*, pages 143–170. University of Amsterdam.

Iatridou, S. (2000). The grammatical ingredients of counterfactuality. *Linguistic Inquiry*, 31(2):231–270.

Ippolito, M. (2006). Semantic composition and presupposition projection in subjunctive conditionals. *Linguistics and Philosophy*, 29(6):631–672.

[6] Thanks to Fabrizio Cariani, Justin Khoo, and an anonymous reviewer for their helpful comments.

Ippolito, M. (2013). *Subjunctive conditionals: A linguistic analysis*. MIT Press.
Karawani, H., Kauf, C., and Zeijlstra, H. (2019). The asymmetry of past tense. In *Proceedings of the 22nd Amsterdam Colloquium*.
Kaufmann, S. (2005). Conditional truth and future reference. *Journal of Semantics*, 22(3):231–280.
Khoo, J. (2015). On indicative and subjunctive conditionals. *Philosopher's Imprint*, 15(32).
Kratzer, A. (1981). The notional category of modality. In Eikmeyer, H.-J., and Rieser, H., editors, *Worlds, words, and contexts*, pages 38–74. Mouton De Gruyter.
Mackay, J. (2017). Past tense and past times in subjunctive conditionals. *Pacific Philosophical Quarterly*, 98:520–535.
Mackay, J. (2019). Modal interpretation of tense in subjunctive conditionals. *Semantics and Pragmatics*, 12:2.
Partee, B. H. (1973). Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy*, 70(18):601–609.
Reichenbach, H. (1947). The tenses of verbs. In *Time: From concept to narrative construct: A reader*, pages 1–12. Walter de Gruyter.
Roberts, C. (1989). Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy*, 12(6):683–721.
Romero, M. (2014). 'Fake tense' in counterfactuals: A temporal remoteness approach. *The Art and Craft of Semantics: A Festschrift for Irene Heim, ed. Luka Crnic and Uli Sauerland*, 2:47–63.
Schulz, K. (2014). Fake tense in conditional sentences: A modal approach. *Natural Language Semantics*, 22(2):117–144.
Schulz, K. (2017). Fake perfect in x-marked conditionals. *Semantics and Linguistic Theory*, 27:547–570.
von Prince, K. (2019). Counterfactuality and past. *Linguistics and Philosophy*, 42(6):577–615.

How Fake Is Fake Past?

Stefan Kaufmann

1 Introduction

Conditionals have long been a topic of interest in the semantic literature. Kratzer's treatment of *'if'*-clauses as restrictors of modal operators is the dominant approach (Kratzer 1977, 1979, 1981a, b, 1986, 1991a, b, 2012). It has its roots in philosophical logic (Goodman 1947; Stalnaker 1968; Lewis 1973, 1975, among others), and for much of its

I am grateful to the audience at the 2019 Paris Summer School on Conditionals, Bridget Copley, John Mackay, and David E. Over for helpful comments on earlier versions of this chapter. All remaining errors, omissions, or misrepresentations are my own. This work was supported in part by the National Science Foundation (#2116972, "Research on conditional and modal language") and a fellowship from the University of Connecticut Humanities Institute.

S. Kaufmann (✉)
University of Connecticut, Storrs, CT, USA
e-mail: stefan.kaufmann@uconn.edu

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_13

existence it shared with that tradition a view of modals and conditionals as propositional operators, with at best a nodding acknowledgment of the sub-clausal structure of their constituents—especially their temporal and aspectual makeup—as a factor in determining and constraining their interpretation. It was only around the turn of the millennium that semanticists working on conditionals began in earnest to pay attention to interactions between temporality and modality below the clause level. Since then, many open questions in this area have been addressed, but many still await conclusive answers.

This chapter addresses one such widely discussed but still unresolved issue: the interpretation of "fake Past" in subjunctive conditionals, as illustrated in (1).

(1) a. If the exam was tomorrow, Jill would be better prepared.
b. If the exam had been tomorrow, Jill would have been better prepared.

Both of (1a,b) are about future events but carry temporal morphology that is normally used for past reference. Iatridou (2000) observed that such a seemingly non-temporal use of Past or Perfect morphology for irrealis or counterfactual marking is rather widespread across languages. The term "fake Past" of "fake Tense" has since established itself for this use. It suggests that this temporal morphology does not have its usual temporal interpretation, but has been coopted as a mood marker of sorts. But there are competing views on the relationship between the ordinary meaning of Past and Perfect and their use in subjunctive marking.

In this chapter I aim to address this issue with an eye toward the ultimate goal of giving a unified account of indicative and subjunctive conditionals. I start by (briefly and selectively) reviewing my views on non-conditional sentences in Sect. 2 and on indicative conditionals in Sect. 3. I then turn to subjunctives and the question of how to extend the analysis to them in Sect. 4.

2 Preliminaries

I start by introducing some background about the interpretation of English tenses and the interplay between time and modality. Both topics are discussed more extensively elsewhere (Kaufmann, 2005b; Kaufmann et al., 2006). I will keep the discussion informal, emphasizing the data motivating the approach.

2.1 Time and Modality

Semantic interactions between temporality and modality often reflect an underlying distinction between an *open* future and a *fixed* past and present. This asymmetry is a property of the external world, or in any case of the way we think and talk about it (cf. Reichenbach's 1956 "most obvious properties of time"), but it has epistemic consequences. Although we may be ignorant about the past and present, that ignorance could be resolved at least in principle: the truth is "out there". In contrast, most of the future—except for the outcomes of (quasi-)deterministic processes like the movements of celestial bodies—is shrouded in uncertainty of a kind that cannot be resolved beforehand.

There are two ways to encode this formally: *branching-time* models and what I will call *filter-funnel* models (the "$T \times W$" models of Thomason 1984). I adopt the latter, schematically shown in Fig. 1. The horizontal lines stand for possible worlds, represented as linearly

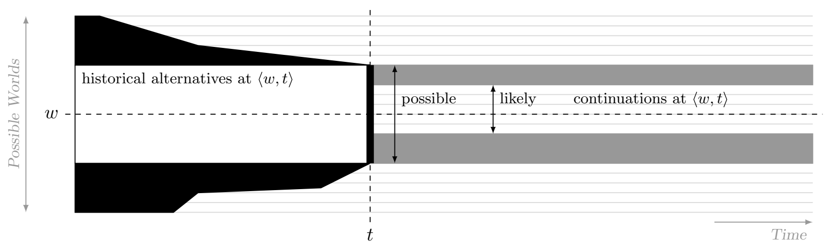


Fig. 1 The filter-funnel model of time

ordered sets of fully specified "snapshots". Each world is accompanied on its trajectory by a dwindling set of *historical alternatives*—worlds that share the same history up to a point but part ways thereafter. The idea is that the historical alternatives of world w at time t are indistinguishable (by the object language) at all points up to and including t, but may come apart at later times. In the figure, the white area encloses the worlds that are indistinguishable from each other and w at all times up to and including t. This indistinguishability is enforced in the semantics by a condition on admissible valuations of atomic sentences at world-time pairs. We say that a sentence is *settled* at $\langle w, t\rangle$ iff it is true at $\langle w', t\rangle$ for all historical alternatives w' of w at t. Thus sentences whose truth value is determined by facts no later than t, are true at $\langle w, t\rangle$ if and only if they are settled at $\langle w, t\rangle$. Not so for sentences whose truth depends on facts later than t: for those, settledness at $\langle w, t\rangle$ is a distinct and stronger notion than truth at $\langle w, t\rangle$. Based on the course of history up to t, some continuations beyond t may be more likely than others, shown in the picture by the brighter beam into the future. The thick vertical line at t marks a metaphysical modal background—the set of world-time pairs at which a sentence must be true if it is to count as settled at $\langle w, t\rangle$.

Fig. 1 shows a *metaphysical* picture. The model can be extended to also represent the changing belief state of an agent by adding, for each point in time, a set of *doxastic* alternatives—world-time pairs which, in view of the agent's beliefs at $\langle w, t\rangle$, may be the one she inhabits. The limitations on our ability to foresee the future are then incorporated by requiring that the set of doxastic alternatives must never "cut across" sets of historical alternatives.[1] To model ignorance about the past and present, the set of doxastic alternatives will typically comprise multiple sets of historical ones.

[1] Formally: if $\langle v, t\rangle$ is a doxastic alternative of $\langle w, t\rangle$ and $\langle u, t\rangle$ is a historical alternative of $\langle v, t\rangle$, then $\langle u, t\rangle$ is also a doxastic alternative of $\langle w, t\rangle$. I assume that the doxastic alternatives of $\langle w, t\rangle$, like its historical ones, all share the same time coordinate t.

2.2 Tense and the Modal *'woll'*

English has two tenses, Past and Present, illustrated in (2) and (3). I spell out the temporal interpretation informally using the framework of Reichenbach (1947), in terms of the three parameters ***S***peech time, ***R***eference time, and ***E***vent time. For simplicity, I take them to range over instants rather than intervals. S is fixed by the utterance situation, whereas R can vary within a certain range that is subject to both pragmatic factors (contextually inherited restrictions on the time the sentence is *about*) and semantic ones (tense, aspectual properties, and adverbial modifiers). The location of E relative to R is constrained by the aspectual properties of the sentence as well as pragmatic factors. I will have less to say about this relationship in this chapter, since my focus is on the reference time. Unless otherwise indicated, I assume that $E = R$.

In (2) and (3), the semantic constraints on the location of R relative to S are given on the right in terms of the *earlier than*-relation $<$.[2] R and E coincide because no aspectual operators are involved.

(2)	a.	Lisa was home yesterday.	$R < S$
	b.	Lisa is home now.	$S = R$
	c.	Lisa is home tomorrow.	$S < R$
(3)	a.	Joe cooked dinner yesterday.	$R < S$
	b.	Joe cooks dinner now.	$S < R$
	c.	Joe cooks dinner tomorrow.	$S < R$

The sentences in (2) and (3) are stative and non-stative, respectively. This difference is behind the contrast between (2b) and (3b): only with Present statives can R coincide with S; for Present non-statives, R must lie in the (near) future of S. Now, whenever this is the case, the sentence states not only that the eventuality in question occurs in the future, but also that this occurrence is already *determined* at S by "some kind of plan, schedule, control, or pattern of events" (Smith 1991, p. 246). This

[2] Here and throughout, the examples are to be read with an *episodic* interpretation in mind—that is, as referring to specific, singular instances of the eventualities in their denotation. This does not remove the habitual reading of (3b) from the purview of the account: Habitual sentences are statives, therefore on this reading (3b) lines up with (2b) in the relevant respects.

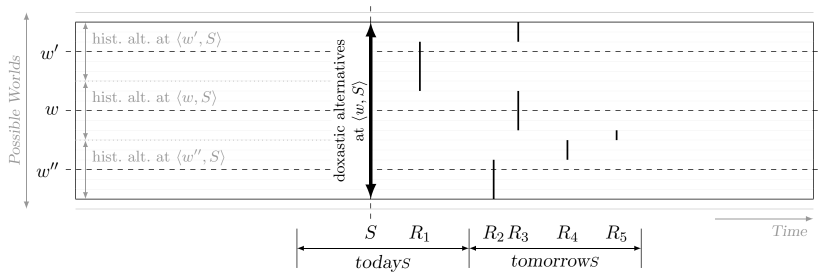


Fig. 2 Interpretation of sentence (4): true at $\langle w, S\rangle$ and $\langle w'', S\rangle$; false at $\langle w', S\rangle$

special connotation has been called the *scheduling reading*; I call it the *Certainty Condition* (CC). It is present in (2c) and in (3b,c), and absent in (2a,b) and (3a).

Thus the presence or absence of the *CC* is directly correlated with future vs. non-future reference. That difference in turn is *not* aligned with the morphological distinction between the tenses;[3] however, it does line up directly with the asymmetry between an open future and a fixed non-future encoded in the filter-funnel model. Kaufmann (2005b); Kaufmann et al. (2006) capitalized on this feature of the model in the linguistic analysis: for a tensed clause to be true at $\langle w, t\rangle$, it must be *settled* (not merely true) that there is an eventuality of the right kind in the right temporal relation to t. Formally, this is achieved by assuming the presence of a covert necessity operator accompanying the tense.[4] To illustrate, consider sentence (4) and the model shown in Fig. 2.

(4) Joe cooks lunch tomorrow.

The existence of a cooking event in the interval *'tomorrow'* is settled at $\langle w, S\rangle$ and $\langle w'', S\rangle$, thus (4) (at speech time S) is true at those worlds

---

[3] Some seek to account for the presence or absence of the CC by postulating a Future tense homonymous with the Present (Dowty 1979; Steedman 2002). But whatever the other merits of such a move may be, it would do nothing to *explain* the distribution of the CC: The choice between Present and Future would depend solely on the relation between S and R, and we will see that that relation alone is sufficient to account for the facts.

[4] See Kaufmann (2005b) for arguments that the Certainty Condition is part of semantics, not pragmatics.

as well as all of their historical alternatives. But (4) is false at $\langle w', S \rangle$ and all of *its* historical alternatives. Thus it is also not *believed* relative to the doxastic modal background comprising those three classes of historical alternatives. However, this doxastic background models an interesting attitude between ignorance and (dis-)belief: the agent believes that the question of (4)'s truth is settled, but does not know which way. Given the way the modalities interact with each other and with time in the model, this attitude is frequently implicated when tensed clauses are interpreted at speech time.[5]

This explains why in English the bare Present with future reference typically gives rise to peculiar connotations.[6] What is not yet settled cannot already be known, so speakers cannot felicitously present themselves as knowing it. Kaufmann derives from this the ready availability of the "scheduling reading". Under this reading, a sentence like (4), for instance, makes an assertion about a plan at speech time, rather than about the next day's events. Formally, the content of such a plan is represented as a modal background which enters the interpretation of the tensed sentence instead of the set of set historical alternatives. The plan is settled at speech time, hence knowable, even if the events in question are not. See Kaufmann (2005b) for details.[7]

---

[5] This last claim rests on certain assumptions about the relative scope of negation and tense, which may not be uniform across all sentences that entail (4)'s falsehood. This issue has not received much attention; but as a tentative observation, it seems that (ia) does and (ib) does not imply that the question of Joe's cooking is settled.

(i) a. Joe does not cook lunch tomorrow.
    b. It is not the case that Joe cooks lunch tomorrow.

The attitude described in the text then is the belief that either (4) or (ia) is true. This is not tautological because the logical negation of (4) is (ib) rather than (ia). Thanks to Bridget Copley for raising this issue.

[6] Not all languages have a similar constraint, even when their tense system is otherwise similar to English. For instance, the bare Present can be used much more freely to talk about the future in German or Japanese, although this use is not altogether unconstrained in either of them. In the present framework, this can be accounted for by assuming that a modal (necessity) operator is always present, but that its strength may be modulated by an ordering source, depending on the language.

[7] Later, and apparently independently, Copley (2008) argued that futurate eventives are derived statives with no future orientation. This idea seems to me to be similar in spirit, if not in detail of implementation, to the one recounted here.

In English, the auxiliary *'will'* can be used for future reference without these connotations. I assume that *'will'* is the Present-tense form of an underlying modal stem *'woll'* whose Past tense form is *'would'* (Abusch 1988, 1997, 1998; Ogihara 1995a, b). The modal *'woll'* does not contribute any temporal meaning, but its tense does. The Present imposes the same constraints on S and R as we saw above (the Perfect in (5a) and (6a) adds the requirement that E precede R; elsewhere, E and R coincide as before).

(5)	a.	Lisa will have been home yesterday.	$S = R$
	b.	Lisa will be home now.	$S = R$
	c.	Lisa will be home tomorrow.	$S < R$
(6)	a.	Joe will have cooked dinner yesterday.	$S = R$
	b.	Joe will cook dinner now.	$S < R$
	c.	Joe will cook dinner tomorrow.	$S < R$

In the modal dimension, *'woll'* is an overt counterpart of the covert necessity operator requiring settledness with bare tense. *'Woll'* differs from that operator in that it requires only *weak necessity* (Kratzer, 1981a; Kaufmann et al., 2006; Portner, 2009). The difference is, simply put, that for a sentence like (6c) to be true, *'Joe cook dinner tomorrow'* need be true only at the most normal or stereotypical alternatives. With this attenuated modal force, statements about the future can more readily be felicitous.[8]

[8] Notice also that with reference to past and present and a metaphysical modal base, sentences with *'will'* are equivalent to their unmodalized counterparts, for in this case truth at the stereotypical alternatives implies truth at all alternatives. Kaufmann (2005b) argues that it is for this reason that (5a,b) and (6a) strongly favor a doxastic reading: Since the modalized form is semantically weaker and morphologically more complex than its non-modalized counterpart, its use is disfavored when both are true, therefore its use implicates that the non-modalized form is false, which (with reference to past or present) can only be the case on a doxastic interpretation.

Condoravdi (2002) previously discussed similar cases of modals for which a metaphysical interpretation is only available with future reference. She accounts for this observation in terms of a "diversity condition" on the modal background, requiring that it contain worlds at which the prejacent is true as well as ones at which it is false. With past and present reference, this can only happen with doxastic modal bases. Condoravdi writes this requirement into the truth conditions, rather than treating it pragmatically as I do. The question of the status of the condition is an open one.

3 Indicative Conditionals

The standard assumption in linguistics is that the *'if'*-clause restricts the modal background of an operator in the consequent (Lewis 1975; Kratzer 1981a).[9] Kratzer assumes that when no overt modal is present, a covert one is inserted for the *'if'*-clause to restrict. Unlike Kratzer, I assume that an operator is always present (either overtly, for instance *'woll'*, or covertly with bare tense). The characteristic semantic property of "indicative" conditionals is that they presuppose that it is possible for the antecedent to be true.[10]

None of this says anything about the *temporal* interpretation. Once this dimension is brought into the picture, new subtleties come into relief. Consider the following simplified paraphrase:

(7) 'If *A*, *C*' is true if and only if '*C*' is true at all (relevant) points in the modal background at which '*A*' is true.

What exactly are *'A'* and *'C'*, and how are they interpreted? In the interest of compositionality it would be desirable to give them the same interpretation in conditionals as in isolation. With this in mind, challenges arise as soon as we apply the rule in (7) by interpreting *'A'* and *'C'* relative to a modal base of the kind outlined above.

3.1 Temporal Perspective

In Sect. 2.1 I argued that the modal background at $\langle w, S\rangle$ consists of world-time pairs $\langle w', S\rangle$, where w' is a historical or doxastic alternative of w at S. All of these points have the same temporal coordinate S (hence

[9] Like most contemporary approaches to conditionals, this is inspired by the *Ramsey Test* (Ramsey 1929): a conditional '*if A, C*' is evaluated against a body of information by first updating with *A*, then evaluating *C* against the result of the update.

Since I am primarily concerned with temporal interpretation, I will have little to say on what Kratzer's account predicts about the logical properties of conditionals. See Kaufmann and Kaufmann (2015) and references therein.

[10] The label "indicative" is misleading as it stands, in that some conditionals with morphologically indicative antecedents are semantically counterfactual. One way to resolve this problem might be to adjust the morphological terminology. See Schulz (2007) for such a proposal.

it is a vertical line in the above figures). But if this is correct, then the well-formedness of conditionals like (8a,b) is surprising:

(8) a. If Lisa stays at home the day after tomorrow, she met her students tomorrow.
b. If Joe cooks dinner tomorrow evening, Lisa bought groceries tomorrow morning.

If the tenses are to have their ordinary temporal meaning, the Past-tensed consequents in (8) cannot be interpreted relative to S. Doing so would result in contradictory constraints on their reference time: The tense would locate it earlier than S, while the adverb would place it later than S. But the sentences are consistent.

Intuitively, the consequent is interpreted as "past in the future", that is, by looking back from the (hypothetical) vantage point of the antecedent's reference time. The sentences in (8) show that the shift into the future does not require the auxiliary *'will'* or any other special marking. Instead, Kaufmann (2005b) argues, the shift is effected by the conditional construction itself. In terms of the Reichenbachian framework, we might say that the conditional introduces an additional parameter, a "hypothetical speech time" which I will refer to as S'. It enters the truth conditions not as a fixed parameter, but as a bound variable ranging over the modal background; see (12) below. The effect is that the background is extended from the vertical line $\langle w', S \rangle$ mentioned earlier to the rectangular area $\langle w', S' \rangle$, where w' is a historical or doxastic alternative to w at S as before, and $S \leq S'$; see Fig. 3 for a visual display. Thus the quantification is no longer purely modal, but modal-temporal.

A look at some more data shows that the introduction of S' explains not only the shift in temporal perspective for the interpretation of the tenses, but also the presence or absence of the Certainty Condition.

3.2 Scheduling Antecedents

A uniform analysis of tensed clauses inside and outside of conditionals would have it that the antecedents of (9) and (10) are the sentences in (2) and (3), repeated on the right.

(9) a. If Lisa was home yesterday, Joe cooked dinner. *[Lisa was home yesterday.]*

b. If Lisa is home now, Joe will cook dinner. *[Lisa is home now.]*

c. If Lisa is home tomorrow, Joe will cook dinner. *[Lisa is home tomorrow.]*

(10) a. If Joe cooked dinner yesterday, Lisa ate it. *[Joe cooked dinner yesterday.]*

b. If Joe cooks dinner now, Lisa will eat it. *[Joe cooks dinner now.]*

c. If Joe cooks dinner tomorrow, Lisa will eat it. *[Joe cooks dinner tomorrow.]*

But there is a difference between, say, *'Joe cooks dinner tomorrow'* in its standalone form and as the antecedent of (10c): on its most natural reading the conditional does not mean "if it is settled (now) that Joe cooks dinner tomorrow. . . " but rather "if and when turns out that Joe cooks dinner tomorrow. . . ". The difference between "if. . . now" and "if and when" is taken care of by the shift of the temporal perspective into the future. But what about the additional change from "settled" to "true"—that is, the fact that the conditional antecedent does not carry the Certainty Condition, despite its future reference? In thinking about this, it is also important to note that the CC is not always absent in future-referring antecedents:

(11) a. If Lisa stays at home tomorrow, she will meet her students later today.

b. If Joe cooks dinner tomorrow, Lisa will buy groceries later today.

Two properties of (11a,b) conspire to force the CC. First, due to the temporal adverbs the reference time of the consequent must precede that of the antecedent: $R_C < R_A$; and since there is no additional aspectual morphology involved, their event times must stand in the same relation. Second, the consequent has Present tense, thus the temporal perspective S' cannot be later than R_C. Thus $S' < R_A$ by transitivity. This triggers the CC, just as $S < R$ would in standalone sentences.

In addition to cases like (11a,b), for many future-directed antecedents a scheduling reading is available, though usually not the most prominent one out of context. It can often be brought out by setting up the context in such a way that S' is restricted to the speech time. In (10c), for instance, this reading becomes available if we assume that Joe and Jim take turns with the dinner preparation according to a schedule and, for plausibility, that Lisa likes Joe's cuisine but detests Jim's. A good paraphrase for the intended interpretation in this case would be *'If it is Joe's turn to cook dinner tomorrow. . .'*. Kaufmann (2005b) calls indicative conditionals with $S = S'$ "non-predictive" and those with $S < S'$ "predictive."

While S' behaves like a speech time with regard to the CC, there are some differences. For instance, recall from Sect. 2.2 that in standalone sentences the reference time of a non-stative cannot be equal to S. This constraint seems to apply only to the actual speech time S, not to S': the reference time of conditional antecedents can coincide with the temporal perspective. Also, not all temporal expressions are sensitive to the shift in perspective. For instance, deictic temporal frame adverbials like *'today'* and *'tomorrow'* typically remain anchored to S even in conditionals.

3.3 Truth Conditions

Kaufmann (2005b) leaves the interpretation of the tenses and the schema for the truth conditions in (7) intact and implements the forward shift in temporal perspective by allowing the conditional marker *'if'* to expand the modal base into the future before restricting it. Formally, at $\langle w, S\rangle$ the modal background becomes the set of points $\langle w', S'\rangle$ such that

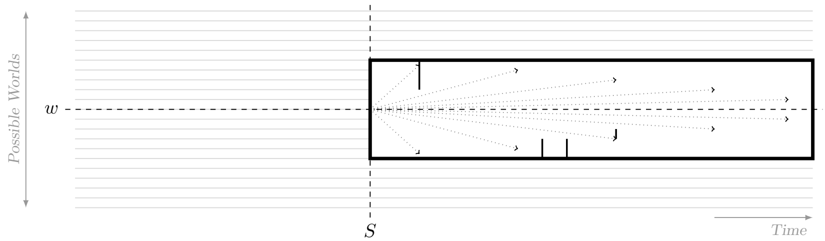


Fig. 3 Forward-extended doxastic modal base at $\langle w, S\rangle$. The set of $\langle w', S'\rangle$ that forms the modal background in evaluating conditionals at $\langle w, S\rangle$ is a subset (subject to an ordering source and possibly further contextual parameters) of the rectangle extending to the right

w' is a (metaphysical or doxastic) alternative of w at S and $S \leq S'$. In the above figures the modal background was a vertical line through $\langle w, S\rangle$; now it is "stretched" into a rectangle as in Fig. 3. The interpretation proceeds as before, by restricting this area to points at which the antecedent is true and evaluating the modal claim relative to those. In this way, both constituents can be evaluated at points in the future of S. To illustrate, consider (10c), repeated here as (12).

