The Problem of Epistemic Relevance

Peter Hawke

The Problem of Epistemic Relevance

ILLC Dissertation Series DS-2017-06

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Science Park 107
1098 XG Amsterdam
phone: +31-20-525 6051
e-mail: illc@uva.nl
homepage: http://www.illc.uva.nl/

The Problem of Epistemic Relevance

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PHILOSOPHY
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Peter Hawke

Original Stanford Version - May 2017
ILLC Version - October 2017

Dissertation Committee

Advisors: Prof. Johan van Benthem
Prof. Krista Lawlor

Readers: Prof. Wesley H. Holliday
Prof. Anna-Sara Malmgren
Prof. Brian Skyrms

Defense: January 30, 2017

Department of Philosophy
Stanford University
Stanford, CA
USA

Abstract

I defend a theory of knowability and a theory of the semantics of knowability attributions, in concert. I call the former the *Dretskean theory of knowability* and the latter *resolution semantics*. The former draws on (and attempts to revitalize) classic formulations of the *relevant alternatives (RA) theory of knowledge* due to Alvin Goldman and Fred Dretske. The RA theorist offers an attractive blend of *epistemic modesty* and *epistemic fallibilism*: she grants that the denial of radical skeptical possibilities lies beyond what is knowable, but denies that a failure to refute such possibilities imperils mundane or scientific knowledge claims. The RA approach has, however, met resistance in the literature, clustered around three wide-ranging objections: a dilemma for RA theory between denying that knowledge is closed under deduction and embracing rampant vacuous knowledge; the threat of missed-clue counter-examples; and concerns that the crucial device of *relevance* is simply ad hoc. I argue that these objections can be defused with a sophisticated semantics for knowability attributions, capitalizing on a growing literature on the interaction between knowledge and meaning. For instance, I appeal to recent advances in the study of subject matter and neo-Fregean two-dimensional semantics, developing novel contributions to these areas as I proceed. In an appendix, I summarize my theory using a precise logical framework.

Contents

Preface to the ILLC Version

I owe thanks to Johan van Benthem: it is by his invitation that my dissertation joins the fine company of the ILLC dissertation series. The content of this document is essentially identical to that which was defended at Stanford in January 2017, and that which appears in the Stanford library. However, I took the opportunity to fix a few typos, add a few recent references, express a few points more smoothly and move an overly long footnote into the main text. Further, I have re-formatted the document to better match the style of the ILLC dissertation series.

Amsterdam, The Netherlands
October, 2017

Preface

Does being in a position to know P require information that rules out *every* possible way in which P is false? Traditional arguments for skepticism apparently assume a 'yes' answer. A *relevant alternatives (RA) theorist* answers 'no'. In this dissertation, I bypass prominent objections to relevant alternatives theory with a novel and precise version thereof called *resolution theory*. Resolution theory marries the old with the new. On the epistemic side, it claims that to be in a position to know proposition P is to have empirical information that discriminates P from not-P (a degenerate case: if P is a priori, discrimination requires *no* empirical information). Thus, resolution theory develops a key claim in the groundbreaking work of early RA advocate Alvin Goldman (in concert with ideas from another key progenitor: Fred Dretske). On the semantic side, it claims that the truth of "a knows that φ" requires that a be positioned to know that each proposition in a certain *set* is false: namely, a set of *defeaters* generated by the *subject matter* of φ in coordination with its *Fregean guise*. Thus, we capitalize on recent insights on how the meaning of "a knows that φ" interacts with the meaning of φ, building mainly on David Chalmers, Jonathan Schaffer and Stephen Yablo. Despite these many debts, resolution theory is a novelty, contrasting with its forerunners in critical ways. I argue that resolution theory overcomes objections that many RA theories fall prey to: Schaffer's problem of *missed clues*; the *closure dilemma*; and worries concerning *ad hocness*. In particular, these objections apply, to varying extents, to the theories we draw inspiration from. Among other consequences, resolution theory motivates a novel framework for epistemic logic, broadly situated in the modal tradition.

Structure. The main story of the dissertation is told over four chapters. Chapter 1 sets the scene, by introducing the RA approach (as we conceive of it), resolution theory and the three wide-ranging problems that RA theories tend to face. We conclude by construing these problems as challenges for resolution theory, and sketch how a resolution theorist might respond. The other core chapters answer the challenges in detail and unearth further motivations for resolution theory. As we proceed, I develop aspects of resolution theory in precise detail.

In chapter 2, we survey the landscape of extant theories of subject matter, and argue that no such theory (in our sample, anyway) jointly satisfies a certain set of prima facie desiderata. In contrast, I introduce the *issue-based theory* of subject matter, and argue that it *does* meet these desiderata. This motivates resolution theory's treatment of subject matter (and, hence, its account of the key notion of *relevance*).

In chapter 3, we intervene in the debate over epistemic closure. Resolution theory rejects epistemic closure in full generality. It is therefore at odds with forceful intuitions concerning the relationship between deductive reasoning and knowledge extension. I argue that a theory that rejects closure is defensible if it meets certain criteria of adequacy, and confirm that resolution theory meets the criteria. Along the way, we cast doubt on rival closure-denying theories due to Schaffer and Yablo.

In chapter 4, I explore an argument from Gibbard for the conclusion that every fact is knowable a priori. I argue that the paradoxical nature of this argument can be defused if one accepts a neo-Fregean semantics, along the lines of the epistemic two-dimensionalism of Chalmers. This response relies on a deflationary account of contingent a priori truth. With this in mind, I discuss how to undermine a version of Gibbard's argument that targets *substantive* knowledge specifically and propose that resolution theory (which incorporates key aspects of the two-dimensionalist picture) provides the requisite tools. Finally, with the neo-Fregean aspects of resolution theory on the table, I show that resolution theory is positioned to respond to *missed clue cases*, which serve as counter-examples for a wide range of RA theories.

I briefly tie things together in chapter 5, and indicate directions for further research.

Two appendices fill out extra details for the interested reader. The first delves more deeply into the motivations for and possible forms of RA theory. The second presents resolution theory as an austere logical framework, in detail.

Methodology. One useful view of philosophy is as an exercise in *theory-building* that is responsive to a wide range of considerations. For example, philosophers tend to be responsive to (and prepared to contrast and weigh) common sense; intuition; scientific orthodoxy; theoretical elegance; explanatory power; and so on. I follow suit.

The theory of knowledge witnesses an interplay between two topics that are sometimes hard to pull apart. One might call the first *pure epistemology* or the *metaphysics of epistemology*: the study of the nature of our epistemic states, tying closely to theories of information (and information flow) and mind (especially the propositional attitudes). The other is the *semantics* of our epistemic vocabulary (in particular, the semantics of knowledge ascriptions).

The standard 'semantical' tools of modern logic are a fruitful means for studying this interplay. Such a framework gives a precise specification of three items: (i) a class of models, described using the powerful tools of set theory; (ii) a formal logical language, generated by a context-free grammar; and (iii) a system of semantical devices (normally a valuation function and a system of truth clauses) that relate the language to the models. As applied to epistemology, item (i) should serve as a precise model of an agent's (or agents') epistemic situation. Item (ii) should serve as a simplified model of natural language that incorporates the epistemic vocabulary of immediate interest. Item (iii) serves as a model of the relationship between epistemic vocabulary and the epistemic properties of the agent under consideration.

There are numerous advantages associated with developing and evaluating frameworks of this type. For one, the tools are standard and well-understood. For another, operating with precision opens the door for rigorous further study e.g. the study of the meta-logical properties of the system, or incorporation into the more complex framework of the formal linguist. For another, mathematical methods help us avoid ambiguity in philosophical theorizing, for instance in the determination of the consequences of a proposal, or in how it contrasts to rival proposals. For another, such frameworks keep semantical and non-semantical issues clearly demarcated.

With this in mind, resolution theory is composed of a theory of the nature of knowability in conjunction with a semantics for knowability ascriptions. I strive to develop both with a precise logical framework clearly in view (culminating in the presentation in appendix B).

Throughout, I make use of standard logical and set theoretic notation. $\neg$ indicates negation; $\wedge$ indicates conjunction; $\vee$ indicates disjunction; $\rightarrow$ indicates the material conditional; $\Box$ indicates the necessity operator; $\vDash$ indicates a satisfaction relation between a model and a sentence; $\in$ indicates set membership; $\subseteq$ indicates the subset relation; $\cap$ indicates set intersection; $\cup$ indicates set union.

Acknowledgments

Writing a doctoral dissertation, it turns out, is less like scaling a mountain and more like escaping a dungeon. The former rewards one's arduous labor with a grand vista. The latter involves finding oneself in a dim, confined space. Escape is unimaginable. Somewhat miraculously, one crawls back into the light several years later and tries to re-adjust to normal life.

Thanks, foremost, to my principal advisors: Krista and Johan. Thanks for listening to all the ideas that worked and (especially) all the ideas that did not. Your patience, encouragement and willingness to let me chart my own course were greatly appreciated.

Thanks to the rest of my committee - Anna-Sara, Brian and Wes - for challenging me to write a clearer, more forceful story.

Thanks to Shane Steinert-Threlkeld for being a great collaborator and offering me a much needed opportunity to, now and then, work on something other than the theory of knowledge.

Thanks to Blake Francis and Katy Meadows for many, many discussions on my work (and other bits of philosophy) over the years.

Thanks to Jill Covington for rescuing me from various administrative mishaps and otherwise over-seeing the logistics of the dissertation.

Thanks to David Martens, Mark Leon and Scott Stapleford, who encouraged me to pursue philosophy in the first place.

Thanks to my family, who suffered through many years of watching me cognize, write and anguish over the current work: Nicolyn, Michaela, Luna and my

parents.

The material in chapter 3 and appendix A, though updated in various ways, is largely based on articles that have appeared in print.

- "Theories of Aboutness", Australasian Journal of Philosophy, forthcoming.
- "Questions, Topics and Restricted Closure", Philosophical Studies, 73(10): 2759-2784, 2016.
- "Relevance in Epistemology and Logic", In Pombo Martins, Redmond and Nepomuceno Fernández, editors, Epistemology, Knowledge and the Impact of Interaction, volume 38 of Logic, Epistemology, and the Unity of Science, pages 207-235. Springer, 2016.

I thank the publishers for permission to include the overlapping material.

Amsterdam, The Netherlands
October, 2017

Chapter 1

The Problem of Epistemic Relevance

The purpose of this dissertation is to defend a novel theory of knowability called *resolution theory*. The term 'resolution' is borrowed from Yalcin [2011] and Yalcin [2016].[1] Metaphorically, to impose a resolution on *logical space* (i.e. the space of possible ways that things can be) is to divide that space up by focusing only on certain distinctions between the possibilities, ignoring more fine-grained distinctions. According to resolution theory, an ordinary knowledge ascription expresses, roughly, that a content is known *in a way that is relative to the alternatives at a given resolution.*

My aim is to contribute to an ongoing and multifarious project in the epistemology literature: that of properly grounding the *relevant alternatives approach* to knowledge. Resolution theory incorporates various influential ideas from the relevant alternatives tradition into a precise new framework. The resulting theory has the resources to evade wide-ranging objections to the relevant alternatives approach. Or so I will argue.

The current chapter spells out my goals in more detail: I here introduce the relevant alternatives approach (section 1.1), resolution theory (section 1.2), the objections to the relevant alternatives approach that concern us (section 1.3) and hint at how resolution theory answers these objections (section 1.4). Finally, section 1.5 lays out a plan for filling out this sketch over the coming chapters.

[1] Yalcin builds on ideas in Stalnaker [1987], chapter 4, and Lewis [1988a]. The idea of a resolution (and its formal implementation as a partition on worlds) also finds utility in the literature on deontic claims: see Cariani [2013].

1.1 The relevant alternatives approach

1.1.1 Knowability, modesty and fallibilism

This dissertation concerns the theory of knowledge. Our approach is slightly indirect. Our broad topic is *knowability* relative to a given body of empirical information: relative to information E, what is possible for an agent to know?[2] That is, rather than studying knowledge *per se*, we abstract away from the contingent cognitive limitations of ordinary agents, and train our focus on the quality of the information available to them. We are thus interested in a necessary condition on having knowledge. Knowability is of central interest (and contingent psychological limitations are often distracting) in at least three important philosophical debates. First, the question of *radical Cartesian skepticism*: how is it that mundane facts are knowable given ordinary empirical information that fails (it seems) to discriminate between genuine sensory perception and systematic sensory deception?[3] Second, the question of *epistemic closure*: in principle, is every joint consequence of a set of known propositions knowable via deduction?[4] Third, the question of *a priority*: how can it be that substantial facts about the world are, in principle, knowable without any empirical information?[5] Our own discussion will touch on all three of these issues.

More specifically, our topic is the *relevant alternatives (RA) approach to knowledge* (even *more* specifically, it will be a particular theory along this line). As we understand her, an RA theorist maintains the following: one knows something only if one's information rules out all of the *relevant* alternatives. It need not be the case that one's information rules out *all* of the alternatives.[6] At its

---

[2] We delay for elsewhere any attempt at relating our results to the debate over Fitch's knowability paradox. For discussion of this paradox, see, for instance, [Williamson, 2000, Ch. 12], van Benthem [2004] and Kvanvig [2006].

[3] For a recent overview of the debate concerning radical skepticism, see Vogel [2014] and Fumerton [2014].

[4] For an introduction to the closure debate, see Luper [2016], Hawthorne [2005] and Dretske [2005]. Note that an appeal to knowability is an increasingly popular way of framing the issue of closure (though the exact notion of knowability that is used is fluid). See, for instance, [Schaffer, 2007a, pg. 235] and [Blome-Tillmann, 2014, section. 6.1].

[5] For a recent overview of the debate concerning the a priori, see Bonjour [2014] and Devitt [2014].

[6] Several classic papers remain the best introduction to the relevant alternatives approach: Dretske [1970], Goldman [1976], Stine [1976] and Dretske [1981]. Bradley [2014] offers a sympathetic and accessible recent discussion at, what seems to me, roughly the right level of abstraction for properly appreciating the approach. For some influential critiques, see Vogel

core, then, the RA theorist proposes a theory of knowability: roughly, that one is in a position to know something just in case one's information rules out all of the relevant alternatives, where 'relevance' determines a non-trivial restriction on the possibilities.

In particular, RA theorists are apt to claim that knowing a mundane (or even scientific) claim does not require ruling out pathological possibilities that are immune to refutation. To know that one has a hand, one needn't rule out that one's senses are being systematically deceived by an all-powerful evil demon. To know that a fair coin (if tossed) will land heads or tails, one needn't rule out that the laws of nature will stall while the coin hovers indefinitely in the air. To know that the tap water is safe, one needn't rule out that a rogue CIA unit is lacing the water supply with a mind-controlling drug. These alternatives seem properly classified as paranoid fantasies, not serious possibilities. Relatedly, RA theory answers the Cartesian skeptic, who argues: knowledge requires information that rules out all alternatives; perceptual experience does not rule out the possibility of systematic sensory deception; thus, ordinary agents lack knowledge of even mundane claims. The literature is littered with inconclusive attacks on the second premise or the coherence of the argument. RA theory offers a fresh approach: deny the first premise.

As this indicates - leaving more specific arguments for appendix A - the appeal of the RA approach boils down to two ideas that have found special currency in contemporary epistemology. First, *epistemic modesty*: a theorist would be remiss to exaggerate the epistemic powers of ordinary individuals.[7] Second, *epistemic fallibilism*: modesty should not lure us into skepticism. Rather, we should posit that whatever epistemic status we point to with ordinary knowledge ascriptions, this status is (in some sense) compatible with being in error.[8]

Over the next eight sections, I clarify some of the finer points concerning our treatment of the RA approach.

[1990],Vogel [1999], Sosa [2004], Ch.2 of Hawthorne [2004] and Ch. 1 of Stanley [2005]. It is notable that many (but not all) of these critics associate the RA approach very closely with contextualism.

[7]Though the idea is implicit in a great deal of discussion on the RA approach, see Schaffer [2007a] for explicit appeal to a notion of epistemic modesty in defense of a RA theory along the RA approach.

[8]For illuminating discussions of fallibilism, see Cohen [1988] and [Fantl and McGrath, 2009, Ch.1].

1.1.2 The RA approach vs. RA theories

At the general level, I prefer to talk about the RA *approach* to knowledge, as opposed to RA *theory* (the latter is the label used most often in the literature). Stated barely, the core RA ideas provide a *sketch* of a theory. To speak suggestively: not only do these bare ideas lack sufficient *form* to constitute a *formal* theory, but they lack sufficient *content* to constitute a *substantive* theory. To arrive at a specific RA theory (i.e. one variant of the RA approach), three core concepts must be fleshed out: *relevance*, *alternative* and *ruling out*. Thus, there are *many* RA theories. Some are merely formal, and some are substantive.

1.1.3 Ruling out and agency

We conceive of *ruling out* as something that a body of information does, not something that an agent *per se* does. If we say that an agent has ruled out a proposition P, we mean that the *agent's information* rules out P. Thus, the agent can be ignored in our models and theorizing. Capturing the available information, on the other hand, is crucial.

1.1.4 Relevance, in general and in particular

In describing the RA approach in generality, we make few commitments regarding the notion of 'relevance'. For instance, we do not commit to relevance being a matter of rationality, or otherwise. Nor should the reader be too quick to equate relevance, as we use the term, with other (possibly related) notions that have appeared in the philosophical literature: for instance, in the study of pragmatics [Grice, 1975]; the study of scientific confirmation [Floridi, 2008]; or the study of logical consequence and relevant logic [Mares, 2014]. (In appendix A, I offer some tentative suggestions for relating some of these areas to our own issues.) Indeed, we see the key task of a *particular* RA theory to be the provision of an account of relevance.

Nevertheless, we understand any satisfactory account of relevance as having an essential feature: relevance is a *non-evidential* restriction on the space of possibilities. That is, however we cash it out, relevance should *not* be understood as capturing how the *available evidence* places limits on the 'serious' possibilities.[9]

[9] Compare this to the discussion in Vogel [1999] of RA accounts that *backslide* by being

For a sense of this feature, note some factors that are frequently posited as determining relevance: psychological salience of an alternative to the attributor and/or the subject of a knowledge ascription;[10] similarity to actuality;[11] presupposition;[12] conversational relevance to the question or topic under discussion;[13] compatibility with the agent's beliefs;[14] a reason (even if far from conclusive) for thinking the alternative is the case;[15] the practical stakes connected with ignoring an alternative.[16]

---

equivalent to a theory that makes no use of a notion of relevance.

[10][Stine, 1976, pg. 256] courts this proposal: "Perhaps the mere utterance of the former sentence is enough to make us loosen up our notion of what counts as a relevant alternative". This thread is picked up ensuing contextualists: for instance, see the discussion of the 'rule of attention' in Lewis [1996].

[11]See the discussion of *remoteness from actuality* in Dretske [1981]. Also see the discussion of the 'rule of resemblance' in Lewis [1996]. Finally, compare the discussion of relevance in Heller [1989] (a position that is perhaps too close to belief-tracking views to count as an RA theory, by our lights).

[12]One can draw a distinction between a *semantic presupposition* and a *pragmatic presupposition*. In the first case, the presupposition is part of the meaning of the sentence; in the second, a presupposition is an attitude that an agent holds towards a claim. Dretske [1970] suggests that semantic presuppositions are irrelevant. Blome-Tillmann [2014] develops the view that pragmatic presuppositions are the root of irrelevance, building on a thread that runs through the contextualist literature, starting with [Goldman, 1976, pp. 776-777] and [Stine, 1976, pp. 255-256]. For an influential discussion of pragmatic presuppositions, see Stalnaker [1972].

[13]For a seminal work in linguistics that ties the notion of a discourse topic very closely to a question under discussion, see Roberts [2012]. For a development of an RA semantics that models the relevant propositions as a question, see Schaffer [2004], Schaffer [2005a], Schaffer [2007a] and Schaffer [2007b]. For a discussion that models the relevant propositions as determined by subject matter, see Yablo [2012] and Yablo [2014].

[14]Cf. the discussion of the 'rule of belief' in Lewis [1996].

[15]See [Stine, 1976, pg. 252], which also traces the view to Goldman [1976]: "Goldman seems to hold what I regard as the correct version of it [i.e. the correct account of relevance], which is that: (1) an alternative is relevant only if there is some reason to think that it is true". This sort of view is also sometimes attributed to Austin [1946], an important fore-runner of the RA approach. See Leite [2012]. Also see the account of *reasonable alternatives* in Lawlor [2013], which is inspired by Austin's discussion. Here is what [Austin, 1946, pg. 87] says: "The doubt or question 'But is it a *real* one?' has always (*must* have) a special basis, there must be some 'reason for suggesting' that it isn't real, in the sense of some specific way, or limited number of specific ways, in which it is suggested that this experience or item may be a phoney. Sometimes (usually) the context makes it clear what the suggestion is: the goldfinch might be a mirage but there's no suggestion it might be stuffed. If the context doesn't make it clear, then I am entitled to ask 'How do you mean? Do you mean it might be stuffed or what? *What are you suggesting?*' The wile of the metaphysician consists in asking 'Is it a real table?' (a kind of object which has no obvious way of being phoney) and not specifying or limiting what may be wrong with it, so that I feel at a loss 'how to prove' it *is* a real one."

[16]Cf. ch. 4 of Hawthorne [2004]; and the introduction and ch. 5 of Stanley [2005].

1.1.5 Diversity in the RA literature

The literature that explicitly and implicitly explores the RA approach is large and diverse, and closely connected to still further complicated debates. For instance, classic discussions of the RA approach include: Dretske [1970], Stine [1976], Goldman [1976] and Dretske [1981]. Recent, explicit support for the view is offered in: Heller [1989], Lewis [1996], Heller [1999], Pritchard [2010], Pritchard [2012], Lawlor [2013] and Holliday [2015b]. Closely related discussions include those concerning conclusive reasons;[17] tracking theories of knowledge;[18] contextualism[19] and, more generally, the interaction between knowledge ascriptions and meaning in context;[20] subject-sensitive invariantism;[21] and the transmission of warrant and other epistemic properties.[22]

I do not claim, therefore, to present the RA approach in a manner that accommodates *every* discussion that sensibly uses (or could use) the label. The dialectic is too unwieldy.[23]

1.1.6 RA and fallibilism

I do not *equate* fallibilism and the RA approach. Doing so would either take the focus off the unique features of the latter, or render the former term objectionably narrow. For instance, a *Bayesian* approach to knowledge can be presented in the *form* of an RA theory: we could say that proposition P is *ruled out* just in case it is inconsistent with the basic evidence E, but that an alternative A is *relevant* only if its probability (degree of rational credence) is sufficiently high given E. This account, however, violates the *substantive* requirement that relevance is a non-evidential constraint, so we do not consider it a *bona fide* RA theory. Further, consider the truth-tracking theories of knowledge advocated by Nozick [1981] and Sosa [1999]: ignoring certain subtleties, these theories claim that to know P is to have beliefs that match the truth value of P across a certain range of counter-

[17] See Dretske [1971], Lando [2016].

[18] See Nozick [1981], Luper [1984],Sosa [1999], Williamson [2000], Kripke [2011].

[19] See Cohen [1988], Cohen [2002], Preyer and Peter [2005], DeRose [2009], Blome-Tillmann [2014].

[20] See Schaffer [2004], Schaffer [2005a], Yablo [2014].

[21] See Hawthorne [2004], Stanley [2005], Fantl and McGrath [2009].

[22] See Wright [2003], Wright [2004], Wright [2014].

[23] For a heroic incorporation of vast swaths of this literature into a coherent discussion, see Holliday [2013]. The unifying issue in this work is that of epistemic closure.

factual situations. The range in question may be termed the *relevant* possibilities, giving things an RA-like construal. However, these theories do not have the *form* of a theory of knowability, in our sense: they do not specify what possibilities are ruled out by the given *empirical information*, and instead focus on the contingent beliefs of the agent.

1.1.7 RA and internalism/externalism

At the general level, the RA approach should not be strongly associated with either externalism or internalism about knowledge or justification. This is not the case for prominent *instances* of the approach, e.g. Dretske [1981] and Goldman [1976] are clearly motivated by an externalistic picture. However, in contrast, consider an RA theory that posits that relevance is a matter of non-evidential, subjectively rational entitlement (cf. Wright [2014]). This seems as internalistic a theory as any.

1.1.8 RA semantics for knowledge attributions vs. RA theory of knowledge

The RA approach can be developed as a theory of the *nature of knowledge* or as a semantic theory governing *knowledge ascriptions.* Intuitively, a knowledge ascription reports that an agent has an epistemic profile of a certain fundamental kind. A theory of the nature of knowledge aims to isolate and model the sort of thing that is being reported on. A semantic theory aims to model the relationship between the state being reported on and the language that does the reporting. I refer to a theory of the former kind as an *RA theory of knowledge*, or *RA theory of knowability* (as the case may be). I refer to the latter as an *RA semantic theory.*

An RA theory of knowledge or knowability requires a notion of *mental content*: a conception of the information that an agent can *receive* (say, some empirical information E) and be *committed to* (say, as the object of a belief). Let C be a content. Then the basic commitment of an RA theory of knowability can be stated thus:

> **RA theory of knowability:** Agent a in situation S is in a position to know C just in case: the empirical information that is available to

> a in S rules out every relevant alternative A to C, where relevance is a function of a and/or S.

On the other hand, the basic commitment of an RA semantic theory is to a kind of *truth clause*. Continuing on our theme, we do not discuss knowledge ascriptions *per se*, but rather knowability ascriptions. The latter is a technical device, though we posit a close relationship between the technical notion of *knowability* and the ordinary concept of *knowledge*. In particular, we expect a knowledge ascription to *entail* certain knowability ascriptions, and our linguistic intuitions concerning knowledge ascriptions often bear on knowability ascriptions. It will often be useful to work with a formal language. As a start, read $K\varphi$ as "the agent in question is in a position to know φ relative to her empirical information". If it aids clarity, we emphasize the use/mention distinction by contrasting φ with $\ulcorner\varphi\urcorner$ (where φ is an interpreted sentence in our language).

> **RA semantic clause:** Relative to a model $\mathcal{M}$ that supplies E, the sentence $\ulcorner K\varphi\urcorner$ is true (verified) just in case: every relevant alternative to φ is ruled out by E, where relevance is determined by the components of $\mathcal{M}$.

One job of model $\mathcal{M}$, in the above setting, is to supply a representation of an agent's epistemic situation. Thus, it is possible to have an account that posits a simple relation between the correct RA theory of knowability and the correct RA semantic clause: relative to a model $\mathcal{M}$ that supplies agent a and situation S (and therefore information E), the sentence $\ulcorner K\varphi\urcorner$ is true (verified) just in case a has sufficient information in S to rule out every relevant alternative to C, where C is the content associated with $\ulcorner\varphi\urcorner$ and relevance is a function of a and/or S. In this case, one settles on an RA theory of knowability first, and accept an RA semantics as a convenient consequence.

However, an RA semantics need *not* generate an RA theory of knowability. For the model $\mathcal{M}$ is usually taken to represent both the circumstances of the agent (the *context of evaluation*) and the circumstances of the interlocutors discussing the agent (the *context of utterance*). An RA semantics may thus posit that relevance is (partly or wholly) a function of the context of evaluation. In this case, relevance is understood as a (partly or wholly) *linguistic* fact. In contrast, an RA theory of knowability understands relevance as a function purely of the circumstances of the agent in question. If relevance is a purely linguistic phenomenon governing how we *talk* about an agent's epistemic status, then it cannot be represented in such a theory.

Thus, an account of relevance may rely entirely on (what are sometimes called) *attributor factors*, or entirely on *subject factors*, or on a combination of the two.[24] If one's account of relevance only appeals to subject factors, then it is often distracting to frame things in a semantic framework. If one's account appeals only to attributor factors, then an appeal to a semantic framework is essential.

Let me illustrate the foregoing with a case study of an RA semantics that is *not* based on an RA theory of knowability: Schaffer's *question-sensitive RA semantics*, in conjunction with his *contrastivist theory of knowability*.[25]

Schaffer proposes *contrastivism* about the knowability relation: instead of a three-place relation $\mathbb{K}(a, S, C)$ between an agent, a situation and a content, it is posited to be a four-place relation $\mathbb{K}(a, S, C, A)$ where A is a content that is inconsistent with C.[26] The leading idea is that the knowledge relation holds -

[24] Appeals to both kind of factor can be found in classic work, such as Dretske [1970] and Goldman [1976]. The distinction is discussed more clearly in [DeRose, 1992, section II], and forms the basis of the contrast between proponents of contextualism (e.g Stine [1976], Cohen [1988]) and subject-sensitive invariantism (e.g. Hawthorne [2004], Stanley [2005]). Here is what [Goldman, 1976, pg.256] has to say: "There are two views one might take on this general problem [of fixing a criterion of relevance]. The first view is that there is a "correct" answer, in any given situation, as to which alternatives are relevant. Given a complete specification of Henry's situation, a unique set of relevant alternatives is determined: either a set to which the facsimile alternative belongs or one to which it doesn't belong. According to this view, the semantic content of 'know' contains (implicit) rules that map any putative knower's circumstances into a set of relevant alternatives. An analysis of 'know' is incomplete unless it specifies these rules. The correct specification will favor either the skeptic or the skeptic's opponent.The second view denies that a putative knower's circumstances uniquely determine a set of relevant alternatives. At any rate, it denies that the semantic content of 'know' contains rules that map a set of circumstances into a single set of relevant alternatives. According to this second view, the verb 'know' is simply not so semantically determinate. The second view need not deny that there are regularities governing the alternative hypotheses a speaker (i.e., an attributer or denier of knowledge) thinks of, and deems relevant. But these regularities are not part of the semantic content of 'know'. The putative knower's circumstances do not mandate a unique selection of alternatives; but psychological regularities govern which set of alternatives are in fact selected. In terms of these regularities (together with the semantic content of 'know'), we can explain the observed use of the term."

[25] Schaffer [2005a] is keen to contrast his position with what he calls 'relevantism', by which he seems to mean the broad RA tradition. However, his account of relevantism builds in commitments that are absent from our own account of the RA approach, so I feel comfortable classing his theory as along the RA line.

[26] Note that Schaffer talks about *knowledge*, not *knowability* (see e.g. Schaffer [2005a]) but this is not crucial for our discussion. Further, note that our three and four-place relations can plausibly be collapsed into a two and three-place relation: one simply takes the agent to be part

i.e. a is in a position a position to know C in situation S, relative to contrast A - just in case the agent's information in S is inconsistent with ('rules out') A. Note that, despite a superficial similarity, this is *not* an RA theory of knowability in our sense. For note that an RA theory of knowability, as presented above, is an explication of an at-bottom *three-place* relation $\mathbb{K}(a, S, C)$, using a notion of relevance determined by a and C.

On the other hand, Schaffer *does* offer an RA semantics. Consider a knowability ascription $K\varphi$ (i.e. "a is in a position to know φ"). This 'binary' ascription makes no mention of a contrast to the content of φ. So, on Schaffer's view, what knowledge relation is thereby expressed? His answer: the contrast is supplied by the discourse context (as a *salient question* associated with φ).[27] This is an RA semantics: knowledge ascriptions are evaluated relative to model that provides a relevant alternative (i.e. a contrast) to every sentence, with relevance representing a fact about the context of utterance.

As we will see, resolution theory offers both an RA theory of knowability and an RA semantics. However, the relationship between the two is not straightforward, so both accounts are essential.

1.1.9 Empirical information and skeptical possibilities

Defusing the threat of Cartesian skepticism is a universal motivation for RA theorists. What's more, RA theorists approach this challenge from a common starting point: that an ordinary agent's empirical information is (always) consistent with certain skeptical possibilities. For instance, brain-in-vat scenarios - i.e. situations of systematic sensory deception - serve as an invariant skeptical possibility, persisting no matter the information available to the agent. I too accept this starting point.

With this in mind, I model empirical information E as follows: E is a set of possible worlds that contains the actual world @. (It won't matter for our discussion if these are considered *centered* worlds, or not.) This is in line with

of the situation, or understand "agent" to refer to a particular agent at a particular time and place. We adopt the latter convention later in the dissertation, but for now keep our discussion as explicit as possible.

[27]See [Schaffer, 2005a, pg. 205]: "Moving finally to declarative sentences (perhaps the rarest form in natural language), these inherit their contrasts from context…In general, context provides the default source of contrasts".

standard accounts of semantic information.[28] Crucially, I assume throughout (sometimes implicitly) that E contains a brain-in-vat world in which the agent's basic empirical beliefs are systematically mistaken.

We can further substantiate our conception of E (at actual world @): E can be thought of as the set of worlds at which the agent has the same *total basic perceptual evidence* as @. That is: the set of worlds where the agent has the same occurrent perceptual experience and *memory* of perceptual experiences. Thus, the information provided by perceptual experiences constitutes the elimination of certain possibilities: those where the agent's experiences differ.[29]

(Above, I called the available experiences of the agent *basic perceptual evidence*. I think this is a natural use of the term 'evidence'. However, I do not put forward a theory of evidence in this dissertation, so this is best understood as a technical usage.)

This is an illuminating idealization. However, it is a mistake to identify the empirical information of an ordinary agent with the information that an agent with *perfect uptake* and *perfect recall* receives from the same history of perceptual experiences. Consider two examples. First, suppose we (briefly) show an ordinary agent a black surface with 1000 white dots. Thus, their raw perceptual experience is incompatible with a situation in which the agent is looking at 1001 white dots. But presumably (unless they are a *savant*) this is not the empirical information that they *receive* from their experience. Rather, their empirical information is compatible with, for instance, the existence of a black surface covered with 1001 white dots, or 999 (though not, say, 30, or 100,000). Second example. Suppose one checks one's very accurate watch at exactly 8pm. Two hours later, one wonders again what the exact time is. Now, presumably, an agent with perfect recall is in a position to know that it is exactly 10pm, since they could (in principle) extract information about the passage of time from their perfectly recorded history of perceptual experiences between 8pm and 10pm. However, an ordinary agent lacks perfect recall, so must check her watch again.

[28] See Floridi [2010]. Cf. [van Benthem, 2011, Ch. 1].

[29] Cf. the treatment of basic perceptual evidence in [Goldman, 1976, section II], Lewis [1996] and throughout the entries in Dodd and Zardini [2014].

1.2 Introducing resolution theory

I now introduce a particular instance of the RA approach: resolution theory. Resolution theory has two components, an epistemic component and a semantic component. I call the first the *Dretskean theory of knowability*. I call the second *resolution semantics for knowability ascriptions*. Both are RA accounts.

After a preliminary, I summarize the position. Then, I elaborate on the epistemic component and semantic component in turn.

1.2.1 Preliminary: propositions

Throughout the dissertation, I use the term *proposition* synonymously with *unstructured proposition*. An unstructured proposition is a set of possible worlds. I want to say little about the scope of the class of all possible worlds (i.e. *logical space*), or what a possible world actually is (it makes little difference to our presentation if we use 'centered possible world' instead). I treat unstructured propositions as our default notion of *content* i.e. the objects of the propositional attitudes, and the 'truth set' of an interpreted declarative sentence. However, I do not mean to say that an unstructured proposition can fill every theoretical role we expect of a proposition, in the fullest sense of the term (indeed, I explicitly deny that a truth set exhausts the meaning of an interpreted sentence, as I embrace the notion of a Fregean guise and the notion of subject matter). Nevertheless, unstructured propositions serve as a simple and flexible model of mental content, and semantic information more generally (cf. Floridi [2010]).

I use $|\varphi|$ to indicate the *truth set* of φ i.e. the set of worlds where this sentence is true (relative to the context of utterance, though I often suppress this detail, for convenience). I say that this is the proposition that φ *expresses*. Let @ be the actual world. I say that proposition P is *true* just in case $@ \in P$. I call true propositions *facts*.

1.2.2 Resolution theory, in overview

The Dretskean theory of knowability can be glossed in an intuitive way.

> **Dretskean knowability:** to be in a position to know fact P given information E (notation: $\mathbb{K}(P, E)$) is for E to *discriminate* the truth

of P from its falsity.

Resolution semantics, in contrast, has a technical air.

Resolution semantics: relative to a model that supplies E, $\ulcorner K\varphi \urcorner$ is true just in case E discriminates actuality from a specific set of *defeaters* generated by the *guise* of $\ulcorner\varphi\urcorner$ at the *resolution* of $\ulcorner\varphi\urcorner$'s subject matter.

Since every way in which $\ulcorner\varphi\urcorner$ is false is a defeater for $\ulcorner K\varphi\urcorner$, the truth of $\ulcorner K\varphi\urcorner$ - relative to E - entails $\mathbb{K}(|\varphi|, E)$. That is, $\ulcorner K\varphi\urcorner$ expresses that the agent is in a position to know $|\varphi|$ *in a particular way.* Thus, according to resolution theory, the epistemic status associated with knowledge is simple and basic; our language for expressing this status is nuanced and flexible.

Early work in the RA tradition introduced two promising ideas: that knowledge is closely aligned with powers of discrimination, and that exactly *what* knowledge one expresses with a knowledge ascription is a surprisingly subtle product of semantics and pragmatics. Recent work in the tradition uses technical tools from the semantics/pragmatics literature and formal logic to startling effect. For instance, Schaffer [2005a] draws on recent work on questions to formulate and support his theory;[30] Holliday [2015b] draws on the tools of formal logic to critique and develop various RA theories with precision. Resolution theory, for its part, aims to resuscitate a straightforward account of the link between knowledge and discrimination; integrate this with a subtle account of knowledge ascriptions that draws on recent advances in the literature on subject matter (e.g. Lewis [1988a], Lewis [1988b], Yablo [2014]) and two-dimensionalism (e.g. Chalmers [2011]); and present the whole package in a precise logical framework.

In the next two sections, I expand on the epistemic component and semantic component of resolution theory more carefully, and show how to take each as an RA theory.

1.2.3 The epistemic picture

The resolution theory of knowability takes inspiration from three insights in Goldman [1976]:[31] first, that perceptual knowledge is a paradigm instance of knowl-

[30] For a sense of the literature on questions, see Hamblin [1958], Hamblin [1973], Belnap and Steel [1976], [van Benthem, 2011, Ch. 6] and Ciardelli et al. [2015].

[31] See Rysiew [2006] for further discussion of Goldman's motivations for an RA approach.

edge. Second, that perceptual knowledge of P is a matter of being able to discriminate actuality from counter-factual possibilities where P is false. Third, that such discrimination does not require perceptual evidence that eliminates *every* counter-factual possibility in which P is false. In normal circumstances, I perceptually know (I can *see*) that there is a dachshund before me when my perceptual experience eliminates the normal possibility that, for instance, there is a labrador before me. I do not require empirical information that eliminates the possibility that there is an extraterrestrial robot before me, cunningly disguised as a dachshund. Here is Goldman [1976], page 772:

> My emphasis on discrimination accords with a sense of the verb 'know' that has been neglected by philosophers. The O.E.D. lists one (early) sense of 'know' as "to distinguish (one thing) from (another)," as in "I know a hawk from a handsaw" (Hamlet) and "We'll teach him to know Turtles from Jayes" (Merry Wives of Windsor). Although it no longer has great currency, this sense still survives in such expressions as "I don't know him from Adam," "He doesn't know right from left," and other phrases that readily come to mind. I suspect that this construction is historically important and can be used to shed light on constructions in which 'know' takes propositional objects. I suggest that a person is said to know that p just in case he distinguishes or discriminates the truth of p from relevant alternatives.

We also take inspiration from a proposal in Dretske [1971]: empirical information E is a conclusive (knowledge-worthy) reason in support of P exactly when: E would not be the case unless P were the case. We re-phrase this subjunctive statement as follows: if P were not the case then E would not be the case. Interpreting this counter-factual conditional in the standard way, we get: at the nearest (most similar) worlds to actuality where P is the not the case, E is not the case.

We combine the proposals from Goldman [1976] and Dretske [1971] as follows: first, we propose that E discriminates P from not-P exactly when: if P were not the case, then E would not be the case i.e. E is a conclusive reason in support of P. Second, we propose: to be in a position to know P is for one's information E to discriminate between P and not-P.[32]

[32]For an alternative approach to our own, see Pritchard [2010] for an RA account that argues

This is an RA theory of knowability (the *Dretskean* theory), relative to an appropriate account of the key RA terms.[33]

- *Alternatives:* an alternative to proposition P is a possible world w such that $w \notin P$.
- *Ruling out:* an alternative is ruled out by information E just in case $w \notin E$.
- *Relevance:* an alternative w is relevant to P at actual world @ just in case w is one of the nearest worlds to @ such that $w \notin P$.

The Dretskean theory has attractive explanatory power. I mention its treatment of a few standard examples (for further discussion, see [Dretske, 1971, section 1.]).

Reliable clock: suppose Jane looks at an extremely reliable clock that (accurately) reads four o'clock. Based on this observation, she is in a position to know that it is four o'clock. For if it were not four o'clock, the clock would read differently. That is: at all of the nearest worlds in which it is not four o'clock, her experience of the clock is different (since it is reliable).

Gettier clock: suppose Jane looks at a broken clock that reads four o'clock. As it happens, it is in fact four o'clock. But Jane is not in a position to know that it is four o'clock: for there is a nearby possible world in which the empirical information at her disposal is the same, but it is not four o'clock.

Lottery: suppose Jane enters a lottery. The probability that any particular ticket will win is miniscule. Still, before the draw (Jane loses), Jane is not in a position to know that she is a loser, since there are nearby worlds in which her empirical information is the same, but she wins. That is: the nearest worlds in which she wins include some in which her observations before winning are identical.

Newspaper: on the other hand, the next day, Jane reads in the New York Times (a very reliable paper) that her number didn't come up, and thereby knows that she lost the lottery. For if she had won, then she would not have received

that the relationship between knowledge and discrimination is *not* straightforward.

[33]Given our inspirations, it might be more appropriate to call this the Goldman-Dretske theory. However, since this is a clumsy moniker, and our theory chiefly follows the *form* of Dretske's proposal, we use 'Dretskean theory'.

the same empirical information: she would have read something different in the paper.

Arithmetic: it is knowable a priori that 2+2 = 4. Hence, this claim is knowable no matter what empirical evidence is available. The current theory of knowability delivers this. For, let E be an arbitrary piece of empirical information. Now, it is vacuously true that E is false at the nearest worlds at which $2 + 2 \neq 4$ - since there are no possible worlds at which $2+2 \neq 4$. Hence, E discriminates $2+2 = 4$ from $2 + 2 \neq 4$, as a degenerate case.

Birdwatching I:[34] an amateur birdwatcher spots a Gadwall duck in the distance, in a situation where there are no birds in the vicinity that have an appearance that is similar to a Gadwall. In this case, her sighting puts her in a position to know that there is a Gadwall duck before her: if there were not, then she would not have had the experience of seeing something of that appearance.

Birdwatching II: on the other hand, suppose that our amateur birdwatcher is in an area (perhaps to her surprise) in which Siberian grebes (which have a very similar appearance to Gadwalls) are plentiful. In this case, spotting a Gadwall in the distance does *not* put her in a position to know that there is a Gadwall duck before her: there is a nearby world where what she is looking at is a grebe, yet her empirical information is the same.

Note that variations on the last two cases make trouble for the Dretskean theory, as it stands. We examine some prime examples in section 1.3: Schaffer's missed clue cases. (As we shall see, the view of resolution theory is that these apparent troubles are an illusion resulting from a *naive* account of the relationship between the Drestkean theory and knowledge ascriptions.)

1.2.4 The semantic picture

Resolution semantics takes its inspiration from two insights in Dretske [1970], Dretske [1972] and Dretske [1981]. The first (ever popular) insight is that *discourse facts* determine a restricted set of contrasts that need to be ruled out in order to know φ. Dretske [1972] provides intriguing linguistic evidence in support of this fact. For instance, there seems to be a difference in exclaiming "a knows that Clyde *sold* his typewriter to Alex" and "a knows that Clyde sold his typewriter to *Alex*", determined by the elements of the claim that are focused. The

[34] Cf. the famous barn cases in Goldman [1976].

first suggests that a can rule out alternatives such as: Clyde gave his type-writer to Alex; Clyde forced his typewriter upon Alex etc. The second suggests that a can rule out: Clyde sold his type-writer to Jane; Clyde sold his type-writer to Carlos etc.

To say that the contrasts are determined by discourse facts is not yet to attribute this phenomenon to either *semantics* or *pragmatics* (or an interaction between the two). The former approach has been pursued by theorists that claim that "knows" is a context-sensitive expression, comparable to, for instance, indexicals or gradable adjectives.[35] The latter approach has been pursued by theorists that claim that, for instance, the contrasts are determined by the pragmatic presuppositions of the discourse [Blome-Tillmann, 2014].

The second insight we lift from Dretske [1970] is that the contrasts to $\ulcorner\varphi\urcorner$ are (at least partly) a matter of semantics, but an aspect of the meaning of the *operand* $\ulcorner\varphi\urcorner$. The *operator* K, meanwhile, has an invariant meaning. Compare:

> So it is with our epistemic operators. To know that x is A is to know that x is A within a framework of relevant alternatives B, C, and D. This set of contrasts, together with the fact that x is A, serves to define *what it is that is known* [my emphasis] when one knows that x is A. One cannot change this set of contrasts without changing what a person is said to know when he is said to know that x is A [Dretske, 1970, pg. 45].

Which semantic mechanisms generate these contrasts? Ideally, we would account for the contrasts via semantic notions that are well-supported by the linguistic data and have an established presence in the theoretical literature. For resolution semantics, we appeal to two such notions. First, that of *subject matter*. Second, that of *Fregean guise*. These are natural choices: not only is each the object of an extensive and fruitful literature, but this literature already draws out a connection between these notions and knowledge ascriptions.[36]

[35]For discussion and critique, see Stanley [2005].

[36]For an extensive overview and critique of the two-dimensionalist approach to the notion of 'guise', see Soames [2007a]. Our own development is most directly influenced by the account in Chalmers [2011]. For further discussion about the neo-Fregean programme, see Perry [2000] and Stanley [2011]. Our treatment of topics/subject matter owes a debt to the tradition started by Lewis [1988a] and Lewis [1988b]. These works draw a strong connection between *questions* and topics. With this in mind, resolution semantics for knowledge ascriptions falls in a tradition that

With this in mind, we present resolution semantics as an RA theory, then elaborate on its components below. Recall the RA truth clause for $\ulcorner K\varphi \urcorner$: relative to a model $\mathcal{M}$ that supplies E, $\ulcorner K\varphi \urcorner$ is true just in case E discriminates actuality from a specific set of defeaters generated by the guise of $\ulcorner \varphi \urcorner$ at the resolution of $\ulcorner \varphi \urcorner$'s subject matter. Thus:[37]

- *Relevance:* a proposition is relevant to $\ulcorner \varphi \urcorner$ just in case it is at the resolution of T_φ, the subject matter of $\ulcorner \varphi \urcorner$.

- *Alternatives:* A proposition A is an alternative to $\ulcorner \varphi \urcorner$ just in case it is a *defeater* for $\ulcorner \varphi \urcorner$: a relevant proposition that is inconsistent with the *epistemic guise* of $\ulcorner \varphi \urcorner$.

- *Ruling out:* A proposition P is ruled out just in case E discriminates not-P from P i.e. $\mathbb{K}(\text{not-}P, E)$ holds.

Relevance is determined by *subject matter.* More precisely, I argue in chapter 2 that a subject matter T is a set of issues/distinctions. A proposition is *at the resolution* of T just in case, intuitively, it speaks to the issues in T. Again, this key notion is considered more precisely in chapter 2. Every interpreted sentence $\ulcorner \varphi \urcorner$ is associated with a subject matter T_φ.

Having epistemic relevance depend on subject matter allows for it to be determined by an elegant mix of context and literal meaning. Intuitively, rational discourse is guided by a *topic* (i.e. a subject matter, a set of salient issues, a question under discussion), against which the subject matter of an *individual claim* can be compared. For instance, a claim φ is *on-topic* (i.e. *conversationally* relevant) exactly when its subject matter is included in that of the discourse topic (or at least *overlaps* with the discourse topic).[38] Thus, if subject matter determines epistemic relevance, *two* sources of subject matter can be combined to maximize explanatory power: subject matter at the level of discourse, and

includes Schaffer [2004], Schaffer [2005a], Schaffer [2005b], Schaffer [2007a], Schaffer [2007b], Yablo [2012] and Yablo [2014].

[37]Resolution semantics is most indebted to the account in Yablo [2012], which similarly understands relevance in terms of subject matter, and ruling out in terms of satisfying a certain counter-factual conditional. However, since both the account of subject matter and the conditional in question differ from that in resolution theory - not to mention the account of 'alternative' at play - the theories are clearly distinguished in the details.

[38]Cf. Roberts [2012].

subject matter at the level of a claim. The former can, for instance, account for alleged linguistic data that *epistemic standards* are discourse-sensitive - perhaps the reason that knowledge is more casually (though correctly) ascribed in the bar than in the courtroom is because court proceedings insist on finer distinctions and tougher issues. Thus, our approach keeps the door open for a form of contextualism. However, in what follows, it will be fruitful to focus only on claim-level subject matter. Thus we focus on *semantics* and delay discussion of the role of *pragmatics* to another time.

Intuitively, a defeater is a reason for withholding belief in φ. To understand our more precise account (which I defend in detail in chapter 4), we need to explain what an 'epistemic guise' is. Start with the notion of a *Fregean guise*. For us, the Fregean guise of $\ulcorner\varphi\urcorner$ is an aspect of its meaning, understood intuitively as a *way of thinking* about the content of $\ulcorner\varphi\urcorner$. To use a classic example, Fregeans agree that "Hesperus is Phosphorus" expresses the content that Venus is identical to itself, but that this content is presented (in certain discourse settings, at any rate) under the guise "the evening star is the morning star", reflecting that the name 'Hesperus' is associated with the role of 'evening star' and the name 'Phosphorus' is associated with the role of 'morning star'. Among other theoretical advantages, this framework explains why "Hesperus is Phosphorus" and "Venus is Venus" have apparently different cognitive significance, despite expressing the same content.

Now, neo-Fregeans in the two-dimensionalist tradition claim that the Fregean guise of $\ulcorner\varphi\urcorner$ determines the set of possibilities that must be eliminated in order to come to know φ.[39] The standard proposal is that these possibilities are the worlds at which the Fregean guise is false. For instance, to know "water is H_2O" one must eliminate the worlds in which the ubiquitous potable liquid (i.e. that substance that serves the role that water serves in the actual world) is not H_2O.

In chapter 4, I argue for a neo-Fregean framework that posits a more complicated relationship between the Fregean guise of $\ulcorner\varphi\urcorner$ and the possibilities that need to be eliminated for knowledge to be achieved. The *epistemic guise* of $\ulcorner\varphi\urcorner$ is, roughly, the conjunction of its Fregean guise with 'acquaintance conditions' that posit that the objects referred to in $\ulcorner\varphi\urcorner$ actually serve the corresponding roles in the Fregean guise. This conjunction expresses a proposition that entails $|\varphi|$. Hence, coming to know the epistemic guise of $\ulcorner\varphi\urcorner$ is a *way* of coming to

[39] Cf. Soames [2007a], Chalmers [2011].

know φ. Resolution semantics claims that the possibilities that need to be ruled out, in order for $\ulcorner K\varphi \urcorner$ to hold, are those that contradict the epistemic guise of $\ulcorner \varphi \urcorner$. (Since it is also posited that the components of the epistemic guise express distinctions that are in T_φ, these possibilities are at the right resolution to be relevant.)

Finally, note that resolution semantics understands 'ruling out A' as: the knowability relation holds between not-A and the given empirical information E i.e. $\mathbb{K}(\text{not-}A, E)$. This connects resolution semantics to the Dretskean theory of knowability. Effectively, a true knowability ascription $\ulcorner K\varphi \urcorner$ expresses that a *set* of propositions are knowable given the available information. Of course, one of the propositions in that set must be $|\varphi|$.

1.3 Problems for the RA approach

We now turn to some important objections to the RA approach, before sketching resolution theory's response in section 1.4. The rest of the dissertation aims to fill out this sketch. First, we note the core challenge that the RA approach faces.

1.3.1 The problem of epistemic relevance

The core challenge for an RA theorist is simple: is there an account of relevance that grounds a satisfactory RA theory?

Of course, as mentioned, an RA theorist also owes us an account of *ruling out* and *alternative*, and the success of the theory hinges on a judicious selection of all three in tandem. However, plausibly, *any* theory of knowability must furnish us with an account of the work of empirical information and the nature of the possibilities that must be eliminated for a claim to be known. Identifying a satisfactory account of relevance is the RA theorist's *unique* burden.

What counts as a satisfactory RA theory? Without pretending to offer a comprehensive list of criteria for adequacy, a good place to start, it seems to me, is to offer an account that defuses the following three generic difficulties.

1.3.2 Three particular problems

RA theories tend to fall victim to three kinds of objection. Before I elaborate in detail, here they are in brief.

- **Closure dilemma:** Epistemic closure is the principle that if one has sufficient information to know φ and φ entails ψ, then one has sufficient information to know ψ. On the other hand, a contingent fact is easy knowledge just in case it is knowable a priori. RA theories fall into two camps: those that reject epistemic closure and those that embrace the possibility of rampant easy knowledge. On the face of it, both features are objectionable.

- **Missed clue cases:** RA theories tend to be subject to compelling counter-examples. *Missed clue cases* come close, I propose, to presenting a *universal* counter-example.

- **Ad hocness:** An RA theory has a number of parameters that can be tweaked. For one, a theorist can propose a *complex* account of relevance, using a diverse battery of sufficient conditions for relevance. Thus RA theorists have a vast scope for refinement in the face of counter-examples and other objections. However, this flexibility has a downside: RA theorists must be on guard to avoid *ad hoc* accounts of relevance.

1.3.3 The closure dilemma

There is an undeniably strong intuition that deductive reasoning from known premises is a sure way to arrive at known conclusions (often thereby *extending* one's knowledge). An *epistemic closure principle* attempts to capture this as a principle in epistemic logic. For us, closure may be phrased at the level of the nature of knowledge, or at the level of knowledge ascriptions. (As usual, read $\mathbb{K}$ as denoting the knowability relation, and $\ulcorner K \urcorner$ as the "it is knowable that" operator in a toy logical language.)

> **Closure w.r.t. the knowability relation:** given information E, if $\mathbb{K}(P, E)$ and P entails Q (i.e. $P \subseteq Q$), then $\mathbb{K}(Q, E)$.
>
> **Closure w.r.t. the knowability operator:** the following formula is valid in the correct logic of knowability claims: $K\varphi \wedge (\varphi \Rightarrow \psi) \rightarrow K\psi$,

where $\varphi \Rightarrow \psi$ is defined as $\Box(\varphi \to \psi)$ with $\Box$ denoting the universal necessity operator and $\to$ the material conditional.

Many philosophers view closure denial as an unequivocal strike against a theory - even a *devastating* one (cf. [Stine, 1976], [Lewis, 1996], [Feldman, 1999, pg. 95], [Kripke, 2011, pg. 163], [Williamson, 2000, pg. 118]). In support of this contention, Hawthorne [2004] observes that closure follows from a pair of particularly simple and intuitive principles: that knowability of a conjunction puts one in a position to know the conjuncts and that replacement of logical equivalents does not alter the knowability of a claim. Further, DeRose [1995] observes that the linguistic data seems to favor closure: how else to explain the sense of contradiction evoked by 'abominable conjunctions' such as "I know that John is a bachelor but do not know that John is unmarried (though it is obvious that being a bachelor entails being unmarried)"? Note that this this claim has the form $(K\phi \wedge \neg K\varphi) \wedge (\phi \Rightarrow \varphi)$, which is *not* a classical contradiction, since it has the propositional form $(\chi \wedge \neg\psi) \wedge (\phi \Rightarrow \varphi)$.[40]

On the other hand, consider the idea of *easy knowledge*. In terms of the knowability relation, we say that knowledge of content C is *easy* just in case C is contingent and C is knowable without basis in *any* empirical information i.e. $\mathbb{K}(C, \emptyset)$ must hold. That is, C is easy knowledge just in case C is a contingent a priori truth. Once again, all of this can be rephrased in terms of knowability ascriptions.

Many philosophers find the idea of easy knowledge objectionable. That is, if a theory allows for easy knowledge, then this is treated as a significant cost (cf. [Cohen, 2002], [Black, 2008]). Various theorists have specifically lodged this complaint against proposals in the RA tradition (cf. [Vogel, 1999, pp. 171-172], [Heller, 1999], [Holliday, 2015b, pp. 113-117]).

Thus, we have two features that a theory of knowability can exhibit (*closure denial* and *easy knowledge acceptance*) that are, at best, intuitively odd. A common view is that they are both highly objectionable.

In fact, this amounts to a dilemma for the RA theorist (cf. [Holliday, 2015b, sects. 2.4-2.5]). For it may be argued that every RA theory *either* rejects closure

[40] On the other hand, a strong contingent of philosophers have defended closure denial, by rejecting closure in full generality in favor of a *restricted* principle. See, for instance, Dretske [1970], Nozick [1981], Schaffer [2007a], Lawlor [2013], Holliday [2015b] and Yablo [2014]. We'll take up these issues in chapter 3.

in full generality *or* embraces the possibility of easy knowledge. At least, this is so relative to minimal assumptions that the RA theorist should be loathe to give up.

To see this,[41] we introduce some notation (we discuss closure at the level of the knowability operator, for convenience): for any sentence φ, let $\mathbf{R}_\varphi$ stand for the set of relevant alternatives relative to φ (and model $\mathcal{M}$). Now, for any further sentence ψ, $K\varphi \to K\psi$ is valid just in case $\mathbf{R}_\psi \subseteq \mathbf{R}_\varphi$. That is: by RA lights, $K\varphi$ guarantees $K\psi$, no matter the available evidence E, just in case ruling out every relevant alternative to φ ensures that every relevant alternative ψ is ruled out. Now, we argue that *if* our RA theory validates $K\varphi \wedge (\varphi \Rightarrow \psi) \to K\psi$, *then* it must endorse the existence of easy knowledge. Assume that our theory validates $K\varphi \wedge (\varphi \Rightarrow \psi) \to K\psi$. Thus, if $\varphi \Rightarrow \psi$ holds, it must be that $\mathbf{R}_\psi \subseteq \mathbf{R}_\varphi$. In particular, let h represent the mundane proposition "I have hands" and b represent the radical proposition "I am a brain-in-vat". We suppose, as is virtually universally accepted among RA theorists, that Kh holds and that $h \Rightarrow \neg b$ holds. It follows, from the assumed acceptance of closure, that both $K(\neg b)$ holds and that $\mathbf{R}_{\neg b} \subseteq \mathbf{R}_h$. But it is also universally accepted among RA theorists that no way of being a brain-in-vat is a relevant alternative to knowing h (for it is further assumed that such alternatives cannot be ruled out by ordinary information. An RA theorist who denies this jettisons a central motivation for the RA approach). It follows that $\mathbf{R}_h$ contains no proper alternatives to $\neg b$ (i.e. no ways of being a brain-in-vat), and so $\mathbf{R}_{\neg b}$ must be empty. Thus, $K(\neg b)$ holds, but vacuously, for it would hold even in the degenerate case where the agent has *no* empirical information (for it is always *vacuously* true that every relevant alternative to $\neg b$ is ruled out). Hence, the theory endorses easy knowledge.

The case for rampant easy knowledge

Can either horn in the dilemma be sensibly pursued? In fact, the recent literature provides intriguing support for the existence of easy knowledge. If these arguments are successful, this is good news for the RA approach in general. It puts pressure however, on RA theories that abandon closure, which already shoulder the burden of defying common sense. Or, at least, this is how their opponents

[41] Spelling out this reasoning with complete precision is surprisingly labor-intensive. See [Holliday, 2015b, sects. 2.4-2.5]. Our own presentation sacrifices some precision for brevity.

would describe the situation.[42]

(As we shall see, resolution theory falls in the closure denial camp. Thus, we must return to the question of whether it grasps the less promising horn.)

One might embrace the possibility of easy knowledge as a consequence of accepting a Wittgensteinian *hinge epistemology.* According to such a view, it is constitutive of rational inquiry that it operate on a basis of *hinge commitments:* propositions which an agent is rational to accept though no empirical information weighs in their favor (an oft-cited candidate: the proposition that one's sensory faculties are not systematically deceptive). Most of the discussion of hinge epistemology focuses on issues of warrant or justification, not knowledge *per se*.[43] But if a case can be made out for the existence of hinge propositions, then the theory of knowledge is ripe for an application of the idea. In particular, one might claim that knowledge hinges on contingent a priori truths.

However, it is far from clear that the notion of a hinge proposition can be satisfactorily fleshed out.[44] I am pessimistic about its prospects, so leave our discussion of hinge epistemology here.

Instead, I concentrate on a recent revival in the fortunes of the *Kripke-Kaplan approach*, according to which the existence of contingent a priori truths reflects subtle interactions between meaning and knowability.[45] This revival argues that contingent a priori truth is defensible, harmless and rampant. *Defensible:* it is claimed that standard objections to classic Kripkean examples of the contingent a priori do not survive scrutiny, and miss the target completely in the case of non-Kripkean examples.[46] *Harmless:* it is claimed that a priori knowledge of contingent truth is cheap but *vacuous*, in the sense that it lacks cognitive significance.[47] *Rampant:* it is claimed that *every* contingent proposition is knowable a priori.

In particular, this last point can be cashed out in *two-dimensionalist neo-*

[42]See Stine [1976] for a groundbreaking and influential defense of the view that closure should be preserved by allowing for easy knowledge (within a context).

[43]See Wright [2003], Wright [2004] and Wright [2014] for the most influential development of these ideas, both in early and refined forms.

[44]See the papers in section II of Dodd and Zardini [2014] for some recent critical discussion. Also see the entries in section IV of Coliva [2012].

[45]For the classic defenses of the contingent a priori in this tradition, see Kripke [1980] and Kaplan [1989].

[46]See [Hawthorne and Manley, 2012, Part I].

[47]See [Hawthorne and Manley, 2012, Ch. 2], especially section 2.6.

Fregean terms: the meaning of sentence φ is posited to be a pair of propositions $\langle P, Q\rangle$, where P is the content, while Q is a way to know P. For φ to be contingent a priori is, on this view, for its content to be associated with a way of knowing it that is cognitively insignificant (the standard representation: $Q = W$, where W is the set of *all* possible worlds). By these lights, the claim that every proposition P is knowable a priori is just to say that there exists a guise Q for each such P that is cognitively insignificant. Put another way: the claim is that for every contingent content P there is a sentence in natural language φ (or, at least, a language could be designed with such a sentence) that has the meaning $\langle P, W\rangle$.

The claim that contingent a priori truth is rampant is supported by at least two interesting arguments. We take the first from Soames [2007b]. Let P be an arbitrary contingent fact, and stipulate that atomic claim p expresses P. Let φ be the sentence "p if and only if it is true at @ that p", where here we treat @ as an indexical that denotes the world at which the utterance is made. First premise: φ is knowable a priori. Second premise: φ has the same truth set as p, and so φ has the content P. Conclusion: fact P is knowable a priori.

The rationale behind the first premise is that one can discern the truth of φ merely by grasping the meaning of its components. For, in every context of utterance, p is true exactly when "it is true at @ that p" is true, by virtue of the indexicality of @. The justification of the second premise is as follows: since P is a fact, "it is true at @ that p" plausibly expresses a necessity (namely, that a certain fact holds at a certain possible world). Hence, in a context where p is true, the truth set (the content) of φ matches that of p.

Here is the second argument, adapted from appendix 1 of Gibbard [2012] (I call it the *cheap trick*, following Gibbard):

> *Gibbard's cheap trick:* First premise: it is knowable a priori that *this* is the way things are (gesturing at the actual way things are). Second premise: since every fact is a logical consequence of the fact that *this* is the way things are, if the latter is knowable a priori then so is every fact. Conclusion: every fact is knowable a priori.

Obviously, the two arguments have distinguishing features. There is also significant overlap in the issues they raise. In chapter 4, I concentrate on evaluating the second, treating it as emblematic of the revived case for the contingent a priori and the rationale for rampant easy knowledge.

For now, we conclude that the closure dilemma presents an acute problem for closure-denying RA theories (like resolution theory). For one, it may seem preferable to offer a theory of knowability that escapes the dilemma completely. For another, if we insist on an RA approach, it may seem that we should steer clear of closure denial.

(Ultimately I will draw the conclusion that the above deflationary rationale for accepting rampant contingent a priori truth is intriguing, but even if successful does not eliminate the need for an account of *substantive knowability* that rejects easy knowledge.)

1.3.4 Missed clue cases

A missed clue case has the following form: there is an agent a with empirical information E. As part of E, the agent knows that P is the case. P, as it happens, is a *clue* that Q is the case - meaning that P *objectively indicates* Q, in some strong sense. But a's information does *not* put them in a position to appreciate that P is a clue for Q. On the assumption that a does not have any other empirical evidence that bears on the truth of Q, it follows that a is not positioned to know Q. The clue is *registered* but not *appreciated*.

The difficulty this presents for the RA approach is that prominent RA theories render the wrong verdict in (at least some) such cases: they rule that a is in a position to know Q relative to E.

Missed clue cases are flagged by Schaffer [2002]. His target seems to be RA theories in general. However, his discussion focuses mainly on the RA theory of Lewis [1996]. This raises the question of the scope of missed clue counter-examples.

In what follows, I consider a concrete missed clue case and reject an influential assessment of its significance due to Brueckner [2003]; assess the impact of the case on a prime example of RA theory: similarity-based accounts; then I present an argument that this impact can be extended to RA theories in general; finally, I consider a more abstract lesson from missed clue cases.

A concrete case

Here is a concrete missed clue case (adapted from Schaffer [2002].)

> *Jane's missed clue.* Suppose that Jane is browsing through the Italian edition of The Bird Almanac. Jane does not read Italian, but is enjoying the book's photos. On page 300, she comes across a photo of a bird with red plumage (the clue) that otherwise seems (to Jane) similar in appearance to a wild canary. Jane is somewhat familiar with canaries, but knows no general facts about the plumage of wild canaries. In particular, she does not know that all wild canaries have *yellowish-green* plumage. All she knows is that the plumage of *domestic* canaries is diverse - and sometimes red. Thus, Jane is in no position to know that the depiction on page 300 is not of a wild canary.

It is helpful to contrast Jane with Professor Byrd, world-renowned ornithologist. If Professor Byrd sees the photo on page 300, she is certainly in a position to know that the depicted bird is not a wild canary. For she has acquired enough empirical information to appreciate the provided clue. This is the fruit of years of study! We might add: Professor Byrd can *see* that the depicted bird is not a wild canary. Whatever Jane can see (e.g. she can see the color of the plumage), she cannot *see* that the depicted bird is not a wild canary.

A precise reading of the notion of *clue* will be useful. I propose: fact P is a clue (in the objective sense) for fact Q just in case: at every possible world w that is nearby (i.e. similar enough) to the actual world @, if P is true at w then so is Q. It is not *necessary* that no wild canary has red plumage. It is a contingent but modally robust *regularity*.

Schaffer [2002] argues that we here have a counter-example to the RA approach, using roughly the following reasoning: an RA theorist should agree that the possibility of a highly 'abnormal' fact holding - e.g. that there are mutant wild canaries with red plumage - is epistemically irrelevant. But it is clear that Jane's information is incompatible with every 'normal' alternative to there not being a wild canary on page 300. For, in such an alternative, the bird depicted on page 300 is a wild canary with yellowish-green plumage. But Jane can *see* that the depicted bird in the actual world has red plumage. Her empirical information rules out that there is a yellow-green plumed wild canary before her. In total, every *relevant* alternative to there not being a wild canary on page 300 is ruled out by Jane's information. Thus, according to the RA approach, she is in a position to know that the bird is not a wild canary.

One objection to this argument is that the notions of 'abnormal' and 'normal' seem closely related to a particular notion of relevance: that based on *similarity* to the actual world. This calls into question the universality of the counter-example. We'll address this issue in the coming sections.

For now, I rebut an influential reply to Schaffer [2002] offered by Brueckner [2003]. Brueckner [2003] suggests that missed clue cases pose no unique problem for the RA approach. What the cases show, in his estimation, is that an RA theorist must include an appropriate belief clause in her full account of knowledge. For though an RA theorist is *right* to conclude that Jane is positioned to know that the bird is not a wild canary (she has registered a clue after all), Jane is not rational to *base a belief* that the bird is not a wild canary on that clue.

Once the terminological issues are clarified, I see no cause to disagree with Brueckner's description. For Brueckner [2003] uses 'in a position to know' with a different meaning to us, something like: an agent that receives the clue is positioned to properly base a belief on the evidence, and so acquire knowledge, *if she has the right background knowledge.* On this usage, something might be 'knowable' for Jane though it is not possible for her to know it without acquiring more information: seeing the clue 'positions Jane to know' that the bird is not a wild canary, though she cannot 'capitalize' on the clue without further information. Thus, Brueckner [2003] should agree with us, it seems, that Jane requires further empirical information if she is to actually know that the bird is not a wild canary. What he adds (and we need not quibble) is that if Jane were to acquire sufficient further information so as to properly *base a belief that* φ on the clue, then she would *know that* φ.

Very well. But by these lights, the appeal to belief and belief-basing seem to me a red herring in the context of our discussion. Why is Jane, in the missed clue example, denied both a rational belief and knowledge that the bird is not a wild canary, on the basis of her observations? All hands agree: she *lacks sufficient empirical information.* If, like Professor Byrd, she had *enough* information (had observed enough birds, spoken to reliable experts etc.), her observations would narrow down the world to one where red plumage indicates not being a wild canary. Thus, Brueckner [2003] should agree that it is not possible for Jane to know, relative to her actual information, what type of bird she sees. But we are here interested in RA theory as an account of knowability in the following sense:

for empirical information E, it is possible for an agent to know φ when their total empirical information is exactly E. Understood in this way, all hands must agree that RA theory (apparently) delivers the wrong result with respect to Jane's predicament. It is irrelevant that if Jane were to *supplement* her total empirical information with further relevant information, then she could know that the bird is not a wild canary.

Missed clues and similarity-based theories

We now show that missed clue counter-examples exist for simple similarity-based RA theories (using a naive semantics for knowledge ascriptions). There are two basic theories along this line. In the first, $\ulcorner K\varphi \urcorner$ is true just in case the empirical information rules out the *nearby* $\neg\varphi$ worlds. In the first, $\ulcorner K\varphi \urcorner$ is true just in case the empirical information rules out the *nearest* $\neg\varphi$ worlds. (As will often be the case in this dissertation, we drop the $\ulcorner\urcorner$ notation in our discussion, using context to distinguish use and mention.)

We work with language $\mathcal{L}$, built up from proposition letters, the usual connectives and the knowability operator K. We work with basic models $\mathcal{M} = \langle W, \{E_w\}_{w \in W}, \mathbf{V}\rangle$, where W is the set of possible worlds, E_w is the information the agent has at world w and $\mathbf{V}$ is a valuation that assigns a proposition to each atomic sentence.

First theory: Enrich our models as follows: for each world w fix a set of worlds N_w that includes w. We call these the *nearby worlds* to w. Now consider a truth condition for $K\varphi$, as follows: $K\varphi$ is true at world w, in enriched model $\mathcal{M}$, just in case: if $u \in N_w$ and $\neg\varphi$ is true at u then $u \notin E_w$.

Now, a missed clue case can be represented by a model $\mathcal{M}$ where: thinking of w as the actual world, we have that p is true at w, but not all worlds in N_w; we have that $p \to q$ is true at every world in N_w, including w; and we have that q is true at w, but not all worlds in N_w. It follows that if q is not true at a world in N_w, then neither is p i.e. every nearby $\neg q$ world is a $\neg p$ world. Further, we stipulate that if p is false at $u \in N_w$, then $u \notin E_w$ i.e. the information at w rules out all of the nearby $\neg p$ worlds. Finally, we stipulate that if $u \in E_w$ then $p \to q$ holds at u i.e. the information does nothing to rule out $\neg(p \to q)$ alternatives.

It follows that Kp is true at w (all of the nearby $\neg p$ worlds are ruled out). It also follows that Kq is true at w (since every nearby $\neg q$ worlds is also a $\neg p$ world, and so is ruled out). Finally, it follows that $K(p \to q)$ is true, but *vacuously* so, since there are *no* nearby $\neg(p \to q)$ worlds to rule out.

Hence, we have constructed the *form* of a missed clue case. Think of p as "bird X has red plumage" and q as "X is not a wild canary". Our account allows that it is knowable that X has red plumage; that it is knowable that X is not a wild canary; yet the agent's information does not eliminate a single world in which: X has red plumage but *is* a wild canary. (What is odd about the account under evaluation, of course, is that it *superficially* avoids a missed clue counterexample, for $K(p \to q)$ holds at w on the current view. But we shouldn't be fooled - this is an expression of vacuous knowability, and does nothing to detract from our intuition that Jane must rule out worlds where X has red plumage but is a wild canary before being able to conclude, from the red plumage of X, that X is not a wild canary.)