(12) If Joe cooks dinner tomorrow, Lisa will eat it.
True at $\langle w, S\rangle$ iff for all stereotypical points $\langle w', S'\rangle$ among the ones at which:

i. w' is an alternative of w at S,
ii. $S \leq S'$, and
iii. it is settled at $\langle w', S'\rangle$ that '*Joe cook dinner*' is true at $\langle w', R_A\rangle$ for some R_A such that $S' \leq R_A$ and R_A is 'tomorrow',

there is a time R_C such that $S' \leq R_C$ and '*Lisa eat it*' is true at $\langle w', R_C\rangle$.

Table 1 shows how Kaufmann (2005b) classifies the readings of indicative conditionals. The major dividing line runs between *non-predictive* and *predictive* conditionals, and the crucial criterion is whether or how

Table 1 Available readings for some Present antecedents from the text. The two middle columns labeled "scheduling" are where the Certainty Condition induces a scheduling reading; these cases comprise both non-predictive and predictive readings. Since *'now'* with non-statives locates the reference time in the *immediate* future, it is hard to tell whether an intermediate reading with $S < S' < R_A$ is available for (10b), hence the '?' in the corresponding cell

		Nonpredictive		*Predictive*	
		$S = S'$		$S < S'$	
			$S' \neq R_A$		
			Scheduling		
(9b)	If Lisa is home now, Joe will cook dinner.	✓	*	*	*
(9c)	If Lisa is home tomorrow, Joe will cook dinner.	*	✓	✓	✓
(10b)	If Joe cooks dinner now, Lisa will eat it.	*	✓	?	✓
(10c)	If Joe cooks dinner tomorrow, Lisa will eat it.	*	✓	✓	✓
(8a)	If Lisa stays home tomorrow, she met her students later today.	*	*	✓	✓
(11a)	If Lisa stays home tomorrow, she will meet her students later today.	*	✓	✓	*

far the range of S' extends into the future of S. This depends on contextual restrictions in addition to constraints imposed by the tenses and/or temporal adverbs.

3.4 Related Work

The observation that indicative conditionals fall into two classes according to the interpretation of their antecedents is not new. Some argue that the most important dividing line through the class of all conditionals places counterfactuals together with predictive indicatives on one side, and non-predictive indicatives on the other (Bennett 1988; Dudman 1984a, b, 1986, 1989, 1991, 1994, 1998, 2000; Funk 1985; Dancygier 1998). Others maintain that all indicatives are semantically more or less homogeneous and distinct from counterfactuals (Bennett

1995, 2003; Lewis 1973; Quirk et al. 1985). But there is little agreement in the formal semantic literature on how to characterize the difference.

Crouch (1993) introduces a distinction between the time at which a sentence is "asserted" and the time at which it is "verified"; the latter may lie in the future. In the logical form, Crouch stipulates that in non-predictive conditionals both constituents are tensed, whereas in predictive conditionals the consequent tense scopes over the antecedent, whose own tense is semantically vacuous. However, since the predictive/non-predictive distinction is not really tied to a difference in linguistic form, but rather to the largely pragmatically determined relationship between S and S', there seems to be no empirical foundation for postulating such a structural ambiguity. Garrett (2001) introduces the "time of enlightenment" as a parameter similar to Crouch's time of verification and my S'.

Schulz (2008) gives a temporal interpretation of predictive indicatives that involves a different allocation of labor between semantics and model theory. A statement about the future has no truth value until its truth value is settled. Thus Schulz's notion of "truth" corresponds to Kaufmann's settledness; in this sense, her account is in effect a supervaluationist variant of Kaufmann's. This allows her to push one layer of modal quantification (the necessity operator Kaufmann assumes for all tensed sentences) into the metalanguage. In the interpretation of conditionals, antecedents shift the interpretation to the first time at which their truth value is defined. In terms of empirical predictions, Schulz's account is similar to Kaufmann's, but it has some shortcomings. For instance, the shift into the future invariably stops at the *first* time at which the antecedent is true, which is too restrictive.[11]

[11] Here is a case in point, due to Rumberg and Lauer (2018). Suppose Sue is considering taking the train home from work tomorrow. The train runs every hour on the hour and takes half an hour. Sue does not yet know whether, in the event that she does take the train, she will catch the one at 5 p.m. or the one at 6 p.m. Then (i) is wrongly predicted to be true, even if she is no more likely to take the earlier train than she is to take the later one.

(i) If Sue takes the train, she will be home at 5:30.

Romero (2014) postulates the following skeletal meaning of the conditional construction, which contains future operators in both constituent positions:[12]

(13) MODAL [if FUT p] [then FUT q]

But this wrongly predicts that the reference times of the antecedent and consequent are independent of each other, as long as both follow the evaluation time of the modal operator. This is not right, as we saw in cases like (11), in which a scheduling reading was forced on a (Present) antecedent whose reference time was later than that of the (Present) consequent.

4 "Fake" Past and "Subjunctive" Conditionals

On their typical uses, subjunctive conditionals presuppose that their antecedent is inconsistent with the modal base (i.e., metaphysically or doxastically impossible).[13] Grammatically, English subjunctives have a Past or Past Perfect form on the modal in the consequent, which is echoed in the temporal morphology of the antecedent. This was illustrated in (1) above, repeated here; the modal is *'would'*, the Past-tense form of *'woll'*.

(1) a. If the exam was tomorrow, Jill would be better prepared.
b. If the exam had been tomorrow, Jill would have been better prepared.

The last couple of decades saw significant advances in our understanding of the compositional semantics of subjunctives, at least in English. But

[12] Romero (2014) is mostly concerned with counterfactual conditionals, to which I turn below; but she justifies the future operators for the constituents explicitly with the forward shifting observed in indicative conditionals (see her Footnote 3).

[13] It is well-known that this does not hold in full generality, but for now I focus on these cases. I briefly turn to "Anderson conditionals," the best-known class of exceptions, below.

open questions remain, and a unified theory of all conditionals is still out of reach. Building on the framework presented in the preceding sections, I am going to sketch what such a unified theory might look like.

4.1 Expanding the Modal Background

If the antecedent is only true at worlds outside the modal background, the modal background must be adjusted to make them accessible.[14] The Past or Perfect marking on subjunctives, in languages which employ this device, presumably plays a role in this operation. There are two major schools of thought on how this works and how to model it in a framework of modal-temporal interaction such as the filter-funnel model introduced above. On the *Past-as-Modal (PaM)* view, the Past is "redirected" from the *temporal* dimension in which it normally enables reference to different times, to the *modal* dimension, now enabling reference to different possible worlds (James 1982; Fleischman 1989; Dancygier 1998; Iatridou 2000; Schulz 2007, 2014; Mackay, 2015). Thus it is non-temporal, hence "fake." On the *Past-as-Past (PaP)* view, antecedent worlds are accessed from an *earlier* time at which they were still alternatives, in effect "re-running" history from that earlier time. Thus the Past is temporal, in a sense, after all. Variants of this view have supporters in philosophy (Dudman 1984a; Edgington 1995, and many others), linguistics (Tedeschi 1981; Kaufmann 2005a; Ippolito 2006, 2013; Arregui, 2007), and psychology (Over et al., 2007).[15]

The two options are depicted schematically in Fig. 4. Two remarks are in order before we move on. First, while the pictures in Fig. 4 resemble the portrayal of historical alternatives in Fig. 1 more than that of doxastic ones in Fig. 2, the PaM/PaP distinction concerns both. Doxastic states have histories, too, and under the idealizing assumption that agents only

[14] Notice that the adjustment at issue here is not the forward extension that was crucial in the analysis of predictive indicatives in the last section. That forward extension was not tied to subjunctive mood or counterfactuality. It is assumed to be operative in all conditionals.

[15] This "re-running" of history does not undo all facts after the relevant time, but only those that are causally affected by the antecedent (Kaufmann 2005a, see also Over and Cruz, this volume). My description glosses over this important detail because I am mainly concerned with temporal reference.

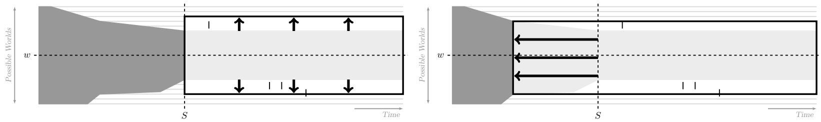


Fig. 4 Past-as-Modal (left) and Past-as-Past (right) expansion of the modal background

ever accumulate information without forgetting or revising any (a highly simplistic assumption that is nonetheless commonly made in dynamic semantics, for instance), the temporal changes in doxastic alternatives resemble those in historical ones, the main difference being that information about events and states is not necessarily acquired in the order in which they occur. Moreover, subjunctive conditionals can have doxastic interpretations,[16] which I assume work quite similarly to the objective or metaphysical ones more widely discussed.

Secondly, while an expansion of the modal background is necessary when the antecedent is not true at any accessible worlds, there are uses of fake Past in which this particular explanation is not (clearly) applicable. Lewis (1973, p. 4) noted that "there are subjunctives pertaining to the future, like [14] that appear to have the truth conditions of indicative conditionals".

(14) If our ground troops entered Laos next year, there would be trouble.

Lewis did not elaborate on the claim that (14) has the truth conditions of an indicative. Presumably he had in mind that its antecedent, while unlikely at the time, could not be entirely ruled out. This would be what has come to be called a *Future-Less-Vivid (FLV)* interpretation (Iatridou 2000).[17] Assuming that fake Past is generally associated with an expansion of accessibility, it would seem that on this use the set that is

16 Although this has been disputed at times in the literature (Veltman 2005, i.a.), it has been established repeatedly and conclusively (Veltman 1985; Rott 1999; Schulz 2007, among others).

17 Iatridou attributes the term to English-language grammars of Classical Greek (Fn. 5). She also notes that the form can be used to (emphatically) express ignorance regarding the event in question, a use which she argues is derived from the counterfactual meaning. I do not explore this matter further in this chapter.

being expanded does not include all (historical or doxastic) alternatives, but perhaps only the most prominent (e.g., stereotypical or likely) ones. This has in fact been proposed (Schulz 2014). The two options schematized in Fig. 4 would then apply to the smaller set of prominent worlds. This needs to be sorted out, but as far as the application of the formal mechanism is concerned, examples like (14) do not seem to pose special challenges.

Another case in which fake Past is not (obviously) justified by a lack of antecedent worlds in the modal background is illustrated by (15) from Anderson (1951).

(15) If Jones had taken arsenic, he would have shown just exactly those symptoms which he does in fact show.

This conditional can naturally be used to argue *for* the truth of the antecedent, thus it cannot be said to presuppose that its antecedent is false. However, an expansion of the modal base is nonetheless involved in its interpretation. Intuitively, (15) implies that the observed symptoms would be *explained* by John's having taken arsenic (if he did). Without going too far into how exactly this idea might be fleshed out, suffice it to say that theories of explanation generally assume that the explanandum is a *consequence* of the explanans, in some (specialized and restricted) sense of the term (Hempel and Oppenheim 1948; Gärdenfors 1988; Salmon 1989; Halpern and Pearl 2005; Halpern 2016, among many others). To establish this relationship, the truth of the explanandum must, in some way or other, be ignored. This seems to be (at least part of) what drives the expansion of the modal background in cases like (15).

With these caveats in place, I turn to a question on which there is so far no consensus: Which of the two expansion operations is involved in the interpretation of subjunctives?

4.2 Two Types of Subjunctives

The literature on English recognizes two kinds of subjunctives, *Simple Past (SP)* and *Past Perfect (PP)*:

(16) a. If Lisa were home {now/tomorrow}, Joe would cook dinner. *[SP]*
b. If Lisa had been home {now/tomorrow}, Joe would have cooked dinner. *[PP]*

These terms are due to Ippolito (2003, 2006, 2013), who observed that while both imply that Lisa is not home at the relevant time (*'now'* or *'tomorrow'*), (16a) also implies that her being home is nonetheless "possible" in a certain sense: that Lisa is alive and therefore *could be* home tomorrow. If Lisa is (known to be) dead, (16a) is odd but (16b) is fine. Under Ippolito's analysis, the contrast has to do with the antecedent's *presuppositions* and the time at which they must be or must have been satisfied: at speech time in SP conditionals and at an earlier time in PP conditionals. For examples like (16), Ippolito assumes that *'Lisa is home'* presupposes that Lisa is alive.

Ippolito's observation shed new light on the finer semantic details of conditionals, but I believe her conclusions from contrasts like (16) are not quite correct. The presuppositions of the antecedent surely play a role, but they do not determine the judgments.

4.3 Time of Intervention

There are two kinds of counterexamples to Ippolito's account, one of which seems to me to pose a serious challenge. To see this, consider one of her own examples. Her account correctly predicts that (17a) is infelicitous in the given context, whereas (17b) is fine.[18]

(17) *John was training for the Boston Marathon last summer when he unexpectedly died.*
a. #If John ran the Boston Marathon next spring, he would win.
b. If John had run the Boston Marathon next spring, he would have won.

[18] The hash marks ('#') indicate infelicity. The sentences thus marked are neither grammatically ill-formed nor semantically deviant in any way, and they may well be true in the given story (when interpreted at an appropriate time). However, they are ill-fitting within their immediate linguistic context (given above).

However, the same antecedent, in the same context, gives rise to different judgments when paired with a different consequent:

(18) a. If John ran the Boston Marathon next spring, that would be a miracle.
b. #If John had run the Boston Marathon next spring, that would have been a miracle.

Perhaps the felicity of (18a) could be explained away in terms of local accommodation of the presupposition that John is alive, or some such mechanism. However, no such story would suffice to explain the oddness of (18b), as far as I can see.

The explanation I would like to propose can be spelled out formally in a couple of different ways, but the basic intuition is simple. John is dead. His running the marathon next spring requires a departure from reality. Exactly how this departure is to be made is not fully specified: the antecedent can be true for different reasons, in different ways. There are worlds at which John did not die last year, and there are others at which he did die but returns to life miraculously in time to run the marathon. The antecedent is true at all of these worlds (and more). Which ones are more similar to ours? That question may be impossible to answer, but fortunately progress does not depend on having the answer, as Stalnaker (1968) and Lewis (1973, 1979) have long since argued. World similarity is vague and context-dependent, yet there is no problem in pairing a precise semantic theory with this murky parameter. Once the vagueness is resolved (or sufficiently reduced), we can get clear judgments about truth and entailment. Naturally, we would expect those judgments to depend on *how* the vagueness is resolved.

What the above sentences show is that judgments about which form is most suitable to express a counterfactual claim can also depend on how the vagueness is resolved. All antecedent worlds differ from ours. The more similar ones have a great deal of overlap with ours and depart from it only in a minimal way, in some sense of "minimal". One can think of this departure as a *miracle* in the sense of Lewis (1979), or as an *intervention* in the sense of Pearl (2000), and the details of the implementation will depend on this choice. Schulz (2017) made a proposal that bears some similarities to my own (but also differs in important respects), and described it in terms of intervention. I am also partial to the interventionist stance (Kaufmann 2001, 2005a, 2013, i.a.) and will use that terminology in the following, although I will not introduce more formal apparatus. The idea is simple. The truth or falsehood of the antecedent is jointly determined by a set of causally relevant variables and causal laws regulating the (in)dependencies between them. Reasoning about what would be if the antecedent were true involves a disturbance in this ensemble of relevant facts—suspending causal laws, changing or un-setting the values of variables. In most cases the variables are tied to the goings-on at specific times (exceptions are timeless truths like those of mathematics); given the asymmetry of causal dependence, those times generally precede the antecedent's reference time.[19] But their relation to the speech time can vary freely, and different interventions before or after the speech time may constitute alternative ways to make the required change. The SP/PP distinction depends on the location of the time of departure relative to the speech time.

As additional evidence for this view, consider again the same sentences, this time in a slightly different context.

John was training for the Boston Marathon last summer when he suddenly fell ill. His health deteriorated gradually, as did his prognosis. Finally no hope was left, and he passed away.

(18) a. If John ran the Boston Marathon next spring, that would be a miracle.
b. If John had run the Boston Marathon next spring, that would have been a miracle.

Now (18b) is much improved. The reason is, I submit, that the kind of departure most naturally associated with this context would be a miraculous recovery before his (actual) death, rather than resurrection after

[19] Pearl's (2000) *'do'* operator involves cutting a variable off from all its parents, then setting it to the desired value. In linguistic and cognitive reality, counterfactual hypotheses often do affect the values of the parents (Sloman and Lagnado 2005; Dehghani et al. 2012); cf. also the notion of "causal ramp" in the philosophical literature (Lewis, 1979; Mårtensson 1999; Bennett 2003).

death. The crucial difference from the above instance is that this hypothetical recovery lies in the past. (18a), meanwhile, is also felicitous but invokes a different type of departure, *viz.* resurrection after death and, more pertinently, after the speech time.

Still further evidence for the proposal comes from "ahistorical" counterfactuals whose antecedents are not clearly associated with any particular past time. Consider (19) from Mizuno and Kaufmann (2018).[20]

(19) a. If 9 were even, it would be divisible by 2.
b. #If 9 had been even, it would have been divisible by 2.

Under Ippolito's account, it is not clear what the relevant presuppositions of the antecedent would be, or why those presuppositions should be satisfied at present but not in the past. Moreover, the infelicity of (19b) is of a peculiar sort: the sentence strongly suggests that there is a past time at which the number 9 *became* odd. The sentence feels odd because such an assumption conflicts with what we believe about mathematics. Note, however, that in the right context such a historical reading becomes available and improves the felicity of (19b): just consider a doxastic reading, speculating on what the speaker—a child, say—would have believed if she had learned that Nine was even. Again, it is unclear how these data could be explained in terms of presuppositions and when they are satisfied.

The upshot is that the time of the intervention is the crucial factor in choosing between the SP and PP forms of counterfactuals. Ippolito was right in observing that PP subjunctives do and SP subjunctives do not call for revisiting the past. But what drives this revisit is not the need to check the antecedent's presuppositions. Rather, it is the need to intervene on the actual history at the right time to make room for the truth of the antecedent. This is always in the past when the antecedent's reference time lies in the past, so it is not surprising that subjunctives about the past always are PP. But future subjunctives may also require a PP form, namely when the intended intervention lies in the past.

[20] Mizuno and Kaufmann are mainly concerned with the Japanese counterparts of (19a,b). Those data are beyond the scope of this paper; I claim without proof that the Japanese patterns of Past marking on counterfactuals fit into the overall framework proposed here.

4.4 Time of Intervention and Temporal Perspective

Stepping back to look at the bigger picture, the next question is how the time of intervention is related to the other temporal parameters. This also requires getting clear on the location of S'. Consider the following context.

> **Context.** *John is supposed to attend a conference this weekend. He has been trying to decide on a travel plan. He likes long road trips, so driving across the country is tempting. Today is Wednesday. The conference starts on Friday.*

Recall that indicatives were subject to the constraint that $S \leq S'$—that is, the temporal perspective must not precede the actual speech time. We will see that this also holds for SP subjunctives. Notice first that the indicatives in (20) are both felicitous, with the (temporal) Past in (20b) enabling reference to the past from the non-past perspective S'.

(20) a. If John leaves on Thursday, he will arrive on Sunday.

$$S \leq S';\ S' \leq R_A;\ S' \leq R_C$$

b. If John left on Tuesday, he will arrive on Friday.

$$S \leq S';\ R_A < S';\ S' \leq R_C$$

The SP counterpart of (20a) is (21a), formed by adding a layer of Past morphology to both constituents. Under a fake-Past analysis of this element, it marks an intervention. I assume that for SP subjunctives this intervention occurs at speech time S.

(21) a. If John left on Thursday, he would arrive on Sunday.

$$S = S_I \leq S';\ S' \leq R_A;\ S' \leq R_C$$

b. #If John left on Tuesday, he would arrive on Friday.

c. If John had left on Tuesday, he would arrive on Friday.

$$S = S_I \leq S';\ R_A < S';\ S' \leq R_C$$

The competing hypothesis that the Past is temporal could be spelled out in two ways. One is that it scopes over the entire conditional construction, which then receives a forward-looking interpretation relative to some past time, call it $S_C < S$. This is, roughly, Ippolito's (2013) account (modulo the checking of presuppositions at speech time, which in terms of its role in the theory corresponds to my intervention no earlier than the speech time). Under this assumption, one could maintain that the relevant constraint on temporal perspective was $S_C < S'$ (the SP counterpart to $S \leq S'$ for indicatives). The other way to implement a temporal Past would embed it in the conditional construction, resulting in a past-in-the-future reference to Thursday from a forward-shifted S'. But both of these ideas face a challenge from (21b), which without further stipulations would be expected to be felicitous. The stipulation could, for instance, be that the reference time of the antecedent must not precede the speech time. This is in fact spelled out in terms of "double access" by Romero (2014). But that proposal has related difficulties with PP subjunctives (see below).

In fact, the SP counterpart of (20b) is (21c), whose antecedent has two layers of Past morphology: a temporal past in addition to (and embedded under) its SP marking. This temporal past, just like the one in (20b), is interpreted relative to S', which in turn does not precede the speech time. That such a temporal Past is needed to realize reference to the past indicates that S' cannot precede S.

Thus the SP subjunctives in (21a,c) are interpreted exactly like the indicatives in (20a,b), modulo the Past marking an intervention at S'. But this intervention-marking Past does not have any temporal import. It is fake.

Incidentally, it is possible, though perhaps somewhat marginal, to have a temporal Perfect in the consequent of a conditional like those in (22).[21] They are interpreted as indicated on the right. The relationships between the speech and reference times are the same as those in (20); the only difference is the additional constraint that the event time

[21] It would be preferable in this case to change the temporal adverb to *'by Sunday'/'by Friday'*. I refrain from making this change to keep minimal pairs.

of the consequent must precede its reference time, yielding a result-state interpretation for the consequent.

(22) a. If John leaves on Thursday, he will have arrived on Sunday.
$S \leq S'; S' \leq R_A; S' \leq R_C; E_C < R_C$

b. If John left on Tuesday, he will have arrived on Friday.
$S \leq S'; R_A < S'; S' \leq R_C; E_C < R_C$

These examples and their SP counterparts do not pose any special challenges. The conditionals in (23a,c) are the SP counterparts of (22a,b), just as (21a,c) were to (20a,b); here as there, the difference is a layer of fake Past.

(23) a. If John left on Thursday, he would have arrived on Sunday.

c. If John had left on Tuesday, he would have arrived on Friday.

PP subjunctives differ from their SP counterparts in having one more layer of Past/Perfect morphology. This raises anew the question of PaM vs. PaP, for there is no reason to assume *a priori* that this question gets the same answer for both layers of Past. We saw above that PP subjunctives are felicitous whenever the relevant intervention targets eventualities in the past; if there is no such construal, the PP form is infelicitous, as in the "timeless" conditional in (19). This indicates that the second layer of Past is not fake, calling for a Past-as-Past analysis.

So if the PP Past shifts a temporal parameter back, which parameter is it? It could be S', resulting in the condition that $S' < S$. But this is unlikely in view of the fact that PP subjunctives about the future do not get a scheduling reading: the antecedent of (24a) is not correctly paraphrased as *'if it had been settled (in the past) that John leaves on Thursday, ...'*. Rather, a better paraphrase is *'if John's leaving on Thursday had not been prevented (by past events) and he had left on Thursday, ...'*.

Following this intuition, I assume that what lies in the past is the time of intervention, as shown in (24a). I also assume that (24b) is interpreted in the same way. Thus unlike (20b) and (21c), (24b) does not have a temporal Past in the antecedent.

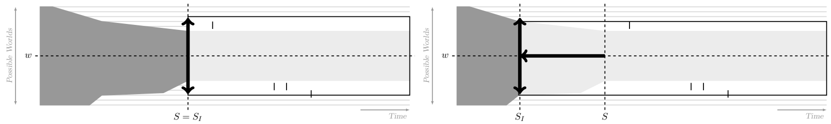


Fig. 5 SP (left) and PP (right) expansion of the modal background

(24) a. If John had left on Thursday, he would have arrived on Sunday.
$S_I < S; S_I \leq S'; S' \leq R_A; S' \leq R_C$

b. If John had left on Tuesday, he would have arrived on Friday.
$S_I < S; S_I \leq S'; S' \leq R_A; S' \leq R_C$

On this account, the second layer of Past is temporal, although it is not interpreted like the other tenses in the sentence, for it regulates the relationship between S and S_I, not between either of those and the reference times of the constituents. This is in line with the commonly made assumption that this Past takes scope above the conditional operator (Ippolito 2013; Romero 2014).[22]

In the wider context of recent research on fake Past, the present proposal is aligned with an overall view according to which the PaM/PaP distinction does not distinguish between classes of sentences—that is, neither are all subjunctives summarily PaM or PaP, nor does the distinction track the SP/PP distinction. Instead, the PaM/PaP distinction draws the line between the different layers of Past morphology used to mark English subjunctives: all subjunctives have a layer of fake Past, and PP subjunctives have an additional layer of temporal Past. These layers lend themselves to a PaM analysis and a PaP analysis, respectively. The first part of this view, a PaM analysis for the layer common to SP and PP subjunctives, has been argued for by Schulz (2014) and Mackay (2019).

[22] The assumption is that there is only one Past operator, even though in English it is expressed on both constituents. This morphological reflex is ensured via feature-passing mechanisms in the syntax. I am not going to go into the details here. In languages which do not mark tense on the antecedent, such as Japanese, there is only one Past marking in PP subjunctives (Mizuno and Kaufmann 2018).

A PaP analysis for the second layer was tentatively suggested by Mackay (2019).[23]

Singling out a Time of Intervention as an additional temporal parameter departs from earlier proposals about the SP/PP distinction. I already discussed Ippolito's in introducing my proposal. I should also mention a refinement of Ippolito's account which seeks to correct some of its problematic predictions. Romero (2014) argues, based on ideas found in Iatridou (2000) and Arregui (2009), that both the future SP (21a) and the past PP (21b) have essentially the same Past-as-Past interpretation, as forward-looking conditionals from the perspective of a past time at which modal accessibility is determined. The difference between them lies solely in the reference times of the constituents relative to the speech time: PP and SP subjunctives are about past and non-past times, respectively. One prediction of this approach is that PP is required when the antecedent refers to the past. This is also predicted by my account, since an antecedent in the past entails a departure in the (still earlier) past. Another prediction of Romero's account is that SP is required when the antecedent does not refer to the past. We already saw in connection with Ippolito's examples that this is incorrect.

4.5 Temporal Past Without Intervention

I argued above that PP subjunctives involve a layer of temporal Past atop a layer of fake Past; the latter is shared with SP subjunctives. One might wonder whether the outer temporal Past is somehow dependent on there being a modal Past in its scope, or whether it can embed indicatives as well. Pursuing this question pushes against the page limit, but I do want to at least mention one class of examples which suggest that embedding indicatives is an option.

Consider the following series of sentences in the given context, which is a slight modification of the earlier one.

[23] This view is also consistent with cross-linguistic evidence: for instance, Mizuno and Kaufmann (2018) show that the Japanese counterparts of PP subjunctives are marked with a Past for which a PaP analysis is called for, whereas the Japanese counterparts of English subjunctives have no Past marking.

Context. *John has been agonizing over his conference travel plans this week. Today is Wednesday. On Monday he was pondering his options.*

(25) a. If he left on Tuesday, he would arrive on Friday.
b. #If he had left on Tuesday, he would { arrive / have arrived } on Friday.

This pattern does not seem to fit well with the earlier discussion: (25a,b) were listed as (21b,c) above but with the opposite felicity judgments. (25a) has the form of an SP subjunctive, similar to (21a), but such a reading cannot be straightforwardly attributed to it as uttered on Wednesday. Moreover, in contrast to typical SPs, nothing is implied about the truth or likelihood of the antecedent, on either Monday or Wednesday. Instead, intuitively (25) relates John's deliberations (on Monday) from the perspective of his inner monologue. Linguistically, this is known as *Free Indirect Speech* (FIS; Eckardt 2014). (26) lists simple sentences for which an FIS interpretation is either the only option or strongly preferred out of context, due to their peculiar combinations of tenses and temporal adverbs.

(26) a. Tomorrow was Tuesday.
b. He realized now that he would leave on Tuesday.

It is characteristic of FIS that it involves a "local" context whose temporal perspective (also called *origo*) coincides with the time at which the thought or speech in question occurred (here: Monday, the time of the "pondering"). In FIS, certain indexical expressions (e.g., adverbs like *'now'* and *'tomorrow'*) are shifted to the local context, while others remain anchored to the actual context (e.g., the pronoun *'I'*). Crucially, tenses are interpreted relative to the actual speech time, not the time of the pondering. This makes (25a) a special kind of Past-tense indicative, true (as FIS) on Wednesday if (27) was what John thought on Monday.[24]

[24] It should be noted that speakers find the variant in (ia), with Perfect in the consequent, also acceptable. This looks like a "mixed" form in which an SP antecedent is paired with a PP consequent. However, that is not necessarily the right analysis: the Perfect in the consequent may simply have its ordinary temporal interpretation. On this analysis,

(27) If I leave on Tuesday, I will arrive on Friday.

The infelicity of the PP subjunctive (25b) is also peculiar. Speakers agree that it is odd in the given context, but also that it would be the preferred way to state the facts from the speaker's perspective at the speech time (Wednesday), as in (24b).

Usages like (25) have not been discussed widely in the literature on conditionals (however, they are taken up by John Mackay in this volume). I set them aside here because they are not directly relevant: their tense marking is peculiar, but it is not fake Past. Still, discussing it here was helpful, I hope, because it shows that non-SP conditionals can be embedded under a temporal Past, and the result is not a PP conditional.

5 Conclusions

Current work in the semantics of conditionals seeks to ground their considerable interpretational versatility in their morphosyntactic makeup in unified and compositional ways. This goal is still some ways off, but much progress has been made toward a better understanding of their temporal interpretation. This paper offered a unified account of indicative and subjunctive conditionals which captures their semantic variability in terms of a few simple parameters. Further work will have to establish whether and to what extent this approach generalizes to other uses of fake Past in English and to conditionals across languages.

---

just as (25a) is the Past-tense counterpart of (27a), so too is (ia) the Past counterpart of (ib).

(i) a. If he left on Tuesday, he would have arrived on Friday.
b. If he leaves on Tuesday, he will have arrived on Friday.

References

Abusch, D. (1988). Sequence of tense, intensionality and scope. In Borer, H., editor, *Proceedings of WCCFL 7*, pages 1–14. CSLI.

Abusch, D. (1997). Sequence of tense and temporal de re. *Linguistics and Philosophy*, 20:1–50.

Abusch, D. (1998). Generalizing tense semantics for future contexts. In Rothstein, S., editor, *Events and Grammar*, volume 70 of *Studies in Linguistics and Philosophy*, pages 13–33. Kluwer.

Anderson, A. R. (1951). A note on subjunctive and counterfactual conditionals. *Analysis*, 12:35–38.

Arregui, A. (2007). When aspect matters: The case of *would*-conditionals. *Natural Language Semantics*, 15:221–264.

Arregui, A. (2009). On similarity in counterfactuals. *Linguistics and Philosophy*, 32:245–293.

Bennett, J. (1988). Farewell to the Phlogiston theory of conditionals. *Mind*, 97(388):509–527.

Bennett, J. (1995). Classifying conditionals: The traditional way is right. *Mind*, 104(414):332–354.

Bennett, J. (2003). *A philosophical guide to conditionals*. Oxford University Press.

Condoravdi, C. (2002). Temporal interpretation of modals: Modals for the present and for the past. In Beaver, D. I., Casillas, L., Clark, B., and Kaufmann, S., editors, *The construction of meaning*, pages 59–88. CSLI Publications.

Copley, B. (2008). Temporal orientation in conditionals. In Guéron, J. and Lecarme, J., editors, *Time and modality*, pages 59–77. Springer.

Crouch, R. (1993). *The temporal properties of English conditionals and modals*. PhD thesis, University of Cambridge.

Dancygier, B. (1998). *Conditionals and prediction*. Cambridge University Press.

Dehghani, M., Iliev, R., and Kaufmann, S. (2012). Causal explanation and fact mutability in counterfactual reasoning. *Mind & Language*, 27(1):55–85.

Dowty, D. (1979). *Word meaning and Montague grammar*, volume 7 of *Studies in linguistics and philosophy*. Reidel.

Dudman, V. H. (1984a). Conditional interpretations of *if*-sentences. *Australian Journal of Linguistics*, 4:143–204.

Dudman, V. H. (1984b). Parsing 'if'-sentences. *Analysis*, 44(4):145–153.

Dudman, V. H. (1986). Antecedents and consequents. *Theoria*, 52:168–199.

Dudman, V. H. (1989). Vive la Révolution! *Mind*, 98(392):591–603.
Dudman, V. H. (1991). Interpretations of 'if'-sentences. In Jackson, F., editor, *Conditionals*, chapter 10, pages 202–232. Oxford University Press.
Dudman, V. H. (1994). On conditionals. *Journal of Philosophy*, 91:113–128.
Dudman, V. H. (1998). On the grammar of conditionals: Reply to Barker. *Analysis*, 58:277–285.
Dudman, V. H. (2000). Classifying 'conditionals': The traditional way is wrong. *Analysis*, 60:147.
Eckardt, R. (2014). *The semantics of Free Indirect Discourse: How texts allow us to mind-read and eavesdrop*. Number 31 in Current Research in the Semantics/Pragmatics Interface. Brill.
Edgington, D. (1995). On conditionals. *Mind*, 104(414):235–329.
Fleischman, S. (1989). Temporal distance: A basic linguistic metaphor. *Studies in Language*, 13(1):1–50.
Funk, W.-P. (1985). On a semantic typology of conditional sentences. *Folia Linguistica*, 19(3/4):365–414.
Gabbay, D. and Guenthner, F., editors (1984). *Extensions of classical logic*, volume 2 of *Handbook of Philosophical Logic*. D. Reidel.
Gärdenfors, P. (1988). *Knowledge in flux: Modeling the dynamics of epistemic states*. MIT Press.
Garrett, E. J. (2001). *Evidentiality and assertion in Tibetan*. PhD thesis, University of California.
Goodman, N. (1947). The problem of counterfactual conditionals. *The Journal of Philosophy*, 44:113–128.
Halpern, J. Y. (2016). *Actual causality*. MIT Press.
Halpern, J. Y. and Pearl, J. (2005). Causes and explanations: A structural-model approach. Part II: Explanation. *The British Journal for the Philosophy of Science*, 56(4):889–911.
Hempel, C. G. and Oppenheim, P. (1948). Studies in the logic of explanation. *Philosophy of Science*, 15(2):135–175.
Iatridou, S. (2000). The grammatical ingredients of counterfactuality. *Linguistic Inquiry*, 31(2):231–270.
Ippolito, M. (2003). Presuppositions and implicatures in counterfactuals. *Journal of Semantics*, 11:145–186.
Ippolito, M. (2006). Semantic composition and presupposition projection in subjunctive conditionals. *Linguistics and Philosophy*, 29(6):631–672.
Ippolito, M. (2013). *Subjunctive conditionals: A linguistic analysis*. The MIT Press.

James, D. (1982). Past tense and the hypothetical: A cross-linguistic study. *Studies in Language*, 6(3):375–403.

Kaufmann, M. and Kaufmann, S. (2015). Conditionals and modality. In Lappin, S. and Fox, C., editors, *The handbook of contemporary semantic theory*, pages 237–270. John Wiley & Sons, Ltd.

Kaufmann, S. (2001). *Aspects of the meaning and use of conditionals*. PhD thesis, Stanford University.

Kaufmann, S. (2005a). Conditional predictions: A probabilistic account. *Linguistics and Philosophy*, 28(2):181–231.

Kaufmann, S. (2005b). Conditional truth and future reference. *Journal of Semantics*, 22:231–280.

Kaufmann, S. (2013). Causal premise semantics. *Cognitive Science*, 37:1136–1170.

Kaufmann, S., Condoravdi, C., and Harizanov, V. (2006). Formal approaches to modality. In Frawley, W., editor, *The expression of modality*, pages 71–106. Mouton de Gruyter.

Kratzer, A. (1977). What 'must' and 'can' must and can mean. *Linguistics and Philosophy*, 1:337–355.

Kratzer, A. (1979). Conditional necessity and possibility. In Egli, U., Bäuerle, R., and von Stechow, A., editors, *Semantics from different points of view*, pages 117–147. Springer.

Kratzer, A. (1981a). The notional category of modality. In Eikmeyer, H.-J. and Riesner, H., editors, *Words, worlds, and contexts*, pages 38–74. Walter de Gruyter.

Kratzer, A. (1981b). Partition and revision: The semantics of counterfactuals. *Journal of Philosophical Logic*, 10:201–216.

Kratzer, A. (1986). Conditionals. In Farley, A. M., Farley, P. T., and McCullough, K.-E., editors, *Proceedings of CLS 22, Part 2: Papers from the Parasession on Pragmatics and Grammatical Theory*, pages 1–15.

Kratzer, A. (1991a). Conditionals. In von Stechow and Wunderlich (1991), pages 651–656.

Kratzer, A. (1991b). Modality. In von Stechow and Wunderlich (1991), pages 639–650.

Kratzer, A. (2012). *Modality and conditionals*. Oxford University Press.

Lewis, D. (1973). *Counterfactuals*. Harvard University Press.

Lewis, D. (1975). Adverbs of quantification. In Keenan, E., editor, *Formal semantics of natural language*, pages 3–15. Cambridge University Press.

Lewis, D. (1979). Counterfactual dependence and time's arrow. *Noûs*, 13:455–476.

Mackay, J. (2015). Actuality and fake tense in conditionals. *Semantics and Pragmatics*, 8(12):1–12.

Mackay, J. (2019). Modal interpretation of tense in subjunctive conditionals. *Semantics and Pragmatics*, 12(2):1–29.

Mårtensson, J. (1999). *Subjunctive conditionals and time: A defense of the classical approach*. Number 10 in Acta Philosophica Gothoburgensia. Department of Philosophy, University of Göteborg.

Mellor, D. H., editor (1990). *Philosophical papers: F. P. Ramsey*. Cambridge University Press.

Mizuno, T. and Kaufmann, S. (2018). Past-as-past in Japanese counterfactuals. In Ronai, E., Stigliano, L., and Sun, Y., editors, *Proceedings of CLS 54*, pages 309–324.

Ogihara, T. (1995a). Double-access sentences and reference to states. *Natural Language Semantics*, 3:177–210.

Ogihara, T. (1995b). The semantics of tense in embedded clauses. *Linguistic Inquiry*, 26:663–679.

Over, D., Hadjichristidis, C., Evans, J., Handley, S., and Sloman, S. (2007). The probability of causal conditionals. *Cognitive Psychology*, 54:62–97.

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge University Press.

Portner, P. (2009). *Modality*. Oxford University Press.

Quirk, R., Greenbaum, S., Leech, G., and Svartvik, J. (1985). *A comprehensive grammar of the English language*. Longman.

Ramsey, F. P. (1929). General propositions and causality. Printed in Mellor (1990), pages 145-163.

Reichenbach, H. (1947). *Elements of symbolic logic*. Free Press, New York, NY.

Reichenbach, H. (1956). *The direction of time*. University of Berkeley Press.

Romero, M. (2014). 'Fake tense' in counterfactuals: A temporal remoteness approach. In Crnič, L. and Sauerland, U., editors, *The art the craft of semantics: A Festschrift for Irene Heim*, volume 71 of *MIT Working Papers in Linguistics*, pages 47–63. MIT. https://semanticsarchive.net/Archive/jZiNmM4N/.

Rott, H. (1999). Moody conditionals: Hamburgers, switches, and the tragic death of an American president. In Gerbrandy, J., Marx, M., de Rijke, M., and Venema, Y., editors, *JFAK. essays dedicated to Johan van Benthem on the occasion of his 50th birthday*. Vossiuspers AUP, Amsterdam, The Netherlands. http://www.illc.uva.nl/j50/contribs/rott.

Rumberg, A. and Lauer, S. (2018). What if, and when? Conditionals, tense, and branching time. Presentation at the Workshop "The Many What Ifs", University of Konstanz.

Salmon, W. C. (1989). *Four decades of scientific explanation.* University of Minneapolis Press, Minneapolis.

Schulz, K. (2007). *Minimal models in semantics and pragmatics: Free choice, exhaustivity, and conditionals*. PhD thesis, University of Amsterdam.

Schulz, K. (2008). Non-deictic tenses in conditionals. In Friedman, T. and Ito, S., editors, *Proceedings of SALT 18*, pages 694–710, Ithaca, NY. Cornell University.

Schulz, K. (2014). Fake tense in conditional sentences: A modal approach. *Natural Language Semantics*, 22:117–144.

Schulz, K. (2017). Fake Perfect in X-marked conditionals. In Burgdorf, D., Collard, J., Maspong, S. and Stefánsdóttir, B., editors, *Proceedings of SALT 27*, pages 547–570.

Sloman, S. A. and Lagnado, D. A. (2005). Do we 'do'? *Cognitive Science*, 29:5–39.

Smith, C. S. (1991). *The parameter of aspect*, volume 43 of *Studies in linguistics and philosophy*. Kluwer.

Stalnaker, R. (1968). A theory of conditionals. In *Studies in logical theory, American philosophical quarterly, Monograph: 2*, pages 98–112. Blackwell.

von Stechow, A. and Wunderlich, D., editors (1991). *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung. [= Semantics]*. Walter de Gruyter.

Steedman, M. (2002). The productions of time. Unpublished ms., Version 4.1.

Tedeschi, P. (1981). Some evidence for a branching-futures semantic model. In Tedeschi, P. and Zaenen, A., editors, *Tense and aspect*, volume 14 of *Syntax and Semantics*, pages 239–270. Academic Press.

Thomason, R. H. (1984). Combinations of tense and modality. In Gabbay and Guenthner (1984), pages 135–165.

Veltman, F. (1985). *Logics for conditionals*. PhD thesis, University of Amsterdam.

Veltman, F. (2005). Making counterfactual assumptions. *Journal of Semantics*, 22:159–180.

Counterfactual Hypothetical vs. Biscuit Conditionals: A Semantic/Pragmatic Analysis of Their Morphological Differences

Eva Csipak and Maribel Romero

1 Introduction

Hypothetical and biscuit conditionals differ intuitively in their meaning. While intuitively *hypothetical* conditionals convey that the truth of the consequent depends on the truth of the antecedent, as exemplified in (1),

This research has been supported by the RU1614 "What if?" funded by the Deutsche Forschungsgemeinschaft (DFG) (RO 4247-1/2). We thank the audiences of the General Meeting of RU1614, the seminar Ling315 at the University of Konstanz (WiSe2019), the workshop UniMod1 at Leiden University, the Linguistics Colloquium at the Institut Jean-Nicod, and SALT 2019.

E. Csipak (✉)
Department of Linguistics, University of British Columbia,
Vancouver, BC, Canada
e-mail: eva.csipak@ubc.ca

M. Romero
Department of Linguistics, University of Konstanz, Konstanz, Germany
e-mail: maribel.romero@uni-konstanz.de

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_14

biscuit conditionals are taken to convey the truth of the consequent in the actual world w_0, regardless of the truth or falsity of the antecedent, as in (2):

(1) If Peter went shopping, there are biscuits on the sideboard. HYP

(2) There are biscuits on the sideboard if you want them. BISCUIT (Austin, 1956)

The above examples illustrate that indicative conditionals appear with the same verbal morphology no matter whether they are interpreted as hypothetical or as biscuit conditionals.[1] More explicitly, in the case of indicative conditionals, English and Spanish have 'normal', non-fake tense morphology in both antecedent and consequent clause, and Spanish additionally has indicative mood morphology, independently of interpretation. This can be seen in (3)–(6): Both (3), a hypothetical conditional, and (4), a biscuit conditional, show simple present tense morphology; the Spanish versions (5) and (6) additionally have indicative morphology.

(3) (On whatsapp: I know you well...)
If you are hungry right now, your stomach **is** growling. HYPIND

(4) If you are hungry right now, there **is** pizza in the fridge. BIIND

(5) Si (tú) tienes hambre, tu estómago está gruñendo. HYPIND
If (you) have.Ind hunger, your stomach **is.Ind** growling.

(6) Si (tú) tienes hambre, hay pizza en el frigorífico. BIIND
If (you) have.Ind hunger, **have.Ind** pizza in the fridge

This parallel in appearance has led many authors to propose a unified syntactic and semantic analysis for hypothetical and biscuit conditionals (Franke 2009; Francez 2015; Lauer 2015; Csipak 2018; Biezma and Goebel 2018; Goebel 2017, a.o.). In a nutshell, they propose that the syntactic and semantic modal template is the same for hypothetical and biscuit conditionals (contra e.g. Ebert et al. 2014). For this line of

[1] This is also independent of whether they appear with overt or covert epistemic or metaphysical modals.

approach, the difference between hypothetical and biscuit conditionals lies purely in the pragmatics: the 'biscuit' interpretation comes about as a pragmatic inference arising from the notion of *conditional independence*. This gives rise to the speaker intuition that the consequent is being asserted of the actual world w_0.

To see how this works in one such analysis, consider Franke (2009). He assumes a standard semantics for an indicative conditional *If A then C*: A $\subseteq$ C. For a hypothetical conditional, A and C are *conditionally dependent*, i.e., upon learning that one of them is true, we may change our belief about the other. For example, learning that Peter in fact went shopping in (1) may cause us to believe that there are biscuits on the sideboard. For a biscuit conditional, A and C are *conditionally independent*, i.e., learning that one of them is true will not change our belief about the other. Consider Franke's example (7).

(7) If you are hungry, there is pizza in the fridge.

Intuitively, learning whether the addressee is hungry or not will not change our beliefs about whether or not there is pizza in the fridge (and vice versa).

The following pragmatic reasoning then ensues for biscuit conditionals, which by definition have conditionally independent antecedents A and consequents C: A cannot possibly affect the truth of C (given conditional independence), yet the speaker uttered both. Therefore the speaker must wish to commit to C in the actual world w_0 regardless of the truth of falsity of A. This gives rise to the 'biscuit' interpretation.[2]

---

[2] The more formal derivation of the pragmatic inference, following Franke (2009), proceeds as follows:

(i) a. The Speaker's epistemic state allows her to utter *If A, C*.

b. But A and C are conditionally independent from each other according to the Speaker's epistemic state.

c. The Speaker must either believe the falsity of A or the truth of C. Otherwise $\Diamond(A \cap \overline{C})$, which contradicts the Speaker's belief that *If A, C* because $[\![\text{if A, C}]\!] = A \subseteq C$.

d. Non-triviality: The Speaker believing the falsity of A would make the statement *If A, C* trivial ($\emptyset \subseteq C$). Since non-triviality is assumed, the Speaker must believe C.

However, the case is different for counterfactual conditionals. Here we observe a difference in the consequent clause verbal morphology between hypothetical and biscuit conditionals. In both English and Spanish, they differ with respect to tense: For hypothetical counterfactuals, we must use an 'extra' layer of past morphology (*fake tense*), giving rise to the form *would* in (8a)/(10a), and crucially we cannot use 'real' tense, as illustrated in (8b)/(10b). But for biscuit counterfactuals, this is (typically) reversed: we have to use non-fake tense, as in (9b)/(11b), and cannot use fake tense, witness (9a)/(11a) (Csipak 2015; pace Franke 2009).[3] Moreover, Spanish biscuit counterfactuals must additionally have indicative mood in the consequent clause, as in (11b).

(8) a. If you were hungry right now, your stomach **would be** growling. HypCF
b. # If you were hungry right now, your stomach is growling.

(9) a. # If you were hungry right now, there would be pizza in the fridge.
b. If you were hungry right now, there **is** pizza in the fridge. BiCF

(10) a. Si (tú) tuvieses hambre, tu estómago estaría gruñendo. HypCF
If (you) had.Subj hunger, your stomach **would.be** growling.
b. # Si (tú) tuvieses hambre, tu estómago está haciendo ruidos.
If (you) had.Subj hunger, your stomach is.Ind growling.

(11) a. # Si (tú) tuvieses hambre, habría pizza en el frigorífico.
If (you) had.Subj hunger, would.have pizza in the fridge.
b. Si (tú) tuvieses hambre, hay pizza en el frigorífico. BiCF
If (you) had.Subj hunger, **have.Ind** pizza in the fridge.

Thus, the puzzle is as follows. We observe an interpretive difference between hypothetical and biscuit conditionals that occurs in both indicatives and counterfactuals. If this difference is purely pragmatic (i.e., due to the posited independence-based inferencing mechanism), why must the two conditional types be expressed with different morphology—i.e.,

[3] We note that the judgments reported here come from native speakers of English. Note that languages like German and Italian allow the form parallel to (9a) to receive a biscuit interpretation for independent reasons.

with different tense and mood in the consequent—when they occur in counterfactual form?

The goal of the present paper is to present a first comprehensive analysis of tense and mood morphology in HypCFs and BiCFs that derives the morphological pattern in (8)–(11) while maintaining the general uniform approach to hypotheticals and biscuits.

To this end, we will follow the temporal remoteness analysis of counterfactual morphology (Dudman 1983, 1984; Ippolito 2003; Grønn and von Stechow 2009; Romero 2017) and extend mechanisms independently needed for breaking Sequence of Tense in attitude reports (Ogihara 1999), as in Romero and Csipak (2019). The main contribution of the present paper is to present pragmatic arguments for why we see the forms that we do, and only those. In particular, we will focus on why the unattested forms are ruled out by pragmatic mechanisms.

The remainder of the paper is organized as follows. Section 2 lays out some necessary background, including treatments of Sequence of Tense, Subjunctive mood and so-called double-access readings where Sequence of Tense is broken. The proposal, partially building on previous work by the authors, follows in Sect. 3, in four parts. In Sect. 3.1, we summarize the implementation of the temporal remoteness analysis of grammatical hypothetical counterfactuals in Romero (2017). Section 3.2 presents Romero and Csipak (2019)'s analysis for breaking Sequence of Tense and 'Sequence of Mood' in conditionals to account for the grammatical biscuit counterfactuals we observe. In Sect. 3.3, we rule out unattested biscuit counterfactuals by appealing to competition between forms. Finally, in Sect. 3.4 we rule out unattested hypothetical counterfactuals by appealing to the Maxim of Manner. Section 4 concludes.