Second theory: Enrich our models as follows: include a similarity ordering $\preceq$ (i.e. a transitive, reflexive relation). We read $w \preceq u \preceq v$ as: "u is at least as similar to w as v". Read $u \prec v$ as: $u \preceq v$ but not $v \preceq u$. Now consider a truth condition for $K\varphi$, as follows: $K\varphi$ is true at world w, in enriched model $\mathcal{M}$, just in case: for all u, if $\neg\varphi$ is true at u and there is no v where which $\neg\varphi$ holds and $w \prec v \prec u$, then $u \notin E_w$. In other words: E_w rules out the nearest $\neg\varphi$ worlds to w.

A missed clue case can be represented by an enriched model $\mathcal{M}$ where: thinking of w as the actual world, we stipulate that p and q are true at w; that the nearest $\neg p$ worlds to w are not in E_w; that the nearest $\neg q$ worlds to w (all of which are stipulate to be $\neg p$ worlds) are not in E_w; and the nearest $\neg(p \to q)$ worlds (which can be constructed to be relatively *remote* i.e. much less similar to w than the nearest $\neg p$ and $\neg q$ worlds) are in E_w. Thus, p and q are both knowable, yet $p \to q$ is not.

We have the form of a missed clue case, with p as "bird X has red plumage" and q as "X is not a wild canary". Our account allows that it is knowable that X has red plumage, that it is known that X is not a wild canary, yet that it is not knowable that red plumage indicates not being a wild canary.

Thus, simple similarity-based RA theories are victims of the missed clue.

The universality of missed clue counter-examples

Do missed clue cases merely show that an RA theorist should resist a simple similarity-based account? We now consider an argument for the existence of missed clue counter-examples for *every* sensible RA theory.

Recall (section 1.1.4) that various factors have been offered in the literature as necessary and/or sufficient conditions on relevance (besides resemblance to actuality): psychological salience of an alternative to the attributor and/or the subject of a knowledge ascription; presupposition; conversational relevance to the question or topic under discussion; compatibility with the agent's beliefs; a reason (even if far from conclusive) for thinking the alternative is the case; and the practical stakes connected with ignoring an alternative.

Note that all of these factors are *subjective* (relative either to the attributor or the subject of the knowledge ascription). That is, each depends on the attitudes of an agent. If those attitudes differ, then the relevant alternatives also differ.

Compare this to *resemblance to actuality*: this is a markedly *objective* criterion (relative to the circumstances of the subject being evaluated for knowledge). It does *not* depend on the attitudes of any agent.

Suppose that the foregoing criteria for relevance exhaust the serious options. Now we can argue as follows.

First premise: key motivating examples for the RA approach can only be accounted for with a sufficient condition for relevance that is objective. I have in mind *Goldmanian barn cases.* We noted a variation of these cases earlier (section 1.2.3):

> *Birdwatching II:* suppose that our amateur birdwatcher is in an area (perhaps to her surprise) in which Siberian grebes (which have a very similar appearance to Gadwalls) are plentiful. In this case, spotting a Gadwall in the distance does *not* put her in a position to know that there is a Gadwall duck before her: there is a nearby world where what she is looking at is a grebe, yet her empirical information is the same.

It is an *objective* matter that the grebes are in the vicinity. The relevance of the grebe alternative is entirely independent (it seems) of the attitudes of any particular agent (certainly the birdwatcher). Thus, barn cases have the consequence that, whatever complex account of relevance we provide, no subjective criterion for relevance is *necessary*. For no change in attitude (of interlocutor or subject) can render the possibility of confronting a grebe irrelevant.

Second premise: the only objective criterion of relevance that we have cause to take seriously is that based on resemblance to actuality.[48] Not only is this criterion directly suggested by barn cases, but we have independent motivation for making sense of it (namely: to ground the orthodox semantics for counterfactual conditionals). What is more, resemblance to actuality neatly incorporates the only *other* objective criterion that is raised in the literature with frequency: that a world is relevant if it represents a deviation from actuality due to a chance process delivering a different result. (However we make sense of resemblance between worlds, it is generally agreed that, say, the different outcomes of a lottery count as 'nearby' possibilities to actuality.)

Third premise: if a missed clue counter-example can be set up for a simple resemblance-based RA theory, then one can be set up for an RA theory that is enriched with some subjective (sufficient, but not necessary) conditions on relevance: one merely stipulates the attitudes of the agents so that the relevant alternatives from the original counter-example are unaltered. For instance, if we build on an account that takes a world to be relevant if it is nearby to actuality, then we stipulate that the agents presuppose that actuality is among such worlds; that they believe that actuality is among such worlds; that only the possibility that actuality is among such worlds is psychologically salient; that only such worlds involve high stakes for the agent(s); that the agent has no reasons (i.e. no supporting beliefs) for thinking that actuality is not among those worlds. And so forth.[49]

Conclusion: every 'serious' RA theory is a victim of a missed clue counter-example.

[48]There is more to say on this point. For instance, I do not here engage with the highly developed objective account of relevance in Dretske [1999].

[49]Note that when it comes to subjective accounts, it is natural to think of relevance of classifying *propositions*, since these are the objects of the attitudes. For resemblance accounts, however, relevance is taken as a property of *possible worlds*. This disconnect needs to be addressed in a precise account that combines subjective and objective factors.

Missed clues in the abstract

Given this alleged universality, one might wonder if missed clue counter-examples represent a violation of a very general constraint on a theory of knowability. Here is a proposal.

We introduce a new operator: K_s, standing for "it is substantively knowable that ...". $K_s\varphi$ is taken to mean that the content of φ is knowable *non-vacuously* given the available information.

Now consider the following principle:

Conservative → intro: $(K_s\varphi \wedge K_s\psi) \rightarrow K_s(\varphi \rightarrow \psi)$

This says: if φ is substantively knowable and so is ψ, then it is also substantively knowable that φ implies ψ (i.e. that is not that φ holds but not ψ). This formula is intuitively valid. What's more, it may not seem that the RA theorist has an independent motivation for rejecting it.

Now, it is notable that missed clue counter-examples exhibit a violation of this principle. The RA theorist seems committed to saying that Jane is positioned to substantively know that bird X has red plumage and that X is not a wild canary, yet that Jane is *not* positioned to *substantively* deny that both X has red plumage and is a wild canary.[50] A proponent of a 'multi-premise' epistemic closure principle for *substantive knowability* can therefore take missed clue cases as a wide-ranging strike against the RA approach.

(As we shall see, resolution theory rejects the validity of conservative → introduction. I will defend this as reflecting a basic anti-skeptical commitment.)

1.3.5 The threat of ad hocness

Recall, again, the diversity of criteria for relevance that have been proposed in the literature: resemblance to actuality: psychological salience; presupposition; conversational relevance; compatibility with belief; some reason in support of an alternative; and practical stakes.

In principle, nothing prevents an RA theorist from offering a *complex, hybrid* account of relevance. That is: she might offer an array of criteria for relevance.

[50] In the terminology of Lewis [1996], Jane knows that it is not that X both has red plumage and is a wild canary, but this is only known *elusively*.

This affords an RA theorist a great deal of scope for tweaking her theory in response to, say, counter-examples that indicate that a certain criterion is either unnecessary or insufficient for relevance (yet otherwise has convenient explanatory power).

What is more, as increasingly sophisticated extant proposals demonstrate, the *logical* aspects of an RA theory can be refined with great ingenuity.[51] An RA theorist therefore has much formal flexibility for tweaking in response to criticisms of the *logical* aspects of a theory (e.g. that it violates closure in a particularly egregious way).

However, the RA theorist must embrace such maneuvers with care, lest she be accused of proposing *ad hoc* refinements to her theory, existing *only* to bypass specific objections to a more elegant, natural and well-motivated RA proposal.

(Resolution theory must take care on this point. I argue in chapter 3 that it exhibits desirable logical properties, resulting from its underlying account of subject matter. This account of subject matter distinguishes itself in the details from existing accounts. I shoulder the burden of providing independent support for it in chapter 2.)

A case study: Lewisian contextualism

I illustrate the threat of ad hocness with a case study: the contextualism of Lewis [1996].

A notable feature of the RA account in Lewis [1996] is its complex account of relevance, based on a large number of so-called *rules of relevance*: the rule of actuality; the rule of belief; the rule of resemblance; the rule of reliability; two rules of method; the rule of conservatism; and the rule of attention.

These rules do substantial work. Lewis claims that skeptical hypotheses are irrelevant in mundane contexts of attribution, but drawing attention to them immediately updates the context so that they *are* relevant. This, it is proposed, explains the force of skeptical arguments. His explanation relies crucially on the rule of attention, which states that if the interlocutors are actively attending to a possibility, then it is relevant. Lewis also claims that, for the purpose of truly ascribing knowledge, we are generally entitled to take for granted the reliability of

[51] See Holliday [2012], Holliday [2013], Holliday [2015a] and Holliday [2015b] for detailed precise developments of extant RA proposals, and refinements of epistemic logic based on an RA approach.

processes like memory and testimony. This relies on the rule of reliability: attributors may defeasibly ignore the possibility that a basic channel for transmitting information is unreliable. He also claims that we are generally entitled to take a sample of objects to be representative of its population. This relies on a rule of method. And so on.

Lewis does not make it clear why he takes this exact set of rules to constitute the right account of relevance. He hints that the purpose of the ordinary knowledge concept is to approximate rigorous Bayesian reasoning. But that is of little help in setting the exact rules of relevance that he proposes. What are we to make of his account?

I see two interpretations. On one interpretation, his rules of relevance are merely a description of the contingent norms that govern our actual, ordinary usage of epistemic vocabulary. On this view, (i) Lewis's general criterion of relevance is flexible enough to accommodate various possible linguistic practices and (ii), as a matter of fact, the practice of ordinary interlocutors is such that they generally ignore and attend to possibilities in accordance with Lewis' rules. Blome-Tillmann [2014] develops a version of this reading according to which relevance is ultimately a matter of *presupposition*.[52] On this view, the rules of relevance reflect largely contingent facts about what tends to be presupposed in ordinary conversation. This approach has a counter-intuitive consequence that Lewis' loose discussion masks: our epistemic practice relies *heavily* on mere presupposition, imperiling any claim to reflect a robust rationality.

However, the second interpretation is of more immediate interest. On this reading, Lewis [1996] does not take the rules of relevance as reflecting a contingent linguistic practice that is rooted in a unified but liberal criterion of relevance (e.g. presupposition). Instead, the rules of relevance offer an irreducibly complex account of relevance.

A virtue of this account is that it conforms with linguistic practice without reducing knowledge attribution to an expression of (potentially arbitrary or otherwise ill-founded) presuppositions. An obvious vice of the account, however, is that it seems *ad hoc*: since no unified, well-motivated theory of relevance *determines* the rules of relevance, we are left with a patchwork quilt that all too

[52]This would explain why Lewis [1996] treats the rule of attention as having such force i.e. as providing a sufficient condition for relevance. For this rule may be interpreted as saying that a possibility is relevant if the presupposition of its negation is explicitly rejected.

conveniently waylays counter-examples.

1.4 Resolution theory defuses the problems

The discussion of the previous section sets the stage for four challenges for resolution theory. I state each challenge, and sketch resolution theory's response. The rest of the dissertation (appendices excluded) fills out these sketches.

1.4.1 Challenge 1: closure denial

Challenge: resolution theory does not validate epistemic closure (single premise or multi-premise) in full generality. For instance, as shown in chapter 3, resolution semantics does not validate the formula $Kp \to K(p \vee q)$.

Reply: resolution theory validates a *restricted* closure principle: knowability is preserved under deductive consequence when no new subject matter is introduced in the conclusion (cf. [Yablo, 2012], [Yablo, 2014, Ch. 7]). I argue in chapter 3 that this restricted closure principle validates and invalidates the *right* patterns of inferences. In particular, it invalidates instances of closure that can be used to ground skeptical paradoxes, without invalidating prominent instances that seem harmless. I show that, in this respect, resolution theory enjoys an advantage over similar theories: namely, Yablo [2012] and Schaffer [2005a].

1.4.2 Challenge 2: the case for rampant easy knowledge

Challenge: we saw in section 1.3.3 that a case can be made for the respectability and ubiquity of easy knowledge, supporting RA theories that secure closure by allowing for easy knowledge (e.g. Stine [1976], Lewis [1996]). The case is roughly this: it can be shown that every fact can be expressed by a contingent a priori truth, but we should hesitate to call this knowledge *substantial*. Rather, semantic considerations point to the existence of rampant *cheap* but *vacuous* knowledge.

Reply: this is a *deflationary* account of easy knowledge, grounding closure-embracing RA theories that embrace closure only *superficially*: if we concentrate on *substantive* epistemic states, closure is *not* preserved by such theories. More importantly, a deflationary approach leaves it open how best to account for substantive knowledge. I argue that resolution theory provides an attractive account of substantive knowability. For one of the arguments in favor of rampant contingent a priori truth - Gibbard's cheap trick - takes on the air of a paradox if we restrict attention to substantive knowability. Resolution theory offers a means for defusing this paradox. Or so I argue in chapter 4.

1.4.3 Challenge 3: missed clue counter-examples

Challenge: resolution semantics takes (what we called in 1.2.3) the Dretskean theory of knowability as its account of *ruling out* A given E: if A were not the case, then E would not be the case. In section 1.3.4, we noted some compelling *missed clue* counter-examples to a naive use of Dretskean knowability. Are these counter-examples inherited by resolution theory? Here is a prima facie reason to think so: in section 1.3.4, we considered the proposal that missed clue cases represent a violation of a general knowability principle, which we called *conservative* $\rightarrow$ *intro*: $(K\varphi \wedge K\psi) \rightarrow K(\varphi \rightarrow \psi)$. Resolution semantics does not validate this principle (see theorem B.5.5 in appendix B). It is not obvious that this is defensible: we suggest in chapter 3 that an instance of closure denial can be motivated if that instance can be used to ground a Cartesian skeptical paradox. But can *conservative* $\rightarrow$ *intro* ground a Cartesian paradox?

Reply: my reply has two parts, the first of which I state definitively here. I suggest that the resolution theorist can get away with biting the bullet on the rejection of *conservative* $\rightarrow$ *intro*, since it allows for the construction of a paradox that is, at the very least, closely related to Cartesian concerns. Suppose that atomic claim e expresses the agent's *total empirical information*, including an experience as of seeing the agent's hand (in good lighting etc.). As usual, we assume that E is compatible with skeptical possibilities, including a brain-in-vat world where the agent is massively deceived. Let h express that the agent has hands. I assume that e is knowable given the information E (this is a triviality, in fact) and that h is knowable given E (our usual assumption that mundane knowledge is a possibility given our limited information). Now suppose we accept $(Ke \wedge Kh) \rightarrow K(e \rightarrow h)$. It follows from all this that the agent is positioned to know that: it isn't that e is true yet h is false. But it is hard to see how *this* could be. After all, the ways in which e is true but h is false are (at best mildly) skeptical scenarios in which the agent's total sensory experience is deceptive on the issue of her hands. Presumably, the agent's total information cannot be used to rule out *such* a possibility. I conclude that the resolution theorist can reject conservative $\rightarrow$ introduction, in accord with her strategy of restricting closure to defuse skeptical threats.

Thus, scenarios exist that have the *form* of a missed clue case relative to resolution semantics, but resolution theory refuses to view this as a problem in itself.

However, there is a residual problem. What of a *concrete, particular* case such as *Jane's missed clue* (section 1.3.4)? Surely, we do *not* want resolution theory to deliver a counter-intuitive result for *this particular* case.

In chapter 4, after I motivate the neo-Fregean aspects of resolution semantics, I show that resolution semantics need *not* deliver a counter-intuitive result. In fact, by attributing a natural *Fregean guise* to the claim that is tested for knowability in the example, resolution theory rules in accord with common sense: Jane must rule out the proposition that the depicted bird has red plumage but *is* a wild canary, as this alternative is indeed relevant.

1.4.4 Challenge 4: independent support

Challenge: I assume that the account of relevance utilized by the Dretskean theory of knowability is *not* ad hoc: not only is it directly inspired by some key motivating examples (i.e. Goldman-Ginet barn cases), but an account of similarity between worlds is required to ground the orthodox approach to the semantics of counterfactuals. However, what of the topic-sensitive account of relevance utilized by resolution semantics? This account will not be ad hoc only if it can be argued that the underlying account of subject matter is independently motivated.

Reply: I argue in chapter 2 that the theory of subject matter in question is a better performer than various prominent rivals: Lewis [1988a], Perry [1989], Yablo [2014], among others. Various compelling intuitive desiderata have been proposed in the literature for a theory of subject matter. I observe that if these desiderata are rounded up and applied uniformly, then no extant proposal (of the prominent ones under consideration, at least) satisfies them jointly. I then introduce a novel proposal - the issue-based theory of subject matter - that *does* jointly satisfy the desiderata. This theory has the right features to ground resolution semantics.

1.5 Plan

Here is the plan for the rest of the dissertation.

Chapter 2 concerns the theory of subject matter. I defend a novel proposal: the *issue-based theory*. I then draw on this proposal in setting up resolution semantics in later chapters (in particular, in fleshing out its account of relevance).

Chapter 3 concerns the debate over epistemic closure. My goal is to motivate the respectability of closure denial. I offer some criteria of adequacy for such a theory, and argue that resolution semantics meets the criteria (in contrast to some closure-denying rivals: Schaffer [2005a] and Yablo [2014]).

Chapter 4 is devoted to studying Gibbard's cheap trick (cf. [Gibbard, 2012, appendix 1]). I have three major aims in the chapter, relative to the rest of the dissertation. First, I aim to argue that the cheap trick should be regarded as a paradox, at least when we focus attention of *substantive knowability* (which, I propose, is of central interest in epistemology, even if one accepts a notion of cheap but vacuous knowledge). Second, I aim to motivate certain aspects of resolution theory (its *neo-Fregean* aspects) as a way of resolving the paradox (along with ensuring that the theory has other pleasant features). Third, I then use the neo-Fregean aspects of resolution theory to show that certain counter-examples to a naive Dretskean theory - including *Jane's missed clue* - can be handled by resolution theory.

Chapter 5 states my conclusion, and observes some of the many possibilities for future work.

Appendix A serves as a background chapter for readers that wish to dig deeper into the basic motivations and various forms of RA theory. I also contrast the general idea of 'relevance' in the RA context to other uses of this term in the philosophical literature. I conclude by offering a high-level logical framework for constructing and comparing RA theories.

Appendix B is dedicated to stating resolution semantics with precision, and in total. In other parts of the dissertation, I introduce aspects of this overall framework as needed. The appendix is for those that are interested to see a precise theory briefly and in its entirety.

Chapter 2

Theories of Aboutness

Our topic is the theory of topics (i.e. subject matter).[1] My goal is to clarify and evaluate three competing traditions: what I call the way-based approach, the atom-based approach and the subject-predicate approach. I develop (defeasible) criteria for adequacy using robust linguistic intuitions that feature prominently in the literature. Then, I evaluate the extent to which various existing theories satisfy these constraints. I conclude that recent theories due to Parry, Perry, Lewis and Yablo do not meet the constraints in total. I then introduce the *issue-based theory*: a novel and natural entry in the atom-based tradition that succeeds in meeting the criteria. Finally, in section 2.6, I categorize a recent theory from Kit Fine as atom-based and contrast it to the issue-based theory, concluding that they are evenly matched relative to our main criteria. I offer tentative reasons to nevertheless favor the issue-based theory.

2.1 Introduction

Descriptive language allows us to say true things about interesting topics. This points to three core semantic concepts: truth; aboutness; topic (i.e. subject matter). Truth and the conditions of truth have attracted ample attention in the philosophy literature. Nominally, aboutness has also received attention, in

[1] Hawke [2017] offers a refined and extended version of the material in this chapter.

the guise of two closely related issues: *reference* and *intentionality*.[2] In contrast, the notion of *topic*, and the sense in which a claim is *about its topic*, have until recently received only passing and sporadic attention.[3]

This new-found attention shadows a concentrated effort to model the *hyper-intensionality* of natural language: that is, the phenomenon of distinct indicative expressions that are true at exactly the same possible worlds, yet are not interchangeable in every linguistic context in which they may be embedded. Thus, a semantic theory overlooks significant dimensions of meaning if it serves *merely* to assign a set of possible worlds (a *truth set*, an *intension*) to indicative expressions. We illustrate with an example from Perry [1989]. Suppose that "Jack brought it about that Jill tumbled down the hill" is true. Apparently, it does not follow that "Jack brought it about that Jill tumbled down the hill and 2+2=4". Nor does it follow that "Jack brought it about that Jill tumbled down and Jones is Jones". Nor that "Jack brought it about that Jill tumbled down and Peter either picked or did not pick a peck of pickled peppers." A tempting explanation: first, there is a difference in topic between "Jill tumbled down the hill" and (for instance) "Jill tumbled down the hill and 2+2=4". Second, the truth of "a brought it about that φ" is sensitive to what φ is *about*, not only its truth set. Thus, the operator "a brought it about that" creates a hyper-intensional context for the operand φ.

A picture of subject matter due to Lewis [1988a] and Lewis [1988b] has steadily grown in influence. Lewis identifies a subject matter with the set of possible ways for the subject in question to be, understood abstractly as a set of (unstructured) propositions that *cover* (i.e. jointly exhaust) logical space.[4] This closely relates subject matter to standard semantic theories of interrogative expressions, which likewise identify a *question* with a set of propositions - namely, the set of answers

[2]As our discussion will intimate, the exact relationship reference, intentionality and subject matter is unlikely to be trivial.

[3]See Ryle [1933], Putnam [1958], Goodman [1961] and Perry [1989] for some important entries in the pre-Lewis discussion of subject matter. Perry [1989] also discusses insights due to Barbara Partee. Linguists have not neglected the topic of topics to the same extent - see Roberts [2011]. How best to relate this tradition to our own discussion must be left for elsewhere. For up-to-date surveys on the issues of reference and intentionality, see Jacob [2014] and Reimer and Michaelson [2017].

[4]Or, at least: a covering of that set of worlds where the subject in question *exists*. However, we will not consider any version of a way-based theory that departs from the basic picture of a topic as a covering of logical space. We leave the subtleties that might motivate a refinement for another time. At any rate, the core Lewisian idea is most minimally and flexibly summarized as: topics are sets of propositions, conceived of as ways that the subject in question can be.

to that question.[5]

Call this broadly Lewisian picture the *way-based conception*. Applications of this approach are now rife, including theories that posit that belief or knowledge are topic-sensitive[6] and theories of partial truth.[7] As a complement, the identification of topics and questions has been leveraged in the pragmatics literature, by modeling a discourse topic as a *question under discussion*.[8]

The Lewisian tradition is not the only game in town, however. At least two other approaches have found traction, with tendencies that are often at odds with the way-based conception.

In the first place, logicians in the "relevantistic" tradition have appealed to subject matter to explain why certain classical argument forms strike many as fallacious, despite preserving truth:[9] it is posited that these forms do not preserve the *subject matter* of the premises in the conclusion, and that such preservation is necessary for sound argumentation.[10] Putting aside the (de-)merits of this explanation, we attend to the picture of subject matter that informs the diagnosis. Framed simply, the leading idea is that the subject matter of φ can be identified, in some sense, with the set of atomic claims from which φ is composed. As a corollary, subject matter is treated as invariant under negation, while conjunction and disjunction merely *merge* the subject matter of their constituents. Stated in more abstract and flexible terms: subject matters are sets of objects of an appropriate kind; every atomic claim can be associated with a set of such objects; the subject matter of φ is determined by (i.e. a function of) the subject matter of the atoms that occur in φ. Call this the *atom-based conception*.[11]

In the second place, philosophers have developed the view that the subject matter of φ is the set of objects *of which something is said* by stating φ i.e. those objects that count as *subjects* of which something is *predicated* by uttering φ. In

[5] See Cross and Roelofsen [2016] for an overview of this tradition.

[6] See Yalcin [2011], Yablo [2014], chapter 7, and Yalcin [2016].

[7] See Yablo [2014], chapter 5.

[8] See Roberts [2012].

[9] For overviews of the relevantistic tradition, and its approach to subject matter, see Read [1988] and [Burgess, 2009, Ch. 5]. For a classic and thorough discussion of the subject, see Anderson and Belnap [1975].

[10] The standard examples of such inferences are *ex quolibet verum*, with a paradigm case being the inference to the validity $p \vee \neg p$ from an arbitrary proposition q; and *ex falso quodlibet*, with a paradigm case being the inference of an arbitrary proposition p from the contradiction $q \wedge \neg q$.

[11] To illustrate how adopting the way-based conception can lead to an account of 'relevant implication' that radically differs from the main tradition, see Lewis [1988b].

general, a subject matter is a set of objects, with no constraints on what objects can so serve. Call this the *subject-predicate conception.* Perry [1989] sketches a sophisticated proposal along these lines.

The goal of the current chapter is to systematically contrast and evaluate the divergent conceptions we have described, and thereby defend a novel version of the atom-based approach: the *issue-based theory.* Put provocatively: the way-based conception enjoys momentum in the recent literature, and the current paper aims to bolster a persistent alternative to this rising trend.

Section 2.3 offers my instrument of evaluation: a set of constraints rooted in robust linguistic intuition. For the most part, noting these intuitions is not original to this chapter, for they appear piecemeal in the literature.[12] The current contribution is to round them up, support them with a *uniform* rationale and apply them *uniformly* to a diverse array of theories. In section 2.3, I justify our general constraints by drawing on ordinary judgments as to whether certain claims are *on-topic* or *off-topic* relative to a given discourse topic. Over sections 2.4.1-2.4.3, I argue that a prominent selection of subject-predicate, atom-based and way-based theories fail to meet our constraints. Section 2.5 argues that the issue-based theory does so succeed. Though my criteria for success are defeasible, I offer this as prima facie support for the issue-based theory.

The issue-based theory identifies a subject matter with a set of *distinctions* or *issues.*[13] While hopefully novel, note that this theory draws liberally from existing ones, across the traditions.[14] The aim is to provide a *synthesis* that preserves strengths and discards weaknesses. The theory allows for the recovery, from an assertion, of a set of subjects of which something is thereby predicated, and a partition of ways things can be with respect to the subject matter of the claim (what I call its *resolution* on logical space). In short, a version of the subject-predicate and way-based approach can be *abstracted* from the issue-based theory. Thus, this paper does not advocate abandoning these conceptions altogether. The suggestion is that they are elegant abstractions that are sometimes illuminating, sometimes misleading.

[12] I chiefly collect these intuitions from the following sources: Goodman [1961], Lewis [1988b], Perry [1989] and Yablo [2014].

[13] Cf. [Yablo, 2014, pg. 27]: "A subject matter - I'll sometimes say topic, or matter, or issue - is a system of differences, a pattern of cross-world variation."

[14] For instance, the theories of Epstein [1994] and Roberts [2012] share many important features of the current approach.

Three caveats. To focus discussion, I concentrate on the austere setting of propositional logic, basic predications and identity statements (putting aside quantification). Further, I cannot claim to engage with every theory of aboutness on offer.[15] A prominent selection must suffice for our purposes. Third, I assume that the ordinary term 'topic' (and 'subject matter') is uni-vocal, and that ordinary judgments about being *on-topic* are systematic. In contrast, some authors embrace the sentiment that 'topic' is too vague or ambiguous to license a unique, unified theory.[16]

As a coda, section 2.6 briefly addresses the more subtle challenges posed by the recent theory of Fine [2016]. I classify this theory as atom-based and argue that it meets the criteria of adequacy in section 2.3. Thus, my main tool of evaluation puts Fine's theory and the issue-based theory in a dead heat. In response, I explore the extent to which they are not competitors, while tentatively framing promising but inconclusive reasons for favoring the issue-based theory.

2.2 Assumptions and notation

We use $\mathbf{s}, \mathbf{t}, \ldots$ to denote subject matters. We use $\mathbf{s} + \mathbf{t}$ to denote the subject matter which is attained by combining $\mathbf{s}$ and $\mathbf{t}$. For instance, the subject matter for a course in cognitive science might be a combination of two topics: neural networks + Bayesian models of cognition. We use $a, b, c, \ldots$ to refer to individual objects; $F, G, \ldots$ to denote properties; and R to denote a relation. When utilizing a formal propositional language, I use p, q as meta-variables that range over atomic sentences and φ, ψ as meta-variables that range over all sentences. We also use standard set-theoretic notation: $\cup$ is set union; $\cap$ is set intersection; $\in$ indicates membership; $\subseteq$ is the subset relation; and $\emptyset$ denotes the empty set. For $P \subseteq W$, P^c indicates the complement of P i.e. $W \setminus P$.

A prevalent feature of everyday talk is that an individual object can, in some sense, serve as a subject matter. For instance, one might say that the topic of

[15] For instance, I pass over proposals in Ryle [1933],Goodman [1961] and Putnam [1958].

[16] Ryle [1933] argues that 'about' has a multiplicity of meanings. Compare Fine [2017b], part II, section 2: "There is an intuitive notion of subject-matter or of what a statement is about. This notion may have a different focus in different contexts. Thus it may be objectual and concern the objects talked about or it may be predicational and concern what is said about them. Our concern here will be with what one might call 'factual' focus, with what it is in the world that bears upon the statement being true or false".

"John is late" is John, or that "John is late" is about John. However, we do not consider any theories that *literally* allow an individual (concrete) object to count as a subject matter.[17] Thus, we assume the correct interpretation of our ordinary talk is that for every object a there is an associated subject matter, which we denote by $\mathbf{a}$ (thus, the topic of "John is late" is not John but, more accurately, **John**).

We use $\leq$ to indicate the *inclusion relation* between topics. Hence, one could write "topology $\leq$ mathematics" to indicate that topology forms part of a larger subject matter: mathematics. Similarly, we use $\bowtie$ to indicate that two subject matters *overlap*: for instance, we write "mathematics $\bowtie$ philosophy" to indicate that mathematics and philosophy are overlapping topics (with neither being inclusive). We understand $\bowtie$ as a defined relation: $\mathbf{s} \bowtie \mathbf{t}$ iff there exists subject matter $\mathbf{u}$ where: $\mathbf{u} \leq \mathbf{s}$ and $\mathbf{u} \leq \mathbf{t}$. That is, overlap amounts to $\mathbf{s}$ and $\mathbf{t}$ having a common part.

We assume that every meaningful sentence φ can be associated with a subject matter $\mathbf{s}_\varphi$ that counts as *the* subject matter of that sentence. We assume that φ is entirely about $\mathbf{t}$ just in case $\mathbf{s}_\varphi \leq \mathbf{t}$. We assume that φ is partly about $\mathbf{t}$ just in case $\mathbf{s}_\varphi \bowtie \mathbf{t}$. These assumptions allow a ready explanation for the role of subject matter in the guidance of discourse. Intuitively, conversation is regulated by a background discourse topic, which determines what assertions count as (conversationally) *relevant* or *irrelevant*. If our topic is Jane's profession, then the claim "Jane is a lawyer" is relevant (i.e. on-topic), whether or not it is true. To say a claim is *somewhat* on-topic is to say that its subject matter overlaps with the discourse topic (e.g. "Jane is a lawyer and loves to procrastinate" in our running example). On the other hand, "John is a lawyer" or "Jane loves to procrastinate" are *not* on-topic. This can be explained as follows: the latter two sentences each has a subject matter, and that subject matter is neither included in, nor even overlaps with, the discourse topic. With these observations in mind, we will not hesitate to make use of intuitions concerning *conversational relevance* as evidence for the subject matter of a particular assertion.

[17]The subject-predicate approach faces the least difficulty in identifying subject matters with concrete objects, if this is insisted upon. According to this conception, a subject matter is a set of objects. However, a slight variation treats subject matters as *plural objects* i.e. instead of using the set of objects S, one takes the subject matter to be that object whose parts are the individuals in S.

2.3 Criteria of adequacy

In this section, we develop some criteria of adequacy for a theory of subject matter: we note apparently unequivocal linguistic intuitions generated by particular examples, and then help ourselves to the obvious generalizations. Section 2.3.2 briefly reflects on this methodology.

2.3.1 Constraints via linguistic data

The connectives

We first build on observations in Perry [1989] of various interactions between subject matter and the connectives.

Suppose that our discourse topic is **Jane's profession**. Clearly, assertions of "Jane is a lawyer" or "Jane is an accountant" are on-topic (and, more broadly, about **Jane**). Now, note further that "Jane is *not* a lawyer" seems equally on-topic, and also seems entirely about **Jane's profession** (and, more broadly, **Jane**). This suggests that subject matter is preserved under negation.

Likewise, an assertion of "Jane is a lawyer or Jane is an accountant" seems to be on-topic, and so is still about **Jane's profession** (and **Jane**). This suggests that shared subject matter is preserved under disjunction. From this, we may note evidence that the preservation of subject matter under negation is not limited to atomic claims: "Jane is neither a lawyer nor an accountant" is also intuitively entirely on-topic. To generalize:

1. If φ is entirely about $\mathbf{s}$ then $\neg\varphi$ is entirely about $\mathbf{s}$
2. If p is entirely about $\mathbf{s}$ and q is entirely about $\mathbf{s}$ then $p \vee q$ is entirely about $\mathbf{s}$

Next, suppose it is asserted that "Jane is a lawyer and John is a lawyer". It is intuitive to say that this assertion is at least *partly* on-topic: it is partly about **Jane's profession** (and partly about **Jane**). One might add that "Jane is a lawyer *and* an accountant" is *wholly* on-topic. To generalize:

3. If p is entirely about $\mathbf{s}$ then $p \wedge q$ is partly about $\mathbf{s}$

Next, note that it is difficult to think of a discourse context where a claim of the form $p \wedge q$ is relevant to the topic at issue, but the claim $p \vee q$ is not (though, of course, the second can be *less informative* than the first). Suppose that our topic is whether Frankie is a bachelor. Obviously, "Frankie is a bachelor" (i.e."Frankie is both a man and unmarried") is entirely on-topic. It also seems hard to deny that "Frankie is not a married woman" (i.e. "Frankie is either a man or unmarried") is on-topic - though not informative enough to resolve the issue completely.

Now reason as follows: suppose that $\mathbf{s}$ and $\mathbf{t}$ are such that for every topic $\mathbf{u}$, we have that if $\mathbf{s} \leq \mathbf{u}$ then $\mathbf{t} \leq \mathbf{u}$. We may conclude that $\mathbf{t} \leq \mathbf{s}$, since this is the special case where $\mathbf{u} = \mathbf{s}$. Substituting the topic of $p \wedge q$ for $\mathbf{s}$, the topic $p \vee q$ for $\mathbf{t}$ and an arbitrary discourse topic for $\mathbf{u}$, we get: the topic of $p \vee q$ is included in that of $p \wedge q$. (The other direction is less obvious. If the topic is "does Jane have a sibling?", it is clearly exactly on-topic to say that "Jane has either a brother or a sister". Further, "Jane has both a brother and a sister" is undeniably on-topic *to some extent*. But does this second assertion include some irrelevant information? Similar questions could be asked about the claim "Jane has a sister".) To generalize:

4. The subject matter of $p \wedge q$ includes the subject matter of $p \vee q$

Finally, note that the statement "either Julia Robinson is an expert in diophantine equations or Raphael Robinson is" is relevant to at least one discourse topic: **the experts in diophantine equations**. On the other hand, if our topic is which philosophers are experts in semantics, then this statement is irrelevant and off-topic. To generalize:

5. Expressions of the form $Fa \vee Fb$ are about something but not necessarily about everything

Validities, contradictions and necessities

Again borrowing from Perry [1989], consider: "Caesar brought it about that Tully fell out of bed". Suppose that this is true. Intuitively, the following claims need *not* be true: "Caesar brought it about that *both* Tully fell out of bed *and* 2 is even"; "Caesar brought it about that *both* Tully fell out of bed *and* either Trump won the 2016 election or he didn't win"; "Caesar brought it about that *both* Tully fell out of bed *and* Jones is Jones"; "Caesar brought it about that *both* Tully fell out of bed *and* everything is self-identical"; "Caesar brought it about that *either* Tully fell out of bed *or* 2 is odd".[18] And so on. Intuitively, the pattern is clear: "a brought it about that φ" reports a relation between an actor a and a particular (actual) state of affairs that is expressed by φ, but exactly *which* state of affairs is so expressed depends heavily on the subject matter of φ, not only its truth conditions. All this suggests that necessities such as "Jones is Jones", "Trump won the election or he didn't win" and "2 is even" (and the impossibilities that result from denying these claims) are about *something.* For if they were not about anything, then presumably conjoining them with "Tully fell out of bed" would produce a claim whose subject matter is, after all, confined to that of "Tully fell out of bed".

Intuition also suggests that validities, contradictories and necessities are not about *every* subject matter. Intuitively, "2+2=4" is about arithmetic (and partly about the number **2**), not about topology. "Jones is not Jones" is about **Jones**, but not about **Jane**, or arithmetic. "Either Trump won or he didn't" is about **Trump**, but not about **Abraham Lincoln**, and not about geometry. (Such intuitions, I take it, are widely shared among generations of beginning logic students and relevant logicians that are struck that an inference from "Trump won and he didn't" to "Jane is a lawyer" offers a conclusion that significantly departs from the premises in what it is about.)

We generalize as follows:

6. If Fa is contingent, then it has different subject matter to $Fa \wedge (b = b)$ and $Fa \wedge (Gb \vee \neg Gb)$

7. If Fa is contingent, then it has different subject matter to $Fa \vee (b \neq b)$ and

[18] One could, of course, use a more standard purported example of a hyper-intensional context to illustrate the above point: belief attributions. However, we follow the lead of Perry [1989] and resist this temptation for the following dialectical reason: the hyper-intensionality of belief ascriptions seems, at least in many cases, intimately interwoven with issues of *mode of presentation* or *guise* i.e. the different ways in which individual objects can be thought about. Now, what is striking about the "brought it about" ascriptions in the main text is that such ascriptions do *not* seem sensitive to issues of mode of presentation. There is no conceivable context where "Caesar brought it about that Tully fell out of bed" is true while "Caesar brought it about that Cicero fell out of bed" is not. Thus, "brought it about" claims allow us to draw conclusions about the sensitivity of a linguistic context to subject matter without the fear of having relabeled or mislabeled a mode-of-presentation phenomenon.

$Fa \vee (Gb \wedge \neg Gb)$

8. A claim of the form $Fa \vee \neg Fa$ is about something (e.g. **a**) but not about everything (at least if Fa is about something but not everything). Likewise for $a = a$ and most cases of Fa where Fa is necessary

9. A claim of the form $Fa \wedge \neg Fa$ is about something (e.g. **a**) but not about everything (at least if Fa is about something but not everything). Likewise for $a \neq a$ and most cases of $\neg Fa$ where Fa is necessary

Denotations

Consider an amusing example from [Yablo, 2014, pg.24]: "Man bites dog" is a more interesting headline than "Dog bites man", since it speaks to a more interesting topic. Put another way, suppose that our discourse topic is **men who have bitten animals**. Then, "Joe bit Rex" is on-topic, but "Rex bit Joe" is not. As usual, we conclude that the subject matter of such sentences must diverge. Or suppose that our topic is **Jane's children**. Then "Jane is the mother of Beth" is on-topic, while "Beth is the mother of Jane" is not.

Further: suppose that our topic is, once again, **Jane's profession**. In this case, an assertion of "Jane is a lawyer" is on-topic, while "Jane is a serial procrastinator" is not. To generalize:

10. Expressions of the form aRb and bRa are not necessarily about the same topic. Nor is the subject matter of Fa necessarily identical to that of an expression of the form Ga

Parts and wholes

Goodman [1961] notes: since Maine is part of New England, "Maine experiences cold winters" is intuitively about **New England**. In general, talking about a part is apparently a way of talking about the whole.

11. If a is part of b, then: if φ is about **a**, then φ is about **b**

Goodman [1961] notes a second intuition: "New England experiences cold winters" seems to be about **Maine** (as well as the other parts of New England). However, as Goodman [1961] points out, if we generalize in the obvious way (a claim about a whole is also about its parts) and combine this with constraint 11, then we can derive absurdities. For instance: since Maine is part of the world, constraint 11 delivers that "Maine experiences cold winters" is about **the whole world**. Now, if a claim about a whole is also about its parts, then "Maine experiences cold winters" is not only about **the whole world**, but also about **Hawaii** in particular, since Hawaii is part of the world. This is obviously wrong-headed. In response, we should reject the principle that a claim about a whole is also about its parts. Indeed, further counter-examples spring to mind: suppose I say "it is illegal to drive over 65 miles per hour on the highway". This is a claim entirely about **the law**. One part of the law deals in copyright infringement. But I have not said anything about **copyright law**. I say "Paris is beautiful". Presumably, I am not talking about **the Paris slums**. The situation is puzzling. What accounts for the intuition that talking about **New England** involves talking about **Maine**?[19] Since we cannot pinpoint the source of these intuitions (if they exemplify a general principle at all), I will not propose a general criterion of adequacy to reflect Goodman's second example.

Questions

It seems that an interrogative utterance expresses a subject matter. To use a favorite example of Lewis [1988b], the question "how many stars are there?" determines a topic of discourse (explicitly: **the number of stars**). Or consider: we can talk about **Jane's profession** by discussing *what Jane's profession is* i.e. by addressing the question "what is Jane's profession?" Our final constraint is thus:

12. A question **Q** serves (in some sense) as a subject matter

It is not obvious that every subject matter can serve as a question. "Jane is a lawyer" is about **Jane**. Can we think of **Jane** as a question? This issue veers us into overtly theoretical territory.

[19] Perhaps such cases call for a Yablovian strategy (cf. [Yablo, 2014, Ch. 5]): posit that the claims we're discussing have contents that can be divided into parts, and that sometimes we deploy them knowing full well that only some of these parts are true (e.g. the part about the Eiffel tower or driving law, not the parts about the Paris slums or copyright law).

2.3.2 Methodological remarks

It is worth emphasizing the nature and limitations of my methodology, since treating the above constraints as our sole criterion of adequacy loads the die against certain approaches.

Our constraints are generalizations based on striking linguistic intuitions, where a "linguistic intuition" is an intuitive judgment as to whether a piece of language is deployed felicitously in a partially described linguistic context. We have mainly been concerned with comparing the *relevance* of an individual claim in a discourse context i.e. the extent to which the topic of the claim overlaps with the discourse topic.

Three caveats. First, I cannot definitively claim I have exhausted the linguistic intuitions that are relevant to determining a theory of subject matter.

Second, I do not claim that linguistic intuition provides the *only* relevant evidence for a theory of subject matter (certainly not in the setting of philosophy). Rational theory selection is carried out by choosing the theory with the highest (expected) theoretical utility. Accommodating linguistic intuition enhances the utility of a theory along one dimension: it exhibits the *explanatory power* of the theory. But one ought not to ignore other theoretical virtues, such as elegance, parsimony and systematicity. Further, explanatory power manifests in diverse ways: for instance, a theory may offer little over its competitors in accounting for basic linguistic data, but have wide applicability for resolving philosophical puzzles. Sometimes we must weigh trade-offs.

Third, accommodating the available linguistic data does not necessitate vindicating straightforward generalizations. Another strategy is to explain why those intuitions are mistaken or misleading, or that the data is more parochial than is first apparent.

In short, the criteria for adequacy that I deploy are best viewed as defeasible, though carrying *prima facie* force.

2.4 Evaluation of existing approaches

We now work through a slew of theories that fall under either the subject-predicate conception, the atom-based conception and the way-based conception. Each such theory provides an account of subject matter and what it is for one subject matter to be included in another ($\leq$), overlap with another ($\bowtie$) or be combined ($+$) with another. Each theory also provides, for an arbitrary φ, an account of *the* subject matter of φ ($\mathbf{s}_\varphi$) and so what φ is about, entirely or partly. In every case, we identify a constraint from section 2.3.1 that is violated by the theory.

2.4.1 The subject-predicate conception

As we understand the subject-predicate conception, the class of subject matters is just the class of all sets. Hence, a subject matter is nothing but a non-empty set of objects, and any non-empty set of objects can serve as a subject matter. $\mathbf{s} \leq \mathbf{t}$ just means $\mathbf{s} \subseteq \mathbf{t}$; $\mathbf{s} \bowtie \mathbf{t}$ just means $\mathbf{s} \cap \mathbf{t} \neq \emptyset$; and $\mathbf{s} + \mathbf{t} = \mathbf{s} \cup \mathbf{t}$.

Consider an atomic claim φ. According to the current conception, $\mathbf{s}_\varphi$ is the set of objects that serve as *subjects* in φ (i.e. of which a property or relation is predicated). For example, the subject matter of "John helped Jack" is $\{\text{John}, \text{Jack}\}$. φ is entirely about $\mathbf{t}$ iff $\mathbf{s}_\varphi \leq \mathbf{t}$. φ is partly about $\mathbf{t}$ iff $\mathbf{s}_\varphi \bowtie \mathbf{t}$. For example, since John and Jack both live in Maine, "John helped Jack" is entirely about **the citizens of Maine**, since $\{\text{John}, \text{Jack}\} \subseteq \{x : x \text{ is a citizen of Maine}\}$. On the natural proposal that **John** $= \{\text{John}\}$, "John helped Jack" is partly about **John**, since $\{\text{John}\} \subseteq \{\text{John}, \text{Jack}\}$.

Perry

Perry [1989] offers a sophisticated elaboration of the subject-predicate conception, drawing on situation theory.[20] For a certain situation s to be the case is for certain objects to stand in certain relations and certain objects to fail to stand in certain relations (at a certain space-time location, one might add). Thus, situation s may be represented by a *partial valuation* ρ_s, assigning either 1 (true), 0 (false) or nothing (undetermined) to every atomic claim, in accord with s. We may then determine whether an arbitrary claim is *verified* or *falsified* by s as follows:

- s verifies p just in case $\rho_\mathsf{s}(p) = 1$. s falsifies p just in case $\rho_\mathsf{s}(p) = 0$.
- s verifies $\neg\varphi$ just in case s falsifies φ. s falsifies $\neg\varphi$ just in case s verifies φ.

[20]See Barwise and Perry [1981] for a classic study of situation theory.

- s verifies $\varphi \wedge \psi$ just in case s verifies φ and verifies ψ. s falsifies $\varphi \wedge \psi$ just in case s falsifies φ or falsifies ψ.

- s verifies $\varphi \vee \psi$ just in case s either verifies φ or verifies ψ. s falsifies $\varphi \vee \psi$ just in case s falsifies φ and falsifies ψ.

The key proposal from Perry [1989]: $\mathbf{s}_\varphi$ is that set of objects that is part of *every situation* that verifies φ. This delivers intuitive consequences for complex claims. $\mathbf{s}_{Fa \wedge Ga}$ is $\{a\}$. $\mathbf{s}_{Fa \wedge Gb}$ is $\{a, b\}$. This meets constraint 3. Thus, "Jane is a lawyer and John is an accountant" is partly about **Jane**. Further, $\neg Fa$ has the same subject matter as Fa, largely meeting constraint 1.

It might at first seem that this theory violates constraint 11. This is so on a flat-footed reading, where the subject matter associated with concrete object a is always $\{a\}$. Consider "Maine experiences cold winters". Perry's theory says: the subject matter of this claim is the set of those objects that are part of every situation that verifies that Maine has cold winters. If this set is {Maine}, then it follows that the claim is entirely about **Maine**, as desired. However, it is then not entirely about **New England**, for {Maine} $\nsubseteq$ {New England}.

However, a Perry supporter has room to maneuver: she can insist that, for any subject matter $\mathbf{s}$, if b is an essential part of a, then $a \in \mathbf{s}$ only if $b \in \mathbf{s}$. This aligns with Perry's core idea: $\mathbf{s}_{Fa}$ plausibly includes the essential parts of a, for it is plausible that these objects must be part of any situation of which a is a part. In this case, so long as Maine is an essential part of New England, we have that **Maine** $\subseteq$ **New England** and "Maine experiences cold winters" is entirely about New England.

However, Perry's theory invites more serious objections.

Objection. Constraint 5 is violated. Suppose that $a \neq b$ (and a is not an essential part of b, or vice versa). On Perry's theory, $Fa \vee Fb$ is associated with the empty set, and so is about every topic (since {} is contained in every set). To see the former, note that $Fa \vee Fb$ is verified by a minimal situation where a has property F, but such a situation does not have b as a part. Similarly, $Fa \vee Fb$ is verified by a minimal situation where b has property F, but such a situation does not have a as a part. Thus, no object is part of every situation that verifies $Fa \vee Fb$.

Objection. Constraint 6 is violated. On Perry's view, "Tully fell out of bed and Tully is Tully" has the same subject matter as "Tully fell out of bed": namely, {Tully}.

Objection. Constraint 9 is violated. For according to this theory, $Fa \wedge \neg Fa$ is about every subject matter. For there is no situation that verifies $Fa \wedge \neg Fa$, and so it is vacuously true (for arbitrary object a) that a is part of every situation that verifies $Fa \wedge \neg Fa$. This leaves Perry with a dilemma. If it is allowed that proper classes can count as subject matters, then $\mathbf{s}_{Fa \wedge \neg Fa}$ is the proper class of all objects, and so $Fa \wedge \neg Fa$ is (partly) about everything. On the other hand, if proper classes are excluded, then there is no such thing as $\mathbf{s}_{Fa \wedge \neg Fa}$, and so $Fa \wedge \neg Fa$ is about nothing.

Objection. Constraint 1 is violated, for, according to Perry's theory, it is not always the case that if φ is about $\mathbf{s}$ then $\neg\varphi$ is about $\mathbf{s}$. For instance: $Fa \vee \neg Fa$ is entirely about $\mathbf{a}$, but $\neg(Fa \vee \neg Fa)$ is not entirely about $\mathbf{a}$ (as in the previous objection, it is either about everything or about nothing).

More fundamentally, consider an objection that applies across the board to theories using the subject-predicate conception.

Objection. Constraint 10 is violated. According to the current conception, aRb and bRa have exactly the same subject matter.

2.4.2 The atom-based conception

In general, a theory along the atom-based conception proceeds as follows: fix a set (or class) $\mathbf{u}$ of distinguished objects (the *universe*). Then: a subject matter $\mathbf{s}$ is any non-empty subset of $\mathbf{u}$. We define inclusion $\leq$ as the subset relation and $\bowtie$ as non-empty intersection. In general, we leave the combination operation $+$ unspecified. Let $\mathbf{T}$ be a *topic function* that assigns a subject matter to every atomic claim p. Then, for an arbitrary sentence φ, the subject matter of φ is just the combination of the subject matters of the atoms in φ, relative to $\mathbf{T}$. A sentence is about any subject matter that includes *the* subject matter of the sentence.

A *particular* theory along this line depends on $\mathbf{u}$, an account of $+$ and any additional constraints on $\mathbf{T}$. In this section, we consider two theories that take $+$ simply as *set union.* (Section 2.6 will present an atom-based theory with a different account of $+$.)

A basic version

A simple atom-based theory is as follows (I resist attributing it to anyone in particular, though this approach seems to me "in the air" in the relevant logic literature). The universe **u** is the class of all sets of possible worlds. In other words, a member of **u** can be viewed as an *unstructured proposition* or a *truth set* or a *piece of information*. Thus, a subject matter **s** is a set of pieces of information. Then, we refine **T** as follows: each atom p is assigned $\{P\}$, where P is just the truth set of p i.e. those worlds at which p is true (thus, this approach assumes that a truth relation is already defined, in the usual manner). Then, for instance, the subject matter for $p \wedge (q \vee r)$ is just $\{P, Q, R\}$, and the subject matter for $p \vee (\neg p \wedge q)$ is $\{P, Q\}$. And so on.

There is also a natural account of **a**, the subject matter associated with object a: let's say that an unstructured proposition P *concerns* a if there are no two worlds w_1 and w_2 such that a is exactly alike in those two worlds but $w_1 \in P$ and $w_2 \notin P$. Then, let **a** be the set of all unstructured propositions that concern a.[21]

This theory goes a long way towards meeting our constraints. Since subject matter is invariant under negation and treats $\vee$ and $\wedge$ symmetrically when it comes to combination, the current account satisfies strong compositionality principles, and so meets constraints 1 through 5. Further, the current account allows for subject matter to contribute a hyper-intensional dimension to meaning, partly satisfying constraints 6 through 9. For instance, the subject matter of $p \wedge (q \wedge \neg q)$ is $\{P, Q\}$, which differs from that for p (namely, $\{P\}$). Further, since the truth sets for aRb and bRa are different, so too are their subject matters, satisfying constraint 10. Further, the account captures basic part-whole intuitions: presumably any difference in how things are for Maine constitutes a difference in how things are for New England. Thus, the set of unstructured propositions that concern New England contains those that concern Maine. Hence, **Maine** $\leq$ **New England**, and so any claim about **Maine** is also about **New England**. Thus, constraints 11 and 12 seem to be accommodated. Finally, since a question can be identified (in many contexts) with a set of unstructured propositions (i.e. the truth sets for the possible *answers* to the question), there is a ready relationship between questions and subject matters on the current view. So constraint 12 is

[21] Cf. the notion of *pw-aboutness* due to Barbara Partee, as discussed in the afterword of Perry [1989].

satisfied.

Objection. Constraints 8 and 9 are violated. To see this, note that the truth set of "2 is even" is W, the set of all possible worlds. Likewise, the truth set of "Jones is Jones" is W. Next, note that W concerns every object a, since it is vacuously true that there are no two worlds w_1 and w_2 such that a is exactly alike in those two worlds but $w_1 \in W$ and $w_2 \notin W$. But, in this case, the subject matter for "2 is even" and "Jones is Jones" is included in that associated with, say, Abraham Lincoln. Thus, the current view has the consequence that "2 is even" and "Jones is Jones" is about **Abraham Lincoln** (in fact, every object).

Similar remarks may be made about necessary falsehoods such as "2 is odd".

Ryle/Parry

Here is another variation on the atom-based theme, mainly following Parry [1968] (Ryle [1933] suggests a theory along similar lines).[22] Take the universe to be the set of all *concepts*. For precision, we understand concepts in an intensional, Carnapian way: as partial functions from possible worlds to individuals that exist in those worlds. Thus, we distinguish *individual concepts*, which map each world (on which that concept is defined) to a singleton, and *general concepts*, which map at least one world (on which the concept is defined) to a set containing at least two objects. Individual concepts may be associated with individual objects, and pieces of language that denote individual objects. General concepts may be associated with properties, and predicates that denote properties. (For simplicity, we ignore relations. The more general picture is what one would expect e.g. the concept $\mathfrak{R}$ associated with n-ary relation R is a partial function from worlds to *n-tuples* of objects.)

It is natural to think of a subject matter as a set of concepts. For instance, mathematics may be thought of the set of all mathematical concepts e.g. the concept of the number 2, the property of being even etc. What of the subject matter $\mathbf{a}$ associated with object a? The natural proposal: $\mathbf{a} = \{\mathfrak{a}\}$, where $\mathfrak{a}$ rigidly designates a. Note that, on this view, Fa is not *entirely* about $\mathbf{a}$: rather it is *partly* about $\mathbf{a}$, since $\{\mathfrak{a}, \mathfrak{F}\}$ overlaps with $\{\mathfrak{a}\}$. This is how it should be (the current theorist can cheerfully say that "Rex is a dog" is partly about **Rex** and partly about **doghood**).

[22]Also see Fine [1986], Parry [1989] and [Burgess, 2009, Ch.5].

We stipulate the nature of the topic assignment $\mathbf{T}$: given atom Fa, $\mathbf{T}(Fa)$ is the set that contains the individual concept associated with the denotation of a, and the general concept associated with the denotation of F. Hence, on the current view, the subject matter of $Fa \vee (\neg Fa \wedge Gb)$ is $\{\mathfrak{a}, \mathfrak{b}, \mathfrak{F}, \mathfrak{G}\}$, where each member is a concept, as described.

This view has many of the advantages of the previous version, but is better positioned to handle constraints 8 and 9. On the current view, the subject matter of "2 is even" is $\{\mathfrak{2}, \mathfrak{even}\}$, and the subject matter of "Jones is Jones" is $\{\mathfrak{Jones}, \mathfrak{Id}\}$. Neither set seems to be included in the subject matter **Abraham Lincoln**, though the former is presumably included **2** and the latter in **Jones**.

Objection. Constraint 10 is violated. On the current view, aRb and bRa have the same subject matter: $\{\mathfrak{a}, \mathfrak{b}, \mathfrak{R}\}$.[23]

Objection. Constraint 11 is violated. (And how to recover it without resorting to ad hocery or artificiality?) Since Maine is not identical to New England, the concept $\mathfrak{m}$ that rigidly designates Maine is not identical to the concept $\mathfrak{ne}$ that rigidly designates New England. Thus, "Maine experiences cold winters" has the subject matter $\{\mathfrak{C}, \mathfrak{m}\}$, which is partly about **Maine** but not partly about **New England**.

(Combining the strengths of our two atom-based variations might strike one as a simple matter: replace the unstructured propositions in the first theory with structures of concepts, drawing on the second theory. To foreshadow, note that this is *exactly* the move that we will exploit for the issue-based theory.)

2.4.3 The way-based conception

According to the way-based conception, a subject matter is a comprehensive set of *ways things can be.* By a "way things can be", I mean an unstructured proposition i.e. a set of possible worlds. Intuitively, these are the worlds that exemplify the way in question. Thus, a singleton $\{w\}$ may be described as a *total* ways things can be. Thus, *way* is for us just an alternative term for *unstructured proposition*, *truth set* or *piece of information.* Further, by "comprehensive" I mean a set of ways $\mathbf{W}$ that *cover* the whole of logical space. That is: the union of all the members in $\mathbf{W}$ is equal to W, the set of all possible worlds. Intuitively, a comprehensive set of ways $\mathbf{W}$ has a way for every possible world: for every

[23]Cf. the criticism of Ryle [1933] in chapter 2 of Yablo [2014].

possible world w, **W** classifies w as being some way or other. (For technical convenience, we stipulate that the empty set $\emptyset$ is a member of every covering. We will not bother to list this element in examples.)

Consider an example from Lewis [1988b]: **the 17th century**. Intuitively, a proposition is a member of **the 17th century** just in case it captures one way for the 17th century to be, out of the comprehensive set of such possible states. (What of the set of worlds where the 17th century does not exist? The simplest maneuver is to count these as constituting one way for the 17th century to be, though this is an awkward usage of the terminology.)

We add a caveat: we exclude the trivial covering $\{W\}$ from the class of subject matters. The reason will be evident shortly.

The way-based conception is driven by the following intuitions: one can classify possible worlds according to any number of distinctions. A subject matter is a system of such distinctions - a way of focusing on certain distinctions, and ignoring others. Thus, one may speak of a way things are *relative to a subject matter* (i.e. relative to the distinctions at issue, and ignoring other possible distinctions that could be drawn). On the way-based conception, we simply *identify* a subject matter with its associated set of ways. (Note that some of these intuitive ideas will re-emerge in service of our issue-based theory, which represents them more directly manner than the way-based conception.)

Inclusion is *not* defined as the subset relation on the way-based conception. Rather, the intuitive idea is that $\mathbf{s} \leq \mathbf{t}$ just in case $\mathbf{t}$ *refines* $\mathbf{s}$ i.e. offers a refined system for dividing up the possibilities. There are at least two important options for making this precise. One might define $\mathbf{s} \leq \mathbf{t}$ as: every way P in $\mathbf{t}$ is a refinement of some way Q in $\mathbf{s}$, in the sense that P entails (i.e. is a subset of) Q. Or one might define $\mathbf{s} \leq \mathbf{t}$ as: every way Q in $\mathbf{s}$ is refined by some way P in $\mathbf{t}$, in the sense that P entails Q.

Thus we exclude the degenerate covering from the class of subject matters. For $\{W\}$ is refined by *every* subject matter, on both definitions. If included as a subject matter, then, it would follow that every subject matter has a common part, and so every claim φ will be partly about *every* subject matter.

As for the subject matter **a**, associated with an object a: we think of this, again, as the set of ways that things can be for a. There are different possible ways that Abraham Lincoln could be (including a degenerate case: not existing).

Abraham Lincoln is the set of all such ways.

As usual, a sentence φ is (somehow) associated with a subject matter $\mathbf{s}_\varphi$. φ is entirely about $\mathbf{s}$ just in case $\mathbf{s}_\varphi$ is included in (i.e. refined by) $\mathbf{s}$. This neatly captures constraints 11 and 12. Clearly, every way for New England to be entails a way for Maine to be, and every way for Maine to be is entailed by a way for New England to be. Hence, **New England** refines **Maine**. Thus, **Maine** $\leq$ **New England**.

The current approach neatly preserves constraint 12. On standard theoretical developments, a question is associated with its set of answers, which in turn can be represented by a set of unstructured propositions. But this is exactly the sort of entity that a subject matter is, on the current view.

(Actually, we need to be more careful than at first meets the eye. It is natural to take a set of answers as a set of unstructured propositions that are *downward closed* i.e. if P entails Q, and Q is an answer to question $\mathbf{Q}$, then P is also an answer to $\mathbf{Q}$. However, we should *not* insist on this feature for a set of unstructured propositions thought of as capturing ways things can be relative to a fixed system of distinctions. For, in this case, if P entails Q, then P may be a region of logical space that can only be captured with distinctions that go beyond what is needed to mark off the region Q. "John is a bachelor" entails "John is a man". The region of logical space where the latter is true can be marked off by focusing only on the distinction between John being a man and not. On the other hand, marking off the region where the former is true requires that we also focus on the distinction between John being married or not.)

Lewis

Lewis [1988b] develops the way-based conception as follows: a subject matter is taken to be a *partition* on the space of possible worlds i.e. a set of mutually disjoint and exhaustive unstructured propositions. $\mathbf{s}$ includes $\mathbf{t}$ just in case $\mathbf{s}$ refines $\mathbf{t}$, where this means that every $P \in \mathbf{t}$ is equal to a union of members of $\mathbf{s}$ (it suffices for Lewis, then, to use the first definition of inclusion mentioned above). $\mathbf{s}_\varphi$ is the binary partition consisting of the truth set of φ and its complement.

Note that this easily satisfies constraint 10, for aRb and bRa have distinct truth sets, and so have distinct subject matters on the current view.

Objection. Constraints 3 and 4 are violated. The root problem is that Lewis' theory entails that the subject matter of φ has no proper parts, and so φ cannot be partly about anything (except its own subject matter). For the subject matter of a sentence is a binary partition, which refines only the degenerate partition $\{W\}$. But we excluded $\{W\}$ from the class of subject matters. Thus, unless they are classically equivalent, the subject matter of two claims can never have a common part.

Objection. Constraints 6 and 7 are violated. For instance, on Lewis' view, $p \wedge (q \vee \neg q)$ has the same subject matter as p, since their truth sets are identical.

Objection. Constraints 8 and 9 are violated. For instance, on Lewis' view, $Fa \vee \neg Fa$ is about every subject matter: for its truth set is W (the set of all possible worlds), and every partition refines $\{W\}$. Furthermore, since $\{W\}$ is not a subject matter, it follows that there is no such thing as *the* subject matter of $Fa \vee \neg Fa$.

Lewis is aware of these difficulties. For instance, Lewis [1988a] explores different conceptions of "partial aboutness" in an effort to land on something fully satisfactory. His proposals encounter many difficulties. From our perspective, we stick with the inviting idea that φ is partially about $\mathbf{s}$ just in case $\mathbf{s}_\varphi$ shares a common part with $\mathbf{s}$.

Further, Lewis notes that his conception requires that necessities such as "2 is even" are about every subject matter, and displays some uneasiness about this result. In response, Lewis [1988b] develops a modification that extends logical space to include *impossible worlds*. He rightly points out, however, that such a maneuver raises difficult and subtle issues, and should be approached with hesitation. As we shall see, the issue-based theory makes no use of impossible situations or worlds.

Yablo

Lewis posits that $\mathbf{s}_\varphi$ divides into two ways things can be: that way according to which φ is true, and that way according to which φ is false. For Yablo [2014], this subject matter is not fine-grained enough. Rather, we should identify $\mathbf{s}_\varphi$ with the basic ways in which φ can be true, and the basic ways in which φ can be false. Altogether: the set of *minimal truthmakers* and *falsemakers* for φ.

More technically, a semantic truthmaker (falsemaker) for φ is associated with a *minimal model* that verifies (falsifies) φ. Conveniently, we can here think of a model as a state description λ: a conjunction of unique literals (a literal being either an atom or the negation of an atom). Thus, $Fa \wedge \neg Fb \wedge Gb$ is a state description, and so a model on the present conception. (More generally, a model in the classical setting is a partial valuation. Cf. section 2.4.1.) Then, λ verifies (falsifies) φ just in case φ ($\neg\varphi$) is a classical implication of λ; and a minimal verifier (minimal falsifier) is one such that there exists no model μ (classically inequivalent to λ) such that λ implies μ and μ implies φ ($\neg\varphi$). Finally, a truthmaker (falsemaker) for φ is an unstructured proposition that is expressed by a model λ that verifies (falsifies) φ. A minimal truthmaker (minimal falsemaker) is expressed by a minimal verifier (minimal falsifier).

For example, consider $p \wedge q$. This formula is itself a state description, and so its sole minimal truthmaker is expressed by itself. On the other hand, two minimal models falsify our formula: $\neg p$ and $\neg q$. Thus, our formula has two minimal falsemakers.

For example, consider $p \vee q$. This formula is verified by two minimal models: p and q. Hence, it has two minimal truthmakers. On the other hand, it is falsified by a unique minimal model: $\neg p \wedge \neg q$, and so has one minimal falsemaker.

Call the set of minimal truthmakers for φ its matter (and denote it $\mathbf{m}(\varphi)$) and the set of minimal falsemakers its anti-matter ($\mathbf{a}(\varphi)$). We then take $\mathbf{s}_\varphi$ to be $\mathbf{m}(\varphi) \cup \mathbf{a}(\varphi)$ on the Yablovian view.[24]

[Yablo, 2014, Ch.3] seems to prefer our second definition of inclusion: $\mathbf{s} \leq \mathbf{t}$ just in case every member of $\mathbf{s}$ is entailed by some member of $\mathbf{t}$. This buys him an advantage over Lewis in satisfying constraint 3: on Yablo's view, $p \wedge q$ has the subject matter $\{P \cap Q, P^c, Q^c\}$, where P is the truth set for p and Q is the truth set for q. p has the subject matter $\{P, P^c\}$. Hence, $\mathbf{s}_p$ is included in $\mathbf{s}_{p\wedge q}$, on Yablo's view. Further, if p is entirely about $\mathbf{a}$, it follows that $p \wedge q$ is partly about $\mathbf{a}$, as desired.

Nevertheless, Yablo's theory does not meet all of our constraints.

Objection. Constraint 2 is violated. On the Yablovian theory, we have $\mathbf{s}_p = \{P, P^c\}$ and $\mathbf{s}_q = \{Q, Q^c\}$. Consider the subject matter $\mathbf{s} = \{P, Q, P^c, Q^c\}$. Since every member of $\mathbf{s}_p$ and every member of $\mathbf{s}_q$ is entailed by some member of $\mathbf{s}$, we have that: $\mathbf{s}_p \leq \mathbf{s}$ and $\mathbf{s}_q \leq \mathbf{s}$. But we also have $\mathbf{s}_{p\vee q} = \{P, Q, P^c \cap Q^c\}$. Since

[24] To preserve more structure, one might instead take $\mathbf{s}_\varphi$ to be the *ordered pair* of the matter and anti-matter of φ. This has an apparent disadvantage: p and $\neg p$ will have different subject matter: $\mathbf{s}_p = \langle p, \neg p\rangle$ and $\mathbf{s}_{\neg p} = \langle \neg p, p\rangle$.

$P^c \cap Q^c$ is not entailed by any member of $\mathbf{s}$, we have $\mathbf{s}_{p\vee q} \not\leq \mathbf{s}$.

Objection. Constraint 4 is violated. Above, we noted that Yablo's view entails that $\mathbf{s}_{p\wedge q} = \{P \cap Q, P^c, Q^c\}$. On the other hand, $\mathbf{s}_{p\vee q} = \{P, Q, P^c \cap Q^c\}$. But it is thus evident that the former does not refine the latter: $P^c \cap Q^c$ is not entailed by any member of $\{P \cap Q, P^c, Q^c\}$.[25]

Objection. Constraints 6 and 7 are violated. On Yablo's view, the subject matter of logically equivalent claims is identical. Thus, in particular, $p \wedge (q \vee \neg q)$ has the same subject matter as p.

Objection. Constraint 8 is violated. Let a be an arbitrary object. On Yablo's view, every atom and its negation is a minimal truth-maker for $p \vee \neg p$. Thus, $\mathbf{s}_{Fa}$ is refined by that of $p \vee \neg p$, on Yablo's view (for Fa and $\neg Fa$ both verify $p \vee \neg p$, and so every minimal truthmaker and falsemaker for Fa is trivially entailed by a member of $\mathbf{s}_{p\vee\neg p}$). On the assumption that Fa is entirely about $\mathbf{a}$, it follows that $p \vee \neg p$ is partly about $\mathbf{a}$.

Our third objection to Yablo [2014] indicates that his theory cannot account for the hyper-intensional consequences of subject matter. Yablo himself seems to consider this a cost in accepting his view (see [Yablo, 2014, sect. 4.4]). On the whole, he wavers between accepting the above theory and a variation on his views based on the work of van Fraassen [1969] (see section 2.4.3).

Van Fraassen-Yablo

Finally, we consider a variation of Yablo's ideas.[26] Yablo, recall, identifies the subject matter of φ with a distinguished subset of truthmakers and falsemakers. Yablo uses *minimality* as the criterion for membership. There are alternatives, however. One, following van Fraassen [1969], builds up a set of distinguished truthmakers ($\mathbf{m}(\varphi)$) and falsemakers ($\mathbf{a}(\varphi)$) for φ in a *recursive* fashion (cf. the account of verification and falsification in section 2.4.1):

- $\mathbf{m}(p) = \{P, \emptyset\}$ and $\mathbf{a}(p) = \{P^c, \emptyset\}$, where P is the truth set for p
- $\mathbf{m}(\neg\varphi) = \mathbf{a}(\varphi)$ and $\mathbf{a}(\neg\varphi) = \mathbf{m}(\varphi)$

[25] Yablo effectively notices this in section 12 of the appendix for Yablo [2014], available at: http://www.mit.edu/∼yablo. The same fact is directed at Yablo's theory as a criticism in [Holliday, 2013, sect. 6.2.1].

[26] See [Yablo, 2014, sect. 4.2]

- $\mathbf{m}(\varphi \wedge \psi) = \{Q \cap R : Q \in \mathbf{m}(\varphi) \text{ and } R \in \mathbf{m}(\psi)\}$ and $\mathbf{a}(\varphi \wedge \psi) = \mathbf{a}(\varphi) \cup \mathbf{a}(\psi)$
- $\mathbf{m}(\varphi \vee \psi) = \mathbf{m}(\varphi) \cup \mathbf{m}(\psi)$ and $\mathbf{a}(\varphi \vee \psi) = \{Q \cap R : Q \in \mathbf{a}(\varphi) \text{ and } R \in \mathbf{a}(\psi)\}$

As with Yablo, we use the following definition for inclusion: $\mathbf{s} \leq \mathbf{t}$ just in case every member of $\mathbf{s}$ is entailed by some member of $\mathbf{t}$. Again, we use $\mathbf{s}_\varphi = \mathbf{m}(\varphi) \cup \mathbf{a}(\varphi)$.

The resulting account has an advantage over Yablo's in terms of accounting for hyper-intensionality, and so in meeting constraints 6 and 7. For on the current theory, p and $p \wedge (q \vee \neg q)$ have different subject matter: $\mathbf{s}_p = \{P, P^c\}$ and $\mathbf{s}_{p \wedge (q \vee \neg q)} = \{P \cap Q, P \cap Q^c, P^c\}$. Further, $\mathbf{s}_p$ is included in $\mathbf{s}_{p \wedge (q \vee \neg q)}$, so $p \wedge (q \vee \neg q)$ is partly about the subject matter of p, as desired.

Further, the current account makes a better show of observing constraint 8: Yablo's theory had the consequence that the truth set for every atom and its negation was part of the subject matter of $p \vee \neg p$. On the current account, $\mathbf{s}_{p \vee \neg p} = \{P, P^c\} = \mathbf{s}_p$.

However, once again, the proposal does not meet all of our desiderata (for similar reasons as Yablo's account.)