2 Background on Tense and Mood

In order to account for counterfactual hypothetical and biscuit conditionals that we observe, we need some formal background on the interpretation of their constitutive morphological ingredients.

First, let us consider the counterfactual hypothetical conditionals below: (12)–(13) are present counterfactuals and (14)–(15) are past counterfactuals.[4]

(12) If you **were** hungry right now, your stomach **would be** growling.

(13) Si (tú) tuvieses hambre ahora, tu estómago estaría gruñendo.
If (you) **had.**SUBJ hunger now, your stomach **would.be** growling

(14) If you **had been** hungry yesterday, your stomach **would have been** growling.

(15) Si (tú) hubieses tenido hambre ayer, tu estómago
If (you) **had.**SUBJ **had** hunger yesterday, your stomach
habría estado gruñendo.
would.have been growling

(At least) two pieces of verbal morphology are involved in these forms[5]: (a) there is a layer of so-called **'fake' past tense** in the antecedent and consequent in English and Spanish; and (b) the antecedent clause appears in the **subjunctive mood** in Spanish.

The layer of 'fake' tense has received two analyses in the literature: It is interpreted modally in the modal remoteness approach (Iatridou 2000; Schulz 2014) and temporally in the temporal remoteness approach (Dudman 1983; Grønn and von Stechow 2009; Romero 2017, a.o.). We follow the temporal approach. The central idea, stemming from Dudman (1983), is that a counterfactual with 'fake' tense involves a back shift in time with a future (metaphysical) conditional interpreted under that back shift, as schematized in (16). 'Fake' tense morphology then follows from Sequence of Tense, independently needed for complement clauses in English and Romance, as we will see in Sect. 2.1.

(16) PAST [MODAL$_{\text{METAPHY}}$ [if (FUT) A] [then FUT C]]

[4] Counterfactuality is a defeasible inference in the Spanish (13) and (15), just as in the English (12) and (14) (Lewis 1973; Anderson 1951). We leave aside Severe Tense Mismatch cases (Ippolito 2003).

[5] See Anand and Hacquard (2009) and Ferreira (2016) on the role of aspectual morphology.

For the subjunctive mood in the Spanish antecedent clauses, we follow Schlenker (2005) and interpret mood as imposing a restriction on the world pronoun, as independently argued for Romance complement clauses. We will briefly introduce the formalism in Sect. 2.2.

Second, let us consider the counterfactual biscuit conditionals in (17)–(18):

(17) If you **were** hungry right now, there **is** pizza in the fridge.

(18) Si (tú) tuvieses hambre, hay pizza en el frigorífico.
If (you) **had.Subj** hunger, **have.Ind** pizza in the fridge.

The most striking morphological features of these examples are the following: (a) while there is a layer of 'fake' past tense in the antecedent clause, there is no 'fake' past tense in the consequent clause; and (b), in Spanish, while the antecedent clause is in the Subjunctive, Indicative mood is found in the consequent clause. In other words, there appears to be a disconnect in the time line and the modal sphere between the antecedent clause and the consequent clause in counterfactual biscuit conditionals. To properly analyse this disconnect, Sect. 2.3 will examine so-called "double-access" readings in complement clauses, where a similar temporal disconnect between the matrix and the complement clause has been previously observed.

2.1 Tense and Sequence of Tense

Consider the attitude report in (19). This sentence is ambiguous between a reading corresponding to the past-over-past direct report in (20a) and a reading corresponding to the past-over-present direct report in (20b). Under the latter reading, the past tense morphology on the embedded verb *was* goes seemingly uninterpreted, a phenomenon known as 'Sequence of Tense' (Abusch 1997; Kusumoto 2005; von Stechow 2009).

(19) Annalea said (last week) that Lucía **was** sick.

(20) a. Annalea said (last week): "Lucía was sick". Past-over-Past
b. Annalea said (last week): "Lucía is sick". Past-over-Present

Let us see how these two readings are derived. Syntactically, the following ingredients have been proposed in the literature. First, (interpretable) tense morphology is treated as a pronoun pro_i (Partee 1973, a.o.) with a temporal feature relative to an anchor time pronoun pro_j (von Stechow 1995; Abusch 1997; Kusumoto 2005, a.o.). In our LFs, the temporal feature and its anchor will appear superscripted after pro_i, e.g. $\text{pro}_i^{[\textsc{past}\ \text{pro}_j]}$. Second, one layer of past temporal morphology may optionally be left uninterpreted when licensed in a chain headed by a temporal pronoun with an interpretable past feature (Ogihara 1995; Kusumoto 1999; Grønn and von Stechow 2009). In our LFs, uninterpreted morphology will appear crossed out, e.g. ~~past~~, and replaced with the default temporal feature [PRES pro_j]. This optionality in dealing with embedded past morphology leads to the two potential LFs in (21):

(21) LFs of (19):
a. $\lambda 0\ \exists_1$ [Annalea think at $\text{pro}_1^{[\text{PAST pro}_0]}$
$\lambda 2\ \exists_3$ [Lucia be sick at $\text{pro}_3^{[\text{PAST pro}_2]}$]] Past-over-Past
b. $\lambda 0\ \exists_1$ [Annalea think at $\text{pro}_1^{[\text{PAST pro}_0]}$
$\lambda 2\ \exists_3$ [Lucia be sick at $\text{pro}_3^{[\cancel{\text{past}}][\text{PRES pro}_2]}$]] Past-over-Present

Semantically, temporal features are interpreted as imposing presuppositions on the value of the pronoun (Heim 1994; Kratzer 1998), as defined in (22)–(24). Furthermore, we treat the value of a temporal/mood pro_i as a world-time pair, i.e., as an index, with temporal and accessibility constraints understood as in (25):

(22) $[\![pro_i^{[\text{PAST } pro_j]}]\!]^g$ is defined only if g(i) < g(j); if defined, $[\![pro_i^{[\text{PAST } pro_j]}]\!]$ = g(i)

(23) $[\![pro_i^{[\text{PRES } pro_j]}]\!]^g$ is defined only if g(i) ∘ g(j); if defined, $[\![pro_i^{[\text{PRES } pro_j]}]\!]$ = g(i)

(24) $[\![pro_i^{[\text{FUT } pro_j]}]\!]^g$ is defined only if g(j) < g(i); if defined, $[\![pro_i^{[\text{FUT } pro_j]}]\!]$ = g(i)

(25) a. For any two indices $<w,t>$ and $<w',t'>$:
$<w,t> < <w',t'>$ iff w=w′ and t is prior to t′.
$<w,t> \circ <w',t'>$ iff w=w′ and t and t′ overlap.

b. For any two indices $<w,t>$ and $<w',t'>$:
$<w,t> \in \text{MOD}(<w',t'>)$ iff t=t′ and w′ is accessible from w via MOD.

The two LFs above then lead to the two sets of truth conditions in (26). In both formulas, $\exists i_3$ ranges over indices i_3 which share the world-coordinate with i_2 and whose time-coordinate is in a particular relation to the temporal coordinate of i_2: it precedes it in (26a), leading to the past-over-past reading, and it overlaps with it in (26b), resulting in the past-over-present reading.

(26) Truth conditions of (19) :

a. $\lambda i_0.\ \exists i_1[\ i_1 < i_0 \wedge \forall i_2 \in Dox_{Annalea}(i_1)$
$\exists i_3\ [i_3 < i_2 \wedge \text{Lucía be sick at } i_3]\ ]$ Past-over-Past

b. $\lambda i_0.\ \exists i_1[\ i_1 < i_0 \wedge \forall i_2 \in Dox_{Annalea}(i_1)$
$\exists i_3\ [i_3 \circ i_2 \wedge \text{Lucía be sick at } i_3]\ ]$ Past-over-Present

2.2 Subjunctive Mood

Spanish and other Romance languages present a mood divide in the complement clauses of attitude verbs: representational verbs like *pensar* 'think' select Indicative, as in (27), while non-representational verbs like *lamentar* 'regret' select Subjunctive, as in (28):

(27) Bea piensa [que Juan enseña / *enseñe semántica]
Bea thinks [that Juan teaches.IND / *teaches.SUBJ semantics]
'Bea thinks that Juan teaches semantics.'

(28) Bea lamenta [que Juan *enseña / enseñe semántica]
Bea regrets [that Juan *teaches.IND / teaches.SUBJ semantics]
'Bea regrets that Juan teaches semantics.'

We follow Schlenker (2005)'s analysis of mood morphology, featuring the following ingredients. First, mood morphology introduces a mood feature on the world pronoun, again represented as a superscript on the pronoun in our LFs, e.g. $\text{pro}_i^{[\text{IND } pro_k]}$. Second, the features IND(icative) and SUBJ(unctive) are relative to a pronoun pro_k that picks up the so-called "local context" (in the sense of Stalnaker 1975): For root clauses, $[\![pro_k]\!]$ equals the Common Ground (CG); for embedded complement clauses, $[\![pro_k]\!]$ (roughly) equals $\text{Dox}_x(w_0)$ of the attitude holder x. Finally, the feature IND imposes a presupposition on the value of the world pronoun whereas the feature SUBJ imposes no presupposition, as defined in (29)–(30):

(29) $[\![pro_i^{[\text{IND } pro_k]}]\!]$ is defined only if $g(i) \in g(k)$;
if defined, $[\![pro_i^{[\text{IND } pro_k]}]\!] = g(pro_i)$

(30) $[\![pro_i^{[\text{SUBJ } pro_k]}]\!] = g(pro_i)$

When we combine these lexical entries with the rest of the complement clause in (27)–(28), we obtain the partial function (31) for the Indicative clause and the total function (32) for its Subjunctive counterpart (where x is the attitude holder):

(31) $[\![\textit{Juan teach semantics at } pro^{[\text{IND} pro_k]}]\!] =$
$\lambda w'\!: w' \in \text{Dox}_x(w_0).$ J teaches sem in w' IND-proposition

(32) $[\![\textit{Juan teach semantics at } pro_i^{[\text{SUBJ } pro_k]}]\!] =$
$\lambda w'\!: \cancel{w' \in \text{Dox}_x(w_0)}.$ J teaches sem in w' SUBJ-proposition

Let us briefly see how the inherent semantics of the relevant attitude verbs leads to the observed selection pattern in (27)–(28).

In the case of *think*, the (standard) lexical entry in (33) simply asks us to check the value of our proposition at the worlds $w \in Dox_x(w_0)$. For that, the partial IND-proposition (31) suffices. To that, we add Heim (1991)'s principle *Maximize Presupposition!* in (34)[6]:

(33) $[\![think]\!](p)(x) = \lambda w_0.\ \forall w \in Dox_x(w_0): p(w)$

(34) *Maximize Presupposition!*: Make your contribution presuppose as much as possible! (Heim 1991)

Given this principle, the maximally presuppositional IND-proposition not only *can* be used, but it also *must* be used. Hence, *think* can take the IND-proposition and cannot take the SUBJ-proposition, as we saw in (27).

In the case of *regret* we have the lexical entry (35) (adapted from Heim 1992's *be glad*). The idea is that, for each world $w \in Dox_x(w_0)$, we compare in terms of desirability the world w^p most similar to w where p is true—which is w itself—and the world $w^{\neg p}$ most similar to w where $\neg p$ is true—namely, $Sim_w(rev_p(Dox_x(w_0))+\neg p)$.

(35) $$[\![regret]\!](p)(x) = \lambda w_0{:}\forall w \in Dox_x(w_0)\ [p(w)].\ \forall w \in Dox_x(w_0)\ [Sim_w(rev_p(Dox_x(w_0))+\neg p) >_{Bou_x(w_0)} w]$$

[6] To see a simple example illustrating Heim (1991)'s *Maximize Presupposition!* at work, consider the choice between the indefinite article *a* in (i) and the definite article *the* in (ii). The indefinite article expresses existence in the truth conditional content and carries no preposition, whereas the definite article expresses existence truth-conditionally but, in addition, carries the uniqueness presupposition that the set denoted by its syntactic sister is a singleton. Since, given world knowledge, the uniqueness presupposition in (ii) is satisfied—there is only one (relevant) sun—, the presuppositionally heavier *the* has to be used, the choice of the non-presuppositional *a* in (i) leading to infelicity.

(i) # A sun is shining.

(ii) The sun is shining.

.

More concretely, $Sim_w(\psi)$ ask us to find the most similar world w′ to w for which $\psi(w')$ yields TRUE, where ψ is the result of revising $Dox_x(w_0)$ with respect to p and updating it with ¬p.

Now, if we take *p* to be total SUBJ-proposition (32), the expression $Sim_w(rev_p(Dox_x(w_0))+ \neg p)$ will be defined.[7] But, if we take *p* to the partial IND-proposition (31) instead, the expression will be undefined.[8] Hence, *regret* must combine with a SUBJ-proposition and cannot combine with an IND-proposition, as we saw in (28).

2.3 Breaking Sequence of Tense

When Sequence of Tense is broken in attitude reports by using an absolute tense, e.g. English present tense in (36), we obtain a so-called "double-access" temporal reading: The time of the embedded proposition must align both with the utterance time t_0, as paraphrased in (36a),

[7] The formal expression $Sim_w(rev_p(Dox_x(w_0))+\neg p)$ instructs us, first, to temporarily revise $Dox_x(w_0)$ with respect to p, as defined in (i). If we take SUBJ-p, the (temporarily) revised SUBJ-$p_{SUBJ\text{-}p}(Dox_x(w_0))$ will contain worlds in which John teaches semantics and worlds in which John does not teach semantics. Then, in a second step, (...)+¬p asks us to update the result of this revision with ¬p, resulting in a set containing only worlds where John does not teach semantics. In a final step, Sim_w ask us to look at the worlds within this updated revised doxastic state and to select the world w′ most similar to w. In general, since the revised and updated doxastic state is not empty, it will be possible to find a world w′ most similar to w_0. Hence, when using SUBJ-p, the formal expression will be defined and the semantic derivation of sentence (28) can proceed.

(i) For any context c and proposition p:
$rev_p(c) = \cup\{X \subseteq W: c \subseteq X \text{ and } X+p \text{ is defined}\}$

[8] Consider again the formal expression $Sim_w(rev_p(Dox_x(w_0))+\neg p)$, this time using IND-p. First, we need to temporarily revise $Dox_x(w_0)$ with respect to IND-p. The (temporarily) revised IND-$p_{IND\text{-}p}(Dox_x(w_0))$ will contain only worlds in which John teaches semantics, as the original $Dox_x(w_0)$ did. Second, we need to update the result of this revision with ¬p, which results in an empty doxastic state (contradiction). Finally, Sim_w ask us to look at the worlds w′ within this empty updated revised doxastic state and to select the world w′ most similar to w. But, since there is no world in that epistemic state, it is impossible to select one. This means that, when using IND-p, the formal expression is undefined and, thus, the semantic derivation of (28) cannot be carried out.

and with the attitude holder's subjective "now" t_1, as in (36b) (Abusch 1997; Ogihara 1999). The same facts holds for Spanish.

(36) John said Mary is pregnant.

a. John said at a past time t_1 that Mary is pregnant at t_0.

b. John said at a past time t_1 that Mary is pregnant at t_1.

However, using the lexical entry (37) for absolute present tense produces the LF (38), which only gives us temporal alignment of t_4 with t_0, as in (39a). To obtain the desired alignment with t_1, Ogihara (1999) proposes an analysis (very much simplified here!) where the temporal property is duplicated and linked to t_2 as well, as underlined in (39b):

(37) $[\![pro_i^{[\text{PRES } pro_0]}]\!]^g$ is defined only if $g(i) \circ g(0)$;
if defined, $[\![pro_i^{[\text{PRES } pro_0]}]\!] = g(i)$

(38) LF: $\lambda 0.\ \exists_1$ [John say at $\text{pro}_1{}^{[\text{PAST pro}_0]}\ \lambda 3\ \exists_4$
[Mary be pregn. at $\text{pro}_4{}^{[\text{PRES pro}_0]}$]]

(39) a. $\lambda t_0.\exists t_1\ [t_1 < i_0 \wedge \forall t_2 \in \text{SAY}_{john}(t_1){:}\ \exists t_4[t_4 \circ t_0 \wedge$
Mary be pregnant at t_4]]

b. $\lambda t_0.\ \exists t_1\ [t_1 < i_0 \wedge \forall t_2 \in \text{SAY}_{john}(t_1){:}\ \exists t_4[t_4 \circ t_0 \wedge$
Mary be pregnant at $t_4 \wedge \underline{t_4 \circ t_2}$]]

We would like to take the time here to point out that breaking Sequence of Tense only works when there is one continuous interval that is talked about. That is, when reporting that John said that Mary is pregnant, it must be *the same* pregnancy that John and the speaker are talking about, even if John was talking about the pregnancy a month ago and the speaker is talking about it now. (36) cannot be used in scenario where five years ago, Mary was pregnant and John talked about this, and now Mary is pregnant again, and the speaker wishes to convey both that John talked about Mary's pregnancy in the past, and that Mary is currently pregnant again. Using Sequence of Tense in this way in order to sneak in a 'by the way' observation is disallowed.

3 Proposal

We are ready to go back to the contrast between hypothetical and biscuit counterfactual conditionals. The crucial differences between the two counterfactual types can be recapitulated as follows.

In hypothetical counterfactuals, the consequent clause must contain 'fake' tense, both in English and in Spanish. This is shown in (40)–(41): While the (a)-versions with 'fake' tense are grammatical, the (b)-versions with no 'fake' tense in the consequent clause are unacceptable under the hypothetical reading:

(40) a. If you were hungry right now, your stomach **would be** growling. HypCF

b. # If you were hungry right now, your stomach **is** growling.

(41) a. Si (tú) tuvieses hambre, tu estómago estaría gruñendo.
If (you) had.Subj hunger, your stomach **would.be** growling. HypCF

b. # Si (tú) tuvieses hambre, tu estómago está(/esté) gruñendo.
If (you) had.Subj hunger, your stomach **is.Ind(/is.Subj)** growling.

In biscuit counterfactuals, by contrast, the consequent clause should contain no 'fake' tense, both in English and in Spanish. This can be seen in (42)–(43): While the (a)-versions with 'fake' tense are deviant as biscuits, the (b)-versions without it are perfect. Furthermore, the consequent clause must appear in indicative mood in Spanish, as shown in (43b):

(42) a. # If you were hungry right now, there would be pizza in the fridge.

b. If you were hungry right now, there **is** pizza in the fridge. BiCF

(43) a. # Si (tú) tuvieses hambre, habría pizza en el frigorífico.
If (you) had.Subj hunger, would.have pizza in the fridge.

b. Si (tú) tuvieses hambre, hay(/*haya) pizza en el frigo BiCF
If (you) had.Subj hunger, **have.Ind** (/***Subj**) pizza in the fridge.

To cover the entire morphological pattern, the following four points need to be accounted for.

First, why does 'fake' tense in the consequent clause make good hypothetical counterfactuals? Here we will adopt the concrete implementation of the temporal remoteness approach proposed in Romero (2017), summarized in Sect. 3.1.

Second, why does the lack of 'fake' tense and the use of indicative mood in the consequent make good biscuit counterfactuals? Following Romero and Csipak (2019), we will propose in Sect. 3.2 that, in these forms, we are breaking Sequence of Tense and, additionally for Spanish, we are breaking what could be called 'Sequence of Mood'; that is, we are doing double access readings at the same time in the temporal and modal domain. In lack of a fully worked-out analysis of double access readings over indices (i.e., <time,world>-pairs), we will extend our simplified version of Ogihara's (1999) idea as a stop-gap solution.

Third, why does 'fake' tense in the consequent clause and, additionally for Spanish, non-Indicative mood make biscuit counterfactuals deviant? In other words, why does maintaining Sequence of Tense and Sequence of Mood in (42a)/(43a) make bad biscuit counterfactuals? We will argue in Sect. 3.3 that pragmatic competition between the relevant forms rules out the unacceptable options.[9]

Fourth and finally, why does the lack of 'fake' tense and the use of indicative in the consequent make hypothetical counterfactuals unacceptable, as in (40b)/(41b)? We will sketch a potential solution in Sect. 3.4 based on the Gricean Principle of Manner. In particular, we will argue that when a speaker is in a position to break Sequence of Tense/Mood when trying to utter a hypothetical counterfactual, the Maxim *Be Brief!* demands that they not utter a conditional at all, but rather plain *q*, the consequent.

3.1 Grammatical Hypothetical Counterfactuals

We start with the grammatical hypothetical conditionals in (44) and (45), which, as we saw, carry a layer of 'fake' tense in English and Spanish and appear in Subjunctive mood in Spanish:

[9] This is true of the typical cases; we discuss exceptions in Sect. 3.3.

(44) If you **were** hungry right now, your stomach **would be** growling. (=(40a))

(45) Si (tú) tuvieses hambre ahora, tu estómago estaría gruñendo.
If (you) **had.**SUBJ hunger now, your stomach **would.be** growling
'If you were hungry now, your stomach would be growling.' (=(41a))

We have now the necessary ingredients for an analysis of the tense and mood morphology in these conditional forms. On the one hand, we have the general LF structure (46) assumed for hypothetical counterfactuals in the temporal remoteness approach (Dudman 1983; Grønn and von Stechow 2009; cf. Ippolito 2003). This includes a back shift in time—represented with PAST in (46)—with a future indicative conditional embedded under that shift. For the sake of concreteness, we assume that the future indicative conditional is headed by a silent modal with a metaphysical modal base METAPHY and a stereotypical ordering source L (cf. Kaufmann 2005), represented as $\text{MODAL}^{\text{L}}_{\text{METAPHY}}$ in (46):

(46) PAST [$\text{MODAL}^{\text{L}}_{\text{METAPHY}}$ [if (FUT) A] [then FUT C]]

On the other hand, we have the lexical entries for the relevant pieces of tense and mood morphology that we saw in Sects. 2.1 and 2.2:

(47) $[\![pro_i^{[\text{PAST } pro_j]}]\!]^g$ is defined only if g(i) < g(j);
if defined, $[\![pro_i^{[\text{PAST } pro_j]}]\!]$ = g(i)

(48) $[\![pro_i^{[\text{FUT } pro_j]}]\!]^g$ is defined only if g(j) < g(i);
if defined, $[\![pro_i^{[\text{FUT } pro_j]}]\!]$ = g(i)

(49) $[\![pro_i^{[\text{IND } pro_k]}]\!]$ is defined only if g(i) $\in$g(k);
if defined, $[\![pro_i^{[\text{IND } pro_k]}]\!]$ = g(pro_i)

(50) $pro_i^{[\text{SUBJ } pro_k]}$ = g(pro_i)

Extending previous analyses, Romero (2017) combines these two sets of ingredients to build the LF below for our examples. The back shift in time is represented by $pro_1^{[\text{PAST pro}_0]}$. This (covert) pronoun introduces an index i_1 whose temporal coordinate precedes that of the utterance index i_0 and at which the future indicative conditional headed by the modal $\text{MODAL}^{L}_{\text{METAPHY}}$ is evaluated. At the same time, the pronoun $pro_1^{[\text{PAST } pro_0]}$, having an interpretable past feature, allows for the past tense morphology in the antecedent and consequent clauses to be left uninterpreted, hence behaving as 'fake' tense morphology in standard Sequence of Tense constructions. This is represented in (51) by crossing out the uninterpreted morphological feature in pro_4, which leaves just the future temporal features of the future conditional to be interpreted in the antecedent and consequent clauses: $\text{pro}_4^{[\cancel{\text{past}}][\text{FUT pro}_8]}$.Additionally for Spanish, the Subjunctive morphology in the antecedent clause is represented by the subjunctive feature on $pro_4^{[\text{SUBJ CG}]}$.[10] Adding ∃-closure to bind pro_1 and pro_4, Romero (2017) delivers the (preliminary) LF (51) for our examples (44) and (45)[11]:

(51) LF: $\lambda 0\ \exists_1[\ \text{MODAL}^{L}_{\text{METAPHY}}\ \text{pro}_1^{[\text{PAST pro}_0]}$
$\lambda 8\ \exists_4[\text{you be hungry at pro}_4^{[\text{SUBJ CG}][\cancel{\text{past}}][\text{FUT pro}_8]}]$
$\lambda 8\ \exists_4[\text{your stomach be growling at pro}_4^{[\cancel{\text{past}}][\text{FUT pro}_8]}]\ ]$

This LF leads to the truth conditions (52). Note the temporal back shift $i_1 < i_0$ above the modal and the lack thereof inside the antecedent and consequent clauses, corresponding to the uninterpreted, 'fake' tense morphology in these clauses. Additionally for Spanish, Subjunctive mood in the antecedent clause imposes no modal presupposition on index i_8: $\cancel{i_8 \in CG}$. The resulting formula correctly matches the truth conditions of hypothetical counterfactual conditionals under the temporal remoteness view:

[10] The Spanish verbal paradigm has only one mood version of 'would+Verb'. Since there is no mood choice for this form in the consequent clause, the mood distinction in the consequent is neutralized.

[11] See Romero (2017) for two adjustments to this LF and truth conditions, one concerning temporal alignment between i_4 and the actual index i_0 and one restricting the methaphysical possibilities quantified over (Morgenbesser cases).