Objection. Constraint 2 is violated, as can be seen with exactly the same counter-example as for Yablo's earlier theory.

Objection. Constraint 4 is violated. Note that $\mathbf{s}_{p \wedge q} = \{P \cap Q, P^c, Q^c\}$. On the other hand, $\mathbf{s}_{p \vee q} = \{P, Q, P^c \cap Q^c\}$. But it is thus evident that the former does not refine the latter: $P^c \cap Q^c$ is not entailed by any member of $\{P \cap Q, P^c, Q^c\}$.

Objection. Constraint 6 is violated. On the current account, $\mathbf{s}_{p \wedge (b=b)} = \{P, P^c\} = \mathbf{s}_p$, since $\mathbf{s}_{b=b} = \{W\}$ where W is the set of all worlds.

2.4.4 Final Tally

Here is a summary of the authors surveyed and the constraints they violate.

- *Perry (subject-predicate):* 1, 5, 6, 9, 10.
- *Parry (atom-based):* 10, 11.
- *Lewis (way-based):* 3, 4 ,6, 7, 8, 9.
- *Yablo (way-based):* 2, 4, 6, 7, 8.

- *van Fraassen-Yablo (way-based):* 2, 4, 6.

We also note that the way-based positions offer an *especially* natural account of constraint 12.

2.5 Positive proposal: the issue-based theory

I now propose a version of the atom-based approach - the issue-based theory - that meets all of our constraints. I emphasize that it tweaks but does not majorly depart from the form of the basic atom-based theories in section 2.4.2. Further, its intuitive rationale is effectively that of the way-based conception (though I propose that it elaborates on these intuitions more straightforwardly). Since these features have proved appealing to various authors, I consider this promising.

In the next two sections, I describe the main tenets of the issue-based theory, leaving some details open. In section 2.5.3, I offer a simplistic but illustrative elaboration of the theory.

2.5.1 The basic proposal

A subject matter, intuitively, is a system of *distinctions.* For instance, the purpose of a *discourse* topic is to focus conversation on certain distinctions, and allow others to recede from view. On this picture, the relationship between questions and topics is intuitively evident. A distinction is associated with a basic *issue*: is the world that way, or not? Thus, a system of distinctions is associated with a system of issues. To resolve each distinction is to answer each associated question, allowing for a *complete* answer to the questions in focus. Hence the notion of *way things can be relative to that subject matter:* each such way is a complete answer that decides every distinction at issue.

A high-level technical elaboration of this picture is as follows. We understand *concepts* as in section 2.4.2: as Carnapian intensions. For simplicity, we assume that every object exists at every world, and ignore the necessity of dividing objects into types. An *individual concept* maps each world to an object. An individual concept that maps each world to the same object is a *rigid designator.* Otherwise, it is a *role.* An n-ary general concept maps each world to a set of tuples of length n. We take 1-ary general concepts to simply map to sets of objects.

Think of a rigid designator as the semantic value of a name; a role as the semantic value of a definite description; and a general concept as the semantic value of a predicate.

A *distinction* (or *issue*) is a tuple of concepts

$$\langle \mathfrak{R}, \mathfrak{o}_1, \ldots, \mathfrak{o}_n \rangle$$

where $\mathfrak{R}$ is an n-ary general concept and each $\mathfrak{o}_i$ is an individual concept.[27] Note that possible worlds *decide* distinctions: given world w and issue $\langle \mathfrak{R}, \mathfrak{o}_1, \ldots, \mathfrak{o}_n \rangle$, it is either the case that

$$\langle \mathfrak{o}_1(w), \ldots, \mathfrak{o}_n(w) \rangle \in \mathfrak{R}(w)$$

or

$$\langle \mathfrak{o}_1(w), \ldots, \mathfrak{o}_n(w) \rangle \notin \mathfrak{R}(w)$$

Thus, a distinction is a means of dividing the space of possible worlds into those where a certain set of objects stand in a certain relation to each other, and those where those same objects do not.

A *subject matter* is a non-empty set of distinctions/issues. Subject matter $\mathbf{s}$ is included in $\mathbf{t}$ when $\mathbf{s} \subseteq \mathbf{t}$ i.e. $\mathbf{t}$ involves the same distinctions as $\mathbf{s}$ and possibly more. Thus, $\mathbf{s} \bowtie \mathbf{t}$ means that $\mathbf{s}$ and $\mathbf{t}$ have non-empty intersection.

Given object a, we say that an issue $\langle \mathfrak{R}, \mathfrak{o}_1, \ldots, \mathfrak{o}_n \rangle$ *concerns a* exactly when there is an i such that $\mathfrak{o}_i$ is a rigid designator that maps to a part of a (possibly a itself). Then: $\mathbf{a}$ is the set of all distinctions that concern a.

Subject matter combination $+$ is just set union.

2.5.2 Resolution, truth and subject predication

A set of issues $\mathbf{s}$ generates a partition of unstructured propositions on logical space. As in the way-based conception, these are best thought of as *ways things can be with respect to* $\mathbf{s}$. In light of the idea that a subject matter focuses on certain distinctions at the expense of others, we call this partition the *resolution generated by* $\mathbf{s}$. Metaphorically, $\mathbf{s}$ divides logical space into contrasting basic possibilities at a certain grain of detail.

[27] Cf. the discussion of issues in Perry [1989].

We fix our partition with an equivalence relation. Two worlds u and v are equivalent with respect to $\mathbf{s}$ (written $u \approxeq v$) just in case: for every issue $\langle \mathfrak{R}, \mathfrak{o}_1, \ldots, \mathfrak{o}_n \rangle$ in $\mathbf{s}$ we have either both

$$\langle \mathfrak{o}_1(u), \ldots, \mathfrak{o}_n(u) \rangle \in \mathfrak{R}(u) \text{ and } \langle \mathfrak{o}_1(v), \ldots, \mathfrak{o}_n(v) \rangle \in \mathfrak{R}(v)$$

or both

$$\langle \mathfrak{o}_1(u), \ldots, \mathfrak{o}_n(u) \rangle \notin \mathfrak{R}(u) \text{ and } \langle \mathfrak{o}_1(v), \ldots, \mathfrak{o}_n(v) \rangle \notin \mathfrak{R}(v)$$

That is: $u \approxeq v$ indicates that u and v decide every issue in $\mathbf{s}$ in exactly the same way.

That is: our issue-based theory allows for the generation of a way-based theory, as a convenient abstraction. (Of course, a way-based theory generated in this manner will satisfy certain constraints, so it is not the case that *every* way-based theory can be generated in this way.)

We turn to the subject matter of a *claim*, and its relationship to that claim's truth conditions. As usual, we assume that every well-formed descriptive sentence φ is associated with a subject matter $\mathbf{s}_\varphi$ as part of its meaning in discourse. In particular, we allow that an atomic claim p can be associated with a complex subject matter i.e. potentially with more than one distinction. We leave open to what extent the subject matter of atoms is a semantic or pragmatic fact (I am inclined to think it is at least partly semantic, though that the context can pragmatically *enrich* the subject matter of an atom e.g. relative to the question under discussion).

In line with the atom-based conception, the subject matter of logically complex expressions is constrained as follows:

- $\mathbf{s}_{\neg\varphi} = \mathbf{s}_\varphi$
- $\mathbf{s}_{\varphi\wedge\psi} = \mathbf{s}_{\varphi\vee\psi} = \mathbf{s}_\varphi + \mathbf{s}_\psi$

Unlike the way-based theories we surveyed, we do not assume that the subject matter of φ is determined by its truth conditions. Rather, the subject matter of $\mathbf{s}$ *constrains* the truth conditions of φ as follows. We say that an unstructured proposition P is *at the resolution of* $\boldsymbol{s}$ just in case P is identical to a union of members of the resolution generated by $\mathbf{s}$. Likewise, we say that an interpreted sentence φ is at the resolution of φ just in case its truth set is at that resolution. Now, we impose the constraint that the truth set of φ must always be at the resolution of $\mathbf{s}_\varphi$. Informally: what a claim *says* must be something *about its subject matter*.

The following result is simple to prove using induction on the structure of formulas.

2.5.1. Proposition. *Given the above constraints on $\neg$, $\wedge$ and $\vee$: the property of φ being at the resolution of its subject matter is preserved under the application of the operations of propositional logic.*

Finally, note that there is a natural way to generate a subject-predicate theory from our issue-based theory. Consider claim φ and its associated set of distinctions $\mathbf{s}_\varphi$. Now, we say that φ *predicates relation R of objects $o_1, \ldots, o_n$* just in case: (i) the distinction $\langle \mathfrak{R}, \mathfrak{o}_1, \ldots, \mathfrak{o}_n \rangle$ is in $\mathbf{s}_\varphi$, with $\mathfrak{R}$ the general concept associated with R and $\mathfrak{o}_i$ the individual concept associated with o_i; (ii) $\langle \mathfrak{o}_1(w), \ldots, \mathfrak{o}_n(w) \rangle \in \mathfrak{R}(w)$ for every w in the truth set of φ.

This accommodates various intuitions. For instance, φ, despite saying something non-trivial about a non-trivial subject matter, might fail to predicate any property/relation. For instance, $Fa \vee Gb$ does not predicate any property. This has appeal: the claim "Jane is a lawyer or Joe is an accountant" says something informative about non-trivial topics (e.g. the professionals that we've met), but one might disagree that it predicates any property of Jane or Joe. It is intuitively *non-committal* about Jane's status and Joe's status.

2.5.3 A toy framework

We now develop the theory explicitly, allowing ourselves the luxury of a simplified framework.

Consider a language $\mathcal{L}$ constructed from one-place predications (e.g. Fa), two-place predications (e.g. aRb) and identity statements ($a = a$) using the logical connectives $\wedge$, $\vee$ and $\neg$.

As a model $\mathcal{M}$ for this language, we fix a set of worlds W called *logical space*; a domain of objects O (considered invariant across worlds) equipped with a transitive, reflexive, anti-symmetric part-hood relation $\preceq$; an assignment function α that a maps each constant symbol a in the language to an individual concept

$\mathfrak{a}$, each predicate symbol F to a general concept $\mathfrak{F}$ and $=$ to the general concept $\mathfrak{Id}$ that maps each world to the set of all pairs of objects; and a topic assignment T that assigns a set of issues to each sentence φ. For simplicity, we assume that no two constants or predicates map to the same concept.

(Note that this model associates a distinction $\langle \mathfrak{F}, \mathfrak{a} \rangle$ with each one-place predication Fa. Likewise, for two-place predications and identity statements.)

Then the truth set $|\varphi|$ (relative to $\mathcal{M}$) for each φ is as follows:

- $w \in |Fa|$ iff $\mathfrak{a}(w) \in \mathfrak{F}(w)$
- $w \in |aRb|$ iff $\langle \mathfrak{a}(w), \mathfrak{b}(w) \rangle \in \mathfrak{R}(w)$
- $w \in |a = b|$ iff $\langle \mathfrak{a}(w), \mathfrak{b}(w) \rangle \in \mathfrak{Id}(w)$
- $w \in |\neg\varphi|$ iff $w \notin |\varphi|$
- $|\varphi \wedge \psi| = |\varphi| \cap |\psi|$
- $|\varphi \vee \psi| = |\varphi| \cup |\psi|$

Further, T obeys:

- $\mathsf{T}(Fa) = \{\langle \mathfrak{F}, \mathfrak{a} \rangle\}$
- $\mathsf{T}(aRb) = \{\langle \mathfrak{R}, \mathfrak{a}, \mathfrak{b} \rangle\}$
- $\mathsf{T}(a = b) = \{\langle \mathfrak{Id}, \mathfrak{a}, \mathfrak{b} \rangle\}$
- $\mathsf{T}(\neg\varphi) = \mathsf{T}(\varphi)$
- $\mathsf{T}(\varphi \wedge \psi) = \mathsf{T}(\varphi \vee \psi) = \mathsf{T}(\varphi) \cup \mathsf{T}(\psi)$

Since each atomic predication corresponds to a unique issue (via assignment function α) we might as well have represented things as follows: T maps from a sentence φ to a set of atomic predications (or an identity statement). For example: $\mathsf{T}(Fa) = \{Fa\}$; $\mathsf{T}(aRb) = \{aRb\}$; $\mathsf{T}(\neg Fa \vee a = c) = \{Fa, a = c\}$; $\mathsf{T}(\neg(aRb \wedge bRa)) = \{aRb, bRa\}$.

For every constant a, the subject matter $\mathbf{a}$ is the set of all distinctions that include $\mathfrak{b}$, where b designates a part of the object designated by a.

The following is again easy to prove.

2.5.2. PROPOSITION. *For every $\varphi \in \mathcal{L}$, the truth set $|\varphi|$ is at the resolution of $\mathsf{T}(\varphi)$.*

As for an example of a resolution of a subject matter: consider subject matter $\mathbf{s} = \{Fa, Ga, Gb\}$. The resolution of $\mathbf{s}$ is expressed by the Carnapian state descriptions built from $\mathbf{s}$. That is, each cell in the resolution is expressed by one of: $Fa \wedge Ga \wedge Gb$; $Fa \wedge Ga \wedge \neg Gb$; $Fa \wedge \neg Ga \wedge Gb$; $Fa \wedge \neg Ga \wedge \neg Gb$; $\neg Fa \wedge Ga \wedge Gb$; $\neg Fa \wedge Ga \wedge \neg Gb$; $\neg Fa \wedge \neg Ga \wedge Gb$; $\neg Fa \wedge \neg Ga \wedge \neg Gb$.

2.5.4 Constraints met

I next argue in detail that the issue-based theory meets our criteria of adequacy.

The roots of this success can be appreciated without the details. Since the issue-based theory ensures that negation does not affect subject matter and that $\wedge$ and $\vee$ combine subject matter in a uniform manner, the theory preserves the intuitive interaction between the connectives and subject matter. Since the theory treats subject matters as composed from *structured* tuples of concepts, it captures the intuition that the structure of a claim affects its subject matter (not only what its parts denote). Since the theory provides a straightforward (set theoretic) account of inclusion and overlap, and an intuitive account of the subject matter $\mathbf{a}$ relative to object a, it neatly captures the intuition that claims about a part are also about the whole. Finally, since there is both a close connection between the notion of a distinction and that of a basic, binary question (both are naturally described with the term 'issue'), the theory draws a close connection between questions and topics.

Now for details. For definiteness, we work with our toy framework.

1. If φ is entirely about $\mathbf{s}$ then $\neg\varphi$ is entirely about $\mathbf{s}$. *Proof.* Suppose that $\mathbf{s}_\varphi \subseteq \mathbf{s}$. We know that $\mathbf{s}_{\neg\varphi} = \mathbf{s}_\varphi$. Hence: $\mathbf{s}_{\neg\varphi} \subseteq \mathbf{s}$.

2. If φ is entirely about $\mathbf{s}$ and ψ is entirely about $\mathbf{s}$ then $\varphi \vee \psi$ is entirely about $\mathbf{s}$. *Proof.* Suppose that $\mathbf{s}_\varphi \subseteq \mathbf{s}$ and $\mathbf{s}_\psi \subseteq \mathbf{s}$. Now, $\mathbf{s}_{\varphi\vee\psi} = \mathbf{s}_\varphi + \mathbf{s}_\psi = \mathbf{s}_\varphi \cup \mathbf{s}_\psi$. Hence, $\mathbf{s}_{\varphi\vee\psi} \subseteq \mathbf{s}$.

3. If φ is entirely about $\mathbf{s}$ then $\varphi \wedge \psi$ is partly about $\mathbf{s}$. *Proof.* Assume $\mathbf{s}_\varphi \subseteq \mathbf{s}$. Now, $\mathbf{s}_{\varphi\wedge\psi} = \mathbf{s}_\varphi \cup \mathbf{s}_\psi$. Since $\mathbf{s}_\varphi$ is non-empty, it follows that $\mathbf{s}_{\varphi\wedge\psi} \cap \mathbf{s} \neq \emptyset$.

4. The subject matter of $p \wedge q$ includes the subject matter of $p \vee q$. *Proof.* $\mathbf{s}_{p\wedge q} = \mathbf{s}_{p\vee q}$.

5. Some disjunctive expressions $Fa \vee Gb$ are about something and some such expressions are not about everything. *Proof.* $\mathbf{s}_{Fa} = \{Fa\}$ and $\mathbf{s}_{Gb} = \{Gb\}$. Thus, $\mathbf{s}_{Fa\vee Gb} = \{Fa, Gb\} \neq \emptyset$. Further, $\{Fa, Gb\} \nsubseteq \{Fb\}$.

6. If Fa is contingent, then it has different subject matter to $Fa \wedge (b = b)$ and $Fa \wedge (Gb \vee \neg Gb)$. *Proof.* In the setting of our toy model, the qualification of contingency is not necessary. At any rate: $\mathbf{s}_{Fa} = \{Fa\}$. Contrast this to: $\mathbf{s}_{Fa\wedge(b=b)} = \{Fa, b = b\}$ and $\mathbf{s}_{Fa\wedge(Gb\vee\neg Gb)} = \{Fa, Gb\}$.

7. If Fa is contingent, then it has different subject matter to $Fa \vee (b \neq b)$ and $Fa \vee (Gb \wedge \neg Gb)$ *Proof.* Similar to the last constraint.

8. A claim of the form $Fa \vee \neg Fa$ is about something (e.g. $\mathbf{a}$) but not about everything (at least if Fa is about something but not everything). Likewise for $a = a$ and most cases of Fa where Fa is necessary. *Proof.* Suppose that Fa is about something but not about everything. In the context of our toy model, this amounts to the assumption that a is not a part of every object b (for, otherwise, it would follow that $\mathbf{a} \leq \mathbf{b}$, for every object b. Thus, if $\mathbf{s}_{Fa} \leq \mathbf{a}$, then $\mathbf{s}_{Fa} \leq \mathbf{b}$). Thus, suppose that b is such that a is not a part of it. It follows that $Fa \notin \mathbf{b}$. Thus, $\mathbf{s}_{Fa\vee\neg Fa} = \{Fa\} \nsubseteq \mathbf{b}$.

9. A claim of the form $Fa \wedge \neg Fa$ is about something (e.g. $\mathbf{a}$) but not about everything (at least if Fa is about something but not everything). Likewise for $a \neq a$ and most cases of $\neg Fa$ where Fa is necessary. *Proof.* Similar to the previous constraint.

10. Expressions of the form aRb and bRa are not necessarily about the same topic. Nor is the subject matter of Fa necessarily identical to that of an expression of the form Ga. *Proof.* According to our toy model: $\mathbf{s}_{aRb} = \{aRb\} \neq \{bRa\} = \mathbf{s}_{bRa}$, if $a \neq b$.

11. If a is part of b, then: if φ is about $\mathbf{a}$, then φ is about $\mathbf{b}$. *Proof.* This follows from our definition of $\mathbf{a}$ and $\mathbf{b}$: since $\mathbf{b}$ contains every distinction concerning a part of b, and the parthood relation is transitive, it follows that $\mathbf{a} \subseteq \mathbf{b}$.

12. A question **Q** can always (in some sense) serve as a subject matter. *Rationale.* Intuitively, a question sets up a system of distinctions/issues. One asks: "how many stars are there?" This generates a set of distinctions: there are no stars (or some stars); there is exactly one star (or not); there are exactly two stars (or not); and so on. Or consider: "who came to the party?" This generates the distinctions: Joe came to the party (or didn't); Jane came to the party (or didn't); and so on. To settle some but not all of these issues to provide a *partial answer* to the question. A *complete answer* decides every issue and corresponds to a cell in the resolution of the associated subject matter.[28]

I conclude that the issue-based theory finds favor over the aforementioned rivals.

2.6 Coda: Fine's state-based theory

A theory of subject matter due to Kit Fine presents a special challenge for the issue-based theory. I outline Fine's theory, drawing mainly on Fine [2016] and Fine [2017b],[29] then note that it too meets the criteria of adequacy in section 2.3. We then observe that Fine's theory can be formulated as an atom-based theory and that, given relatively mild assumptions, a version of the issue-based theory can be recovered from Fine's theory. Hence, to some extent, both can be embraced. Nevertheless, I offer tentative reasons for breaking the tie in favor of the issue-based theory.

2.6.1 Fine's theory of subject matter

Fine's theory embeds into his *truthmaker semantics.* Like Perry [1989], he offers a formal account of when a situation verifies or falsifies a sentence, but with two crucial differences: he makes room for *impossible situations* and provides semantic clauses in the style of van Fraassen [1969]. Following Fine, we label the associated relations *exact verification* (denoted $\vdash$) and *exact falsification* (denoted $\dashv$).

Let Σ be a set of *states* partially ordered by (transitive, reflexive, anti-symmetric) part-hood relation $\sqsubseteq$. We leave open the possibility that some such states are

[28] Cf. the account of questions in Roberts [2012].

[29] Also see Fine [2014] and Fine [2017a].

properly described as *possible* and some such are described as *impossible*. We assume that every subset of states $A \subseteq \Sigma$ has a *fusion*, denoted by $\bigsqcup A$ (or $\sigma \sqcup \tau$ when $A = \{\sigma, \tau\}$). Mathematically, $\bigsqcup A$ is the lowest upper bound for A, relative to ordering $\sqsubseteq$. Conceptually, we think of $\sqcup A$ as the situation that results from fusing together the possible 'chunks of reality' that compose the members of A into a 'larger chunk' with the members of A as parts. This might result in an impossible situation: fusing a situation where John is a cat with a situation where John is not a cat results in an impossible situation in which John is both a cat and not. Finally, we assume, as usual, a background domain D of (actual or merely possible) objects, ordered by a part-hood relation $\preceq$.

We again work with a simplified propositional language $\mathcal{L}$ - in particular, we restrict the atomic claims to one-place predications Fa. As our semantic primitives, we have a *truthmaker assignment* $\mathfrak{t}$ and *falsemaker assignment* $\mathfrak{f}$, each of which maps each atomic claim in the language to a set of states in Σ. The semantic clauses are then as follows:

- $\sigma \vdash Fa$ iff $\sigma \in \mathfrak{t}(Fa)$. $\sigma \dashv Fa$ iff $\sigma \in \mathfrak{f}(Fa)$.
- $\sigma \vdash \neg\varphi$ iff $\sigma \dashv \varphi$. $\sigma \dashv \neg\varphi$ iff $\sigma \vdash \varphi$.
- $\sigma \vdash \varphi \wedge \psi$ iff there exist states τ, υ such that

  $$\sigma = \tau \sqcup \upsilon \text{ and } \tau \vdash \varphi \text{ and } \upsilon \vdash \psi$$

  $\sigma \dashv \varphi \wedge \psi$ iff either $\sigma \dashv \varphi$ or $\sigma \dashv \psi$.
- $\sigma \vdash \varphi \vee \psi$ iff either $\sigma \vdash \varphi$ or $\sigma \vdash \psi$. $\sigma \dashv \varphi \vee \psi$ iff there exist states τ, υ such that

  $$\sigma = \tau \sqcup \upsilon \text{ and } \tau \dashv \varphi \text{ and } \upsilon \dashv \psi$$

Let $[\varphi]$ denote the union of the set of exact verifiers and set of exact falsifiers for φ.

Now for Fine's basic account of subject matter: the set of subject matters is the set of states. That is: possible or impossible 'chunks of reality' serve as subject matters. The subject matter of expression φ - as usual denoted $\mathbf{s}_\varphi$ - is the fusion of (all of) the exact verifiers and exact falsifers of φ. In particular, $\mathbf{s}_{Fa}$ is the fusion of the truthmakers and falsemakers for Fa i.e. $\bigsqcup[Fa]$. Subject matter combination $+$ is defined as: $\mathbf{s} + \mathbf{t} = \mathbf{s} \sqcup \mathbf{t}$. Finally, subject matter inclusion $\leq$ is defined as: $\mathbf{s} \leq \mathbf{t}$ iff $\mathbf{s} \sqsubseteq \mathbf{t}$.[30]

Fine does not, as far as I know, provide an account of $\mathbf{a}$, the subject matter associated with object a. Here is a natural proposal:

$$\mathbf{a} := \bigsqcup\{\sigma \in \Sigma : \text{there exists } Fa \in \mathcal{L} \text{ s.t. } \sigma \in [Fa]\}$$

It also natural to then impose the following constraint on our model: $a \preceq b$ entails that $\mathbf{a} \sqsubseteq \mathbf{b}$.

2.6.2 Fine's theory as atom-based

The apparent novelty of Fine's theory of subject matter raises two questions. Can we classify it under one of the three conceptions we have explored? Second, does the theory meet our criteria of adequacy? We answer the first question using three simple results.

2.6.1. PROPOSITION. *On Fine's theory:*

$$\mathbf{s}_\varphi = \bigsqcup_{Fa \text{ in } \varphi} [Fa]$$

That is: on Fine's theory, the subject matter of φ is the fusion of the truthmakers and falsemakers of the atoms that appear in φ. The proof is by induction on the structure of φ.

Next, call $A \subseteq \Sigma$ an *ideal* just in case it is closed under parts and fusions. We use A^* to denote the smallest ideal that contains set A. Now, note that every $\sigma \in \Sigma$ can be associated with a unique ideal, called the *principal ideal generated by* σ and denoted $\mathtt{I}[\sigma]$. Namely:

$$\mathtt{I}[\sigma] := \{\sigma\}^* = \{\tau \in \Sigma : \tau \sqsubseteq \sigma\}$$

Furthermore, let I be an arbitrary ideal. Note that $\bigsqcup I$ exists and is a member

[30] Fine [2016], section 5, calls this an account of the *bi-lateral* subject matter of φ. In contrast, the *positive subject matter* of φ is the fusion of its exact verifiers, while its *negative subject matter* is the fusion of its exact falsifiers. While potentially useful technical notions, note that (by our lights) these will not do as general accounts of subject matter, for they potentially assign different subject matter to φ and $\neg\varphi$, in tension with constraint 1.

of I. Thus, $I = \mathtt{I}[\bigsqcup I]$. Hence, there is a one-to-one correspondence between the set of ideals for the state space and Σ. Now, consider a standard result.

2.6.2. PROPOSITION. *Let $\langle \Sigma, \sqsubseteq, \mathfrak{t}, \mathfrak{f} \rangle$ be a state space s.t. every subset of Σ has a fusion. Consider $\sigma, \tau, \mu \in \Sigma$. Then:*

1. $\sigma \sqsubseteq \tau$ *iff* $\mathtt{I}[\sigma] \subseteq \mathtt{I}[\tau]$

2. $\sigma \sqcup \tau = \mu$ *iff* $(\mathtt{I}[\sigma] \cup \mathtt{I}[\tau])^* = \mathtt{I}[\mu]$

Putting our results together, we get:

2.6.3. PROPOSITION. *Let $\langle \Sigma, \sqsubseteq, \mathfrak{t}, \mathfrak{f} \rangle$ be a state space s.t. every subset of Σ has a fusion. Then:*

1. $\mathbf{s}_\varphi \leq \mathbf{s}_\psi$ *iff*
$$\textit{if } \sigma \sqsubseteq \bigsqcup_{Fa \textit{ in } \varphi} [Fa] \textit{ then } \sigma \sqsubseteq \bigsqcup_{Fa \textit{ in } \psi} [Fa]$$

2. $\mathbf{s}_\varphi + \mathbf{s}_\psi = \mathbf{s}_\chi$ *iff*
$$\sigma \sqsubseteq \bigsqcup_{Fa \textit{ in } \varphi \textit{ or } \psi} [Fa] \textit{ iff } \sigma \sqsubseteq \bigsqcup_{Fa \textit{ in } \chi} [Fa]$$

As noted in section 5 of Fine [2016], this indicates a second, equivalent formulation of Fine's theory of subject matter: relative to $\langle \Sigma, \sqsubseteq, \mathfrak{t}, \mathfrak{f} \rangle$, the set of subject matters is the set of ideals. Subject matter inclusion $\leq$ is set inclusion. Subject matter combination $+$ is set union followed by closure under parts and fusions. The subject matter $\mathbf{s}_\varphi$ is the smallest ideal that includes all truthmakers and falsemakers for atoms in φ. Altogether: we have an atom-based theory.

2.6.3 Fine's theory meets the constraints

Fine's theory meet our criteria of adequacy. Here is the thrust. Fusing together verifiers and falsifiers washes out the difference between the verification/falsification conditions for φ and $\neg\varphi$, as well as between $\varphi \wedge \psi$ and $\varphi \vee \psi$. Thus, constraints 1 through 5 are met. Next, the current theory accommodates hyper-intensional phenomena, since necessary claims like $1 + 1 = 2$, $1 = 1$ and $Fa \vee \neg Fa$ are intuitively verified by different facts. Likewise, impossibilities like $1 + 1 = 3$ and

$Fa \wedge \neg Fa$ are falsified by different facts. Thus, constraints 6 through 9 are met. Next, the claims aRb and bRa are intuitively verified (and falsified) by different situations, so their subject matter diverges, meeting constraint 10. The restriction that $a \preceq b$ implies $\mathbf{a} \sqsubseteq \mathbf{b}$ ensures that constraint 11 is observed: a statement about a part is also about the whole. Finally, Fine has a neat (though slightly artificial way) to generate a subject matter from a question, and a question from a subject matter. Consider a subject matter $\mathbf{s}$, and consider its *maximal possible parts*.[31] If we understand (naturally enough in the current setting) a question to be a set of possible situations (understood as possible answers), then our subject matter thereby generates a question. On the other hand, one can generate a subject matter from a question by fusing its possible answers.

2.6.4 Breaking the tie

The issue-based theory and Fine's theory are evenly matched, relative to the criteria of section 2.3. Are they competitors? Not if the only theoretical goal for a semantics is to accommodate robust linguistic intuitions. In this case, the theories stand as equally serviceable tools, until we find discriminating linguistic data.

Indeed, whether the theories substantially differ from a *formal* perspective depends on one's theoretical commitments. For, given relatively mild assumptions, it is possible to generate an issue-based theory from Fine's theory. To see this, consider an *atomistic* state space $\langle \Sigma, \sqsubseteq, \mathfrak{t}, \mathfrak{f} \rangle$: every situation in Σ is the fusion of a set of *atomic states*, where σ is atomic just in case σ has no proper parts (besides the degenerate fusion of the empty set of states). Suppose we also insist that every atomic expression Fa be assigned a unique atomic state as its sole truthmaker (denoted $+\texttt{Fa}$) and a unique complementary atomic state as its sole falsemaker (denoted $-\texttt{Fa}$). Then, on the Finean picture:

- $[Fa] = \{+\texttt{Fa}, -\texttt{Fa}\}$
- $\mathbf{s}_{Fa} = \mathbf{s}_{\neg Fa} = +\texttt{Fa} \sqcup -\texttt{Fa}$
- $\mathbf{s}_{Fa \wedge Gb} = \mathbf{s}_{Fa \vee Gb} = +\texttt{Fa} \sqcup -\texttt{Fa} \sqcup +\texttt{Gb} \sqcup -\texttt{Gb}$

[31]See the notion of 'factoring' in the appendix of Fine [2017b].

More generally: if φ is composed from all and only the atoms $F_1a_1, \ldots, F_na_n$, then $\mathbf{s}_\varphi = +\mathtt{F_1a_1} \sqcup -\mathtt{F_1a_1} \ldots +\mathtt{F_na_n} \sqcup -\mathtt{F_na_n}$. Now, note that the pair $+\mathtt{Fa}$ and $-\mathtt{Fa}$ provide a serviceable representation of a basic distinction (between a certain property holding of a certain object, or not), and that the following is easily shown (by induction) in the current setting:

- $\mathbf{s}_\varphi \leq \mathbf{s}_\psi$ iff $[\mathbf{s}_\varphi] \subseteq [\mathbf{s}_\psi]$
- $\mathbf{s}_\varphi + \mathbf{s}_\psi = \mathbf{s}_\chi$ iff $[\mathbf{s}_\varphi] \cup [\mathbf{s}_\psi] = [\mathbf{s}_\chi]$

Hence, in the current setting, subject matters may equivalently be defined as sets of atomic states; the subject matter $\mathbf{s}_\varphi$ as $[\varphi]$; subject matter combination as $\cup$; and subject matter inclusion $\leq$ as $\subseteq$. The result is an issue-based theory.

At any rate, my view is that accommodating ordinary linguistic data is *not* the only worthwhile goal for semantic theory. In this spirit, I offer two suggestive, but inconclusive, reasons to prefer the issue-based theory over Fine's theory. In the first place, it is hard to identify a pre-theoretic rationale for Fine's account that meshes with its details. In particular, Fine himself hints at a pre-theoretic rationale that seems an ill-fit with his proposal. In the second place, Fine's insistence that the subject matter of a claim be determined by its verification conditions robs his theory of useful explanatory power. In particular, it seems ill-placed to account for various hyper-intensional contexts.

On the first point. An account of a fundamental semantic notion should gel with our pre-theoretic views on its nature. This goes beyond accommodating our *use* of that notion in discourse; rather, the desideratum is to avoid departing dramatically from 'folk theory'. It is unreasonable to expect a folk theory to be comprehensive, precise or free of confusions. It is not unreasonable to ask our precise theory to explicate an existing notion, not invent a new one.

By this measure, the issue-based theory is attractive. It is guided by an intuitively appealing idea: a subject matter acts as a system of distinctions or issues, with claims in discourse judged as relevant exactly when they are sensitive to the distinctions that the discourse topic brings into focus.

Compare the core idea that, it seems, underpins Fine's approach:

> This is a 'fact'-based conception of subject-matter; the subject-matter of a statement is given, in effect, by those parts of a possible world which the statement is about [Fine, 2014, pg.209].

This starting point has some appeal: it is natural to take meaningful claims as making pronouncements, accurate or inaccurate, about an actual situation (a *fact*). To address the subject matter **Jane's profession** is to pronounce on the facts concerning Jane's profession. To discuss mathematics is to pronounce on the facts of the natural numbers, or the measurable spaces, or whatever.[32] (These remarks echo the advocacy of *Austinian topic situations* in Austin [1950] and Barwise and Etchemendy [1987], positing that the hallmark of a meaningful claim is to ascribe a property to a particular situation.)[33]

But how to reconcile this starting point with the surface details of Fine's account? On one presentation of Fine's theory, the subject matter of φ is the fusion of its verifiers and falsifiers. For many innocuous claims, this fusion is an impossible situation. But it defies intuition to claim that, say, "Jane is a lawyer" is about the impossible situation in which Jane is both a lawyer and not, in every conceivable way. And since impossible situations cannot be actualized, so much for the guiding rationale that meaningful claims are directed at *facts*.

According to the second presentation, the subject matter of φ is the ideal $\mathtt{I}[\varphi]$ i.e. the set of exact verifiers and falsifiers for φ, closed under fusions and parts. Again, this is a counter-intuitive when applied to innocuous cases. Observing your new car, I utter a truth: "your new car has plush leather seats". It is natural to say that this utterance is about a certain state of affairs: a particular car has seats with a particular quality. It is *less* natural to add that the utterance is (equally) about merely possible situations in which that same car has, say, fabric seats, or no seats. Further, it offends the idea that a claim's topic is an actual state of affairs (i.e. a concrete particular) to take topics as, fundamentally, *sets* of situations (i.e. an abstraction, or *type* of situation, at best).

Now, the second point. Beyond straightforward linguistic data, we can contrast the scope of two theories for illuminating *puzzle cases*. Now, consider a crucial distinguishing feature: according to Fine's theory, but not the issue-based theory, the subject matter of a claim is a function of its (exact) verification and falsification conditions. With this in mind, I below survey some controversial examples from the philosophical literature. In each case, we observe (i) a claim

[32]The issue-based theory can accommodate this only indirectly, as follows: an actual situation corresponds to a distinction - that situation obtaining or not - and it is this *distinction* that is, strictly speaking, the subject matter.

[33]Cf. the subject-predicate conception.

φ that (apparently) showcases aspects of meaning beyond truth/verification conditions, and (ii) that the issue-based theory can accommodate this by positing that φ's subject matter involves distinctions that are independent of what makes it true or false. I do not here claim that this is the *only possible* apparatus for accounting for these puzzle cases, nor that I have shown that it is the best. My point is modest: in contrast to theories that tightly link subject matter and verification, the issue-based theory accommodates these cases without additional resources. Indeed, failure on this front seems especially egregious if a key goal for a theory of subject matter is to accommodate hyper-intensional phenomena, for the cases in question seem prime examples of such phenomena.