(52) $\lambda i_0.\ \exists i_1\ [i_1 < i_0 \wedge \forall i_8 \in \text{Metaph}^L(i_1):$
$\exists i_4\ [\cancel{i_8 \in CG \wedge}\ i_8 < i_4 \wedge \text{you be hungry at } i_4] \rightarrow$
$\exists i_4\ [i_8 < i_4 \wedge \text{your stomach be growling at } i_4]]$

This accounts for the grammaticality of 'fake' tense in the antecedent and consequent of hypothetical counterfactuals both in English and Spanish and for the grammaticality of subjunctive mood in the antecedent of hypothetical counterfactuals in Spanish.

3.2 Grammatical Biscuit Counterfactuals

We turn now to grammatical biscuit counterfactual forms like (53)–(54), whose consequent clause has no 'fake' tense in English and Spanish and bears Indicative mood in Spanish:

(53) If you were hungry right now, there **is** pizza in the fridge. =(42b)

(54) Si (tú) tuvieses hambre, hay pizza en el frigorífico.
If (you) had.Subj hunger, **have.Ind** pizza in the fridge. =(43b)

Following Romero and Csipak (2019), we propose that these forms involve broken Sequence of Tense and broken "Sequence of Mood", leading to a double access reading of the temporal and modal parameters of the evaluation index.

To implement this idea, some formal apparatus will be needed. Next to temporal intervals overlapping with two times à la Ogihara (1999), we need modal 'intervals'—i.e., stretches of logical space—overlapping with two modal contexts.[12] We construe an interval as a plural sum T of time points and, following Schlenker (2004), as a plural sum W of possible worlds. We put these pluralities into a pair to form an i(nternally)-plural

[12] As noted by a reviewer, temporal intervals are convex: For any two time points t_1 and t_2 belonging to an interval, all points temporally ordered between t_1 and t_2 also belong to that interval. To have convex modal intervals, we would need an ordering of worlds, e.g. à la Stalnaker (1968), Lewis (1973) or Kratzer (2012). We leave for future research what ordering system would be best suited.

index <W,T>. Temporal precedence < and overlap ∘ between i-plural indices are defined in (55) and a parallel definition for modal overlap • is given in (56). Note that the condition on the latent parameter has been relaxed: While (25) required the equality $w{=}w'$ for atomic worlds, (55) requires a non-empty intersection $W \cap W' \neq \emptyset$ between plural worlds (and similarly for (56)):

(55) For any two indices $<W,T>$ and $<W',T'>$:
$<W,T> < <W',T'>$ iff $W \cap W' \neq \emptyset$ and (the entire) T is prior to (the entire) T′.
$<W,T> \circ <W',T'>$ iff $W \cap W' \neq \emptyset$ and T and T′ overlap.

(56) For any two indices $<W,T>$ and $<W',T'>$:
$<W,T> \bullet <W',T'>$ iff $T \cap T' \neq \emptyset$ and W and W′ overlap.

The contribution of mood is redefined in (57): $\text{pro}_i^{[\text{IND pro}_k]}$ presupposes modal overlap • between index g(i) and the maximal i-plural index—imax defined in (58)—corresponding to the local context g(k). For example, if our local context is $\{<w_1,t_7>, <w_2,t_7>, <w_3,t_7\}>$, its i-max is $<w_1 \oplus w_2 \oplus w_3,\ t_7>$.

(57) $[\![pro_i^{[\text{IND } pro_k]}]\!]$ is defined only if $g(i) \bullet \text{imax}(g(k))$;
if defined, $[\![pro_i^{[\text{IND } pro_k]}]\!] = g(pro_i)$

(58) For any set I of (atomic) indices:
$\text{imax}(I) = <\max(\{w': \exists t'\ [<w',t'> \in I]\}),$
$\max(\{t': \exists w'\ [<w',t'> \in I]\})>$

Finally, we assume that, if a proposition is predicated of an i-plural index <W,T>, that proposition must hold true throughout that entire modal-temporal space, that is, through all the pairs <w,t> such that $w \in W$ and $t \in T$.

Let us apply the idea of temporal/modal double-access and this formalization to our examples. The present and indicative morphology in (53)/(54) leads to LF (59). This gives us the temporal and modal alignment of index i_4 with the (atomic) utterance index i_0 and with the CG in the last ∃-subformula in (60), but no temporal or modal alignment of i_4 with the (atomic) counterfactual index i_8 and $\text{Metaph}^L(i_1)$:

(59) LF: $\lambda 0\ \exists_1[\text{MODAL}^{L}_{\text{METAPHY}}$ at $\text{pro}_1{}^{[\text{PAST pro}_0]}$
$\lambda 8\ \exists i_4[$you be hungry at $\text{pro}_4{}^{[\text{SUBJ CG}]\ \cancel{\text{past}}\ [\text{FUT pro}_8]}]$
$\lambda 8\ \exists i_4[$be pizza at $\text{pro}_4{}^{[\text{IND CG}]\ [\text{PRES pro}_0]}]]$

(60) $\lambda i_0.\ \exists i_1\ [i_1 < i_0 \wedge \forall i_8 \in \text{Metaph}^L(i_1):$
$\exists i_4[\cancel{i_8 \bullet \text{imax(CG)} \wedge}\ i_8 < i_4 \wedge \text{you be hungry at } i_4] \rightarrow$
$\exists i_4[i_4 \bullet \text{imax(CG)} \wedge i_0 \circ i_4 \wedge \text{there be pizza at } i_4]]$

To supply the desired alignment, we extend Ogihara's idea and propose to duplicate the temporal and modal relations as $i_8 \circ i_4$ and $i_4 \bullet \text{imax}(\text{Metaph}^L(i_1))$ to allow for local binding, resulting in (61), with the duplication underlined:

(61) $\lambda i_0.\ \exists i_1\ [i_1 < i_0 \wedge \forall i_8 \in \text{Metaph}^L(i_1):$
$\exists i_4[\cancel{i_8 \in \text{CG} \wedge}\ i_8 < i_4 \wedge \text{you be hungry at } i_4] \rightarrow$
$\exists i_4[i_4 \bullet \text{imax(CG)} \wedge i_0 \circ i_4 \wedge \text{there be pizza at } i_4 \wedge$
$\underline{i_4 \bullet \text{imax}(\text{Metaph}^L(i_1)) \wedge i_8 \circ i_4}]]$

Crucially, i_4 in the last ∃-subformula is an i-plural index overlapping temporally with the time parameters of i_0 and i_8 and overlapping modally with the world parameters of imax(CG) and $\text{imax}(\text{Metaph}^L(i_1))$. That is, for each atomic i_0 of shape $<w_0,t_0>$ and each atomic i_8 of shape $<w_8,t_8>$, there is an i_4 of shape $<w_0 \oplus \ldots \oplus w_8, t_0 \oplus \ldots \oplus t_8>$ temporally and modally overlapping with them. For each such combination of i_0 and i_8, there being pizza in the fridge is predicated of the entire index i_4. As a result, by breaking Sequence of Tense and Sequence of Mood in BiCFs, the truth conditions (61) guarantee not only that there is pizza in each hypothetical hungry-index i_8, but also at each potential actual index i_0. This hard-wires the 'biscuit' effect: the feeling that the consequent is being asserted (of i_0) regardless of the truth of the antecedent.

In sum, conditionals that combine 'fake' tense and—for Spanish—subjunctive mood in the antecedent clause with no 'fake' tense and—for Spanish—indicative mood in the consequent clause lead to truth conditions that explicitly deliver the 'biscuit' effect. This renders them appropriate forms to express biscuit conditional meanings.

3.3 Unacceptable Biscuit Counterfactuals

After having presented an account of why counterfactual biscuits with a mood mismatch, such as (53)/(54), are grammatical, we now turn to the less acceptable variants:

(62) # If you were hungry right now, there would be pizza in the fridge. =(42a)

(63) # Si (tú) tuvieses hambre, habría pizza en el frigorífico.
If (you) had.Subj hunger, would.have pizza in the fridge. =(43a)

We observe that (62) and (63), which do not break Sequence of Tense/Mood, are unacceptable. Notice that this is the case in English and Spanish, respectively, pace Franke (2009), who claims both of the following sentences are acceptable:

(64) If you had needed some money, there was some in the bank.
(= Franke (2009)'s (113a), cited from Johnson-Laird (1986))

(65) If you had been hungry, there would have been pizza in the fridge.
(= Franke (2009)'s (114e))

Our informants agree that (64) is in fact acceptable, but (65) is not. We point out that language variation plays an important role here: Csipak (2018) shows that while English is restricted in this way, German is not, and Csipak (2015) argues that languages in which the subjunctive has a 'politeness' use in unembedded clauses typically allow it to appear in the consequents of biscuit conditionals such as (65). We do not discuss these languages further.

Let us return to the unacceptable (62) and (63) and compare their truth-conditions with those of the acceptable forms (53)–(54). Recall the truth-conditions of the acceptable forms in (61), repeated here as (66):

(66) $\lambda i_0.\ \exists i_1\ [i_1 < i_0 \wedge \forall i_8 \in \text{Metaph}^L(i_1)$: (=(61))
$\exists i_4[\cancel{i_8 \in CG \wedge}\ i_8 < i_4 \wedge$ you be hungry at $i_4] \rightarrow$
$\exists i_4[i_4 \bullet \text{imax}(CG) \wedge i_0 \circ i_4 \wedge$ there be pizza at $i_4 \wedge$
$i_4 \bullet \text{imax}(\text{Metaph}^L(i_1)) \wedge i_8 \circ i_4]]$

The index i_4 in (66) must stretch temporally and modally to include the time and world parameters of (past hypothetical) i_8 and (present actual) i_0. Recall that, under the double access reading, the relevant proposition has to hold of *all* the points in the temporally and modally stretched i_4. This means that there is pizza in the fridge at all indices i' of shape <w',t'> such that $t_8 \leq t' \leq t_0$ and $\text{imax}(\text{Metaph}^L(i_1)) \leq w' \leq \text{imax}(CG)$. Hence, (66) entails that there is pizza in the fridge at a hypothetical index t' that shares the world parameter with i_8 but is temporally posterior to i_8.

Consider now the truth conditions of the unacceptable forms (62) and (63), given in (67):

(67) $\lambda i_0.\ \exists i_1\ [i_1 < i_0 \wedge \forall i_8 \in \text{Metaph}^L(i_1)$:
$\exists i_4\ [\cancel{i_8 \in CG \wedge}\ i_8 < i_4 \wedge$ you be hungry at $i_4] \rightarrow$
$\exists i_4\ [i_8 < i_4 \wedge$ there be pizza at $i_4]]$

The index i_4 in (67) is only specified to be temporally posterior to the hypothetical index i_8 (and, following definition (25a), to share its world parameter with i_8); no requirement that i_4 stretches to overlap with i_4 is imposed by these truth-conditions. Thus, the acceptable (53)–(54) have stronger truth conditions than the unacceptable (62)–(63).

Following Franke (2009), we predict that both sets of biscuit counterfactuals, (53)/(54) and (62)/(63), are actually grammatical, and both receive a biscuit interpretation. This is irrespective of tense and mood, since conditional independence is defined independently, and *p* and *q* are conditionally independent in both sets of conditionals. But this means that (53)/(54) and (62)/(63) compete for signalling the same message, namely the conditional *if p, q* as well as the speaker's commitment to *q* in i_0. This latter message is signalled differently by the two sets of conditionals: the semantically stronger form (53)/(54) breaks Sequence of Tense/Mood to allow a double access reading, thus explicitly signalling the overlap of i_4 with i_0. The semantically weaker form (62)/(63) on the other hand relies purely on pragmatic inferencing (deriving the 'biscuit' reading from conditional independence of *p* an *q*).

In a context where the speaker wishes to signal a counterfactual biscuit meaning, i.e., where she wants to signal both her commitment to the counterfactual conditional and to the truth of the consequent in the actual world, the stronger form should be chosen, and the weaker form should be dispreferred.

To explain how this preference comes about, we appeal to two recent works: the work on cessation implicatures by Altshuler and Schwarzschild (2013) discusses similar effects in another domain, and we use Lauer (2014)'s analysis of Need-a-Reason implicatures to explain why the pragmatic inferencing that takes place in (62)/(63) is non-optional and thus leads to contradictory messages.

Altshuler and Schwarzschild (2013) propose that competition between tenses leads to cessation implicatures. This phenomenon is illustrated in (68)–(69). Under certain conditions (with a stative predicate and when no topical past time is salient), we observe a cessation implicature when the past tense is used instead of the present tense: the implicature that the stative property does not hold at the utterance time.

(68) John is sick.

(69) John was sick. $\rightsquigarrow$ John is no longer sick.

To derive this cessation implicature, Altshuler and Schwarzschild (2013) argue that present tense stative predicates entail being true not just of the utterance time but also of prior times, due to the *Open Interval Hypothesis*.

(70) Open Interval Hypothesis (Altshuler & Schwarzschild 2013)
The run-time of a state is an open interval. That is, if e is a stative eventuality and $t' \subseteq \mathrm{T}(e)$, then there is a temporal instant t'' such that $t'' < t'$ and $t'' \subseteq \mathrm{T}(e)$.

This means that, by using present tense, the speaker of (68) conveys the stronger message that the interval of John being sick includes both the speech time and times prior to speech time. In contrast, using the past tense in (69) only commits the speaker to John being sick at some times prior to speech time, thus giving rise to the implicature that the speaker does *not* want to commit to John being sick at speech time.[13] We note that this implicature is easily cancellable, as in (71).

(71) John was sick last week. In fact, he still is.

A similar, but slightly different mechanism is at work in our examples. We assume that a speaker has a choice between two forms, (72) and (73). By uttering (72), the speaker breaks Sequence of Tense, whereas (73) observes it.

(72) If you were hungry now, there is pizza in the fridge. =(42b)

(73) # If you were hungry now, there would be pizza in the fridge. =(42a)

We first consider the semantics of the stronger form (72), seen above in (66). By breaking Sequence of Tense, the speaker signals that she is committed to there being pizza in the fridge both at the counterfactual hungry-indices and at the actual index. Note that inclusion of the actual index happens semantically.

Now turning to (73), we remember that the semantics only commit the speaker to there being pizza at the counterfactual hungry-indices, *not* at the actual index. This of course gives rise to the implicature that the

[13] We note that the purpose of Altshuler and Schwarzchild's proposal is to argue against Sequence of Tense as presented in Ogihara (1999). This does not diminish its similarity to our example.

speaker does not want to commit to there being pizza at the actual index. Let us call this the non-actuality implicature.

Given conditional independence and the reasoning resulting from it (i.e., that the speaker must have some evidence for *q* in the actual world), we would expect that the non-actuality implicature should be cancelled. But we argue that this implicature is mandatory: it is a Need-a-Reason implicature in the sense of Lauer (2013, 2014). Lauer argues that when speakers choose an otherwise 'dispreferred' form (i.e., one that is more complex and less informative than a competitor), their interlocutors draw the inference that they must have a reason for choosing this form. Such implicatures are not cancellable. Consider Lauer's example (7) below.

(74) Somewhere in San Francisco, A and B are planning a dinner party, talking about who they should invite.

A: Is John in town?

B: No, he is in Paris or in London.

B's utterance gives rise to an ignorance implicature: B does not know which of the two cities John is in. Lauer calls this a Need-a-Reason implicature and observes that is almost impossible to cancel. For the purpose of the conversation in (74) it is irrelevant whether John is in Paris or in London. Furthermore, the form using *or* is also more complex than its alternatives (*John is in Paris* and *John is in London*), as well as being less informative than the alternatives. In sum, B has chosen a form that is—seemingly needlessly—less informative and more complex than its alternatives. The interlocutors therefore infer that B must have 'had a reason' for choosing this form, namely to convey the content of the implicature (that B doesn't know whether John is in Paris or in London). This makes the implicature very hard to cancel. And in fact it is very difficult to follow up B's utterance with 'In fact, he is in London.'

Returning to our example, we saw that maintaining Sequence of Tense and choosing the subjunctive—i.e., leaving it open whether there is pizza at the actual index—results in a less informative statement than breaking Sequence of Tense and choosing the indicative—i.e., committing oneself to there being pizza at the actual index.[14] Thus the implicature that is derived from using the weaker form ('the speaker does not want to commit whether there is pizza in the actual world') is a non-cancellable, Need-a-Reason implicature. But remember that the hearer is also invited to follow a Franke-style reasoning about the independence of p and q: since p and q are conditionally independent, the speaker should have reason to assume that q does hold at the actual index. Thus, the Need-a-Reason implicature and the reasoning triggered by independence send conflicting messages: the speaker is signalling both that she doesn't want to commit to there being pizza in the actual world (qua subjunctive) and that she does have reason to believe that there is pizza in the actual world (qua conditional independence). Compared to its competitor (72), (73) is thus not only weaker semantically, but also gives rise to contradictory inferences. It is therefore dispreferred.

Note that there are contexts where the dispreferred form becomes available when the desired interpretation is conditional independence (a 'biscuit' reading), but no commitment of the speaker to q in i_0. For example, consider modal subordination contexts such as (75) due to Swanson (2013). Here, the speaker is only committed to there being biscuits at her desire indices (conditionally independently of p), but crucially not at i_0. Since the BiCF appears in a modal subordination context, the grammar does not allow for a morphological choice and the consequent clause must feature the Subjunctive. This means that that there is no explicit morphological cue leading to the biscuit interpretation. Hence, conditional independence alone derives the 'biscuit' interpretation.

(75) I want to vacation at a posh hotel in London. We would have tea every afternoon, and there **would be** biscuits on the sideboard if one were so inclined. (Swanson 2013)

[14] We leave it open whether it is more complex to use the subjunctive or to break Sequence of Tense/Mood.

Importantly, the speaker of (75) has *not* committed to there being biscuits on the sideboard of the posh London hotel *at* i_0.[15]

3.4 Unacceptable Hypothetical Counterfactuals

We have now accounted for the acceptability of counterfactual biscuits that break Sequence of Tense/Mood and the unacceptability of counterfactual biscuits that do not break it (in the typical case). What remains is to explain why hypothetical counterfactuals which break Sequence of Tense/Mood are unacceptable:

(76) # If you were hungry right now, your stomach is growling. = (40b)

(77) # Si (tú) tuvieses hambre, tu estómago está gruñendo.
If (you) had.Subj hunger, your stomach is.Ind growling. = (41b)

To give forms like this a better chance, we consider a scenario where a speaker might wish to convey both the hypothetical counterfactual and the consequent proposition.

(78) **Context:** A is a detective at a crime scene. Her partner, B, is stuck in traffic and can thus not see the crime scene yet. A is describing the scene.
A: #If the victim had gotten poisoned with arsenic, he is showing the following symptoms: X, Y, and Z.

How come the utterance by A is not acceptable? It seems like this would be a convenient way of communicating two things at once: both what symptoms the victim would be showing in case of arsenic poisoning, and the fact that the victim is actually showing these symptoms. Yet even in

[15] We will have nothing to say on the status of (i), which seems to oscillate between the speaker's dream worlds and what she has read about the actual world.

(i) I want to vacation at the Waldorf-Astoria in New York. I read the brochure and know all about their features. We would have tea every afternoon, and there **is** a sauna if we were so inclined.

the context where A is trying to convey exactly those two facts, (78) is unacceptable.

To explain this, we appeal to a suggestion by DeRose and Grandy (1999) about why speakers use conditionals. They propose that there are two main reasons why speakers choose to utter a conditional of the form *if p, q* rather than plain *q*, and that when neither of these conditions is met, via the Gricean Principle of Manner (Grice 1975), plain *q* should be uttered. We tentatively follow their argument, outlined below, but leave open the possibility that there may be other reasons to utter *if p, q*.

The first reason for which a speaker may utter *if p, q* rather than *q* is that they are uncertain about the truth of *q* in the actual world w_0, but sufficiently certain about the truth of *if p, q*. In such cases, *p* and *q* will be conditionally dependent, and the resulting conditional *if p, q* will be interpreted as a hypothetical conditional. DeRose and Grandy argue that there is a competition between *q* and *if p, q* such that if the speaker is sufficiently certain that *q* is true at w_0, then she should utter plain *q* rather than *if p, q*. Conversely, if the speaker is only sufficiently sure about *if p, q* but not about plain *q*, then she should utter the conditional.

The second reason for uttering *if p, q* is in cases where the speaker is uncertain whether plain *q* is sufficiently relevant (to the conversation) at w_0. In that case, *p* and *q* will be conditionally independent and the speaker will utter a biscuit conditional. And again there is competition between plain *q* and *if p, q* such that if the speaker is sufficiently certain that *q* is relevant to the interlocutors, she should simply utter plain *q* instead of the conditional.

Taken together, the two reasons why one should *not* utter a conditional when a simple *q* might suffice falls under Grice's Principle of Manner: in particular, the submaxim to *be brief*—uttering plain *q* rather than the longer *if p, q* is preferable when the conditions for uttering *q* are given.

Since we are interested in deriving a hypothetical counterfactual reading for (78), we would expect the first line of argument to apply. But we can see that the speaker of (78) is violating these rules: what she wants to convey is, first, a conditional dependence between *p* and *q*, and, second, that *q* holds in the actual world. But by uttering *if p, q* instead of plain q, she is signalling that she is not in a position to utter plain q (remember that q and *if p, q* are in competition such that if a speaker is reasonably certain that q holds in w_0, she should utter that instead of the conditional). So on the one hand, by uttering the conditional, the speaker signals that she is *not* trying to convey that q holds in w_0. But by breaking Sequence of Tense/Mood, she is explicitly trying to do the opposite: breaking Sequence of Tense/Mood ensures that the speaker communicates their belief that q *does* hold at w_0. This contradiction is what causes (78) to fail.

We now consider an interesting related case.

(79) If Tom had come, Mary still would have come.

(80) *If Tom had come, Mary (still) came.

The acceptable (79) conveys both that the counterfactual *if p, q* is true and that q is true at w_0, whereas (80), which attempts to break Sequence of Tense/Mood, is unacceptable. The reason for (80) failing is parallel to the one given above.

The crucial difference between the pair in (79)–(80) and (78) is that for (79) it is already common ground that Mary did come in the actual world (signalled by the presupposition trigger *still*). So the speaker of (79) is actually not trying to newly establish that q holds in w_0, but is only trying to convey the conditional link between p and q. The relation between q and w_0 is already settled.

4 Outlook

We have illustrated how to extend the unified, independence-based approach for indicative hypothetical and biscuit conditionals to *counterfactual* hypothetical and biscuit conditionals to derive (40)–(43). We have done this by combining the temporal remoteness approach to counterfactual conditionals with breaking Sequence of Tense/Mood in an extension of Ogihara (1999). Then, we have shown how this extension derives the attested combinations of conditional interpretation and tense/mood and why pragmatic principles rule out the unattested combinations. Competition between more informative and less informative forms rules out biscuit counterfactuals that do not break Sequence of Tense/Mood (except in cases of modal subordination), and the Gricean principle of Manner rules out hypothetical counterfactuals that do, via an insight by DeRose and Grandy (1999) of what the communicative purpose of conditionals is.

There are a number of open issues that we leave for future research.

First, the analysis of double access readings as presented in Romero and Csipak (2019) needs to be further refined and unified across constructions, possibly as a *de re* analysis. This includes purely temporal double access like (36), index double access like our counterfactual biscuits (42b)/(43b) and potentially other double-access-looking data. For example, in (81), translated into Spanish from Schlenker (2004), the (a)-version breaking Sequence of Tense/Mood can be used to convey that, if Juan thought of the actual rainy weather pattern that it counts as good weather, Juan would be crazy, while the (b)-version does not allow for this interpretation.

(81) Context: It is raining outside and the speaker sees that. [Spanish]

a. Si Juan pensase que hace.Ind buen tiempo, estaría loco.
'If John thought that the weather **is.Ind** nice, he would be crazy.'

b. # Si Juan pensase que hiciese.Subj buen tiempo, estaría loco.
'If John thought that the weather **was.Subj** nice, he would be crazy.'

Second, on the empirical front, we only consider counterfactual biscuit conditionals in languages like English and Spanish, which both use verbal morphology to convey counterfactuality and have attested double-access readings on that morphology. How does this work in languages that use other means to express counterfactuality, and do they allow a 'biscuit' interpretation of counterfactuals? If so, do they also permit a double-access indexing option, or is the 'biscuit' interpretation derived from pragmatic inference as in Swanson (2013)'s (75)?

Finally, a challenge is build into our analysis by virtue of using the temporal remoteness approach to counterfactuality. Counterpossibles like (82) have been argued to posit a problem for the overall temporal remoteness line:

(82) If two plus two were five, this addition would be correct.

The temporal remoteness approach depends on being able to go back to a point in time where it was possible for the antecedent to be true, and it is not clear whether this is possible for the antecedent of (82). One way to solve this might be to relativize indicative and counterfactual conditionals to a given epistemic state (cf. Leahy 2018). We leave this possibility for future research.

References

Abusch, Dorit. 1997. Sequence of tense and temporal de re. *Linguistics and Philosophy* 20, 1–50.

Altshuler, Daniel & Roger Schwarzschild. 2013. Moment of change, cessation implicatures and simultaneous readings. In E. Chemla, V. Homer & G. Winterstein (eds.), *Sinn und Bedeutung* 17, 45–62.

Anand, Pranav & Valentine Hacquard. 2009. The role of imperfect in romance counterfactuals. In V. Schmitt, M. Prinzhorn & S. Zobel (eds.), *Proceedings of sinn und bedeutung 14*, 37–50. University of Vienna.