The key resource for the issue-based strategy is that it can assimilate Fregean thinking into its framework. In what follows, I adopt an essentially Fregean perspective: the meaning of a name n is associated (somehow) with a pair $\langle \mathfrak{o}, \mathfrak{r} \rangle$, where $\mathfrak{o}$ is a rigid designator that picks out a certain object o and $\mathfrak{r}$ is a role that is associated with that object (i.e. a *guise*). I use a standard running example: "Clark Kent" rigidly designates Kal El, in the guise of the mild-mannered reporter: $\langle \mathfrak{c}, \mathfrak{m} \rangle$. This is contrasted with "Superman", which rigidly designates Kal El, in the guise of the super-powered hero: $\langle \mathfrak{c}, \mathfrak{s} \rangle$.

Frege: Compare "Clark Kent is late" and "Superman is late". It seems that these have the same truth set and verification conditions, but they nevertheless differ in their meaning: the first seems keyed to the distinction between the mild-mannered reporter being late or not, while the second is keyed to the distinction between the super-powered hero being late or not. The issue-based theory easily accommodates this, by positing that the subject matter of the first is $\{\langle \mathfrak{L}, \mathfrak{c} \rangle, \langle \mathfrak{L}, \mathfrak{m} \rangle\}$, and the subject matter of the second is: $\{\langle \mathfrak{L}, \mathfrak{c} \rangle, \langle \mathfrak{L}, \mathfrak{s} \rangle\}$.

Austin:[34] I say "Clark Kent is having a good night at the poker table", gesturing at someone who is, in fact, Jimmy Olsen. If Jimmy is actually having a bad night, but unbeknownst to me, Clark Kent is having a good night at a different poker table across town, then my claim seems to be false. Thus, it seems to be about Jimmy Olsen, and verified if *Jimmy* is having a good night. But, presumably, its meaning is different to "Jimmy Olsen is having a good night". An explanation: the subject matter of the claim is $\{\langle \mathfrak{G}, \mathfrak{j} \rangle, \langle \mathfrak{G}, \mathfrak{m} \rangle\}$, where $\mathfrak{j}$ rigidly designates Jimmy Olsen, and $\mathfrak{m}$ is the role of the mild-mannered reporter.

[34]See Austin [1950] and its development in Barwise and Etchemendy [1987].

Dretske:[35] I say "Clark Kent is an *award-winning* reporter" (focus on "award-winning"). This seems to have the same truth set and verification conditions as "Clark Kent is an award-winning *reporter*". But the meanings of the claims seem to differ. The first seems keyed to, say, the distinction between Clark Kent being a lousy reporter, or not. The second seems keyed to, say, the distinction between Clark Kent being an award-winning novelist, or not. The issue-based theory can easily capture this difference: roughly, the subject matter of the first claim is $\{\langle\mathfrak{A}, \mathfrak{c}\rangle, \langle\mathfrak{L}, \mathfrak{c}\rangle\}$, where $\mathfrak{A}$ corresponds to being an award-winning reporter and $\mathfrak{L}$ corresponds to being a lousy reporter; the subject matter of the second is $\{\langle\mathfrak{A}, \mathfrak{c}\rangle, \langle\mathfrak{N}, \mathfrak{c}\rangle\}$, where $\mathfrak{N}$ corresponds to being an award-winning novelist.

Donnellan:[36] I say "the mild-mannered reporter from the Daily Planet is working hard tonight" gesturing (unbeknownst to me) at someone who is *not* Clark Kent, but rather Jimmy Olsen, the boorish reporter from the New York Times (though he is indeed working hard). Plausibly, this is a referential use of a definite description, about the New York Times reporter, and therefore true if he is working hard. Neverthless, its meaning presumably differs from "the boorish reporter from the New York Times is working hard tonight". The issue-based theory can capture the difference: the subject matter of this last claim is $\{\langle\mathfrak{W}, \mathfrak{j}\rangle, \langle\mathfrak{W}, \mathfrak{b}\rangle\}$ (where $\mathfrak{j}$ rigidly designates Jimmy Olsen and $\mathfrak{b}$ is the role of being the boorish reporter for the New York Times), while that of our original claim is $\{\langle\mathfrak{W}, \mathfrak{j}\rangle, \langle\mathfrak{W}, \mathfrak{m}\rangle\}$.

2.7 Conclusion

We conclude that there is a case for the issue-based theory of subject matter. In the absence of any compelling counter-considerations, or rival accounts with equivalent explanatory power, we accept the issue-based account, and build resolution semantics upon its base.

[35] See Dretske [1972].
[36] See Donnellan [1966]

Chapter 3

Questions, Topics and Restricted Closure

Single-premise epistemic closure is the principle that: if one is in an evidential position to know that φ where φ entails ψ, then one is in an evidential position to know that ψ.[1] In this chapter, I defend the viability of *opposition* to closure. A key task for such an opponent is to precisely formulate a *restricted* closure principle that remains true to the motivations for abandoning unrestricted closure but does not endorse particularly egregious instances of closure violation. I focus on two brands of epistemic theory (each the object of sustained recent interest in the literature) that naturally incorporate closure restrictions. The first type holds that the truth value of a knowledge ascription is relative to a relevant *question*. The second holds that the truth value of a knowledge ascription is relative to a relevant *topic*. For each approach, I offer a formalization of a leading theory from the literature (respectively, that of Jonathan Schaffer and that of Stephen Yablo) and use this formalization to evaluate the theory's adequacy in terms of a precise set of desiderata. I conclude that neither theory succeeds in meeting these desiderata, casting doubt on the viability of the underlying approaches. Finally, I argue that resolution theory fares better.

[1] This chapter refines and extends Hawke [2016a].

3.1 Introduction

Is knowledge closed under deductive entailment? If ψ follows from φ, and one knows that φ, is one essentially in a position to know that ψ (putting aside one's contingent cognitive limitations)?

The issue remains contentious.[2] On one hand, closure denial provides an antidote to philosophical paradox. Here is a striking example that we outline in detail in section 3.2.2:[3] a closure denier can resolve the Cartesian skeptical paradox by allowing for agents that know mundane empirical truths, yet are not positioned to know they are not victims of systematic sensory deception.

On the other hand, closure strikes many as virtually indisputable in the face of the powerful intuition that deductive reasoning is essentially[4] a sure-fire means for knowledge extension, and so part of the bedrock of rational inquiry.[5]

In this chapter, I explore new frontiers for the sensible denial of closure.[6] One who denies closure has an important task: to offer a *restricted* closure principle that, in a disciplined manner, affords deduction an intuitively satisfying epistemic scope.[7] Stephen Yablo states this point vividly:

---

[2]The current debate traces back to Dretske [1970], Goldman [1976] and Stine [1976].

[3]Though not the only one: other epistemic paradoxes that crucially involve an appeal to unrestricted epistemic closure include Saul Kripke's dogmatism paradox and Stewart Cohen's easy knowledge paradox. See Bradley [2014] and [Yablo, 2014, Ch. 7] for a survey of relevant paradoxes. The discussion in this chapter can easily be adapted to these other cases.

[4]One can admit *some* exceptions to deduction's capacity to extend knowledge without betraying the title of "closure advocate". For instance, if a conclusion is reached by competent deduction from known premises, but was *already* known, then the deduction did not succeed in *extending* knowledge.

[5]Richard Feldman: "To my mind, the idea that no version of the closure principle is true - that we can fail to know things that we knowingly deduce from other facts we know - is among the least plausible ideas to gain currency in epistemology in recent years" [Feldman, 1999, pg.95]. John Hawthorne: "[Overstatement aside], I am inclined to side with Feldman. The intuitive consequences of denying Single-Premise Closure seem to be extremely high" [Hawthorne, 2004, Sect .1.5]. Saul Kripke: "...I am sympathetic to those philosophers who regard this idea [i.e. closure rejection] as intrinsically implausible or even preposterous" [Kripke, 2011, pg.163]. (I emphasize, however, that Hawthorne and Kripke take closure rejection seriously enough to construct a careful refutation.) Timothy Williamson: "We should in any case be very reluctant to reject intuitive closure, for it *is* intuitive. If we reject it, in what circumstances can we gain knowledge by deduction?" [Williamson, 2000, pg.118].

[6]Dretske [1970] and Nozick [1981] are the most well-known closure deniers. More recent authors that embrace restricted closure in lieu of unrestricted closure include Black [2008], Lawlor [2013] and Holliday [2013].

[7]See Lawlor [Lawlor, 2013, sect. 4.7] for further discussion of restricted versus unrestricted closure.

> What would be a "good way of denying" [closure]? Closure cannot just be thrown under the bus. A good way of denying it would tell us what is right in the principle - call that the defensible core - and explain how the remainder can be done without [Yablo, 2014, pg.114].

Saul Kripke makes a similar point in similarly evocative language:

> It is incumbent on any author who rejects the deductive closure of knowledge to state such conditions [i.e. conditions for when a valid deduction preserves knowledge] ... Without them, and with a mere rejection of the deductive closure of knowledge, anyone who proves anything from known premises could be criticized for the well-known fallacy of giving a valid argument for a conclusion! [Kripke, 2011, pg.200]

Locating a satisfying restriction is tricky.[8] On one hand, it must be compatible with the failure of knowledge-transmission across certain key (skepticism-inducing) inferences: that one knows one has hands ought not put one in a position to know that one is not a handless brain-in-a-vat. On the other hand, an opponent to closure must avoid endorsing obvious absurdities, such as that a correct mathematical proof does not ground knowledge,[9] or that knowing a conjunction need not put one in a position to know each conjunct. For instance, that Kripke [2011] shows that the tracking theory of Nozick [1981] exhibits this latter trait is generally taken as a fatal criticism. Notice that a closure denier cannot vindicate these *instances* of closure failure by direct appeal to the basic motivation for denying closure: it is hard to see a skeptical paradox as arising *merely* from the principle that knowing a conjunction puts one in a position to know each conjunct, or that mathematical proof generates knowledge.

I focus on two approaches to the semantics of knowledge attributions that, I argue, allow for closure failure yet indicate natural restrictions on closure. The

[8]For revealing discussion of the instances of closure that we particularly resist denying, see Kripke [2011], Hawthorne [2004], Lawlor [2005], Holliday [2013] and Yablo [2014].

[9][Schaffer, 2005a, sect. 5] notes this objection succinctly (in the voice of the closure advocate): "Surely deduction transmits knowledge. ... How could it not, given that mathematical proof is deductive and mathematical proof yields knowledge?" Though preserving mathematical inquiry is a serious challenge for closure deniers, the challenger must resist overstating their case by mistaking closure denial for a bolder thesis: that deduction (and so mathematical proof) is *always* epistemically inert. The challenge, rather, is to show how it is possible to account for the epistemic power of mathematical proof while denying closure in full generality.

first is the *question-sensitive* approach, according to which the truth of a knowledge attribution is question relative, a version of which is defended by Schaffer [2004] and Schaffer [2005a]. The second is the *topic-sensitive* approach, according to which the truth of a knowledge attribution is subject matter relative, a version of which is suggested by Yablo [2012] and [Yablo, 2014, ch. 7]. After making the case that Schaffer and Yablo's positions give rise to closure failure (section 3.3), I formalize these positions, and evaluate the resulting theories for adequacy (section 3.5). My conclusion: both fail to achieve an intuitively satisfying balance between closure violation and preservation. I then turn (section 3.6) to a variation of the topic-sensitive approach that I show fares better: namely, *resolution theory*. A resolution theorist distinguishes herself from the Yablovian through her accounts of the subject matter of a sentence (in particular, for her, subject matter is a function of the *atomic predications* that occur in that sentence), and what it is *for things to be a certain way* regarding a subject matter.

To prime the reader, I emphasize four crucial features of the current chapter's approach.

First: I argue (section 3.2.1) that the substantive issues of the closure debate emerge best when phrased in terms of *knowability*. relative to the available empirical information (cf. Dretske [2005]).[10]

Second: I deliberately remain silent on further important closure questions concerning justification, reason, evidence and warrant (cf. Wright [2003]).

Third: I understand the issues raised by the closure debate as best addressed by consideration of the semantics and logic of knowability attributions. In particular, I often regiment discussion by providing a precise semantics for an artificial logical language, so as to pin down systematic principles that may otherwise be elusive (for instance, it allows us to postpone the complicating effects of focus, indexicals and proper names in natural language). Throughout, I focus entirely on propositional logic, ignoring matters related to, for instance, higher-order knowledge or quantification. Our topic is the *internal propositional logic of knowability attributions*.

Fourth: for our purposes, the adequacy of a closure-restricting theory amounts to it delivering the right results for certain formal principles detailed in section

[10]The literature on epistemic closure has steadily shifted to setting up the discussion in terms of knowability instead of knowledge *per se*. Observe the treatments in Schaffer [2007a] and Blome-Tillmann [2014].

3.4, some of which I propose a closure opponent should accept as valid, some reject as invalid.[11] To justify this pattern of validity and invalidity, I rely on the following sorting device: a closure denier ought to buck intuition and reject the validity of an instance of closure when that instance can be used to ground a skeptical paradox. (The identification of any *further* criteria for the adequacy of a closure-denying theory I leave for elsewhere.) Note that my aim is not to convince a hardened closure advocate to accept this pattern of (in)validity as a desideratum for their own theory, nor is my aim to conclusively refute all closure-validating theories. Rather, it is to fortify the *respectability* of closure denial, by showcasing a closure-restricting theory that avoids certain unwelcome trade-offs.

3.2 Closure and its discontents

3.2.1 The form of closure

Following one orthodox line [Nozick, 1981], I capture epistemic closure formally as follows:

Unrestricted Closure: $(K\varphi \wedge \varphi \Rightarrow \psi) \rightarrow K\psi$

where φ and ψ are sentence meta-variables, $\wedge$ is conjunction and $\rightarrow$ is the material conditional. Call this the *formal principle of unrestricted single-premise closure* (or *unrestricted closure* or *closure* when the context is clear). The intended purpose of this formula is to succinctly express - in a symbolic form amenable to precise logical techniques - what is at stake in the closure debate. In the current section, I provide a reading of this formula that, I argue, best crystallizes this import.

Read $\varphi \Rightarrow \psi$ as "φ (deductively) entails ψ". I understand entailment in standard terms, as equivalent to the *necessity* of $\varphi \rightarrow \psi$. I remain silent on how best to understand the kind of necessity at issue. It is convenient to assume that if φ entails ψ then $\varphi \rightarrow \psi$ is knowable *a priori*. Indeed, I find it natural to think of the relevant necessity as *epistemic necessity*, understood against the backdrop of a space of *basic epistemic possibilities*. Note, however, that the main results of

[11] Compare the discussion of "egregious violations" of closure in Nozick [1981], Hawthorne [2004], Kripke [2011] and Holliday [2013]

this chapter can be replicated if we replace $\varphi \Rightarrow \psi$ with $K(\varphi \to \psi)$ in the closure formula, and drop talk of entailment and *a priority* altogether.

Some philosophers (e.g. [Nozick, 1981, pg.204]) complete the interpretation by reading $K\varphi$ as: "the subject knows that φ". But, obviously, an ordinary agent may know φ, yet not recognize (let alone *know*) many of φ's consequences. If denial of closure is to have significance, it *cannot* simply amount to the denial of: if ordinary agent a knows φ and φ entails ψ, then necessarily ψ is known by a. Note that it does not help to replace $\varphi \Rightarrow \psi$ with $K(\varphi \to \psi)$ in the statement of closure: an ordinary agent may well know two facts that *together* entail ψ, yet fail to put two-and-two together.

Instead, one might be tempted to join John Hawthorne and propose the following interpretation for the closure formula: if one knows φ and thereby comes to believe a consequence ψ of φ using a competent deduction from φ (all the while continuing to know φ), then necessarily one knows that ψ [Hawthorne, 2004, sect. 2.3]. Thus the proposal is to read $K\varphi$ as "a knows that φ", but enrich the reading of $\Rightarrow$ to not only express entailment, but extension of belief via competent deduction.

Though on the right track, I am doubtful that this interpretation expresses a valid principle and (more importantly) does full justice to the intuitions that motivate adherence to a closure principle. Consider an example due to [Lawlor, 2005, pp.32-33]. Imagine *Edward*, who knows φ (*facts from chemistry*), competently thereby comes to believe consequence ψ (*this homepathic cold medicine is too diluted to contain active ingredients*) - all the while knowing φ - yet maintains throughout a further belief χ (*my mother has always sworn that homeopathy is efficacious*) which weighs against ψ. Edward, Lawlor suggests, may well (even in the face of belief in χ) believe ψ on the basis of φ. Nevertheless, we can suppose, his confidence in ψ is sufficiently suppressed so that he does not *know* ψ. Moral: ordinary human psychology is complex, so valid principles governing occurrent mental states are hard to pin down.

Lawlor's counter-example suggests a better interpretation for closure. Intuitively, there is a tension in Edward's belief system that cries out for reflective resolution (as Lawlor notes). *Upon reflection* Edward must question his grounds for believing χ (given his knowledge of φ), and so come to know ψ. Edward's knowledge of φ ensures that he has sufficient *informational resources* to know ψ.

In his way is a dearth of time, attention, discipline or other cognitive resources.

As usual, I take it as uncontroversial that our empirical knowledge is rooted in the sensory reception of *empirical information* that narrows down the state of the actual world. To say more is to quickly stray into controversial territory. So I treat this information only in very abstract terms: as a proper subset of logical space that contains the actual world (i.e. a true contingent unstructured proposition). I commit to an important assumption: that an ordinary agent's empirical information cannot distinguish between a 'perceptually normal world' and a corresponding 'brain-in-vat world'. I call the empirical information that a has received the *basic empirical evidence* belonging to a (or *evidence* for short). I say that a is in an *evidential position* to know φ just in case it is possible that an agent with precisely the same evidence as a knows that φ. Briefly: φ is knowable for a. Finally, φ is knowable *a priori* just in case φ is knowable in every possible evidential position.

Now, I propose that the following principle deserves our keen attention:

> if a is in an evidential position to know φ and φ deductively entails ψ, then a is in an evidential position to know ψ.

That is: if a has received sufficient empirical information to know φ then a has sufficient empirical information to know ψ (though perhaps it remains for a to further *process* that information in order to deduce ψ). Call the above the *informal principle of unrestricted closure*. Our principle is in the spirit of the framing singled out by Dretske [2005].

I therefore use the formula we started with as a succinct expression of informal unrestricted closure, reading $K\varphi$ as "φ is knowable to a". K is a *knowability operator*.

We say a theory of knowability attributions has *closure failure* (or *closure rejection*) just in case it invalidates unrestricted closure. A theory upholds *restricted closure* with respect to restriction $\texttt{Restr}(\varphi, \psi)$ just in case it has closure failure but, nevertheless, validates the following schema for some reading of $\texttt{Restr}$:

$$\textbf{Restricted Closure:}\ (K\varphi \wedge \varphi \Rightarrow \psi \wedge \texttt{Restr}(\varphi, \psi)) \rightarrow K\psi$$

3.2.2 Against closure: the skeptical paradox

To scrutinize the motivation for closure denial, we attend to the *skeptical paradox* (cf. [Cohen, 1988]). In what follows, read ‘a brain-in-a-vat’ as shorthand for, roughly, ‘a mere brain that is envatted and thereby systematically deceived”.

P1. a is in an evidential position to know that she has hands.

P2. a is not in an evidential position to know that she is not a brain-in-a-vat.

P3. That a is a brain-in-a-vat is an alternative to a having hands.

P4. For any φ and ψ, if a is in an evidential position to know φ and ψ is an alternative to φ, then a is in an evidential position to know $\neg\psi$.

Or formally:

P1. Kh

P2. $\neg K(\neg b)$

P3. $h \Rightarrow \neg b$

P4. For any φ, ψ: $(K\varphi \wedge \varphi \Rightarrow \neg\psi) \rightarrow K(\neg\psi)$.

There seems an immediate and persuasive rationale for accepting each claim. P1: an utterly banal ordinary knowledge claim, and as such treated as a datum in philosophical theorizing. P2: a brain-in-vat scenario is *by selection* the kind of scenario for which one cannot accrue disconfirming empirical evidence. P3: it is not (conceptually, metaphysically) possible that one is, simultaneously, both handed and a brain-in-vat. P4: represents the seemingly plain fact that knowledge can be extended through deductive reasoning.

Yet the premises are jointly inconsistent. We have a paradox.

Denial of P1 offers a *skeptical solution*. Denial of P2 offers a *moorean solution*. To many (myself included), both solutions seem a significant betrayal of intuition.[12] I know of no opposition to P3.[13] So our attention becomes trained on solutions that deny P4.

Resistance to denying P4 has crystallized as three related objections. First, as mentioned, the intuitive connection between closure and the extension of knowledge by competent deduction is emphasized. Second, it is claimed that denying closure is at odds with the linguistic data: we *do not* comfortably utter $K\varphi \wedge \neg K\psi$ when $\varphi \Rightarrow \psi$ is an obvious truth (as Lewis [1996] observes, it "sounds contradictory" to exclaim that one knows one has hands but do not know that one is a brain-in-a-vat). Such conjunctions are labeled "abominable" by DeRose [1995]. Third, against *particular* closure-denying theories, critics locate *particular* instances of closure whose acceptance is purportedly common ground and demonstrate that the theory in question violates these instances [Hawthorne, 2004]. Here, an *instance of closure* is a schema that results from replacing the meta-variables φ and ψ in the unrestricted closure schema with schema with more structure. An example: $K(\varphi \wedge \psi) \wedge ((\varphi \wedge \psi) \Rightarrow \varphi) \rightarrow K\varphi$ (or equivalently and more briefly: $K(\varphi \wedge \psi) \rightarrow K\varphi$).

All three objections are defused by locating a satisfactory restricted form of closure. No doubt, closure has intuitive appeal, abominable conjunctions sound abominable and certain instances of closure failure seem particularly egregious. But the cost of faulting these intuitions is easier to bear if the closure opponent has a sensible *explanation* for our faulty intuitions. So a restriction on closure should indicate why *ordinarily* deductive reasoning extends knowledge, thereby explain why unrestricted closure is intuitive: our intuitions are largely molded

[12]This, of course, is too quick a dismissal of the moorean approach, especially given the sophisticated versions on the market. Again, my purpose is not to conclusively refute every rival to closure denial.

[13]It pays to be careful about this point, since there is an author - namely, Roush [2010] - that, in the first place, seems to explicitly deny P3 in the skeptical paradox and, in the second place, develops this as a response to the paradox. We need not, however, disagree with Roush [2010] on any substantive issues connected to the first point (we delay discussion of the second point for elsewhere). Roush [2010] points out that it is possible to be a brain-in-vat that is handed. For instance, think of the systematically deceived subjects depicted in the film *The Matrix*: the entire body of these subjects is envatted, in contrast to the (perhaps standard) image of a mere brain floating in a vat. Thus, strictly speaking, having a hand does not entail not being a brain-in-vat, so P3 can be rejected on a sensible reading of h and b. This indicates that, for our purposes, b should *not* merely be read as "a is a systemically deceived brain-in-vat", in the loosest sense of 'brain-in-vat'. Nor should it be read as "the sum total of visual appearances that a has experienced are systematically deceiving". Rather, to be a *brain-in-vat*, in our sense, is to be envatted as a *mere brain* and systematically deceived. With the meaning of the predicate so stipulated, it is seemingly indisputable that having hands entails not being a brain-in-vat.

by everyday cases, and closure failure chiefly rears its head beyond this domain (e.g. in the philosophy classroom). This same strategy explains our aversion to "abominable conjunctions": they are "conversational abominations" - not "logical abominations" - that thwart our ordinary expectations (cf. Dretske [2005]). Third, a satisfying restricted closure principle finds the right balance of closure failure and violation, validating instances of closure that are genuinely beyond sensible dispute.

The strategy of *dissolving* the skeptical paradox has received particular attention as an alternative to closure rejection, at least among theorists that are focused on semantic considerations. Most notably, *contextualists* have posited an equivocation that creates the illusion of paradox. A contextualist proposes that the contents of a knowledge ascription can change from one context of utterance to another.[14] In particular, they think that the epistemic standards associated with "knows that" can vary across discourse contexts. On their view, the seeming paradox of P1-P4 is a product of the fact that - though the circumstances of the subject of the knowledge attribution remains invariant - there are contexts in which it is apt to deny to P1 (for instance, *philosophical contexts*) and contexts in which it is apt to deny P2 (for instance, *mundane contexts*). However, there is no one context in which it is apt to deny *both* and so no context where P4 is false. Thus, if the context is explicitly fixed then the equivocation, and sense of paradox, disappears.

Contextualism has been robustly debated.[15] I do not here offer final judgement on the (de-)merits of contextualism, nor deny that many forms of contextualism stand in clear opposition to closure denial. Nevertheless, I *do* wish to observe an over-looked fact: the line between certain important versions of contextualism and closure rejection blurs when the closure debate is properly focused on issues of *knowability on the evidence*. Indeed, I claim it is no longer clear for such contextualists that they are best understood as offering a strategy of paradox dissolution *as opposed to* rejecting P4. The grand opposition between contextualism and closure denial is exaggerated. I illustrate this in the next section with a brief appraisal of David Lewis' paradigmatic contextualist stance.

[14]For a sample of the prominent defenders of contextualism, see Stine [1976], Cohen [1988], DeRose [1995] and Lewis [1996].

[15]For penetrating discussion see any of Hawthorne [2004], Stanley [2005], or DeRose [2009].

3.3 Three existing paths to closure rejection

I showcase three positions from the recent literature that are best understood, I propose, as sophisticated forms of closure denial. As developed here, each position amounts to a version of *relevant alternatives semantics* for binary knowability attributions, differing in their accounts of *relevance* and *ruling out*:[16]

> "φ is knowable to a" is true in context c just in case every *relevant* alternative proposition to φ is ruled out by the evidence E available to a. **(RA semantics)**

An *alternative proposition* to sentence φ is a proposition that is mutually exclusive to the proposition expressed by φ. I represent the propositions *relevant* to sentence φ in context c by a function $\mathbf{R}_c$ from sentences to sets of propositions. Effectively then, one who upholds unrestricted closure is one who holds that necessarily, when φ entails ψ, the relevant alternatives supplied by $\mathbf{R}_c(\psi)$ are a subset of $\mathbf{R}_c(\varphi)$. A closure denier rejects this constraint.

3.3.1 Lewis' context-sensitivity

David Lewis holds that a discourse context supplies a fixed set of *relevant propositions*,[17] constant across all sentences [Lewis, 1996]. The relevant alternatives to φ in that context, then, are the relevant propositions that are alternatives to φ. Thus, for Lewis, if φ entails ψ, then $\mathbf{R}_c(\psi)$ is a subset of $\mathbf{R}_c(\varphi)$.

Lewis diagnoses our puzzlement over the skeptical paradox as follows: a discourse context shifts if neglected possibilities become *psychologically salient* to the discourse participants. When a skeptical paradox is raised in conversation such a shift is typically witnessed. In a mundane context (skeptical possibilities ignored), P1 and P4 are true, while P2 is false. However, explicit consideration of P2 or P4 raises awareness of skeptical hypotheses, shifting to a philosophical context where P1 is false, but P2 and P4 are true.

[16] Schaffer, in particular, distances his views from what he calls *relevantism* [Schaffer, 2005a, pg.267]. His opposition is grounded in (i) understanding relevantism as endorsing a binary conception of knowledge-itself and (ii) closely associating 'relevance' with David Lewis' particular elaboration. Neither assumption is built into my statement of relevant alternatives semantics.

[17] Lewis uses the term *possibility*, instead of *proposition*, somewhat ambiguous between talk of *possible worlds* and *propositions*. The difference does not matter for our discussion of Lewis.

Call this the *attention explanation.* Lewis notes that his account affords claims like "I am not a brain-in-vat" a curious status in ordinary contexts: "a knows that $\neg b$" is true, but its truth cannot be explicitly expressed (or even *thought*) without becoming false. Lewis calls this *elusive knowledge* (in fact, call it *strongly* elusive, since there is *no* context or evidential status relative to which this "knowledge" can be truly expressed). In terms of knowability: $\neg b$ is known on the evidence in ordinary context c, but is not knowable *non-elusively.*

Lewis alleges to provide a cunning defense of closure in the face of the skeptical paradox. But there is a clear sense in which he is an (unacknowledged) closure denier. Define the function $\mathbf{R}_c^e$ as supplying, in context c, the alternatives to sentence φ that need to be ruled out in order for "a knows that φ" to be true *explicitly* i.e. non-elusively. That is, take $\mathbf{R}_c^e(\varphi)$ to be the set of propositions that *would become salient* when φ is explicitly considered in the discourse, incorporated with the background of relevant propositions supplied by context c. Using RA semantics, one may then define a knowability operator K_e (*explicit knowability*) relative to $\mathbf{R}_c^e$. Then, in ordinary context c, $K_e h$ holds, but $K_e(\neg b)$ does not.

That K_e does not validate closure puts Lewis' credentials as a serious closure advocate in grave doubt. Closure advocacy, if properly motivated, is committed to deduction's unimpeachable status as a tool for knowledge extension. Now, on Lewis' theory, suppose I self-attribute the true claim "I know I have hands" in ordinary context c, with ordinary evidence E. Does deductive reasoning then place me in a position to self-attribute "I know I am not a brain-in-vat"? No: if I attempt such deductive reasoning, uneliminable brain-in-vat possibilities become salient to myself (nor can I self-attribute "I know I have hands"). Of course, this just focuses on what can be *said* following deduction. Putting aside the expression of knowledge, is there a significant sense in which Lewis maintains closure? No. For Lewis, ordinary evidence does not make the denial of a skeptical hypothesis knowable in any significant sense, despite this denial falling (in mundane contexts) under his category of "elusive knowledge". For strong elusiveness amounts to a *merely* technical category in Lewis' theory, representing little epistemic significance. We can see this in two ways. First, since we are incapable of expressing this "knowledge" with ordinary language, it is effectively independent of ordinary talk (and thought). Second, let us semantically descend from knowledge talk to knowledge-itself, and consider strongly elusive knowledge in this light. It is natural to understand Lewis as proposing that knowledge-itself (independent of linguistic expression) is a ternary relation $\mathbb{K}(s, P, \mathbf{A})$ between an agent a, proposition P and set of alternatives $\mathbf{A}$ to P. So let P be the proposition expressed by "I am not a brain-in-vat". Then, for Lewis, given a's ordinary evidence, $\mathbb{K}(s, P, \mathbf{A})$ holds (only) when $\mathbf{A}$ is empty. But, on Lewis' theory, *every* proposition is known in *this* sense, no matter the evidence.

3.3.2 Schaffer's question-sensitivity

We turn to Jonathan Schaffer's *contrastivist* account of knowledge (cf. our treatment of Lewis). Schaffer holds that knowledge is *question-sensitive* in the following sense: from a metaphysical point of view, knowledge is a ternary relation $\mathbb{K}(s, P, Q)$, where a is a subject, P is the known proposition and Q is a *contrast* proposition. P and Q together constitute a *question* - *P or Q?* - with respect to which a is in an epistemic position to supply the correct answer P. Various considerations support this picture, Schaffer proposes. For one, it gels with the bulk of our epistemic language: consider explicitly contrastive language such as *a knows that P rather than Q* or knowledge-wh expressions such as *a knows who stole the painting*, where the object of knowledge seems to explicitly be a question. For another, if inquiry is a question-directed activity [Hintikka, 2007], and knowledge 'keeps score' of inquiry, then the question-sensitivity of knowledge seems natural.

For Schaffer, particularly *explicit* knowledge claims make the question at issue plain. But how are we to understand the semantics of a *binary* knowledge ascription "a knows that P" on Schaffer's view? His answer: when we utter sentences of this type, the *context* supplies an implicit contrast:

> Moving finally to declarative sentences (perhaps the rarest form in natural language), these inherit their contrasts from *context*... In general, context provides the default source of contrasts [Schaffer, 2005a, pg.249].

I develop Schaffer's suggestion as follows: on his approach, the space of relevant propositions, relative to a sentence φ, is a set of *answers to a question* $\mathbf{Q}_c(\varphi)$ supplied (partly) by context. Then: φ is knowable in context c on the subject's evidence E just in case every answer to $\mathbf{Q}_c(\varphi)$ that contrasts with φ is ruled out by E.

This account naturally suggests closure rejection: there seems no reason that context could not supply quite different questions to φ and its consequence ψ. The question associated with h in context may be "hand or stump?", and the question associated with $\neg b$ may be "ordinary perceptual profile or massive sensory deception?". Note that the respective contrast propositions so expressed are *independent*. Thus there seems no reason to insist that the alternatives in $\mathbf{R}_c(\psi)$ necessarily be a subset of those in $\mathbf{R}_c(\varphi)$.

The account also suggests natural *restrictions* on closure, as [Schaffer, 2005a, sect. 5] essentially observes . For one: ψ is knowable if φ is knowable when φ entails ψ *and* the contrast to ψ (in context) entails the contrast to φ.

3.3.3 Yablo's topic-sensitivity

Finally, I sketch Stephen Yablo's topic-sensitive account knowledge attribution [Yablo, 2014]. For him, the truth of a knowledge claim is *subject matter sensitive*: whether φ is known depends not only on what proposition is expressed by φ (that is, its content, or information), but also what φ is *about*. This move is natural. Subject matter is a subject of independent interest in the theory of meaning (one arguably left neglected when meaning is too closely tied to mere truth conditions [Perry, 1989]). *Aboutness properties* are a theoretical tool already available to, and seeking application for, the epistemic theorist. That this application exists is suggested by an intuitively close connection between *knowing that* and *knowing about*: knowing that John is a lawyer implies knowing something *about* him. If I know *about* John, then there is something *that* I know about him. As Yablo puts it:

> Knowledge attributions care about subject matter, over and above truth-conditions. They take note of how P is true or false in various worlds, not only which worlds it is true or false in. [Yablo, 2014, pg.121]

The second sentence broadly indicates Yablo's conception of what subject matter *is* (an evolution of the conception of Lewis [1988a]). The subject matter $\mathbf{T}(\varphi)$ of a sentence φ may be identified with the class of *ways things can be* with respect to that subject matter, which, for Yablo, may be divided into the (minimal) semantic *truth-makers* and *false-makers* for φ. A *semantic truth-maker* for φ logically necessitates the *truth* of φ. A *semantic false-maker* for φ logically necessitates the *falsity* of φ. (Where does the subject matter for φ come from? For our part, we may take it as a combined consequence of semantics and context.)[18]

This suggests an RA semantics for knowability attributions: for φ to be knowable on the evidence is to be in a position to rule out the (minimal) *false-makers* for φ determined by $\mathbf{T}(\varphi)$. Indeed, [Yablo, 2014, Ch.7] endorses exactly a theory along these lines.[19][20]

This account naturally suggests closure rejection: there seems no reason that semantics and context could not supply vastly different subject matter (and so false-makers) to φ and its consequence ψ. Intuitively, unlike "I am not a brain-in-vat", the subject matter of "I have hands", in an ordinary context, concerns ordinary distinctions concerning the state of my extremities, and does *not* involve the skeptical subject matter of the epistemology classroom. So there seems no reason to insist that the alternatives in $\mathbf{R}_c(\psi)$ necessarily be a subset of those in $\mathbf{R}_c(\varphi)$.

The account also suggests a natural *restriction* on closure, as Yablo essentially observes: ψ is knowable if φ is knowable when φ entails ψ *and* the subject matter of ψ (in context) is included in the subject matter of φ.

[18]Yablo [2014] in fact offers two accounts of subject matter - what he calls the *recursive* account and the *reductive* account. He is hesitant to fully commit to one model at the exclusion of the other [Yablo, 2014, sect. 4.11]. Either can be integrated into an RA semantics for knowability attributions. We focus here on the reductive account. The recursive account produces an RA theory that shares important similarities with compositional S-theory (see sect. 3.5.1).

[19]Also see Yablo [2012], available on Yablo's website at http://www.mit.edu/ yablo/home/Papers.html.

[20]This ignores a subtle feature that Yablo incorporates into his picture: namely, that the agent not only be in a position to discard the falsemakers of φ, but also have a suitable grip on the minimal truthmaker that is in fact responsible for the truth of φ: one might be "right to regard Q as true, but, if you are sufficiently confused about *how* it is true - about how things stand with respect to its subject matter - then you don't know that Q" [Yablo, 2014, pg. 119]. I here put aside this extension for two reasons: (i) Yablo does not elaborate on his proposal and (ii) it seems to me that however the details are worked out, the critical results of proposition 3.5.4 will hold (since the pertinent counter-examples depend only on falsemaker structure).