Anderson, Alan Ross. 1951. A note on subjunctive and counterfactual conditionals. *Analysis* 11, 35–38.

Austin, John L. 1956. Ifs and cans. In *British Academy*, vol. 42, 109–132.

Biezma, Maria & Arno Goebel. 2018. The pragmatic ingredients to get perfect biscuits. In R. Truswell, C. Cummins, C. Heycock, B. Rabern & H. Rohde (eds.), *Sinn und Bedeutung 21*, 179–198.

Csipak, Eva. 2015. *Free factive subjunctives in German*: University of Göttingen PhD dissertation.

Csipak, Eva. 2018. Discourse-structuring conditionals and past tense. In R. Truswell, C. Cummins, C. Heycock, B. Rabern & H. Rohde (eds.), *Sinn und Bedeutung 21*, 301–315.

DeRose, K. & R.E. Grandy. 1999. Conditional assertions and 'biscuit' conditionals. *Nous* 33, 405–420.

Dudman, V. H. 1983. Tense and time in english verb clusters of the primary pattern. *Australian Journal of Linguistics* 3, 25–44.

Dudman, V. H. 1984. Conditional interpretations of if-sentences. *Australian Journal of Linguistics* 4, 143–204.

Ebert, Christian, Cornelia Ebert & Stefan Hinterwimmer. 2014. A unified analysis of conditionals as topics. *Linguistics and Philosophy* 37, 353–408.

Ferreira, Marcelo. 2016. The semantic ingredients of imperfectivity in progressives, habituals, and counterfactuals. *Natural Language Semantics* 24, 353–397.

Francez, Itamar. 2015. Chimerical conditionals. *Semantics & Pragmatics* 8, 1–35.

Franke, Michael. 2009. *Signal to Act: Game Theory in Pragmatics*: University of Amsterdam PhD dissertation.

Goebel, Arno. 2017. Laws for biscuits: Independence and dependence in conditionals. In D. Burgdorf, J. Collard, S. Maspong & B. Stefánsdóttir (eds.), *Semantics and Linguistic Theory 27*, 377–396.

Grice, H. P. 1975. Logic and conversation. In P. Cole & J. Morgan (eds.), *Syntax and semantics 3: Speech arts*, 41–58. New York: Academic Press.

Grønn, Atle & Arnim von Stechow. 2009. Temporal interpretation and organization of subjunctive conditionals. Ms. University Oslo.

Heim, Irene. 1991. Artikel und definitheit. In A. von Stechow & D. Wunderlich (eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, 487–535. Berlin: de Gruyter.

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9, 183–221.

Heim, Irene. 1994. Comments on Abusch's theory of tense. In H. Kamp (ed.), *Ellipsis, tense and questions*, 143–170. U. Amsterdam.

Iatridou, Sabine. 2000. The grammatical ingredients of counterfactuality. *Linguistic Inquiry* 31, 231–270.

Ippolito, Michela. 2003. Presuppositions and implicatures in counterfactuals. *Natural Language Semantics* 11, 145–186.

Johnson-Laird, P. 1986. Conditionals and mental models. In Elizabeth Closs Traugott (ed.), *On conditionals*. Cambridge University Press.

Kaufmann, Stefan. 2005. Conditional truth and future reference. *Journal of Semantics* 22, 231–280.

Kratzer, Angelika. 1998. More structural analogies between pronouns and tenses. In D. Strolovitch & A. Lawson (eds.), *Semantics and Linguistic Theory 8*, 92–109. Ithaca, N.Y.: CLC.

Kratzer, Angelika. 2012. *Modality and conditionals*. Oxford: Oxford University Press.

Kusumoto, Kiyomi. 1999. *Tense in embedded contexts*: University of Massachusetts at Amherst PhD dissertation.

Kusumoto, Kiyomi. 2005. On the quantification over times in natural language. *Natural Language Semantics* 13, 317–357.

Lauer, Sven. 2013. *Towards a dynamic pragmatics*: Stanford University PhD dissertation.

Lauer, Sven. 2014. Mandatory implicatures in Gricean pragmatics. In Judith Degen, Michael Franke & Noah Goodman (eds.), *Proceedings of formal and experimental pragmatics 2014*, 21 – 28.

Lauer, Sven. 2015. Biscuits and provisos. In E. Csipak & H. Zeijlstra (eds.), *Sinn und Bedeutung 19*, 357–374.

Leahy, Brian. 2018. Counterfactual antecedent falsity and the epistemic sensitivity of counterfactuals. *Philosophical Studies* 175, 45–69.

Lewis, David. 1973. *Counterfactuals*. Malden: Blackwell.

Ogihara, T. 1999. Double access sentences generalized. In T. Matthews & D. Strolovitch (eds.), *Semantics & Linguistic Theory 9*, 224–236. Ithaca: CLC.

Ogihara, Toshiyuki. 1995. The semantics of tense in embedded clauses. *Linguistic Inquiry* 26, 663–679.

Partee, Barbara. 1973. Some structural analogies between tenses and pronouns in English. *Journal of Philosophy* 7, 601–609.

Romero, Maribel. 2017. Tense and mood in counterfactual conditionals: The view from Spanish. In T. van Gessel A. Cremers & F. Roelofsen (eds.), *Amsterdam Colloquium 21*, University of Amsterdam.

Romero, Maribel & Eva Csipak. 2019. Counterfactual biscuit conditionals: A temporal remoteness account. In K. Lamp K. Blake, F. Davis & J. Rhyne (eds.), *Proceedings of semantics and linguistic theory 29*, 441–458. LSA Journals.

Schlenker, Philippe. 2004. Sequence phenomena and double access readings generalized. In J. Lecarme & J. Guéron (eds.), *The syntax of time*, 555–595. Cambridge: MIT Press.

Schlenker, Philippe. 2005. The lazy frenchman's approach to the subjunctive. In T. Geerts, I. van Ginneken & H. Jacobs (eds.), *Romance languages and linguistic theory 2003*, 269–309. Amsterdam: Benjamins.

Schulz, Katrin. 2014. Fake tense in conditional sentences: A modal approach. *Natural Language Semantics* 22, 117–144.

Stalnaker, R. 1968. A theory of conditionals. In W. L. Harper, R. Stalnaker & G. Pearce (eds.), *Ifs*, 41–55. Dordrecht: Springer.

Stalnaker, R. 1975. Indicative conditionals. *Philosophia* 5, 269–286.

von Stechow, A. 1995. On the proper treatment of tense. In M. Simons & T. Galloway (eds.), *Semantics and Linguistic Theory 5*, Ithaca, NY: CLC.

von Stechow, A. 2009. Tense in compositional semantics. In Wolfgang Klein & Ping Li (eds.), *The expression of time*, 129–166. Berlin/New York: De Gruyter.

Swanson, Eric. 2013. Subjunctive biscuit and stand-off conditionals. *Philosophical Studies* 163(3), 637–648.

Events Are the Source of Causal Readings in the Simplest English Conditionals

Bridget Copley

1 A Causal Puzzle About Statives and Eventives in Conditionals

Suppose that there is a country called Clavarel[1] where the queen is selected from among the eligible candidates by a series of coin tosses. In the particular situation we are interested in, it so happens that if the coin comes up heads, Yolanda becomes the queen; and if the coin comes up tails, she does not. Suppose also that if someone becomes the queen, any daughters she may have consequently become princesses; moreover,

[1] Thanks to Rachel and Charlie Seymour for Clavarel (https://thedowntownmerrylegs.bandcamp.com/track/clavarel).

B. Copley (✉)
SFL (CNRS/Université Paris 8/UPL), Paris, France
e-mail: bridget.copley@cnrs.fr

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6_15

there is no other way to become a princess of Clavarel. Further suppose that Yolanda has exactly one daughter, named Xanthippe.

In such a context, we can see that (1a) is true, while (1b) is false. What's wrong with (1b)? It certainly seems like it conveys a causal relation which is at odds with what we know—it's backwards in fact. (1b) seems to say that if Xanthippe becomes the princess, then *by consequence* Yolanda becomes the queen. (1a), on the other hand, has no such flavor.

(1) a. If Xanthippe is the princess, Yolanda is the queen.
    b. If Xanthippe becomes the princess, Yolanda becomes the queen.

The puzzle here is, why? Taking seriously Bennett's comment that "The search for a deep understanding of conditionals would go better if the taxonomy were got right first" (Bennett 1988), I will make two methodological choices in teasing out (merely) where to start with the taxonomy.

1.1 Causation Rather Than Time

The first choice has to do with time. In addition to feeling causally backward, (1b) feels temporally backward; it seems to say that Xanthippe becomes the princess of Clavarel *first*, and only then does Yolanda become the queen; contrary, again, to what we know about Claverelian dynastic politics. Now, the distinction between stative predicates (e.g., *be the princess*) and eventive predicates (e.g., *become the princess*) can be cast in terms of times; see, for example, Hallman (2009). We might therefore consider recruiting the concept of time to try and solve this puzzle. This is certainly a plausible move.[2]

However, I am going to avoid making reference to time in explaining the contrast in (1). There are two reasons for this. One reason to start

[2] The fact that antecedents like *If Xanthippe becomes princess* are grammatically present but conceptually future-oriented was noticed by Dudman (1983), if not earlier; see also Kaufmann (2004) and Kaufmann et al. (2006). The significance of this fact is still controversial to philosophers, see Edgington (1995, 2003), Bennett (2003: 15).

from causation is that temporality can be in part derived from causation from the fact that causes don't happen after their effects, so that taking causation as primitive gets us some temporality for free as it were. On the other hand, taking time as primitive gets us nothing in the way of causal meaning. A second reason is that eventivity and stativity are phenomena internal to the verb phrase; approaches to the verb phrase across various frameworks have concluded that much, if not all, of eventive verbal meaning is organized in terms of causal relations (e.g., Dowty 1979; Pustejovsky 1995; Ramchand 2008; Croft 2012; Copley and Harley 2015). For these reasons, I will defer discussion of times and instead start from causation, relying on a simple (structural equation) version of causal models (Pearl 2000).[3]

A causal model is a directed acyclic graph used to represent causal influences (Pearl 2000; Pearl and Mackenzie 2018). For example, our causal knowledge about the situation in Clavarel can be represented as the following, where C is whether the coin turns up heads, Q is whether Yolanda is the queen, and P is whether Xanthippe is the princess.

(2) $C \longrightarrow Q \longrightarrow P$

There can be different interpretations of causal models like the one above (notably, the probabilities of the nodes are often used). In the interpretation we will use, we will use simply the nodes' truth values.

(3)
$C :=$ 1 if heads, 0 if tails
$Q :=$ 1 if Yolanda is the queen, 0 if Yolanda is not the queen
$P :=$ 1 if Xanthippe is the princess, 0 if Xanthippe is not the princess

Supposing that Yolanda becomes queen upon the coin's landing one way or another, regardless of whether the Clavarelian nobles have seen which side the coin has landed on, the table in (4) shows the combinations of truth values that are possible given the causal information in (3).

[3] Linguists are beginning to make use of causal models; see Kaufmann (2002, 2005, 2013), Baglini and Francez (2016), Nadathur (2016), Lauer and Nadathur (2018), Baglini and Bar-Asher Siegal (2019), Bar-Asher Siegal and Boneh (2019).

(4)

	C	Q	P
a.	1	1	1
b.	0	0	0
c.	1	0	0
d.	1	1	0

Each line of the table is a possible *situation*; I assume in this very limited scenario that 1 and 0 are the only possible values and that each node has a value in each situation. It should be uncontroversial that lines (a) and (b) are possible in our Clavarel scenario, but those in (c) and (d) may or may not be possible, depending on further details. The idea is that, even though there is strictly speaking no time between the coin coming up heads and, e.g., Yolanda becoming queen, it is still possible to be in a state where one has had the thought that the coin has come up heads but not yet concluded that Yolanda is now queen, as in line (c). However, don't get hung up on the issue of whether lines (c) and (d) belong in the table or not. The puzzle, and our solution to the puzzle, are the same regardless of whether they are there or not. In Sect. 4, we will look at a more concise way to generate the lines of the table according to the arrows in the causal model graph, which will hopefully shed light on this issue.

1.2 Compositionality and the Stative–Eventive Distinction

So, the first methodological choice was to use causal models. The second methodological choice is to take a compositional perspective. Semantic composition allows us to theorize about unknown meanings of parts of a sentence if we know the meaning of the whole. In this, it is a kind of algebra. And, as in algebra, it is helpful to start with the simplest expressions in order to solve for unknowns. This is why we are using the morphologically simplest prejacents in (1) to understand how *if* works; we will not even be looking at conditionals with *will* in the consequent.

Very quickly we see that the relevant distinction between (1a) and (1b) is one of stativity (as in *being*) and eventivity (as in *becoming*). The distinction between eventivity and stativity is perhaps the major division in the lexical aspect (or "Aktionsart") of verb phrases. We might think of eventive predicates as processes going on in time, with phases succeeding one another; states on the other hand, at least as far as we will talk about them here, can be thought of as "a single, undifferentiated period" (Smith 1991: 19). Using stative predicates, as in (1a), there is no causal reading, and using eventives, as in (1b), there is a causal reading. We would like to know why this is so.

1.3 *Do* Will Not Do

Pearl and Mackenzie's (2018) discussion around causal models suggests a possible way to link event semantics to causal models. They make a distinction between "seeing" and "doing". For example, one should distinguish *If I see that Xanthippe is the princess, I see that Yolanda is the queen* and *If I make Xanthippe the princess, I make Yolanda the queen.* Their point is that seeing or inferring correlations between values of variables is not at all the same as changing or "wiggling" the value of those variables. Only the latter gives us truly causal information, in the way that turning a light switch on and off tells us which appliance it is connected to.

This idea enjoys pride of place in "interventionist" theories of causation (e.g., Woodward 2006); for Pearl, it also marks the difference between talking about correlations and talking about causation at all (Pearl and Mackenzie 2018: Chap. 1). They note that in the field of statistics, there has been a real reluctance to use causal language for lack of a formalized way to talk about it. And indeed it is illegitimate to talk about causation using only tools that express correlation. However, as they point out, once we have a formal way to talk about causation, it is no longer illegitimate to do so.

Pearl's (2000) formal tool for talking about causation is the "*do*-operator". The *do*-operator erases all causal influences on the node in question and changes the value of the node to the specified value, as shown in (5):

(5) $do(X = x)$: erase all incoming arrows into X and change the value of X to x.

Looking at the difference between seeing and doing in this system, one could wonder whether this distinction might correspond to the distinction in English between statives and eventives, respectively. After all, seeing is stative and doing is eventive. It would be very nice if this system developed for other uses could turn out to explain the stative/eventive distinction and thereby explain the contrast in (1).

The causal model in (2), together with treating conditionals as strict, works well enough for (1a). That is, given the model and the possible values, it is true that all cases where Xanthippe is the princess of Claverel, Yolanda is the queen. There is one situation where Xanthippe is the princess, namely the first line of the table in (4), and on that line, Yolanda is queen.

But what we would like to know is whether the *do*-operator, "off the shelf", as it were, might help us understand the meaning of eventive predicates *become the princess* or *become the queen* as in (1b). For suppose it did. Then if the conditional is again treated as strict implication, the meaning of (1b) would be expected[4] to be as follows:

(6) All situations where $do(P(x) = 1)$ are situations where $do(Q(y) = 1)$

That is, all cases where we erase the arrows into the $P(x)$ node and set the value of the $P(x)$ node to 1 are cases where we erase the arrows into the $Q(y)$ node and set it to 1.

But even then, we do not explain our puzzle. We do get (1b) to be false, but for the wrong reason. Making the reasonable assumption that we have the freedom, as erasers of arrows and setters of values, to apply the *do*-operator to whatever node we want, it is indeed false, as desired, that any case where we $do(P(x) = 1)$ is a case where we $do(Q(y) = 1)$. But although we get the right result, this reason for getting that result is exactly wrong. The problem with (1b) is not that there is no arrow

[4] To underline what we are doing here: (6) is merely *our* expected meaning if we are treating each eventive verb as involving a *do*-operator, not Pearl's expectation. For one, he does not distinguish between eventive and stative verbs.

between $Q(y)$ and $P(x)$; it is that there *is* an arrow going into $P(x)$ in the causal model, but somehow the conditional seems to require an arrow of causal influence to go the other way. In short, the *do*-operator cannot help us understand here how eventive predicates relate to causal readings in conditionals, because it erases one of the very things that we need to use, namely that incoming arrow.

1.4 How We Will Solve the Puzzle

The *do*-operator is not the only possible way to model "doing", i.e., intervention. We can retain something like the intuitive idea that *seeing* and *doing* are to be distinguished, and that this distinction matters for statives and eventives. I will rely on another tradition to make sense of this idea, namely dynamic semantics (Heim 1982; Kamp 1981; Groenendijk and Stokhof 1990). More closely hewing to our characterization of the puzzle in (1), statives (as in *being*) and eventives (as in *becoming*), will be treated as distinct modes of updating the world. Asserting a proposition in dynamic semantic frameworks changes the context it is uttered in; propositions are "context change potentials". Statives, we will see, behave like propositions. I will argue here that to account for the behavior of eventive predicates, we need events to be, in effect, "world change potentials". Crucially, learning or accepting a proposition does not require it to be false (or for that matter true) before the learning or accepting, whereas an event's happening does put a condition on the initial state; for Xanthippe to become a princess, she has to not already be a princess.

The analogy between propositions as context change potentials and events as world change potentials suggests that we might want to model eventives as predicates of ordered pairs of atoms, not predicates of atoms. That is, although in the neo-Davidsonian tradition we typically think of causal relations between subevents, here the events will actually be themselves represented by an ordered pair between two situations. The types[5] we will be interested in are in (7)[6]:

(7) a. stative predicate: $\langle s, t\rangle$
b. eventive predicate: $\langle s{\times}s, t\rangle$

Accordingly, there will be two ways for the speaker to update the current situation: *learning*, which is an update with a predicate of situations, and *happening*, which is an update with a predicate of relations between situations.

(8) a. learning = update with a predicate of situations
b. happening = update with a predicate of a relation between situations

These two kinds of updates, associated with the verb phrases themselves, will be situated with a dynamic treatment of *if* φ, ψ.

Working from these components, I will argue that the event in the consequent is, literally, the causal relation that yields the causal reading in (1b). To argue this, first we will set out to understand what eventives are, both in terms of the tests we use to distinguish eventives and statives, and in the sense of how they appear in conditionals. It turns out that not all apparent eventives in conditional antecedents are really, truly, eventives.

[5] As is usual in formal semantics, I will assume that the denotations of phrases are functions from typed objects to other typed objects, where these typed objects are perhaps themselves such functions. Type is a formal property, which allows us to compose meanings with each other according to a small collection of compositional rules, including Functional Application Heim and Kratzer (1998). In short, if two denotations are not of the right type to compose, they will be unable to compose. The basic types used here are s, the type of situations, and t, the type of truth values.

[6] The type of the eventive predicate could also be written as $\langle\langle s, s\rangle, t\rangle$, as in Copley and Harley (2015, 2021). The (equivalent) cross notation used here underlines the fact that the relation between situations, as far as type theory and semantic composition are concerned, is not treated like a compound type.

The ones that are not get a stative reading, so they must be excluded if we want to say something about the behavior of eventives. These can be thought of as eventives with a stative component on top of them, so they may not be truly semantically "bare" eventives even though they are morphologically bare; or they can be thought of as stative themselves. This point relates to our methodological decision to consider only the simplest conditionals, namely those with no aspectual, modal, future, tense, or counterfactual morphology on the prejacents. Recall that this choice is an important one, as it allows us to see how event semantics interacts with conditional semantics when there is nothing semantically active (in the relevant sense) insulating the verb phrase (the locus of event semantics) from *if* (the locus of conditional semantics).

Subsequently, we will see evidence from assertability facts to the effect that "true" eventives (i.e., "truly" bare eventives in the sense we have seen) have a type difference from statives and larger phrases with aspect, modals, futures, and/or tense. I will use these to argue for the type difference in (7) above, where events take us from one situation to another in much, but not entirely, the same way that utterances do in dynamic semantics.

My proposal to solve the puzzle in (1) is given in Sect. 4, with three components: (a) the causal model is used to create a causal sequence of situations, (b) eventives update the situation using a relation between situations instead of a situation, and (c) *if* φ, ψ provides sequential updates. With this framework, we see that the combination of the eventive predicate in the consequent and the causal model itself ultimately provides the causal reading of the conditional and thereby explains the puzzle.

A last note before we dive in. This line of inquiry is nowhere near the point of proposing a theory of all conditionals. It turns out to be quite enough for now to argue that event semantics and causal models can be linked to account for the puzzle.

2 True Eventives Are a Different Type Than Statives

2.1 Tests for Eventivity vs. Stativity

To understand what statives and eventives are doing in the simplest English conditionals, we have to have tests to distinguish statives from eventives. Such tests are notoriously language-specific and can be difficult to interpret, but two of the most robust tests are shown in (9) and (10). The first test involves *must*. In (9a), which has a stative predicate, the epistemic reading of the *must* sentence is available; with an eventive, as in (9b), the epistemic reading is not available.

(9) a. Xanthippe must be sick. epistemic reading possible ⇒ stative
b. Xanthippe must get sick. no epistemic reading possible ⇒ eventive

The second test is variously referred to as a biscuit conditional, an Austinian conditional, or a relevance conditional (Austin 1961; DeRose and Grandy 1999; Krifka 2014, and references therein). The idea behind this test is that there can be no causal connection between the prejacents; rather, there is a kind of discourse relation (which one it is has been argued rather closely but is not important to its use here as a test for stativity). So, (10a) comes out something like "if you care to know, then it is relevant to you/I assert that there is beer". To get this reading on a simple conditional it is required to have a stative in the consequent, as in (10a). Using an eventive in the consequent, as in (10b), does not permit this reading.[7]

(10) a. If you care to know, there's beer. relevance reading ⇒ stative consequent
b. If you care to know, we get beer. no relevance reading ⇒ eventive consequent

[7] That is, it does not permit this reading as long as we are not dealing with a generic or habitual reading; this is exactly the point to be discussed just below.

While there is agreement in formal semantics that a broad eventivity/stativity distinction is crucially visible to a number of grammatical phenomena, mostly related to aspect and modality, there are various ways to analyze the distinction. Common to many of them is Davidson's (1967) argument that the meanings of sentences such as *Jones buttered a piece of toast* make reference to an abstract event argument. In further "neo-Davidsonian" developments of this idea, beginning with Higginbotham (1983, 1985, 1986) and Parsons (1990), we can think of the verb phrase "butter the toast" as predicating buttering the toast of an event variable *e*. This idea can be extended to states; how far to extend it is a live issue of contention (see, for instance Kratzer 1995; Maienborn 2005, 2007). For now the assumption we will make is that the verb phrases, whether intuitively eventive or stative, are to be treated as predicates. The question at issue will be what they are predicates of. We will see that while statives make sense as predicates of situations, it makes sense for eventives to be predicates of relations between situations.

2.2 Some Apparent "Eventives" in English Are Actually Stative

It turns out that very often in English, predicates that we expect to be eventive actually behave like statives instead. These stative readings of apparently eventive predicates fall into several categories: generic/habitual readings, futurate readings (where there is a future-oriented, "planned" or "settled" reading but no future morphology [Lakoff 1971; Vetter 1973; Dowty 1979; Kaufmann 2005; Copley 2008, 2018], and what I will call "storytelling readings", where the reference is to what happens in a story, play, etc. (cf. the "director's reading" in Ritter and Rosen [1997]). The fact that generic/habitual readings are stative is well-known, while the behavior of futurate and storytelling readings has not been much talked about (but see Copley [2018]).

So, for instance, an "eventive" predicate such as *drink beer* permits a generic/habitual and epistemic reading for (11a) in a context where one is looking around the throne room and one sees lots of empty beer bottles. Likewise, (11b) gets a relevance reading.

(11) Generic/habitual readings behave like statives
 a. Clavarelian nobles must drink beer.
 b. If you care to know, Clavarelian nobles drink beer.

Futurate readings behave exactly the same way. A context that favors the futurate reading in (12a) is one where, for instance, we see on Xanthippe's calendar "Z: Green Linnet" for tomorrow, and we know that her best friend is Zelinda and the Green Linnet serves only beer. In this context an epistemic reading is possible, and we might also felicitously use the relevance conditional in (12b).

(12) Futurate readings behave like statives
 a. Xanthippe must drink beer with Zelinda tomorrow.
 b. If you care to know, Xanthippe drinks beer with Zelinda tomorrow.

Finally, storytelling readings behave exactly the same way, permitting epistemic *must* as in (13a) and (when in the consequent) relevance conditionals as in (13b).

(13) Storytelling readings behave like statives
 a. (I haven't read to the end of the book, but from what I've read,) Abelard must drink beer at the end of the book.
 b. If you care to know, Abelard drinks beer at the end of the book.

We can ask whether stativity is really the property that makes futurates and storytelling examples behave like statives—indeed, we will return to this question. But it is incontrovertible that in (12) and (13), respectively, futurate and storytelling readings do behave like statives.

So, several readings of "eventive" predicates behave like statives do with respect to the stativity tests introduced above. Readings that are not generic/habitual readings, not plannable (so do not license futurate readings), and not set in a story (so do not have storytelling readings), do not behave this way.

(14) "None-of-the-above" readings do not behave like statives
  a. Yolanda must get sick tomorrow. no epistemic reading
  b. #If you care to know, Yolanda gets sick tomorrow.

(15) a. It must rain tomorrow. no epistemic reading
  b. #If you care to know, it rains tomorrow.

What follows from this is that habitual/generic readings, futurate readings, and storytelling readings are just statives; or more precisely, that the highest predication in their structure is stative. This could happen in one of (at least) two ways. Either a "derived stative" method, where they have an unpronounced stative operator that takes the eventive predicate as an argument; or a sui generis method, where there never is truly an eventive predicate in the semantics and the root itself, or something about the predication itself, is allowed to pick out a stative meaning, and there is still morphological "bareness", i.e., no extra operator. Both possibilities have been suggested for generic/habitual readings: either there is a Gen operator that takes the eventive predicate as an argument, or there is simply a stative predication of something like a kind argument (see Krifka et al. 1995; Carlson 2009). For futurates, many, including Copley (2018), have proposed that futurate meaning is supplied with an unpronounced operator; Kaufmann (2005) is unique as far as I know in leaving the door open for the other possibility. There is a very interesting question here about what exactly morphology tells us about the complexity of denotations; it is immaterial, however, to the present paper.

Returning to the context where a fair coin toss determines whether or not Yolanda becomes the queen, we see that *become the queen* behaves like a (true) eventive, not a stative. This predicate is episodic, cannot be planned so is not futurate,[8] and it is not—to the people living it—in a story.

---

[8] Futurates require temporal adverbials; if (and only if) we drop the assumption that it is a fair election and add a temporal adverbial, we do indeed get a futurate reading. In that case *Yolanda must become the queen tomorrow* has a possible epistemic reading and *If you care to know, Yolanda becomes queen tomorrow* is a good relevance conditional.

(16) "None-of-the-above" readings do not behave like statives (to a speaker within the story, so not a storytelling reading)

a. Yolanda must become the queen.
b. #If you care to know, Yolanda becomes the queen.

Just to remember why the above facts are important: Recall that I am arguing that the causal reading of (1b) depends on its consequent being eventive. It is important to the argument therefore that we only consider truly eventive readings.

2.3 True Eventives Are Not Assertable

Now that we can distinguish true eventives from merely apparent eventives, we can start to argue that statives and true eventives have a type difference. The first step is to see that true eventives in English are not assertable.

(17) True eventive

a. If it rains tomorrow…
b. #It rains tomorrow.

(18) True eventive

a. If Yolanda gets sick tomorrow…
b. #Yolanda gets sick tomorrow.

This observation has been noted by Dudman (1983), Edgington (2003), a.o. Of course, in contrast to (17) and (18), many of the simplest antecedents are assertable on their own, as shown in (19) through (22) below. These correspond, again, to statives and generic/habitual, futurate, and storytelling readings, as in (19), (20), (21), and (22), respectively:

(19) a. If Xanthippe is there. . .
b. Xanthippe is there.

(20) a. If Zelinda drinks beer. . .
b. Zelinda drinks beer.

(21) a. If Xanthippe drinks beer with Zelinda tomorrow. . .
b. Xanthippe drinks beer with Zelinda tomorrow.

(22) a. If Abelard drinks beer at the end of the book. . .
b. Abelard drinks beer at the end of the book.

Finally, although in antecedents, true eventives are possible, it seems that in the consequents of the simplest conditionals (recall, these are conditionals with no aspectual, modal, future, tense, or counterfactual morphology on the prejacents), as long as we resist the temptation to add *will* or similar, true eventives are not possible. Instead, we always get a generic/habitual, a futurate or a storytelling reading. Adding *will* is what makes the episodic reading possible. For example, there is no reading of (23) below which means that if Zelinda should happen to drink beer, Xanthippe will also happen to drink beer. Instead, we get generic/habitual readings either low, in each prejacent, or high, scoping over the whole conditional:

(23) Generic/habitual reading
(If Zelinda drinks beer,) Xanthippe drinks beer.

a. 'If Zelinda habitually drinks beer, Xanthippe habitually drinks beer.'
b. 'Generally, if Zelinda drinks beer, Xanthippe drinks beer.'
c. #'If Zelinda happens to drinks beer, Xanthippe will happen to drink beer.'

The same is true for futurate readings; a low and a high scope for the plan are both possible, but without *will*, there's no merely episodic, non-futurate reading.

(24) Futurate reading
(If Zelinda drinks beer tomorrow), Xanthippe drinks beer tomorrow.

a. 'If there's a plan for Zelinda to drink beer tomorrow, there's a plan for Xanthippe to drink beer tomorrow.'
b. 'There's a plan such that if Zelinda drinks beer tomorrow, Xanthippe drinks beer tomorrow.'
c. #'If Zelinda happens to drink beer tomorrow, Xanthippe will happen to drink beer tomorrow.'

(24a) and (24b) are possible, but what we emphatically do not get is an episodic reading as in (24c) where Zelinda's happening to drink beer tomorrow has something to do with Xanthippe happening to drink beer tomorrow.

2.4 Assertability Has to Do with Type

So, what is it that distinguishes unassertable antecedents from assertable antecedents in English? I want to suggest that it is *type*, in line with the proposal in (7), repeated here below as (25):

(25) a. stative predicate: $\langle s, t \rangle$
b. eventive predicate: $\langle s \times s, t \rangle$

That is, the reason that true eventives are unassertable in English is a formal reason; specifically, they are unable to take a situation argument, because they are of the wrong type to do so.

One argument for this is provided by the fact that only in the simplest phrases does the unassertable–assertable distinction correlate with a conceptual eventivity–stativity distinction. What I mean by this is that it's easy to say that (true) verb phrases about events in English are unassertable, and verb phrases about states in English are assertable; yet we see very similar behavior in larger phrases which we might be more hesitant to say are about events or states. That is, these larger phrases—aspectual, modal, or tensed—absolutely behave like "eventives" or like "statives" on the eventive–stative tests and on assertability. Thus, we must conclude that if these larger phrases are to be thought of as denoting stative or eventive predicates, it is in a very abstract sense, at some distance from any *conceptual* distinction between what states are and what events are. The more abstract distinction seems well-accounted for by a formal property, such as type, as I am proposing here.

Let's take a brief look at some of these larger phrases. With either progressive or perfect aspect added, we get stativity, and also assertability, as in (26). Yet there has been some head-scratching over the nature of the states introduced by aspects (e.g., Hallman 2009); they pass tests for stativity, but what conceptual state do they represent?

(26) a. Xanthippe is writing a letter.
b. Xanthippe must be writing a letter.
c. If you care to know, Xanthippe is writing a letter.

(27) a. Xanthippe has been writing a letter.
b. Xanthippe must have been writing a letter.
c. If you care to know, Xanthippe must have been writing a letter.

Futures provide another example of how larger phrases can have assertability distinctions.[9]

(28) a. #Oh look—it'll rain.
b. Oh look—it's going to rain.

(29) a. #If you care to know, it'll rain.
b. If you care to know, it's going to rain.

The similarity to aspect is such that Copley (2009) analyzed it in terms of aspectualized and unaspectualized futures; as Klecha (2011) points out, however, necessary vs. optional modal subordination is another way to think about it. Either way, if there is a state involved at the highest level of *be going to*, again, it is quite an abstract one.

[9] The epistemic *must* test is not available with futures, presumably for morphological reasons, and the *oh look* in (28) is required to rule out an assertable reading of *will* that relies on evidence of long standing; see Copley (2009) for details.

Past tense poses an interesting case. With past tense in English we only get assertability, and the "stativity" tests are passed handily (*have* here expresses anteriority [Condoravdi 2002]).

(30) a. Xanthippe left.
b. Xanthippe must have left. epistemic reading possible
c. If you care to know, Xanthippe must have left.

Yet, no one would say that past tense sentences are stative. And actually, discourse facts indicate clearly that they are not, since they advance the narrative time while statives do not (ter Meulen 1995: a.o.), as shown in (31). In (31a), Xanthippe leaves after the time the speaker walks in, while in (31b) Xanthippe is there when the speaker walks in.

(31) a. I walked in. Xanthippe left.
b. I walked in. Xanthippe was there.

So we might say that past tense is transparent to the eventivity/stativity distinction; it just passes up the eventivity/stativity of its complement.

The data for combinations of aspect, modals, futures, and tense are of course complex; we have only scratched the surface. But to recap what I want to take away from this, which we can see already in the few examples above: Assertability and non-assertability correlate nicely with eventivity and stativity with the simplest (verb) phrases, but as more material is added, the assertability/non-assertability distinction remains, while any relevant notion of "stative" or "eventive" gets more and more abstract.

This general effect looks something like a grammatical bleaching effect (Meillet 1912; Traugott 1980; Sweetser 1988). The way I would like to think about it here is that the cognitive ontology and the grammatical ontology do not always have to be the same (Borer 2005; Copley and Harley 2015). One way to demonstrate this point is from mismatches between the two. For instance, grammatical gender and conceptual gender do not always match up (e.g., German *das Mädchen* "DET$_{neuter}$ girl"), or grammatical mass/count vs. conceptual mass/count (e.g., *There's* *dog all over the couch*). Typically, the conceptual categories are rich, while the grammatical categories are discrete and narrow.[10]

What I want to suggest here is such a split between grammatical and conceptual categories, in the service of arguing that type is responsible for the eventive–stative difference, in terms of the eventive–stative tests. The grammatical types "state" (for us, s, a situation) and "event" (for us, $s \times s$, a relation between situations) are presumably what the tests diagnose. These are indeed associated, in the verb phrase, with rich conceptual "states" and "events"; but the farther we get from the verb, the less the rich concepts of "state" and "event" matter to "stativity" and "eventivity" in terms of the eventive–stative tests and assertability.

If we now consider, very briefly, what assertion might be, we can see that the notion of type is a good candidate for the kind of formal property required by assertion.

Intuitively, asserting a proposition requires the speaker to present a proposition in a certain way. The proposition itself is something that can have a truth value, and depending on one's ontology, this can be achieved in a number of ways. A popular way to do this, especially in linguistic semantics, is to treat propositions as sets of possibilia such as worlds or situations. In this view, propositions are essentially cases of predication (or equivalently, "description"): they partition a domain of possibilia into a set and its complement, corresponding respectively to the truth values true and false.

There is broad consensus, however, that assertability requires more than a mere predication; it additionally requires an act of something like endorsement or commitment from a mind. Conversely, it is quite possible to have a predication without this act. This point can be seen perhaps most clearly in work by Recanati (e.g., Recanati 2007), who argues that assertion requires the speaker to apply the proposition—itself a predicate of situations—to the current situation. We can call this act a "Recanati reckoning".

On this hypothesis, what goes wrong with our unassertable eventives is that they are not predicates of situations, i.e., they are not type $\langle s, t\rangle$, so assertion, which applies a predicate of situations to the current situation,

10 Both, however, can be explicated using formal techniques.

is not available. Consider the word *ball*, the denotation of which, let's assume, is a predicate of entities, type $\langle e, t \rangle$. The utterance of *ball* on its own is not an assertion. It can certainly be uttered, and there are certainly conditions under which its utterance is more or less pertinent, such as when the speaker is urgently pointing at a ball—which suggests that there are conditions on the *utterance* of the word *ball* by itself—but these utterance conditions cannot be assertability conditions in the sense of a Recanati reckoning. The idea is that *ball* is of the wrong type to be asserted.

I want to suggest that English unassertable eventives run along the same lines. This would require eventives to be predicates of something other than situations, while statives would have to be predicates of situations (because they are assertable). The prime candidate for this other kind of object is that they are "events", and thus that these eventives are indeed predicates of "events", just as it says on the tin. But how do we know what the type of events is?

3 True Eventives Are Relational

We have now seen evidence that English true eventives have a different type from English statives. English true eventives cannot be asserted, which, I argued above, suggests that they are not predicates of situations. We must look at a different kind of argument to try to understand what the type of eventives actually is.

The usual reason given in the aspectual literature (e.g., Comrie 1976; Giorgi and Pianesi 1997; Smith 1991) for the unassertability of true eventives in English is that there is a problem with having an eventive non-aspectualized predicate occur at the present instant because eventive predicates represent change and it is not possible to represent change at one instant. In other words, eventive predicates *at some level at least*, whether the grammatical level, the conceptual level, or both, should involve a relation, change, or comparison between two things; while stative predicates should make reference to only one thing. Such a claim is made succinctly by Hallman (2009), who gives a formal analysis of this kind.

Many mainstream formal Davidsonian approaches adopt something like this principle to explain the unassertability of English eventives, but the principle is not itself formalized into what the grammatical system sees. That is, there is nothing representing a relation, a comparison, or a change in an expression of the form *predicate*(e), though there certainly can be in the concept of "event" that e refers to. Telic predicates may have subevents, but for atelic eventives, generally there aren't two things but only one thing, namely an event e, and thus the expression *predicate*(e) is insufficient on its own to explain why atelic eventives are also impossible in English at the present moment, i.e., *Xanthippe plays baseball* can't mean that Xanthippe is playing baseball. If we want the principle to fall out of English's grammatical system—and we definitely do, since not all languages are like English—it is necessary that our type difference between eventives and statives make eventives involve two of something and statives involve one of something.

Let's call eventives that are predicates of a relation "relational eventives". Despite the popularity of Davidsonian event arguments, where the eventive is ultimately a predicate of an event argument, the idea of relational eventives has been proposed a number of times, rather independently each time. The (at least) two indices necessary for a relational eventive are represented either as temporal variables (which includes Hallman's interval-based analysis of eventives) or as situations. Theories where the indices are conceptually treated as times include, for instance, Croft (2012) and Verkuyl (2019). Theories where the two indices are treated more like situations include Fernando (2004, 2005), van Lambalgen and Hamm (2008), and Copley and Harley (2015, 2021). (See also Krifka [2014] for similar indices used for certain conditionals.)

For our event semantics, we will use a simplified version of what is proposed in Copley and Harley (2015, 2021). In these, event arguments are replaced with function terms—very dumb functions from one situation to another. Again making use of the grammatical–conceptual distinction, the (grammatical) function is associated with a (conceptual) input of energy that may or may not provoke a change from the initial situation to the final situation. Thus here the "two things" are two situations related causally.

In this paper, we will not represent the input of energy itself. That is, we will treat events as mere ordered pairs. The reason we *can* make events in our toy semantics ordered pairs is because we are dealing with perhaps the simplest eventive predicate of all, namely *become P*. While many predicates place requirements on the manner in which we get from one situation to another, *become P* is one of the ones that does not; in particular, it only says that P does not hold in the first situation and holds in the second. Thus, an ordered pair is sufficient to account for this requirement in our toy semantics here.

With this framework in mind, we can return to the difference between stative predicates (type $\langle s, t\rangle$) and eventive predicates (type $\langle s \times s, t\rangle$, as well as understanding how this fits into what we know about assertion.

True eventives in English can't be asserted, I have argued, because they are the wrong type; while they are of type $\langle s \times s, t\rangle$, assertion is looking for something of type $\langle s, t\rangle$. But even though assertion is not possible for eventives we can imagine a different way to use a type $\langle s \times s, t\rangle$ predicate to update the situation—in effect, one is not adding a proposition, as in assertion, but noting that an event occurs (or has occurred, or is occurring). So, formally, assertion and this kind of noting (I will call it "happening") will share something even though they are different. We will update in two ways—one way with type $\langle s, t\rangle$ (i.e., adding a proposition) and one way with $\langle s \times s, t\rangle$.[11]

4 Solving the Puzzle

We are ready to solve the puzzle. The main idea is that the relation inherent to true eventives is, literally, the causal relation from which springs the causal readings of conditionals like that in (1b). The proposal involves a mapping from a causal model to the event semantics of Copley

[11] Usually a distinction is made, in theory as in language, between an entity's information state about the world, and the state of the world itself. The system I will be presenting here seems to do its job without such a distinction; the situations here represent only the way the world is. However, inferences about these situations can be made by using information from the causal model, without reifying information states. I am not sure yet whether this setup amounts to a feature or a bug; the question turns on whether it is necessary or otherwise helpful to reify information states *for these particular examples*.

and Harley (2021). We will do this by indexing the nodes in the causal model and similarly indexing a sequence of situations that will easily map to a relational meaning for eventives. Finally, to show how this mapping between linear causal models and event semantics resolves the puzzle in (1), we will get by with a very simple (perhaps too simple?) dynamic denotation for conditionals.

Incidentally, the indexing is why, in this paper, we can only use a "toy", non-branching (linear) causal model. We are greatly aided by the fact that our causal model has causality going in a single direction ($C \rightarrow Q \rightarrow P$); in other words, we don't have to deal with the complications of a branching network of causal relations with "colliders" ($A \rightarrow B \leftarrow C$) or "forks" ($A \leftarrow B \rightarrow C$). It's simple to inductively define a sequence of situations from the sequence of nodes; harder to say how to index situations from a network of nodes. Do we need to use differently defined situations? Times? And moreover, the answers to these questions may only come, or may come most interestingly, from cross-linguistic data. Thus, in this paper we will stick to a linear causal model. However, the definitions below are in principle useable for branching models, if only the indexing definition can be appropriately changed.

4.1 Causal Models with Indexed Situations

The next two definitions are simplified, sometimes radically, from Pearl (2000: 44) and inspired by similar adaptations from Baglini and Francez (2016), Baglini and Bar-Asher Siegal (2019).

Definition 1 A *causal structure* $\mathcal{D}$ *of a set* $\mathcal{V}$ *of variables* is a directed acyclic graph (DAG) in which each node corresponds to a distinct element of $\mathcal{V}$. Links between the nodes are represented as ordered pairs of elements of $\mathcal{V}$. Each link represents a direct functional relationship among the corresponding variables.

The usual kinship terms (notably "parent" and "child") can be used on these structures. We can also define (in a simplified fashion) *endogenous* and *exogenous* variables:

Definition 2 In a causal structure of a set $\mathcal{V}$ of variables, a variable A is *endogenous* if there is a variable X such that (X, A) is a link. Conversely, A is *exogenous* if there is no variable X such that (X, A) is a link.

Causal models are to be defined as follows[12]:

Definition 3 A *causal model* is a pair $\mathcal{M} = \langle \mathcal{D}, \Theta_{\mathcal{D}} \rangle$ consisting of a causal structure $\mathcal{D}$ and a set of parameters $\Theta_{\mathcal{D}}$ compatible with $\mathcal{D}$. The parameters in $\Theta_{\mathcal{D}}$ assign a function $d_A = f_A(\text{parents}_A)$ to determine the value of each node V, where parents_V is the set of parents of A in $\mathcal{D}$.

The idea behind Definition 3 is simply that by definition, in a causal model, the value of any node A depends on the value of its parent(s); in the example we are dealing with, of course, Q and P each only have one parent and C is endogenous, so has no parent. We will skip over the question of what exactly the set $\Theta_{\mathcal{D}}$ of parameters corresponds to. It can be thought of as a placeholder for whatever it is that determines the value of the child given the value of the parent.

The relations in the causal model behave as in Pearl (2000), where a relation between nodes such as (A, B) conveys that the second node "listens" to the first node. That is, the second node's value is sensitive to first node's value. It is very important to remember that this relation is not always paraphraseable by the main verb *cause*; better words are *influence* or *affect*. The absence of such a relation between nodes conveys that there is no influence from one node to another.

Now, for our toy model, we can further winnow down causal models to those that are linear. The model represented graphically above in (2) is a linear causal model.

[12] Definition 3 is really very simplified indeed with respect to structural causal models. Usually one would include two other factors: an "error" variable such that the child depends on *both* the set of parents and the error variable, as well as a condition that when conditioned on the parents, the child is independent of its non-descendants (the "causal Markov condition"). The latter is a constraint on which models are useful for calculating causal influence; those that do not obey this condition do not seem to be appropriate for causal reasoning. Because the argumentation for this is involved, we will elide it here, but see Pearl (2000) for extensive discussion.

Definition 4 A *linear causal model* is a causal model such that for all its links with arbitrary A, B such that (A, B), and for arbitrary X: if $(X, B) \in$ model, then $A = X$, and if $(A, X) \in$ the model, $X = B$.

Taking a step back, let's remind ourselves what we need causal models to do. Causal models need to constrain the set of possible situations that we are considering, that can be the case. That is, they need to give us at least the information in the table we saw above:

(32)

	C	Q	P
a.	1	1	1
b.	0	0	0
c.	1	0	0
d.	1	1	0

Causal models have another job to do, however. They also need to constrain possible *transitions* between situations in a causal sequence, i.e., possible (relational) events. This is going to be the key to linking causal models to event semantics. In doing this, they will also give us the information in (32)—that is, the information in the table in (32) above will be generated by what we are about to do in (33) below. First we will define our fundamental types and some expressions having to do with predication.

Types. We will use several types: entities (type e, variables $x, y, \ldots$); situations (type s, variables $s, s', \ldots$); and type t which will represent truth values. The situations are, intuitively, to be identified with lines on the truth table, in other words, partial assignments of values to basic predications. The degree type will correspond to truth values for our toy semantics in this paper. We will use the type e to represent the complex type $s \times s$ which represents events, with variables $e, e', \ldots$.

Basic predications. For the nodes of our causal model, we will first distinguish between what we will call *basic predications* (cf. *measure functions* [cf. e.g., Hay et al. (1999)]), and *relations*. Basic predications are of the form $p(\times)(s))$, where p is a function from entities and situations to degrees (values). A *basic predication equation* gives the value of this function evaluated at these arguments. The word "basic" emphasizes that there is no predication more basic than that which is evaluated at a situation; that is, we will not be writing anything like "$p(x)$".

Definition 5 For any p, x, s where p is a function from entities and situations to degrees, x is an entity, and s is a situation, an expression of the form $p(\times)(s)$ is a *basic predication expression*.

Definition 6 An equation of the form $p(x)(s) = d$ is a *basic predication equation*.

Definition 7 An expression of the form $\lambda s.p(x)(s)$ is a *basic predication lambda expression*; it names a function that takes a situation s and returns the degree to which x is p in s.

Nodes map to basic predication lambda expressions. Here we get into the crucial part that allows us to link causal models with event semantics. The nodes in our causal model are going to be interpreted, via an assignment function $\mathcal{G}$, as basic predication lambda expressions, as in Definition 7. So, where we are used to seeing a simple variable with a value, i.e., something like $X = x$, we have now "replaced" the variable with a basic predication lambda expression.

It may not be clear yet why we would want to put situation arguments into the interpretations of nodes. The reason why stems from our working assumption that eventives are predicates of relations from situations to situations; these are the situations at which basic predications are evaluated. So, since we know that eventive verbs are in a sense instances of the arrows in a causal model but with the values of the nodes filled in, and in causal models the arrows are relations between nodes, it makes sense that we need some way to get from nodes to situations. The assignment function $\mathcal{G}$ will do this job for us. It assigns to every node a corresponding expression of type $\langle s, t\rangle$.

(33) $\mathcal{G}(A) = \lambda s.p(x)(s)$

Now our linear causal model needs to build a situation sequence for us. We can use the directions of the arrows to construct by induction a set of basic predication expressions, using the basic predication expressions from each node but indexing situations in each equation. In particular, indices $n, n+1, \ldots$ are added along the direction of the arrows. Then we will assign the same indices to the variables. The indices themselves have no significance except to provide a successor.

Definition 8 An indexing on a *linear causal model* $\mathcal{M} = \langle \mathcal{D}, \Theta_{\mathcal{D}} \rangle$ is defined inductively as follows:

An exogenous variable is assigned the index n (write: A_n). For all B is such that (A, B), B is assigned the index $n+1$ (write: B_{n+1}). With such an indexing, $\mathcal{M}$ is an *indexed linear causal model*.

Definition 9 To define a *causal set of basic predication expressions* for $\mathcal{M}$, let $\mathcal{V}$ be a set of variables; $\mathcal{M}$ be an indexed linear causal model for $\mathcal{V}$; and $\mathcal{G}$ be an interpretation function assigning each variable in $\mathcal{P}$ a basic predication lambda expression. Then, for any node A_n such that A_n is exogenous in the indexed linear causal model M, add $[\mathcal{G}(A)](s_n)$ to the set. For any node B_{n+1} such that B is endogenous in the indexed linear causal model M and add $[\mathcal{G}(B)](s_{n+1})$ to the set.

(34) Causal set of basic predication expressions
$\{c(h)(s_n), q(y)(s_{n+1}), p(x)(s_{n+2})\}$

Along with the few assumptions we made about Clavarelian dynastic politics that gave us the model in the first place, there are some further common-sense assumptions to make.

(35) a. Persistence: Once a queen/princess, always a queen/princess
b. Closed World: If not mentioned, then 0

(35a) is a fact about what it's like to be a queen or a princess in Clavarel. (35b) is a more general constraint. We will need to apply them in the order given (first Persistence, then Closed World) in our derivations to get the right result.

Persistance gives us a different set of basic predications:

(36) Entailed set of basic predications for persistence
$\{q(y)(s_{n+2}), p(x)(s_{n+3})\ldots\}$

4.2 Denotations of Verbal Predicates

Now we can turn to the language side of things, beginning with verbal predicates.

We will assume that the meaning of a stative predicate such as *be the princess* is essentially a basic predication.