3.4 Criteria of adequacy for closure rejection

I concentrate on Yablo and Schaffer's theories, and put aside Lewis. Nevertheless, our conclusion that the latter inadvertently *embraces* closure rejection serves two important dialectical purposes that background the coming discussion. First, the blurring of Lewisian contextualism with closure denial dispels the myth that there is *no question* that the impulse to accept both P1 and P2 in the skeptical paradox drives one to contextualism *rather than* closure rejection. Second, both Schaffer and Yablo express amenability to the main features of Lewis' approach to the skeptical paradox, and attempt to incorporate these into their own accounts (cf. [Schaffer, 2005a, sect.5] and [Yablo, 2014, pg.127]). I don't explicitly incorporate such features into my formalizations, but their addition would not affect the substance of the coming discussion, for we would then focus our attention on *non-elusive (explicit) knowability* and proceed to parallel, equally significant conclusions.

We have various theories on the table that render closure rejection natural. How to choose? I now propose four precise criteria for adequacy for a theory with closure rejection, boiling down to a list of dialectically significant principles that, I propose, a theory that rejects closure ought to validate, and a list of principles that such a theory ought to invalidate (cf. Nozick [1981], Hawthorne [2004], Kripke [2011], Holliday [2013]).

Since *validity* is key, an explicit formal framework will aid us. We work with a propositional language built up using atoms $p, q, r, \ldots$ (intuitively describing *basic predications*), the usual propositional connectives, knowability operator K, necessity operator $\Box$ and expressions of the form $\texttt{Restr}(\varphi, \psi)$. We postpone any issues connected with nested K, $\Box$ or $\texttt{Restr}$ operators by simply excluding such formulas from our language.

An *RA model* $\mathcal{M}$ is a tuple $\langle W, \mathbf{R}, \{E_w\}_{w \in W}, \mathbf{V}\rangle$. W is a set of *possible worlds*; $\mathbf{R}$ is a *relevancy function* that accepts a sentence and returns a set of *propositions relevant to that sentence* (intutively, this function is partly set by context); E_w is the subject's basic empirical evidence at world w; and $\mathbf{V}$ is a valuation function that assigns a proposition to each atom (equivalently: a set of atoms to every world). When I write *proposition* I have in mind an *unstructured proposition*: a set of possible worlds. Assume that E_w is consistent with w i.e. $w \in E_w$.

Intuitively, an RA model (excluding $\mathbf{R}$) describes the *evidential situation* of a subject a.

We deploy a relevant alternatives semantics to define satisfaction relation $\vDash$. Given model $\mathcal{M}$ and world w, an atom p is true at w just in case $\mathbf{V}$ assigns p to w. Expressions with boolean connectives are as one would expect. $\Box\varphi$ is true at w just in case φ is true at *every* world in W. We abbreviate $\Box(\varphi \to \psi)$ by $\varphi \Rightarrow \psi$. $K\varphi$ is true at w just in case every relevant alternative to φ in $\mathbf{R}$ is ruled out by E_w (we'll return to what *ruling out* means). For now, we leave the semantics of $\texttt{Restr}(\varphi, \psi)$ as unsettled. We express that φ is true at w in $\mathcal{M}$ with $\mathcal{M}, w \vDash \varphi$.

In what follows, we essentially study sub-classes of RA models that can be associated with restrictions on the relevancy function $\mathbf{R}$. Relative to a sub-class, one may identify the *valid* sentences, together constituting a *theory of knowability*: that set of sentences such that each is true at *every* world in *every* model in that sub-class of models. If I say that a sentence containing meta-variables such as φ and ψ is valid, I mean that that sentence *schema* is valid i.e. that sentence is valid with φ and ψ replaced with arbitrary sentences.

3.4.1 Criterion 1: avoidance of egregious violations

A theory witnesses an *egregious violation* of closure if it renders as invalid a specific *instance* of closure whose validity is uncontroversial common ground between closure proponents and opponents.[21] I take this condition as fulfilled when the principle in question is both (i) highly intuitive and (ii) cannot be used in any obvious way to construct a skeptical paradox along the lines of section 3.2.2. Our exact criterion for success will be that a theory that rejects unrestricted closure must *validate* the following dialectically important test cases:[22]

[21] I owe the terminology of "egregious violation" and "egregious non-violation" to Yablo [2014].

[22] Is there room to sensibly resist the validity of conjunctive distribution, supposing one has already embraced closure denial (on the basis of skeptical paradoxes)? Consider the following instance (pointed out by an anonymous reviewer): $K(h \wedge (h \vee \neg b)) \to K(h \vee \neg b)$, where, as usual, h is mundane and b is a skeptical hypothesis (note that expressions of this form are particularly pertinent to our evaluation of Yablo's theory: see the proof of proposition 3.5.4). Note the following: $h \wedge (h \vee \neg b)$ is logically equivalent to h and so expresses a mundane, "lightweight" proposition. On the other hand, $h \vee \neg b$ is logically equivalent to $\neg b$ and so expresses an explicitly anti-skeptical, "heavyweight" proposition. A closure denier may be tempted, for these reasons, to propose that $h \wedge (h \vee \neg b)$ is knowable, while $h \vee \neg b$ is not. To stifle this temptation, consider first an example where our intuitions are clearer: $K(h \wedge \neg(\neg h \wedge b)) \to K\neg(\neg h \wedge b)$. This is the claim that if it is knowable that one has hands and is not a handless brain-in-vat, then one is in a position to know that one is not a handless brain-in-vat. Note that $h \wedge \neg(\neg h \wedge b)$ is equivalent to the mundane h and $\neg(\neg h \wedge b)$ is equivalent to the heavyweight $\neg b$. Should we therefore conclude,

- **Conjunctive distribution:** $K(\varphi \wedge \psi) \rightarrow K\varphi \wedge K\psi$ (cf. Kripke [2011]).
- **Conjunctive weakening**: $K(\varphi \wedge \psi) \rightarrow K(\varphi \vee \psi)$ (cf. [Holliday, 2013, sect. 6.2.1]).

3.4.2 Criterion 2: avoidance of egregious non-violations

An instance of closure is an *egregious non-violation* (for the closure opponent) just in case it can be used to construct a skeptical paradox that intuitively is essentially equivalent to that of section 3.2.2. Our exact criterion for success will be that a theory *invalidates* the following dialectically important test cases:

- **Conjunctive negation:** $K\varphi \rightarrow K\neg(\neg\varphi \wedge \psi)$
- **Conjunction addition:** $(K\varphi \wedge \varphi \Rightarrow \psi) \rightarrow K(\varphi \wedge \psi)$

(In what follows h= *a has hands* and b = *a is a brain-in-a-vat.*) For the closure opponent, the first counts as an egregious non-violation since she should presumably accept the following counter-example: set $\varphi = h$ and $\psi = \neg b$. Consider the paradox: *I know I have hands*; *I do not know that I am not a handless brain-in-vat*; *having hands entails not being a handless brain-in-vat.*

The second principle counts as an egregious non-violation for we have, for the closure opponent, the following purported counter-example: $\varphi = h$ and $\psi = \neg b$. Consider the paradox: *I know I have hands*; *I do not know both that I have hands and am not a brain-in-vat*; *having hands entails not being a brain-in-vat.*

I emphasize: while a convergence of intuition concerning validity served as one basis for the prior list of egregious *violations*, I do *not* appeal to convergence

as closure deniers, that while $(h \wedge \neg(\neg h \wedge b))$ is knowable, $\neg(\neg h \wedge b)$ is not? No. Rather, we should not accept $K(h \wedge \neg(\neg h \wedge b))$. For consider the paradox (cf. the motivating paradoxes for criterion 2): *I know I have hands*; *I do not know both that I have hands and am not a handless brain-in-vat*; *having hands entails having hands and not being a handless brain-in-vat.* This is on a par with the standard skeptical paradox, so the closure denier ought to treat like as like, and therefore accept $\neg K(h \wedge \neg(\neg h \wedge b))$. Now, return to the proposed counter-example to conjunctive distribution that we started with. On the uncontroversial supposition that knowability claims are closed under applications of De Morgan's law, it follows that $h \wedge (h \vee \neg b)$ and $h \wedge \neg(\neg h \wedge b)$ are equivalent with respect to knowability. Thus, the closure denier should conclude that the former is not knowable after all. Incidentally, I see no motivation for a closure denier to reject closure under applications of De Morgan's laws: for what skeptical paradox can be generated merely through such applications?

of intuition concerning *non-validity* as the basis for our list of egregious *non*-violations. There is no such immediate intuition: at least prior to the construction of skeptical paradox, most, I believe, would find a claim of validity for these principles to be wholly unremarkable. However, a closure opponent should should treat such principles as on a par with unrestricted closure: their initial appeal must be weighed against the apparent (pre-theoretic) fact that they can be used to construct skeptical paradoxes. The approach of the closure opponent is to propose that the latter trumps the former. This conclusion motivates our list.

3.4.3 Criterion 3: explicit closure restriction

The invalidity of unrestricted closure is an immediate consequence of satisfying criterion 2. However, we require further that a *promising* form of *restricted closure* must be satisfied by an adequate theory:

Restricted closure: $(K\varphi \wedge \varphi \Rightarrow \psi \wedge \texttt{Restr}(\varphi, \psi)) \rightarrow K\psi$

Settling the requirement of *promise*, I propose, amounts to offering a truth clause for $\texttt{Restr}(\varphi, \psi)$ such that (i) this truth clause has an intuitive reading that plausibly affords deduction a satisfying scope in ordinary discourse, or, more generally, discourse concerning contingent empirical propositions (I leave this desideratum deliberately vague) and (ii) the validities of criterion 1 are a *consequence* of the restricted closure principle in question (hence, it has *explanatory value*, accounting for the desirable behavior of the theory under evaluation).

3.4.4 Criterion 4: preservation of mathematical inquiry

Consider the following two principles.

- **Closure via necessity I:** $(K\varphi \wedge \Box\varphi \wedge \varphi \Rightarrow \psi) \rightarrow K\psi$
- **Closure via necessity II:** $(K(\Box\varphi) \wedge \varphi \Rightarrow \psi) \rightarrow K(\Box\psi)$

The first principle says that knowability extends to the consequences of necessary truth. Note that it has the form of a restricted closure principle (with restriction $\Box\varphi$).[23] The second principle says that if a necessary truth is knowable, then so too is it knowable that its consequences are necessary. This principle meets our definition of an *instance* of closure. Of course, theories that validate such principles may involve different philosophical commitments depending on how we interpret the necessity at issue. Continuing the trend of reading $\Box$ as *epistemic* necessity (i.e. *a priority*), these principles may be understood, in the context of the closure-denying theories we next discuss, as stating that *a priori* knowability is closed under entailment. In harmony with this understanding, I leave it to the reader to check that, for each of these theories, $\Box\varphi$ holds at w only if $K\varphi$ holds (even when $E_w = W$ i.e. the agent has *no* empirical information). Further, it may be shown that the above two principles are *equivalent* in the context of these theories. Further, it may be shown that each theory *validates* these principles.

I propose that it is a mark in favor of a theory if it validates closure via necessity, a happy fact for the closure-denying theories discussed in this chapter. For not only are the above principles intuitive, but they imply that mathematical knowledge via deduction is secure (so long as mathematical truth is necessary). I leave discussion of criterion 4 at that.[24]

[23]This restriction need not match up with the most promising restrictions for satisfying criterion 3.4.3 i.e. it need not be that $\texttt{Restr}(\varphi, \psi)$ holds just in case $\Box\varphi$ holds. Indeed, the left to right direction is undesirable, since $\texttt{Restr}(\varphi, \psi)$ is intended to maintain deductive reasoning as a resource for extending knowledge of contingencies. At any rate, I see no difficulty in positing multiple restricted closure principles, so long as each is robust enough to unify knowledge by deduction in some significant domain.

[24]The validity of the above two principles is no surprise if $\Box\varphi \to (K(\varphi) \wedge K(\Box\varphi))$ is valid, as it is in the theories that we next discuss. This last validity deserves two comments. First, those influenced by Kripkean examples to reject modal rationalism will balk at this validity if the necessity at issue is, say, *metaphysical necessity*. For they will think: though Hesperus is necessarily identical to Phosphorus, it does not follow that it is always knowable that Hesperus is identical to Phosphorus. Fortunately, then, the necessity at issue in this chapter is more naturally read as *epistemic necessity* or *logical necessity*. Second, this validity points to an interesting divergence in how *logical omniscience* is regarded in a logic of knowability, rather than knowledge. For: the validity in question may be regarded as capturing a type of logical omniscience and is therefore undesirable in a logic of knowledge. For a logic of knowability, on the other hand, this validity is a natural desideratum (unless, of course, one is prepared to deny that logical or epistemic necessities are always knowable in principle, no matter the empirical information).

3.4.5 Roush on conjunctive negation

Above, I argued that the closure denier should reject the following principle, for reasons that parallel her rejection of unrestricted closure (namely, the avoidance of skeptical paradox).

> **Conjunctive negation:** $K\varphi \rightarrow K\neg(\neg\varphi \wedge \psi)$

Our assessment is at odds with Roush [2010] and, along related lines, [Wright, 2014, pp.234-235]. Here is an extensive quote:

> If we assume I know that I have a hand, then we should not have the slightest hesitation to credit me with knowledge that I am not a handless brain in a vat. No appeal to the closure principle is needed to support this conclusion. The claim is independently obvious because that you are not a handless brain in a vat is just not much to know. If we know that someone has hands then it follows that she is not a handless person with high blood pressure, or a handless victim of child abuse, but this would not give us any assurance that she need not go to a doctor for these conditions. To a person who already knows she has hands these claims say nothing at all about how far she might or might not be susceptible to heart disease or suicide. For this reason they are statements that it is trivially easy to know if you know that you have hands. If I know that I have hands, then in virtue of that I know I am not a handless anything [Roush, 2010, pg.245].

Thus, Roush [2010] holds that knowing that one is not a handless brain-in-vat is a simple matter for ordinary agents, with little epistemic import. This stance offends our immediate intuitions. Intuitively, it is of *great* philosophical significance to know that one is not a handless brain-in-vat. Of course, being a handless brain-in-vat should not be confused with the more general property of being a brain-in-vat *simpliciter* (which allows for the possibility of one's entire body being envatted. Though note again that our own usage of 'brain-in-vat' in this dissertation is generally as shorthand for 'a mere brain that is envatted and systematically deceived'). Nor should it be confused with being systematically deceived *simpliciter*, for there is a diversity of such scenarios (compare being deceived by an evil demon to being deceived by aliens through a brain-in-vat mechanism). Nevertheless, to know that one is not a handless brain-in-vat is presumably to have ruled out one important *type* of skeptical scenario. Indeed, our intuitive sense as ordinary knowers, it seems to me, is that the prospect of ruling out *any* scenario of this type is dim. This is why, presumably, Cartesian skepticism can be introduced to a classroom of first-time epistemology students via *either* an evil demon scenario, *or* a *limbless* brain-in-vat scenario *or* a brain-in-vat scenario *simpliciter*. The import of these scenarios is identical.

Indeed, Roush [2010] *agrees* that the following intuitively strikes one as a skeptical paradox: 'I know I have hands', 'having hands entails not being a handless brain-in-vat', 'I do not know that I am not a handless brain-in-vat'. Her strategy is to explain the intuition away:

> If I am right, then why have we been under the impression all this time that the adjusted conclusion "I am not a handless brain in a vat" is nontrivial? One reason is that philosophers are like all human beings in being susceptible to associational "thinking", that is, in drawing conclusions that have not been stated, purely on the basis of the proximity of words to one another. All people are sometimes victims, for example, of the devices of highly trained advertising agencies that do psychological research on how we are moved by associations. There was an ad recently that said, above a vivid picture of a train, "Legally, we can't say you can throw it under a train", of the TOUGHBOOK laptop computer. The ad did not assert that you can throw it under a train (and have it survive), but because precisely that clause was inscribed - see the original sentence - an exaggerated impression was created, in just about everyone I would venture, of just how tough the TOUGHBOOK is. Similarly, the words of our adjusted conclusion are "I am not a . . . brain in a vat", and this created a strong impression that this sentence without the ellipses had been asserted, or at least that some information was conveyed about this matter. Philosophers are not immune to such unconscious mistakes; we are all apt to make them when our conscious attention is directed elsewhere [Roush, 2010, pg.246].

In short, Roush [2010] claims that the unease generated by the skeptical paradox is produced by seductive but fallacious reasoning. We conclude ignorance of

$\neg(\neg h \wedge b)$ from ignorance of $\neg b$. This is an instance of concluding ignorance of $\neg(\varphi \wedge \psi)$ from ignorance of $\neg\psi$. But to reason like this, in general, is to abuse modus tollens: the knowability of $\neg(\varphi \wedge \psi)$ does *not*, in general, imply knowability of $\neg\psi$, since $\neg\psi$ is not a tautological implication of $\neg(\varphi \wedge \psi)$ (cf. [Wright, 2014, pg. 235]).

I find this diagnosis unconvincing, for four reasons.

First: philosophers and their students are, in many cases, peculiarly astute when it comes to logic, so it is implausible that this community should uniformly fall prey to a basic error of propositional reasoning. For my part, the invalidity of $\neg(\varphi \wedge \psi) \to \neg\psi$ is striking and immediate. But noticing this does nothing to quell the intuition that no ordinary agent can know that they are not a handless brain-in-vat.

Second: the unknowability of 'I am not a handless brain-in-vat' is supported by more than bare intuition. Indeed, one may deploy the standard rationale for treating skeptical possibilities as *live*: my basic empirical information (i.e. my experiences and my memories of those experiences) is compatible with both a handless brain-in-vat scenario and with the actual scenario (in which I have hands).[25]

Third: the discussion in this chapter is, I take it, representative of typical discussions of epistemic closure in the literature, insofar as a 'brain-in-vat' ascription is shorthand for being a *mere* brain that is envatted and thereby systematically deceived. Thus, if the goal is to intuitively evaluate whether $K(\neg(\neg h \wedge b)) \to K(\neg b)$ is valid, it is misleading in our context to emphasize that $\neg(\varphi \wedge \psi) \to \neg\psi$ is not a tautology. Given the reading of 'brain-in-vat' that is relevant for this chapter, it is *obvious* that one can deduce that 'a is not a brain-in-vat' from 'a is not a handless brain-in-vat'. Indeed, these claims are *logically equivalent* on this reading. Of course, this equivalence is not a matter of propositional logic (i.e. a product of the meaning of the connectives). Rather, it is a product of the meaning we have attached to 'brain-in-vat'.

Fourth, and finally: Roush [2010] claims that "no appeal to the closure principle is needed" to be convinced that being a "handless brain in a vat is just not much to know". This is correct: one need not accept unrestricted closure (i.e. closure in full generality) to be convinced of the aforementioned. However, it is

[25] Avnur et al. [2011] make this point in a critical discussion of Roush [2010].

clear that the intuitive appeal of "[i]f I know that I have hands, then in virtue of that I know I am not a handless anything", and so that being a "handless brain in a vat is just not much to know", amounts exactly to the intuitive appeal of the following *instance* of closure: $K\varphi \to K\neg(\neg\varphi \wedge \psi)$ i.e. conjunctive negation. I agree with Roush [2010] that intuition counsels that this principle is valid. However, to *thereby* conclude that the principle must be accepted without further ado is to beg the question against the advocate of restricted closure, who is prepared to trade off the preservation of such intuitions in order to chart an escape from skepticism. Put another way: the restricted closure advocate allows the intuitive validity of an instance of closure to be trumped by the intuitive force of a skeptical paradox (i.e. by a seeming counter-example to that instance).

In total, our discussion of Roush [2010] reinforces the point in the main text: that there is a striking (and unsurprising) symmetry between the skeptical paradox generated by conjunctive negation and the paradox generated by unrestricted closure (sect. 3.2.2). Conjunctive negation, framed in the abstract, appears to us as valid, yet we have independent support for holding both that Kh and that $\neg K(\neg(\neg h \wedge b))$. *Mutatis mutandis*, the closure denier ought to reject conjunctive negation.

3.5 Evaluating Schaffer and Yablo

I now evaluate Schaffer and Yablo's proposals. For each, I first sketch a formal framework (I understand each as proposing a *constraint* on the class of "legitimate" RA models).

3.5.1 Schaffer

Define a *S-model* to be a tuple $\langle W, \{E_w\}_{w \in W}, \mathbf{Q}, \mathbf{V} \rangle$ with elements as in an RA model, except $\mathbf{Q}$ is a function accepting a sentence φ and returning an ordered pair of propositions $\langle T_\varphi, A_\varphi \rangle$. Call T_φ the *thesis* and A_φ the *anti-thesis* (or *contrast*) to φ. I stipulate that $T_\varphi \cap A_\varphi = \emptyset$, i.e. that the thesis and anti-thesis associated with φ are disjoint. Further, I stipulate that $\mathbf{V}$ has the property that $\mathbf{V}(p) = T_p$ for every propositional atom p.

Intuitively: $\mathbf{Q}$ assigns to each sentence the *natural question* associated with that sentence, in context. An S-model generates an RA model: we simply take

$\mathbf{R}(\varphi) = \{T_\varphi, A_\varphi\}$ for every φ. We then deploy RA semantics: $K\varphi$ holds at w in $\mathcal{M}$ just in case the anti-thesis A_φ is *ruled out* by E_w.

Following suggestions in Schaffer, I understand *ruling out* in the context of S-models as follows: proposition P is ruled out by evidence E just in case P and E are mutually exclusive i.e. there is no world at which both P and E are true. Call this *comprehensive ruling out.*[26]

A special rationale can be given for adopting comprehensive ruling out in the question-sensitive setting (though note that Schaffer does not offer this rationale, nor, as far as I know, any argument to similar effect). Following a rich literature in linguistics, we may think of a question as a *set of propositions* (answers) [Hamblin, 1958, Ciardelli et al., 2015]. A compelling trend in the recent literature is to take this set as *downward-closed:* if P is an answer to question $\mathbf{Q}$, then any Q that entails P is also an answer to $\mathbf{Q}$. If I ask "In which cardinal direction lies the emerald city?" you can equally well provide a basic answer "The emerald city lies to the east" or one more specific than is required "One that follows the eastwards-bearing yellow-brick road will find the emerald city". Thus, we should, strictly speaking, understand the answers to $\mathbf{Q}(\varphi)$ that are *relevant alternatives* to φ as the *downward closure* of A_φ. Thus, it is natural to propose a notion of ruling out such that *all* relevant alternative answers - even overly specific ones - are ruled out when the contrast A_φ is ruled out. Now note: if E_w is disjoint from A_φ, then E_w is disjoint from every proposition that entails A_φ. That is: if E_w comprehensively rules out A_φ, then E_w comprehensively rules out every P that entails A_φ, and so every answer to the question of φ that is incompatible with φ.

As it stands, the theory generated by the class of S-models is bereft of much content: since we have not offered any general constraints on what questions can be assigned to sentences, it is easy to find a counter-model to just about any proposed validity (including, say, conjunctive distribution). We could leave our evaluation there. But to better explore Schaffer's general approach, natural restrictions on the $\mathbf{Q}$ function could be proposed (though, in doing so, we go beyond Schaffer's explicit commitments). In particular, it is natural to think that the *logical form* of a sentence will play a role in determining the *natural*

[26] What is the effect of replacing comprehensive ruling out with counterfactual ruling out (as introduced in sect. 3.5.2) in the question-sensitive setting? I note here only that, though the details deserve proper scrutiny, I do not think that this move produces a theory that escapes objection. For instance, it can be shown that compositional S-theory with counterfactual ruling out does not validate $K(p \wedge q) \to K(p \vee q)$.

question associated with that sentence. In support of these intuitions, note that constraints along the lines I will propose are fruitfully deployed in leading recent versions of the compositional semantics of questions [Ciardelli et al., 2015].[27]

Suppose (throughout the next few examples) that "p or q?" is the natural question associated with p, where is q is the natural contrast to p. Suppose further that "r or s?" is the natural question associated with r. What then is the natural question associated with $\neg p$? (Likewise: $\neg r$.) Intuitively, the selection is not arbitrary, but rather identical to that of p: "p or q?". Or consider $p \wedge r$. What is the natural question associated with this sentence? Intuitively, it is "$p \wedge r$ or q or s?". Or consider the claim $p \vee r$. What is the natural question here? Intuitively:

[27] There is a second reason for a Schafferian to be attracted to our account of the 'natural question' associated with φ in context: this account allows for a natural interaction between Schaffer's account of epistemic closure and a Schafferian account of the semantics for knowability expressions. This, in turn, provides the Schafferian with tools for accounting for an *apparent* penchant for observing unrestricted epistemic closure with ordinary expressions of knowledge and knowability.

I elaborate. Start with Schaffer's take on the closure debate (see especially Schaffer [2005a] and Schaffer [2007a]). Consider the knowledge relation $\mathbb{K}$. According to the standard treatment, this is a binary relation between an agent and a proposition. On such a conception, unrestricted closure (understood as a principle governing knowledge itself, rather than a validity of our ordinary epistemic *language*) is usefully framed as:

$$\text{If } P \subseteq Q \text{ then } \mathbb{K}(a, P) \text{ entails } \mathbb{K}(a, Q)$$

(Putting aside the contingent cognitive limitations of the agent. As usual, we work with unstructured propositions, largely for simplicity). Schaffer, of course, rejects a binary interpretation of knowledge, and his treatment does *not* validate unrestricted closure framed in the ternary setting:

$$\text{If } P \subseteq Q \text{ then, for any contrasts } R \text{ and } S,\ \mathbb{K}(a, P, R) \text{ entails } \mathbb{K}(a, Q, S)$$

However, Schaffer advocates a restricted closure principle that is better suited for his view:

Contrastive Closure: If $P \subseteq Q$ and $S \subseteq R$ then: $\mathbb{K}(a, P, R)$ entails $\mathbb{K}(a, Q, S)$

In other words, if the truth of Q is guaranteed by that of P, and ruling out every world that contrasts with P ensures that every world that contrasts with Q is ruled out, then the (ideal) agent knows Q (rather than its contrast) when she knows P (rather than its contrast).

(Note that Schaffer tends to separate contrastive closure into two principles, which he calls *Expand-P* and *Contract-Q*. It is easily shown that the conjunction of these principles is equivalent to contrastive closure.)

Ideally, contrastive closure should explain the extent to which binary knowability attributions observe closure. Our account of the 'natural question in context' for φ aids in this. For instance, consider conjunctive distribution: $K(\varphi \wedge \psi) \to K\varphi \wedge K\psi$. As noted, this formula is uncontroversially valid. Now, according to our account, the 'natural' contrast assigned to $\varphi \wedge \psi$ in context (call it R) is the union of the contrast assigned to φ (call it P) with that assigned to ψ (call it Q). Thus, $P \subseteq R$ and $Q \subseteq R$. In this case, contrastive closure ensures that $K(\varphi \wedge \psi) \to K\varphi \wedge K\psi$ is true in context.

"$p \vee r$ or $q \wedge s$?".

To illustrate: consider two questions "Do I have a hand or a stump at the end of my right arm?" and "Is Barack Obama or George W. Bush current president of the US?". Compactly: *Hand? Or: stump?* and *Obama? Or: Bush?*. Having fixed that these are the questions of interest, the natural question associated with *Not hand* is still the question *Hand? Or: stump?*. The natural question associated with *Hand and Obama* is *Hand and Obama? Or: stump? Or: Bush?*. The natural question associated with *Hand or Obama* is *Hand or Obama? Or: stump and Bush?*.

With this in mind, a *compositional S-model* is an S-model satisfying the following *compositional structure*.

- $\mathbf{Q}(\neg\varphi) = \langle A_\varphi, T_\varphi \rangle$, where $\mathbf{Q}(\varphi) = \langle T_\varphi, A_\varphi \rangle$.
- $\mathbf{Q}(\varphi \wedge \psi) = \langle T_\varphi \cap T_\psi, A_\varphi \cup A_\psi \rangle$, where $\mathbf{Q}(\varphi) = \langle T_\varphi, A_\varphi \rangle$ and $\mathbf{Q}(\psi) = \langle T_\psi, A_\psi \rangle$.
- $\mathbf{Q}(\varphi \vee \psi) = \langle T_\varphi \cup T_\psi, A_\varphi \cap A_\psi \rangle$ where $\mathbf{Q}(\varphi) = \langle T_\varphi, A_\varphi \rangle$ and $\mathbf{Q}(\psi) = \langle T_\psi, A_\psi \rangle$.

The resulting *Compositional S-theory* holds attractions for a closure opponent:

3.5.1. Proposition. *Compositional S-theory:*

1. *validates conjunctive distribution:* $K(\varphi \wedge \psi) \rightarrow K\varphi \wedge K\psi$*;*
2. *validates conjunctive weakening:* $K(\varphi \wedge \psi) \rightarrow K(\varphi \vee \psi)$.

Proof:

1. Assume that $\mathcal{M}, w \vDash K(\varphi \wedge \psi)$. Note that $\mathbf{Q}(\varphi \wedge \psi) = \langle T_\varphi \cap T_\psi, A_\varphi \cup A_\psi \rangle$. Thus, $E_w \cap (A_\varphi \cup A_\psi) = \emptyset$. Hence, $E_w \cap A_\varphi = E_w \cap A_\psi = \emptyset$. Thus: $\mathcal{M}, w \vDash K\varphi \wedge K\psi$.
2. Assume that $\mathcal{M}, w \vDash K(\varphi \wedge \psi)$. Note that $\mathbf{Q}(\varphi \wedge \psi) = \langle T_\varphi \cap T_\psi, A_\varphi \cup A_\psi \rangle$. Note also that $\mathbf{Q}(\varphi \vee \psi) = \langle T_\varphi \cup T_\psi, A_\varphi \cap A_\psi \rangle$. Thus, $E_w \cap (A_\varphi \cup A_\psi) = \emptyset$. Hence, $E_w \cap (A_\varphi \cap A_\psi) = \emptyset$. Thus: $\mathcal{M}, w \vDash K(\varphi \vee \psi)$.

$\square$

Unfortunately, it does not meet our criteria for success.

3.5.2. Proposition. *Compositional S-theory validates conjunctive negation:*

$$K\varphi \to K\neg(\neg\varphi \wedge \psi)$$

Proof:
For arbitrary compositional S-model $\mathcal{M}$: $A_{\neg(\neg\varphi\wedge\psi)} = A_\varphi \cap T_\psi$. Since $E_w \cap A_\varphi = \emptyset$ implies that $E_w \cap A_\varphi \cap T_\psi = \emptyset$, we have that $\mathcal{M}, w \vDash K\varphi \to K(\neg(\neg\varphi \wedge \psi))$. □

3.5.2 Yablo

Define a *Y-model* to be a tuple $\langle W, \{E_w\}_{w\in W}, \mathbf{T}, \mathbf{V}\rangle$ with elements as in an RA model, except $\mathbf{T}$ is a function accepting a sentence φ and returning a *Y-topic*: the set of *minimal truth-makers and false-makers* for φ. A minimal truth-maker for φ is a sentence λ that expresses a *basic state of affairs that logically necessitates the truth of* φ: λ is to be understood as a conjunction of *literals* (where a literal is either an atomic proposition or its negation) such that (i) λ tautologically entails φ and (ii) λ is minimal in sense that there exists no conjunction of literals μ ($\mu \neq \lambda$) such that λ entails μ and μ entails φ. A *minimal false-maker* for φ is a minimal truth-maker for $\neg\varphi$. For example: a minimal truth-maker for $p \vee q$ is p. A minimal false-maker for $p \vee q$ is $\neg p \wedge \neg q$.

The minimal truth-makers and false-makers for φ determine a set of propositions *expressed by* those sentences. I willfully overload terminology and call the proposition expressed by a minimal truth-maker a *minimal truth-maker*, and that expressed by a minimal false-maker a *minimal false-maker*. A Y-model generates an RA model: set $\mathbf{R}(\varphi)$ to be the set of minimal truth-makers and false-makers of φ. We now deploy RA semantics: $K\varphi$ holds at w in $\mathcal{M}$ just in case each minimal false-maker for φ is ruled out by E_w.

As Yablo notes, *comprehensive ruling out* will not do as a notion of *ruling out* for this semantics. For the *union* of the minimal false-makers for φ is identical to the set of worlds at which φ is false. In this case, for E_w to *comprehensively* rule out the minimal false-makers of "a has hands" would involve E_w being inconsistent with skeptical alternatives such as "a is a brain-in-vat" (since the worlds at which the latter hold are subset of those at which "a does not have a hand" holds). But we have assumed that we do *not* have basic empirical evidence of *that* kind.

I therefore follow Yablo in deploying a *subjunctivist* notion of *ruling out* P with evidence E: if P were the case, then the evidence E would *not* be the case.[28] Call this *counterfactual ruling out.* Following standard Stalnaker-Lewis semantics, we paraphrase this condition as follows: the nearest P-worlds to actuality are inconsistent with E.

This is an appropriate notion of *ruling out* in a topic-sensitive setting. A subject matter is intended to divide logical space into those propositions that capture the *distinctions* with which that subject matter is concerned: two worlds are in one such proposition just in case they are identical *with respect to the subject matter.* In this case, propositions that carve up logical space at *finer grain* than the subject matter in question - i.e. propositions that involve distinctions not captured by the subject matter - are, in some sense, to be ignored as *irrelevant.* Thus, if P entails Q, where Q is on-topic but P is not, we do not necessarily require P to be ruled out when Q is. Counter-factual ruling out delivers this result.

Thus we work with *ordered Y-models* $\mathcal{M}$ (likewise: ordered RA models), equipped with an ordering $\preceq_w$ for each world w in W (with the stipulation that w is the nearest world to itself). Call the theory generated by an RA semantics on this class of models *Y-theory.*

Y-theory holds attractions that set it apart from compositional S-theory.

3.5.3. Proposition. *Y-theory:*

1. *does not validate conjunction addition:* $(K\varphi \wedge \varphi \Rightarrow \psi) \rightarrow K(\psi \wedge \varphi)$*;*

2. *does not validate conjunctive negation:* $K\varphi \rightarrow K\neg(\neg\varphi \wedge \psi)$*.*

Proof:

[28] This departs from Yablo's own treatment (at least as developed in Yablo [2012], in contrast to a non-committal stance in chapter 7, footnote 6 of Yablo [2014]), which is similarly subjunctivist, but rather requires that the truth of P be tracked by the agent's *belief*, according to a Nozickian *sensitivity* condition. Our own treatment is inspired by the notion of a *conclusive reason* due to Dretske [1971], and better fits discussion of knowability on the evidence. The technical results for Yablo's theory can be maintained if we switch to a belief-tracking notion of ruling out.

1. Counter-example: $(Kp \wedge p \Rightarrow q \rightarrow K(p \wedge q)$. The minimal false-maker for p is $\neg p$, and those for $p \wedge q$ are $\neg p$ and $\neg q$. Counter-model: the nearest $\neg q$ world is compatible with the evidence; the proposition expressed by $\neg q$ is a proper subset of the proposition expressed by $\neg p$; the nearest $\neg p$ world is inconsistent with the evidence (in this case, this world must be a $\neg p \wedge q$ world).

2. Counter-example: $Kp \rightarrow K\neg(\neg p \wedge q)$. The minimal false-maker for p is $\neg p$, and for $\neg(\neg p \wedge q)$ is $\neg p \wedge q$. So, a counter-model is one where the nearest $\neg p \wedge q$ world is compatible with the evidence, but the nearest $\neg p$ world is not.

□

Again, however, our criteria for success are not met (the following results are essentially shown to hold by [Holliday, 2013, sect. 6.2.1]).

3.5.4. Proposition. *Y-theory:*

1. *does not validate conjunctive distribution:* $K(\varphi \wedge \psi) \rightarrow K\varphi \wedge K\psi$;

2. *does not validate conjunctive weakening:* $K(\varphi \wedge \psi) \rightarrow K(\varphi \vee \psi)$.

Proof:

1. Counter-example: $K(p \wedge (p \vee q)) \rightarrow K(p \vee q)$. The minimal false-maker for $p \vee q$ is $\neg p \wedge \neg q$, and for $p \wedge (p \vee q)$ is $\neg p$. A counter-model: set the nearest $\neg p$ world to be a $\neg p \wedge q$ world that is incompatible with the evidence. Set the nearest $\neg p \wedge \neg q$ world to be compatible with the evidence.

2. Counter-example: $K(p \wedge q) \rightarrow K(p \vee q)$. The minimal false-maker of $p \vee q$ is $\neg p \wedge \neg q$, and those for $p \wedge q$ are $\neg p$ and $\neg q$. A counter-model: set the nearest $\neg p \wedge \neg q$ world to be more distant than the nearest $\neg p$ and $\neg q$ worlds.