(37) $[\![\text{be the princess}]\!] = \lambda x \lambda s.[\![\text{the princess}]\!](x)(s)$

Such a predicate takes an entity as its argument to yield something of type $\langle s, t\rangle$, i.e., a basic predication.[13]

Since events are ordered pairs of situations, eventive predicates are predicates of ordered pairs of situations. Let us consider an arbitrary ordered pair $e = (1_e, 2_e)$ where 1_e and 2_e are both situations. Then *become the princess* has the following denotation, keeping in mind that this is interpreted as involving energetic causation, although we are only reifying the change, not the energy:

(38) $[\![\text{become the princess}]\!] = \lambda x \lambda e \in D_{s\times s}.\neg[\![\text{ the princess}]\!](x)(1_e)$
$\& [\![\text{the princess}]\!](x)(2_e)$

[13] For readers unfamiliar with compositional semantic notation (and see Heim and Kratzer (1998) for an excellent introduction), the evaluation function $[\![\ ]\!]$ relates linguistic expressions to their meanings (or "denotations"). In lambda (λ) expressions as in the right half of the equation in (38), the lambdas allow us to correctly compose meanings together according to the syntactic structure they are in. So, in (38), the meaning of *be the princess* first takes an entity argument x (in our example, x refers to Xanthippe) and then takes a situation argument s. This order of operations is also reflected in the syntactic structure, where compositional operations are only possible between "sisters" in the tree. It is possible to leave the lambdas out here and still understand the proposed solution to the puzzle, but semanticists very much wish to see them there to understand how to compose the meanings.

Again, I stress that the relational argument of eventive predicates here is not going to correspond to one of the arrows in a causal model, but rather to a relation between situations chosen from a sequence of situations that is constrained by that causal model. Causal models provide the structures that are further constrained by eventive predicates. Recall too that the nodes in causal models do not need values in order for that formal structure to be a causal model. To reiterate: **events are not to be identified with the arrows in causal models**, even though they are similarly represented as ordered pairs. While a causal model provides information as to the possible situations and ordered pairs of situations that are compatible with it, it does not pick out a single ordered pair of those situations; this is what an event is. So rather than corresponding to arrows in the causal diagram, events correspond rather to transitions between situations, where the set of possible transitions between situations at any point is constrained by the causal model.

The interpretation of eventive predicates as involving energy tells us that we will be able to say that we can move between situations (update them) through the input of energy. But this is not the only way to move between situations.

Events and assertions thus both take us from one situation to another, but with vastly different interpretations of how we get from one situation to the other: either through energetic causation, which happens in the world, and which puts a condition on the initial situation; or through a mental process of concluding/learning a proposition, which does not put such a condition on the initial situation.

4.3 *If*

For *if*, we need a meaning that will meet two conditions. First, it needs to allow each prejacent to update in its own way (learning or happening). This can be accomplished through compositionality. Second, it needs to ensure that any possible "routes" where φ survives but ϕ does not are removed.

Our toy causal model, however, is linear, and so is the situation sequence associated with it. The situation sequence gives us a single name for the immediately causally accessible situation from s_n—there is only one, namely s_{n+1}. Where $A \to B$ is in the model, then provided that we know whether A holds at s_n, we know whether B holds at s_{n+1}, because that is what the causal model tells us. Effectively, the work that might instead be done by a universal quantifier is here done by the causal model. Here, we don't need to use a quantifier to remove from s any possibilities where φ survives but φ does not, because we already know from the causal model that if φ survives, there is no possible outcome where ψ does not survive. (If alternatively we were to indeed consider outcomes where φ survives yet ψ does not, we would need to assume a different causal model from the one we are assuming.)

Thus, all we *need* for the toy system is the conjunctive update in (39) and this is the update I will use.

(39) $$s[\text{if } \varphi, \psi]^{\mathcal{M}} := s[\varphi]^{\mathcal{M}}[\psi]^{\mathcal{M}}$$

A situation is updated with φ and then the result of that is updated with ψ to yield another situation. All of this takes place with respect to the causal model M. As usual, a successful update is one where the update results in a situation in the model and this yields truth (or if one prefers, acceptability). Note that s must be in the domain of φ as we would expect with indicative conditionals.

This move does however raise some questions.

First, it raises the question of whether even a non-linear causal model would remove the routes where φ survives and ψ does not. The answer here depends on whether knowing the value of an influencing node A determines the value of the influenced node B, where $A \to B$ is in the model. This in turn depends on whether the arrows correspond to functions. Typically this is not taken to be the case; the value of a node is a function of the values of all its influencing nodes. However, models that do have this property have been explored (Copley and Kagan 2021).

Second, even if the conjunctive update were to be generally appropriate, it would immediately raise the question of how to express the difference between conjunction and implication. I can point to the limited use of conjunctive structures having conditional meaning as a suggestive fact linking conjunctions to implications, but it is merely a suggestive fact at this point. I do think that an answer to that question could have something to do with *if* encoding reference to a causal model and *and* not encoding it, but I will not try to make this thought more precise here.

Moving on, we can further define a relation for compatibility/accessibility. This will be used in the "learning" kind of update below. This definition ensures that one can learn that P (I assume that truth values do not change once learned).

(40) Compatibility/accessibility relation R: for arbitrary situations s, s', sRs' just in case all basic predications mentioning s of the form $p(x)(s)$ are such that the truth value of $p(x)(s)$ is equal to the truth value of $p(x)(s')$.

Our two ways to update, one way for statives and one way for eventives, are in (41) below.[14]

(41) Two kinds of update:
  a. learning: $s[\varphi_{\langle s,t\rangle}]^{\mathcal{M}} := \iota s' : sRs'$ and $\varphi(s')$
  b. happening: $s[\varphi_{\langle s\times s,t\rangle}]^{\mathcal{M}} := \iota s' \in \mathcal{S}^{\mathcal{M}} : [\iota e \in \mathcal{S}^{\mathcal{M}}\times\mathcal{S}^{\mathcal{M}} :$ $[2_e = s'$ and $\varphi(e)]]$

Thus, depending on whether φ is a stative (type $\langle s, t\rangle$, (41a)) or an eventive (type $\langle s\times s, t\rangle$, (41b)), different updates are used. Note that *e*, which represents the event, is in a very real sense analogous to the accessibility/compatibility relation in (41a). Another important point is that both of these put a constraint on the starting situation *s*. In (41a), *s* has to be in the domain of *R*, and in (41b), *s* has to be in the domain of *e*. However, in practice, the accessibility/compatibility relation is more

[14] There are in principle a couple of alternative ways to achieve the same effect; we could have instead put it in a dedicated assertion operator, or in a verbalizing head.

permissive than *become*. What this corresponds to is the fact that with *being* in s', s' must only be accessible to/compatible with s. So, in s the value of φ is either true or its value is unknown. However, *become the princess*, as in (38), requires that in the (causally) previous situation s, φ is false. This point turns out to be crucial to solving the puzzle.

4.4 Putting It All Together

Recall that we want to explain (1a) and (1b), repeated here as (42a) (which is true) and (42b) (which is false):

(42) a. If Xanthippe is the princess, Yolanda is the queen.
b. If Xanthippe becomes the princess, Yolanda becomes the queen.

Informally, in both, we want the antecedent to take us from the current situation to another one, and the consequent takes us from that other situation to yet another one; we update first with the antecedent φ, then with the consequent ψ.

The procedure for evaluation is to search the causal direction set of basic predication expressions given by the causal model for the terms needed. This set, whose elements we last saw populating the table in (34), is represented again in (43):

(43) Basic predication expressions, causal direction ($\mathcal{M}$)

$$\begin{aligned} s_n &: c(h)(s_n) \\ s_{n+1} &: q(y)(s_{n+1}) \\ s_{n+2} &: p(x)(s_{n+2}) \end{aligned}$$

So, in (1a), to evaluate the antecedent we search for something in M that tells us the value of $p(x)(s)$ for any s. We have $p(x)(s_{n+2})$, so we start at s_{n+1}. Updating that situation with the antecedent returns s_{n+2}. Now we check to see if we can update s_{n+2} with something that tells us the value of $q(y)(s_{n+2})$. In fact, because of our assumption of persistence (once the queen always the queen), we can learn that $q(y)(s_{n+2})$ is true.

Let $\varphi = p(x)$ and $\psi = q(y)$.

(44)

a. $s[\text{if } \varphi_{\langle s,t\rangle}, \psi_{\langle s,t\rangle}]^{\mathcal{M}}$
$= s[\varphi_{\langle s,t\rangle}]^{\mathcal{M}}[\psi_{\langle s,t\rangle}]^{\mathcal{M}}$ definition of *if*

b. $s[\varphi_{\langle s,t\rangle}] = \iota s' \in S^{\mathcal{M}} : sRs'$
and $\varphi_{\langle s,t\rangle}(s')$ definition of learning

c. $\iota s' \in S^{\mathcal{M}} : [sRs' \text{ and } \varphi_{\langle s,t\rangle}(s')]$
$= s_{n+2}$ since $\varphi_{\langle s,t\rangle} = p(x)$ and $p(x)(s_{n+2})$

d. $s[\text{if } \varphi_{\langle s,t\rangle}, \psi_{\langle s,t\rangle}]^{\mathcal{M}} = s_{n+2}[\psi_{\langle s,t\rangle}]^{\mathcal{M}}$ from (44a) and (44c)

e. $s_{n+2}[\psi_{\langle s,t\rangle}] = \iota s' \in \mathcal{S}^{\mathcal{M}} : s_{n+2}Rs'$
and $\psi_{\langle s,t\rangle}(s')$ definition of learning

f. $\iota s'' \in \mathcal{S}^{\mathcal{M}} : [s_{n+2}Rs'' \text{ and } \psi_{\langle s,t\rangle}(s'')]$
$= s_{n+2}$ since $q(y)(s_{n+1})$ and persistence

The result of these two updates is s_{n+2}, which is in $\mathcal{S}^{\mathcal{M}}$, so this utterance is judged true/acceptable, as desired.

For (1b), unlike (1a), we have to have it that Xanthippe is not princess in the starting situation. So we find ourselves starting in situation s_{n+1}. Then the antecedent takes us to s_{n+2}. Can we say anything about our consequent evaluated at that situation, namely $q(y)(s_{n+2})$? Due to persistence, we can; it is true. This makes it true that Yolanda is the queen in s_{n+2}, but this is not what we were looking for. We were looking for Yolanda to *become* the queen. For this to be true, we would need to know that $q(y)(s_{n+2})$ were *false*, and that $q(y)(s_{n+3})$ were true. This is not the case so (1a) is false. And since we know that $q(y)(s_{n+1})$ is false and $q(y)(s_{n+2})$—that is, that Yolanda becomes queen from s_{n+1} to s_{n+2}—we can see where the impression that (1b) is exactly causally backward comes from.

(45) a. $s[\text{if } \varphi_{\langle s \times s,t \rangle}, \psi_{\langle s \times s,t \rangle}]^{\mathcal{M}}$
$= s[\varphi_{\langle s \times s,t \rangle}]^{\mathcal{M}}[\psi_{\langle s \times s,t \rangle}]^{\mathcal{M}}$ definition of *if*

b. $s[\varphi_{\langle s \times s,t \rangle}]^{\mathcal{M}} :=$
$\iota s' \in \mathcal{S}^{\mathcal{M}} : [\iota e \in \mathcal{S}^{\mathcal{M}} \times \mathcal{S}^{\mathcal{M}} :$
$1_e = s \;\&\; 2_e = s' \;\&\; \varphi(e)]]$ def. happening

c. $\iota s' \in \mathcal{S}^{\mathcal{M}} : [\iota e \in \mathcal{S}^{\mathcal{M}} \times \mathcal{S}^{\mathcal{M}} :$
$[1_e = s \;\&\; 2_e = s' \;\&\; \varphi(e)]] = s_{n+2}$

d. $s[\text{if } \varphi_{\langle s,t \rangle}, \psi_{\langle s \times s,t \rangle}]^{\mathcal{M}}$
$= s_{n+2}[\psi_{\langle s \times s,t \rangle}]^{\mathcal{M}}$ from (45a) and (45c)

e. $s_{n+2}[\psi_{\langle s \times s,t \rangle}]^{\mathcal{M}} =$
$\iota s'' \in \mathcal{S}^{\mathcal{M}} : [\iota e' \in \mathcal{S}^{\mathcal{M}} \times \mathcal{S}^{\mathcal{M}} :$
$[1'_e = (s_{n+2}) \;\&\; 2'_e = s''$ and $\psi(e')]]$ def. happening

f. $\iota s'' \in \mathcal{S}^{\mathcal{M}} : [\iota e' \in \mathcal{S}^{\mathcal{M}} \times \mathcal{S}^{\mathcal{M}} :$
$[1'_e = (s_{n+2}) \;\&\; 2'_e = s'' \;\&\; \psi(e')]] = s_{n+3}$

The result of these two updates is s_{n+3}. But even if we accommodate the existence of an s_{n+3} (we need to redefine $\mathcal{S}^{\mathcal{M}}$ to allow for this but this should not be problematic), we still have a problem. The problem is the contradiction between (46a) and (46b):

(46) a. Yolanda is not queen in s_{n+2} because s_{n+2} must be in domain of f'
b. Yolanda is queen in s_{n+2} because of persistence

So, since the consequent update we are trying to make results in contradictory demands on s_{n+2}, the conditional is judged false/unacceptable. The toy semantics here thus explains both (1a) and (1b) and thereby solves our puzzle.

4.5 Related Examples

There is not too much more we can do with the linearity restriction on our causal models. Classic tests such as strengthening the antecedent and Sobol sequences, for instance, will require non-linear causal models, which we cannot yet link to event semantics. However, there are a few related examples that we can more or less easily treat with a linear causal model and which are likely to be questions in the mind of the reader at this point.

One question is what happens if we mix and match the eventives and statives as in (47a) and (47b).

(47) a. If Xanthippe becomes the princess, Yolanda is the queen. true
b. If Xanthippe is the princess, Yolanda becomes the queen. false

The framework correctly accounts for these judgments. In (47a), Xanthippe becomes the princess from s_{n+1} to s_{n+2}, and by persistence, Yolanda is indeed the queen in s_{n+2}. For (47b), Xanthippe is the princess in s_{n+2}, but from s_{n+2} we can't do a happening update with Yolanda's becoming the queen, for the same reasons as for (1b).

A second question picks up some loose ends from above. How does this framework deal with the futurate, generic/habitual, and storytelling readings, which I argued above are really stative? Given that they are stative, they are predicates of situations and therefore they update via learning. So, for instance, the example in (48), gets the reading that on updating via learning that Xanthippe has a plan to leave tomorrow, we update via learning that Zelinda calls her today.

(48) If Xanthippe leaves tomorrow, Zelinda calls Yolanda today.

A third question that we can answer here has to do with past tense; what happens if we slightly complicate our simplest conditionals by adding past tense to one or more of the prejacents? We have avoided time until now (and correctly, I think, for the examples we were dealing with) but we cannot avoid it forever. So, without getting into questions of temporal anaphora which might make some of the mixed and matched conditionals with past tense feel infelicitous, here is a felicitous conditional which is true relative to our causal model M.

(49) If Xanthippe became princess, Yolanda became queen.

Recall that matrix sentences with past tense on them are generally assertable. That means that both prejacents here are predicates of situations, in our toy semantics; the result in (50) is the same (using here an existential quantifier theory of tense, as in (50)). The assertability of the past tense phrases tells us we need to look at a single situation, not a relation between situations, to evaluate them. They are predicates of situations, not predicates of ordered pairs.

(50) $\textsc{Past}(\mathrm{p}) = \lambda s.\exists s' \text{ BEFORE } s : p(s')$

Consequently, for (50), both prejacents must involve a learning update, rather than a happening update. The conditional has us first learning that, in a past situation relative to *s*, Xanthippe became princess, so we update *s* with that fact; consequently we learn that in a past situation relative to the updated situation, Yolanda became queen. This is congruent with our model, so the sentence is judged true.

Likewise, the framework can also handle "backtracking" conditionals such as (51). The antecedent, which has a true eventive prejacent, is updated via happening, while the past tense consequent is updated via learning.

(51) If Xanthippe becomes princess, Yolanda became queen.

So, the predicted reading is that the antecedent takes us from *s* where Xanthippe is not princess to s' where she is ($= s_{n+3}$) and then we learn, as for (50), that in a past situation relative to the updated situation, Yolanda became queen. This is the case, so (51) is judged true.

References

Austin, John Langshaw. 1961. Ifs and cans. In *Philosophical papers*, 153–180. Oxford University Press.

Baglini, Rebekah & Elitzur A. Bar-Asher Siegal. 2019. *Direct causation: A new approach to an old question.*

Baglini, Rebekah & Itamar Francez. 2016. The implications of managing. *Journal of Semantics* 33. 541–560.

Bar-Asher Siegal, Elitzur & Nora Boneh. 2019. Sufficient and necessary conditions for a non-unified analysis of causation. In *36th West Coast Conference on formal linguistics*, 55–60. Cascadilla Proceedings Project.

Bennett, Jonathan. 1988. Farewell to the phlogiston theory of conditionals. *Mind* 97(388). 509–527.

Bennett, Jonathan. 2003. *A philosophical guide to conditionals*. Oxford University Press.

Borer, Hagit. 2005. *Structuring sense: Volume I: In name only (structuring sense).* Oxford University Press.

Carlson, Greg. 2009. Generics and concepts. In Francis Jeffry Pelletier (ed.), *Kinds, things, and stuff: Mass terms and generics*, 16–35. Oxford University Press.

Comrie, Bernard. 1976. *Aspect*. Cambridge University Press.

Condoravdi, Cleo. 2002. Temporal interpretation of modals. In David Beaver, Stefan Kaufmann, Brady Clark, & Luis Casillas (eds.), *Stanford papers on semantics*. Palo Alto: CSLI Publications.

Copley, Bridget. 2008. The plan's the thing: Deconstructing futurate meaning. *Linguistic Inquiry* 39(2). 261–274.

Copley, Bridget. 2009. *The semantics of the future*. Routledge. *Outstanding dissertations in linguistics*. Routledge.

Copley, Bridget. 2018. Dispositional causation. *Glossa: A Journal of General Linguistics* 3(1).

Copley, Bridget & Heidi Harley. 2015. A force-theoretic framework for event structure. *Linguistics and Philosophy* 38. 103–158.

Copley, Bridget & Heidi Harley. 2021. What would it take to tame the verbal Hydra? (under review). CNRS/Paris 8 and University of Arizona ms.

Copley, Bridget & Olga Kagan. 2021. The Russian perfective and maximality of causal influences. Talk presented at the Converging On Causal Ontology Analyses (COCOA) come-as-you-are online workshop, June 2021.

Croft, William. 2012. *Verbs: Aspect and causal structure*. Oxford University Press.

Davidson, D. 1967. The logical form of action sentences. In N. Rescher (ed.), *The logic of decision and action*. University of Pittsburgh Press.

DeRose, Keith & Richard E Grandy. 1999. Conditional assertions and "biscuit" conditionals. *Noûs* 33(3). 405–420.

Dowty, David. 1979. *Word meaning and Montague Grammar*. Dordrecht: Reidel.

Dudman, Victor H. 1983. Tense and time in English verb clusters of the primary pattern. *Australian Journal of Linguistics* 3(1). 25–44.

Edgington, Dorothy. 1995. On conditionals. *Mind* 104(414). 235–329.

Edgington, Dorothy. 2003. What if? questions about conditionals. *Mind & Language* 18(4). 380–401.

Fernando, T. 2004. A finite-state approach to events in natural language semantics. *Journal of Logic and Computation* 14(1). 79–92.

Fernando, Tim. 2005. Comic relief for anankastic conditionals. In P. Dekker & M. Franke (eds.), *Proceedings of the 15th Amsterdam colloquium*, 71–76.

Giorgi, Alessandra & Fabio Pianesi. 1997. *Tense and aspect: From semantics to morphosyntax*. Oxford University Press on Demand.

Groenendijk, Jeroen & Martin Stokhof. 1990. Two theories of dynamic semantics. In *European workshop on logics in artificial intelligence*, 55–64. Springer.

Hallman, Peter. 2009. Instants and intervals in the event/state distinction. Unpublished ms., UCLA.

Hay, J., C. Kennedy & B. Levin. 1999. Scalar structure underlies telicity in degree achievements. In *Proceedings of SALT*, vol. 9, 127–144. Citeseer.

Heim, Irene. 1982. *On the semantics of definite and indefinite noun phrases*. UMass dissertation.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Blackwell.

Higginbotham, James. 1983. The logic of perceptual reports: An extensional alternative to situation semantics. *The Journal of Philosophy*: 100–127.

Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16(4). 547–593.

Higginbotham, James. 1986. Linguistic theory and Davidson's program in semantics. In E. Lepore (ed.), *Inquiries into truth and interpretation*, 29–48. Oxford: Blackwell.

Kamp, H. 1981. A theory of truth and semantic representation. In J.A.G. Groenendijk, T.M.V. Janssen & M.B.J. Stokhof (eds.), *Formal methods in the study of language*, vol. 135, Mathematical Centre Tracts, 277–322. Amsterdam.

Kaufmann, Stefan. 2004. Conditioning against the grain. *Journal of Philosophical Logic* 33(6). 583–606.

Kaufmann, Stefan. 2005. Conditional truth and future reference. *Journal of Semantics* 22(3). 231–280.

Kaufmann, Stefan. 2013. Causal premise semantics. *Cognitive Science* 37(6). 1136–1170.

Kaufmann, Stefan, Cleo Condoravdi & Valentina Harizanov. 2006. Formal approaches to modality. In *The expression of modality*, 71–106. De Gruyter Mouton.

Kaufmann, Stefan Heinz. 2002. *Aspects of the meaning and use of conditionals*. Stanford University dissertation.

Klecha, Peter. 2011. Optional and obligatory modal subordination. In *Proceedings of sinn und bedeutung* 15. 365–379.

Kratzer, Angelika. 1995. Stage-level and individual-level predicates. *The generic book*. 125–175.

Krifka, Manfred. 2014. Embedding illocutionary acts. In *Recursion: Complexity in cognition*, 59–87. Springer.

Krifka, Manfred, Francis Jeffry Pelletier, Gregory Carlson, Alice Ter Meulen, Gennaro Chierchia & Godehard Link. 1995. Genericity: An introduction. In Gregory Carlson & Francis Jeffry Pelletier (eds.), *The generic book*, 1–124. University of Chicago Press.

Lakoff, George. 1971. Presupposition and relative well-formedness. In Leon A. Jacobvits Danny D. Steinberg (ed.), *Semantics: An interdisciplinary reader in philosophy, linguistics, and psychology*, 329–340. Cambridge, UK: Cambridge University Press.

van Lambalgen, Michiel & Fritz Hamm. 2008. *The proper treatment of events*, vol. 6. Wiley.

Lauer, Sven & Prerna Nadathur. 2018. Sufficiency causatives. *Manuscript, in revision.*

Maienborn, C. 2005. On the limits of the Davidsonian approach: The case of copula sentences. *Theoretical Linguistics* 31(3). 275–316.

Maienborn, C. 2007. On Davidsonian and Kimian states. In I. Comorovski & K. Von Heusinger (eds.), *Existence: Semantics and syntax*, vol. 84. Springer Science and Business Media, 107–130. Springer.

Meillet, Antoine. 1912. *L'évolution des formes grammaticales*. Z. Zanichelli.
ter Meulen, Alice. 1995. *Representing time in natural language*. Cambridge, MA: MIT Press.
Nadathur, Prerna. 2016. Causal necessity and sufficiency in implicativity. *Semantics and Linguistic Theory* 26. 1002–1021.
Parsons, Terence. 1990. *Events in the semantics of English: A study in subatomic semantics*. Cambridge, MA: MIT Press.
Pearl, Judea. 2000. *Causality: Models, reasoning and inference*. Cambridge University Press.
Pearl, Judea & Dana Mackenzie. 2018. *The book of why: The new science of cause and effect*. Basic Books.
Pustejovsky, James. 1995. *The generative lexicon.* MIT Press.
Ramchand, G. 2008. *Verb meaning and the lexicon: A first-phase syntax*. Cambridge University Press.
Recanati, François. 2007. *Perspectival thought: A plea for (moderate) relativism.* Oxford University Press.
Ritter, Elizabeth & Sara Thomas Rosen. 1997. The function of *have*. *Lingua* 101(3). 295–321.
Smith, Carlota. 1991. *The parameter of aspect*. Dordrecht: Kluwer.
Sweetser, Eve E. 1988. Grammaticalization and semantic bleaching. In *Annual meeting of the Berkeley linguistics society*, vol. 14, 389–405.
Traugott, Elizabeth Closs. 1980. Meaning-change in the development of grammatical markers. *Language Sciences* 2(1). 44–61.
Verkuyl, Henk J. 2019. Event structure without naïve physics. In Robert Truswell (ed.), *The Oxford handbook of event structure*, 171–204. Oxford University Press.
Vetter, David C. 1973. Someone solves this problem tomorrow. *Linguistic Inquiry* 4(1). 104–108.
Woodward, James. 2006. Sensitive and insensitive causation. *The Philosophical Review* 115(1). 1–50.

Index

S. Kaufmann et al. (eds.), *Conditionals*, Palgrave Studies in Pragmatics, Language and Cognition, https://doi.org/10.1007/978-3-031-05682-6