□

3.6 Positive proposal: resolution theory

I turn to resolution semantics, which may be classed as a novel variation on the topic-sensitive approach. Instead of identifying the subject matter of φ with the minimal truth-makers and false-makers for φ, it is identified with a set $\mathsf{T}(\varphi)$ of atomic predications determined by the atoms that appear in φ, and represented by the basic *ways things can be with respect to those atomic predications* i.e. the set of those propositions represented by a conjunction of literals that contains all and only atoms in $\mathsf{T}(\varphi)$.

(I present a simplified version of resolution semantics that ignores the neo-Fregean aspects of the full account, as discussed in section 1.2 of chapter 1. We motivate and incorporate these aspects in chapter 4, section 4.5.)

3.6.1 Basic ideas and forerunners

Consider two leading ideas: (i) the subject matter of a sentence φ is, in context, solely a function of the atomic predications that are expressed by φ (thus, distinct sentences have the same subject matter if they contain the same atoms); (ii) metaphorically, a subject matter imposes a *resolution* on logical space: 'focus' on a subject matter involves a 'focus' on the distinct ways things can be with respect to that subject matter, 'ignoring' distinctions that are 'invisible' (i.e. irrelevant) to that topic.

A diagnosis of the odd consequences of Yablovian subject matter suggests that (i) is correct. Y-theory, recall, invalidates $K(p \wedge q) \rightarrow K(p \vee q)$. A strange result, for a topic-sensitive approach to knowledge should allow that closure is maintained in cases where no new subject matter is introduced in the inference from premise to conclusion. But, intuitively, $p \wedge q$ and $p \vee q$ involve the same subject matter.[29] In particular, if I say "either Mary is a lawyer or Mary is a professor", I am, intuitively, speaking *on the same topic* as if I were to say

[29] For those unsatisfied with a direct appeal to our intuitions concerning disjunctions, consider the following argument based on three unremarkable assumptions. First, p and $\neg p$ have the same subject matter, as do q and $\neg q$, on the assumption that negations preserve subject matter. Thus, $p \wedge q$ and $\neg p \wedge \neg q$ have the same subject matter, on the assumption that conjunctions combine the subject matter of their conjuncts in a uniform manner. Thus, $p \wedge q$ and $\neg(\neg p \wedge \neg q)$ have the same subject matter. Thus, $p \wedge q$ and $p \vee q$ have the same subject matter, on the assumption that disjunction can be *defined* in terms of negation and conjunction, in the standard manner: $\varphi \vee \psi ::= \neg(\neg\varphi \wedge \neg\psi)$ (that is, the meaning of a disjunctive expression is exhausted by a disjunction-free expression containing only conjunctions and negations).

"Mary is a lawyer and Mary is a professor" (namely: Mary's profession). Only my *commitments* concerning that topic differ. But Yablo's account entails that claims of the form $p \wedge q$ and $p \vee q$ must have *different* subject matter.[30][31]

(ii) has similar intuitive appeal. If my chosen subject matter is Mary's profession, then I am 'alive' to the distinction between Mary being a lawyer and being a doctor, but 'ignore' the distinction between Mary being married or divorced (or, for that matter, between Fido being a dog and being a cat). If our conversation centers on Mary's profession, and you suddenly bring up Mary's marital status (or Fido's species), I will accuse you of either being off-topic, or changing the topic.[32] Put metaphorically, certain ways things could be are 'visible' (i.e. relevant) to a subject matter, while other ways things could be are 'invisible' (irrelevant) to that subject matter. This suggests that subject matter generates a *partition on logical space*: a set of mutually disjoint and exhaustive propositions (the *cells* of the partition) capturing distinct ways that things can be *with respect to that subject matter*. A sentence is then exactly 'on-topic' only if it expresses a cell in the partition, or a union of such cells; it is (to some extent) 'off-topic' if it expresses a proposition properly contained in a cell, or properly contained in a union of cells.[33] Thus a 'resolution' is imposed by that subject matter on logical

[30] According to Yablo's account of inclusion [Yablo, 2014, pg.46], the latter's subject matter is not even *included* in that of the former, for it is not the case that the minimal falsemaker for the latter (i.e. $\neg p \wedge \neg q$) is implied by a minimal falsemaker for the former (neither $\neg p$ nor $\neg q$) .

[31] Yablo's account may not be alone in this regard. For instance, for Lewis [1988b], a subject matter is a partition of logical space (or, at least, a member of a special class of partitions), and a proposition P is about a subject matter just in case the truth value of P supervenes on that subject matter i.e. the subject matter in question is a refinement of the partition comprised of P and the complement of P. It follows that the propositions expressed by "Mary is a lawyer and Fido is a dog" and "Mary is a lawyer or Fido is a dog" concern different subject matters (at least on the assumption, say, that the former and its complement make up a subject matter, for then the former proposition is trivially about this subject matter, while that expressed by "Mary is a lawyer or Fido is a dog" is not).

[32] In some cases, this change may be an unobjectionable enrichment: for example, I might say "Mary is a lawyer" and you add "Mary is a *successful* lawyer".

[33] Some utterances are correctly described as *partially* off-topic: if we are discussing Mary's profession and you utter "Mary is a *successful* lawyer", then it is natural to say that your statement is on-topic to the extent that it speaks to the matter of Mary's profession, and off-topic to the extent that it speaks to the matter of Mary's success (though if your refinement of the topic of conversation is welcome, we are more inclined to describe your utterance as *enriching* the topic, rather than departing from it). Another example: suppose that the topic is Mary's success as a professional, and you utter "Mary is a *married* lawyer". Here, the subject matter of your utterance seems to be an *enriching* of the sub-topic of Mary's profession to the extent that it speaks to both the matter of her profession and her marital status, and an *ejection* of the sub-topic of Mary's success.

space.

This suggests an RA semantics: suppose we classify a *relevant alternative* to φ as a cell in the partition generated by φ's subject matter that entails $\neg\varphi$; that is, an alternative to φ that is *visible* to that subject matter. Then to be in a position to know φ is to have evidence that rules out every such 'on-topic' alternative to φ.

Natural ideas inevitably have forerunners. Consider *relatedness logic* as studied by [Epstein, 1994, Part III]. Epstein develops a formal account of subject matter where shared subject matter is a consequence of shared propositional variables, and binary logical connectives share the single function of *combining* subject matter. So, Epstein joins us in accepting (i).

Another forerunner: David Lewis [1988a] first proposes that we identify a subject matter with a partition on logical space, with two worlds in the same cell just in case they are indistinguishable with respect to the topic in question. So, Lewis joins us in accepting (ii).

3.6.2 Resolution semantics

Define an *R-model* to be a tuple $\langle W, \{E_w\}_{w\in W}, \mathsf{T}, \mathbf{V}\rangle$ like an RA model, except T is a function returning an *R-topic* for a sentence φ: a set of atomic propositions from our language.

Such a representation of subject matter is natural. Consider "Mary is a lawyer". The subject matter of this sentence is *Mary's profession*, which, in ordinary contexts, corresponds to a set of 'on-topic' basic predications: *Mary is a lawyer; Mary is a medical doctor; Mary is a cashier; . . . .* (So a subject matter generates a *question*: *is Mary a lawyer or a doctor or a cashier or . . . ?*).

T has the following constraints: if p is an atom, then $p \in \mathsf{T}(p)$ i.e. atomic predication p is always part of its own subject matter. Furthermore: $\mathsf{T}(\varphi)$ is the *union* of the R-topics assigned to the *atoms* and *knowability expressions* in φ. For example: if $\mathsf{T}(p) = \{p\}$ and $\mathsf{T}(q) = \{q, r\}$, then $\mathsf{T}(p \vee q) = \{p, q, r\}$. Connectives merely *combine* subject matter.

A *state description with respect to* T is a conjunction of literals λ, such that all and only the atoms in T appear. For instance, the state descriptions for $\mathsf{T} = \{p, q\}$ are $p\wedge q$, $p\wedge\neg q$, $\neg p\wedge q$ and $\neg p\wedge\neg q$. For any R-model, the set of state descriptions for T forms a *partition* on logical space W: the set of propositions expressed by those state descriptions. Call a cell in this partition a *way for* T *to be.* Now, an R-model generates an RA model as follows: set the relevant alternatives $\mathbf{R}(\varphi)$ for φ as the ways for $\mathsf{T}(\varphi)$ to be that entail that φ is false.

For similar reasons to Y-semantics, we work with *counterfactual ruling out* and *ordered* R-models. Altogether, we arrive at resolution semantics for knowability attributions: $K\varphi$ holds in $\mathcal{M}$ at w just in case, for every way for $\mathsf{T}(\varphi)$ to be, if that way entails $\neg\varphi$, then E_w is inconsistent with the nearest worlds to w at which things are that way.

3.6.3 Evaluation

3.6.1. THEOREM. *For resolution semantics: if φ tautologically entails ψ, then $K\varphi \to K\psi$ is valid if $atoms(\psi) \subseteq atoms(\varphi)$. Further, if φ tautologically entails ψ, φ is consistent and ψ is not a tautology then: $K\varphi \to K\psi$ is valid only if $atoms(\psi) \subseteq atoms(\varphi)$.*

Proof:
The first part is trivial. We prove the second by contraposition: assume that $atoms(\psi) \nsubseteq atoms(\varphi)$. We construct resolution model $\mathcal{M}$ such that $\mathcal{M}, w \vDash K\varphi \wedge \neg K\psi$. Let $\lambda^+ = \lambda^+_\varphi \wedge \lambda^+_\psi \wedge \lambda^+_{\varphi,\psi}$ be some conjunction of literals that contains all and only the atoms that occur in one or both of φ and ψ, and such that λ^+ tautologically entails φ (and so also entails ψ). In particular, let λ^+_φ contain all and only the atoms that occur in *only* φ; let λ^+_ψ contain all and only the atoms that occur in *only* ψ; let $\lambda^+_{\varphi,\psi}$ contain all and only the atoms that occur in *both* φ and ψ. Further, let $\lambda^- = \lambda^-_\varphi \wedge \lambda^-_\psi \wedge \lambda^-_{\varphi,\psi}$ be some conjunction of literals that contains all and only the atoms that occur in one or both of φ and ψ, and such that λ^- tautologically entails $\neg\psi$ (and so also entails $\neg\varphi$). In particular, let λ^-_φ contain all and only the atoms that occur in *only* φ; let λ^-_ψ contain all and only the atoms that occur in *only* ψ; let $\lambda^-_{\varphi,\psi}$ contain all and only the atoms that occur in *both* φ and ψ. Now, let $W = \{w_1, w_2, w_3\}$. Set valuation $\mathfrak{v}$ so that λ^+ holds at w_1 (and so φ holds at w_1); λ^- holds at w_2 (and so both $\neg\psi$ and $\neg\varphi$ hold at w_2); and $\lambda^-_\varphi \wedge \neg\lambda^-_\psi \wedge \lambda^-_{\varphi,\psi}$ holds at w_3. Note that since λ^- tautologically entails $\neg\varphi$ and λ^-_ψ contains no atoms from φ, it must be that $\lambda^-_\varphi \wedge \lambda^-_{\varphi,\psi}$ entails $\neg\varphi$, and so $\neg\varphi$ holds at w_3. Note then that $\{w_2, w_3\}$ is a (and the only) relevant alternative proposition to φ in this model (since it is the proposition expressed

by $\lambda^-_\varphi \wedge \lambda^-_{\varphi,\psi}$) but $\{w_2\}$ is not. Note further that $\{w_2\}$ is a relevant alternative to ψ (expressed by $\lambda^-_\psi \wedge \lambda^-_{\varphi,\psi}$). Finally, set $E_{w_1} = \{w_1, w_2\}$ and set $\preceq_{w_1}$ so that $w_3 \preceq_{w_1} w_2$. Thus, the nearest $\lambda^-_\psi \wedge \lambda^-_{\varphi,\psi}$ world is compatible with the evidence at w_1, while the nearest $\lambda^-_\varphi \wedge \lambda^-_{\varphi,\psi}$ is not. In total: $\mathcal{M}, w_1 \vDash K\varphi \wedge \neg K\psi$. □

Some immediate corollaries of interest:

3.6.2. Corollary. *Resolution theory:*

1. *validates conjunctive distribution:* $K(\varphi \wedge \psi) \rightarrow K\varphi \wedge K\psi$*;*
2. *validates conjunctive weakening:* $K(\varphi \wedge \psi) \rightarrow K(\varphi \vee \psi)$*;*
3. *invalidates conjunctive negation:* $K\varphi \rightarrow K(\neg(\neg\varphi \wedge \psi))$*.*

Note that the third result establishes that unrestricted closure is *not* valid for resolution theory. To satisfy criterion 2, it remains to check one last desideratum.

3.6.3. Proposition. *Resolution theory invalidates conjunction addition:*

$$K\varphi \wedge \varphi \Rightarrow \psi \rightarrow K(\varphi \wedge \psi)$$

Proof:
Counter-model: construct $\mathcal{M}$ so that $p \Rightarrow q$ holds (so no $p \wedge \neg q$ worlds); so that the nearest $\neg p$ world is a $\neg p \wedge q$ world incompatible with E_w; and so that the nearest $\neg p \wedge \neg q$ world is compatible with the evidence. Then: Kp holds but not $K(p \wedge q)$. □

Altogether, resolution theory satisfies criteria 1 and 2 of sect. 3.4. All that remains is to give a promising *restriction* on closure (criterion 3). Set the truth clause for $\texttt{Restr}(\varphi, \psi)$ as: $\mathsf{T}(\psi) \subseteq \mathsf{T}(\varphi)$. It is then immediate that resolution semantics validates

$$(K\varphi \wedge \varphi \Rightarrow \psi \wedge \texttt{Restr}(\varphi, \psi)) \rightarrow K\psi.$$

Informally: according to resolution theory, knowability is extended via deductive consequence if the subject matter of the consequent is necessarily contained in that of the antecedent. Closure holds when we fix the subject matter (cf. [Yablo, 2014, sect.7.8]).

3.6.4 Coda: further principles of interest

Resolution theory's character is further revealed by considering three further principles.

- **Disjunction Introduction:** $K\varphi \to K(\varphi \vee \psi)$
- **Modus Ponens:** $K(\varphi \wedge (\varphi \to \psi)) \to K\psi$
- **Conjunction Introduction:** $(K\varphi \wedge K\psi) \to K(\varphi \wedge \psi)$

Resolution theory invalidates disjunction introduction and validates modus ponens (consequences of Theorem 3.6.1). Further, it invalidates conjunction introduction (by reductio: if both conjunction introduction and modus ponens were valid, then unrestricted closure would be too. But we have already established the latter invalidity).

A closure opponent ought to handle the above principles with care. Disjunction introduction is an instance of closure that many - including [Nozick, 1981, pg.236] - feel carries particularly acute intuitive weight, so discarding it is a fatal error [Hawthorne, 2004, Kripke, 2011, Holliday, 2013]. Modus ponens is a version of 'closure under implication', so it is not immediately obvious what attitude a closure opponent should take to it. Conjunction introduction is related to the perplexities of Makinson's *preface paradox* [Makinson, 1965, Hawthorne, 2004].

It is a point in favor of a theory with closure rejection if it not only settles the status of the above principles, but deflates the surrounding controversy with natural explanations for this status. Resolution theory accomplishes this. A topic-sensitive theorist identifies disjunction introduction as a paradigmatically worrisome instance of closure, for the disjunct ψ may well introduce subject matter beyond that of φ. That modus ponens is preserved by resolution theory can be understood as an advantage, capturing "as much closure" as anyone could reasonably desire.[34] Finally, that conjunction introduction fails offers an interesting

[34] It need not rattle the resolution theorist that, in an ordinary context, she must conclude that Kh, $K(h \Rightarrow \neg b)$ and $\neg K(h \wedge (h \Rightarrow \neg b))$ all hold (where h is, once again, an ordinary empirical claim and b is a skeptical hypothesis). In fact, this represents a subtle result that the resolution theorist should offer as illuminating: though both h and $h \Rightarrow \neg b$ are knowable on the evidence - the latter *a priori* knowable - it does not follow that these potential pieces of knowledge can be *integrated* and so potentially known *at once*. For the introduction of the subject matter of b focuses on distinctions that the agent's evidence cannot track.

route to resolving the preface paradox: just because various claims are individually knowable does *not* mean, for the resolution theorist, that these claims can be integrated and potentially known as a single, conjunctive claim. For the subject's evidence may not be sensitive to the fine-grained distinctions that arise by *combining* the subject matter of the individual claims.

3.7 Conclusion

We have located a theory of knowability on the evidence that rejects unrestricted closure, and does so in a manner that meets important criteria for success. That theory - resolution theory - may be seen as a version of the topic-sensitive approach to knowability, and may be summarized as follows: φ is knowable on the subject's evidence, in context, just in case that evidence tracks (is *sensitive* to) the network of distinctions captured by the subject matter of φ, in context.

Chapter 4

Gibbard's Cheap Trick

Gibbard's cheap trick is an argument for the conclusion that every fact is knowable a priori, using two essential premises: that the sentence "things are exactly *this* way" conveys, in context, an a priori truth and, second, that a priori knowability is closed under deduction. These premises resist easy rejection, even in the face of well-known strategies for denying the existence of the contingent a priori, or restricting the scope of deduction for extending knowledge. In this chapter, I do four things. First, I present the cheap trick with precision, against a backdrop of minimal assumptions. Second, I chart some of the landscape of possible objections and cast doubt on an interesting sample, bolstering the cheap trick's status as a paradox. Third, I develop a (surprisingly forceful) deflationary strategy for biting the bullet, and argue that it is more convincing than a related response that is advocated by Gibbard (as I read him). Finally, I argue that it is nevertheless valuable to locate an account of *substantive* knowability that does not bite the bullet, and show that resolution theory (equipped, as it is, with a Fregean apparatus) fills this role. To bolster the case for resolution theory, I explain how the Fregean tools motivated by our discussion aid in addressing challenging cases such as *missed clue cases*.

4.1 Introduction

Philosophers typically accept a sharp distinction between a priori and a posteriori knowledge.[1] Further, it is typically accepted that a priority is a significant, rarefied status. In this chapter, I examine an argument - described in Appendix 1 of Gibbard [2012] - for the conclusion that *every* fact is knowable a priori. In light of its logical simplicity, Gibbard notes that the argument has the appearance of a "cheap trick".[2] I thus refer to it as *Gibbard's cheap trick*.

> *The cheap trick:* It is knowable a priori that *this* is the way things are (gesturing at the actual way things are). Further, that *this* is the way things are entails every true proposition (every 'fact'), and a priori knowability is closed under deduction. ∴ Every fact is knowable a priori.

My aim is to explore the significance of the cheap trick: clarify it, illustrate that it is not easily dismissed and propose how best to respond to it. The chapter proceeds in four stages.

In the first stage (section 4.2), I aim to clarify the cheap trick, by presenting it and its supporting rationale against a backdrop of minimal assumptions. We conclude that the cheap trick is valid, focusing our subsequent discussion on the plausibility of its premises. I also introduce a flexible abstract framework for representing and contrasting theories of knowability (section 4.2.1); relate our discussion to the broader debate concerning the contingent a priori (section 4.2.3); and remark on Gibbard's take on the cheap trick (section 4.2.4).

In the second stage (section 4.3), I consider three initial objections to the cheap trick and offer replies, bolstering the cheap trick's status as a paradox. The objections are inspired by standard strategies in the literature for challenging the existence of the contingent a priori (sections 4.3.2 and 4.3.3) or the closure of knowability under logical implication (section 4.3.1). The objections fail in an interesting way: one can accept the force of these worries for the standard

[1] For some push-back on this point, see Devitt [2014], who argues that there is no such thing as a priori knowledge, and Williamson [2013], who argues that the line between a priori and a posteriori knowledge is blurry.

[2] See [Gibbard, 2012, pp.252-254]. Note that I do not slavishly recreate Gibbard's presentation of the cheap trick argument. My intention is to present it in a way that lays the core issues bear.

cases (i.e. purported instances of the contingent a priori generated by rigidifying a description, and purported instances of closure failure revealed by standard skeptical paradoxes), but on inspection this provides little reason to reject the premises of the cheap trick.

In the third stage (section 4.4), we consider a different sort of response: a surprisingly forceful rationale for biting the bullet and accepting the cheap trick's conclusion. To start, I identify an important 'victim' of the cheap trick: *epistemic (neo-Fregean) two-dimensionalism*, according to which a meaning (and the object of a cognitive attitude) is modeled as a pair of unstructured propositions (cf. Chalmers [2011]).[3] Epistemic two-dimensionalism is thus, foremost, a semantic theory, but one that generates a theory of knowability. I call this the *2D theory*. Following Gibbard's lead, I prove that epistemic two-dimensionalism is committed to the premises (and conclusion) of the cheap trick (section 4.4.2), and provides a conducive setting for at least two distinct deflationary rationales for biting the bullet. Of these, I find the second rationale more compelling: namely, that the cheap trick merely shows that cheap but *vacuous* knowledge is in ample supply (section 4.4.4). I contrast this with Gibbard's diagnosis (sketched in sections 4.2.4 and 4.4.3), understood here as a development of the first, less satisfactory rationale for biting the bullet. I conclude this stage with reasons for *nevertheless* developing an account of knowability that resists the cheap trick (section 4.4.5).

In the final stage (section 4.5), I present an account of knowability (namely, resolution theory) that rejects the first premise of the cheap trick and thereby resists its conclusion. Like the 2D theory, this theory is generated by a semantical theory: namely, *resolution semantics*. I argue that such an account must match epistemic two-dimensionalism's capacity to accommodate tricky linguistic data. Hence, resolution semantics may be regarded as a sophisticated variant on epistemic two-dimensionalism that forbids contingent a priori knowledge without jettisoning its forerunner's most attractive features. For transparency, I develop the account in two stages: an initial stage that minimally tweaks epistemic two-dimensionalism, but does not quite meet our desiderata (section 4.5.1); and a more nuanced final proposal that draws in further distinctive features of resolu-

[3]The history of the two-dimensionalist framework is a complicated matter. For some key texts, see Davies and Humberstone [1980], Stalnaker [1978] and Jackson [1998]. One reason to focus on the framework of Chalmers [2011] is that it best withstands the criticisms identified in the comprehensive analysis of two-dimensionalism offered by Soames [2007a].

tion theory (section 4.5.2). I conclude by answering a possible objection to my positive proposal (section 4.6). The answer has the pleasant consequence that the threat of missed clue counter-examples (cf. Chapter 1) is defused for the resolution theorist.

4.2 Clarifying the cheap trick

The current section aims for a perspicuous presentation of the cheap trick and its supporting rationale. To accomplish this, I first present a precise, abstract, flexible framework for comparing and developing theories of knowability. This framework serves three main functions in this chapter. First, it allows us to establish the validity of the cheap trick using only minimal assumptions. This puts the spotlight on the cheap trick's premises, our focus for the rest of the chapter. Second, since every discussion must take *some* claims for granted, working with a precise framework has the benefit that my own assumptions are laid bare. Third, epistemic two-dimensionalism and resolution theory have subtle features that are best framed precisely. A sufficiently abstract framework accomplishes this without irrelevant technical details.

4.2.1 Preliminary: epistemic scenarios

We work with a simple model of an agent's epistemic situation. The model is deliberately abstract: the bare framework involves minimal assumptions about the features of epistemic situations, and is therefore silent on the many further constraints that a theorist might impose. Thus, an argument that is valid with respect to our *bare* model remains valid for every enrichment.

An *epistemic scenario* $\mathcal{S}$ is, for us, a tuple $\langle \mathcal{W}, @, \mathbf{A} \rangle$, with the following components: a set $\mathcal{W}$ of *worlds*; a designated actual world @; and a function $\mathbf{A}$ that assigns a set of propositions to each proposition P. Relative to $\mathcal{S}$, we call a subset of $\mathcal{W}$ an *unstructured proposition*. As usual, I use 'proposition' as shorthand for 'unstructured proposition'. A proposition is true (a *fact*) when $@ \in P$. Suggestively, I write $P \vee Q$ for $P \cup Q$; $P \wedge Q$ for $P \cap Q$; $\neg P$ for $\mathcal{W} - P$; and $P \Rightarrow Q$ for $P \subseteq Q$.

I write $\mathbf{A}_P$ instead of $\mathbf{A}(P)$. If $A \in \mathbf{A}_P$ then we (suggestively) call A a *way of knowing* P. We insist that if $A \in \mathbf{A}_P$ then $P \not\subseteq A$ and if $@ \in P$ then $@ \notin A$

i.e. if P is a fact then A is an *alternative* to the truth (we do not insist, at our current level of abstractness, that A be inconsistent with P).

$\mathbf{A}$ models the truism that knowledge of P requires enough information to eliminate sufficient alternatives to P. If $A \in \mathbf{A}_P$, interpret this as: eliminating every world in A is sufficient for being in a position to know P. A simple suggestion for enriching this picture is to posit a unique such A: namely $\neg P$. However, we leave open the possibility that there are potentially *multiple* ways of knowing P. Nor do we assume that knowability of P requires that *every* world in $\neg P$ be ruled out, nor that *only* worlds in $\neg P$ need to be ruled out.

Given an epistemic scenario $\mathcal{S}$, we write $\mathbb{K}(P, E)$ to indicate a precise explication of "P is knowable given empirical information E". Namely: for some $A \in \mathbf{A}_P$, the evidence E is inconsistent with A i.e. $E \cap A = \emptyset$.

Given $\mathcal{S}$, we say that P is knowable a priori when $\emptyset \in \mathbf{A}_P$. Note that if P is knowable a priori, then $\mathbb{K}(P, E)$ holds for any empirical proposition E. In particular: $\mathbb{K}(P, \emptyset)$. Thus, we write $\mathbb{K}(P)$ to more briefly indicate a priori knowability. P is knowable a posteriori just in case there exists $A \in \mathbf{A}_P$ such that $A \neq \emptyset$.

I emphasize, once again, the *abstractness* of our account of 'knowable a priori'. This is a *technical* notion, designed to emphasize crucial information-theoretic, non-psychological aspects of the ordinary philosophical notion. Note that our account is *not* committed to claiming that the content of a propositional attitude is an unstructured proposition. Rather, we trade on the idea that whatever account of an attitude and its content is correct, it is possible to abstract an unstructured proposition from that content, and so engage with a core *epistemic* issue: what is the truth set of that content, and is any empirical information required to know a content with that truth set?

In particular (delaying details for Section 4.4), the current framework is compatible with epistemic two-dimensionalism in the style of Chalmers [2011], which holds that the content of an attitude is an ordered pair of (structured) propositions. From this pair, one can abstract a pair of unstructured propositions $\langle P_1, P_2 \rangle$, and in turn an unstructured proposition P_1. This abstraction is of epistemic interest since the Chalmersian holds that P_2 represents the *epistemic dimension* of $\langle P_1, P_2 \rangle$: knowing this content requires sufficient information to eliminate the worlds at which P_2 is false (i.e. that fall outside P_2). Thus, eliminating the $\neg P_2$ worlds is usefully construed as *a way of knowing* P_1. By these lights, every way in which P_1 is knowable corresponds to a two-dimensional proposition $\langle P_1, Q\rangle$. We can model this in our framework by treating $\mathbf{A}_P$ as the set containing exactly $\neg Q$ for every Q where $\langle P, Q\rangle$ is a possible content. Then, $\mathbb{K}(P)$ essentially reports the existence of a two-dimensional content: $\langle P, \mathcal{W}\rangle$.

We indicate the class of all epistemic scenarios by $\mathfrak{S}$. We understand a *theory of knowability*, technically, as a subset of $\mathfrak{S}$. That is, a theory of knowability amounts to a *constraint on what counts as a 'legitimate' epistemic scenario.* To arrive at the correct theory of knowability, on this approach, is to locate that subset that is neither too constrained nor unconstrained.[4] Put another way: the correct theory of knowability is one that appropriately enriches the basic assumptions of our model.

To further appreciate our framework's flexibility, we exhibit various commitments that *could* be adopted to enrich the model (thereby shouldering a greater philosophical burden), echoing major competitors in the philosophical literature.

To start, our basic model leaves it open how best to flesh out the notion of a 'world'. One could, for instance, consider worlds as total ways things could be; or situations (i.e. partial ways things could be);[5] or centered worlds (a way things could be supplemented with an agent and time that constitutes a *perspective* on that way);[6] or as ways things could be *simpliciter*, some possible, some impossible.[7] Though we can ignore such issues, note that a choice along this dimension could have an effect on how we understand propositions as relating to ordinary language, or how we design a formal language.[8]

[4]Note how this would relate to a *logical approach* to studying such theories: here, one proposes a suitable formal language; then proposes a way of interpreting that language on a model; and then identifies a theory with a set of sentences in the language that are closed under logical implication. One can then investigate rival theories by either (i) proposing a set of axioms (which correspond to a set of models - those that validate the axioms) or (ii) by proposing a constrained set of models, and generating a theory by finding out what sentences are validated by that set. Effectively, our own approach is that of (ii), but we delay proposing a formal language for another time.

[5]Cf. Barwise and Perry [1981], Kratzer [2012].

[6]Cf. Lewis [1979].

[7]Cf. Nolan [1997], Jago [2014].

[8]For instance, if we understand worlds as situations, we may *not* wish to understand ordinary language denials such as "it is not raining in London" as expressing $\neg P$ (i.e. the complement of the P-worlds), where P is the proposition expressed by "it is raining in London". For it may be that there are situations that neither determine that London is raining nor that it isn't. Of course, accepting this view about ordinary language does not prevent us from explicitly introducing and making use of the classical connectives $\neg, \vee, \wedge$ in, say, a formal language, or an

Our account of knowability can also be enriched in various ways. Consider *infallibilism*, the doctrine that knowability of P requires information that entails that P is false. This can be represented by an epistemic scenario where $A \in \mathbf{A}_P$ guarantees that $\neg P \subseteq A$.

Consider *naive infallibilism*: this can be represented by the restriction that $\mathbf{A}_P = \{\neg P\}$.

Consider *Bayesian fallibilism*, according to which eliminating (alternative) A is sufficient for knowing P just in case this renders P sufficiently probable. This can be represented by a model where (relative to some pertinent probability function $\mathbb{P}$ and threshold τ) if $A \cap P = \emptyset$ and $\mathbb{P}(P|\neg A) > \tau$ then $A \in \mathbf{A}_P$. Notice that, in the case of Bayesian fallibilism, the set $\mathbf{A}_P$ generally contains more than one member. Hence, our basic framework's allowance for *multiple members of* $\mathbf{A}_P$ (i.e. multiple ways of knowing P) is a sensible abstraction, given our desire to work with a largely neutral basic framework.[9]

Consider *classical relevant alternatives (RA) theory*, according to which $\mathbf{A}_P$ contains exactly one proposition A, though it is allowed that A is merely a proper subset of $\neg P$ in some cases.[10]

Consider the *Dretskean information-tracking theory*, inspired by Dretske [1971], according to which P is knowable just in case: if P were not the case, then E would not be the case. On this view, $A \in \mathbf{A}_P$ is the set of nearest $\neg P$ worlds. Three important constraints follow: $\mathbf{A}_P$ is a singleton $\{A_P\}$; $A_P \subseteq \neg P$; and if $P = P_1 \vee \ldots \vee P_n$, then $A_P \subseteq A_{P_1} \cup \ldots \cup A_{P_n}$.

Consider the *Goldmanian belief-tracking theory*, inspired by Goldman [1976], according to which P is knowable just in case: if P were not the case, then the agent would not believe P on the basis of E. Two relevant constraints: $\mathbf{A}_P$ is a singleton $\{A_P\}$ and $A_P \subseteq \neg P$.

4.2.2 The cheap trick

Let Q be an arbitrary fact - for instance (suppose), that there are right now exactly 3, 894, 411, 561 eggs in China. On the other hand, let P be the fact that

enriched natural language.

[9]See Holliday [2015b] for an extended defense of a multi-path approach to fallibilism, based on abstract considerations.

[10]Cf. Dretske [1970], Stine [1976], Lewis [1996], Lawlor [2013], Pritchard [2012], Holliday [2013].

the world is exactly *this* way (suppose that I gesture at the actual world in order to settle the referent of "this" in context). Since I apparently *succeed* in referring to the actual world-state using the designator "this", notice that $P = \{@\}$. Now consider this argument (cf. [Gibbard, 2012, pg. 253]):

P1. P is knowable a priori i.e. $\mathbb{K}(\{@\})$.

P2. a priori knowability is closed under deductive consequence i.e. if $\mathbb{K}(P)$ and $P \Rightarrow Q$ then $\mathbb{K}(Q)$.

C1. $\therefore$ Q is knowable a priori i.e. $\mathbb{K}(Q)$

C2. $\therefore$ Generalization: every fact is knowable a priori.

Support for P1: "things are exactly *this* way" is plausibly a contingent a priori truth. First, a statement of this sentence will be true in every context of utterance (or, at least, in contexts that are not dysfunctional). Second, that this is so is a product of the linguistic role of the demonstrative term "this". That is, the invariant truth of this claim can be (easily) recognized by a competent user of the language who reflects on the linguistic function of its components. In terms of our abstract framework, accepting **P1** amounts to imposing the following constraint on the *legitimate* epistemic scenarios: $\emptyset \in \mathbf{A}_{\{@\}}$.

Support for P2: An instance of the intuitive principle that deduction is a route to knowledge extension (this principle is often termed *epistemic closure*). In terms of our abstract framework, accepting **P2** amounts to imposing the following constraint on the *legitimate* epistemic scenarios: if $\emptyset \in \mathbf{A}_P$ and $P \subseteq Q$, then $\emptyset \in \mathbf{A}_Q$.

Validity: assume **P1** and **P2** hold for epistemic scenario $\mathcal{S}$. **P1** says that $\mathbb{K}(\{@\})$ holds. Let Q be any fact. Thus, $@ \in Q$, by definition. Thus, $\{@\} \Rightarrow Q$. So, by **P2**, $\mathbb{K}(Q)$.

4.2.3 Remark: the contingent a priori and cheap knowledge

It is illuminating to position the cheap trick in the storied dialectic concerning contingent a priori truths (i.e. contingent facts that are knowable without any empirical information).[11] Purported examples of such truths fall into two families.

[11]The origin of the dialectic is Kripke [1980].

On one hand, it is claimed by Kripke [1980] and Kaplan [1989] that such truths can be located by 'rigidifying' a description. For instance, one might stipulate that 'Julius' refers to the person who actually invented the zipper (if anyone did), and so claim that "Julius invented the zipper (if anyone did)" is a contingent a priori truth.[12] We say that such examples are of the *description type*. Alternatively, Kaplan [1989], Soames [2007a] and Soames [2007b] use indexicals to construct purported instances, such as "I am here now" and "φ if and only if, actually, φ".[13] Such examples (including "*this* is the way things are") are of the *indexical type*.

Are these sentences genuine examples of contingent a priori truth? A typical ground for resistance is that, if so, the techniques for their construction wildly over-generate a priori knowledge. Consider some oft-discussed instances of the descriptive type: "Bob is the tallest spy in China (if anyone is)" and "Neptune is the perturber of the planets (if anything is)".[14] The first concerns a matter of extreme and deliberate secrecy; the second concerns a major astronomical fact. In neither case is it plausible that one can know the reported facts by linguistic stipulation alone.[15] Thus, the purported examples of the descriptive type have wound up fueling interest in *acquaintance constraints* on singular thought, of the form: one can only grasp a singular thought concerning object o if one is suitably acquainted with o.[16] On the other hand, accepting examples of the indexical type has a similarly egregious consequence: that *every* fact is knowable a priori. To see this (other than by recalling the cheap trick), one may note that "φ if and only if, actually, φ" is typically analyzed as having the same truth set as φ.[17]

However, the contingent a priori has experienced a recent revival. On one hand, the proposal that singular thought is constrained by acquaintance has been

[12]Kaplan [1989] develops this approach in technical detail, through his use of the *dthat* operator, which accepts an individual concept α (i.e. a function from worlds to objects) and returns a constant individual concept that has the value of $\alpha(@)$ at every world w.

[13]Discussion of the latter was initiated by Davies and Humberstone [1980].

[14]Where 'Bob' is stipulated to refer to the unique tallest spy in China (if there is one), and 'Neptune' is stipulated to refer to the unique perturber of the planets (if there is one).

[15]Cf. Donnellan [1977], Evans [1982], Soames [2007a].

[16]Cf. Russell [1910], Evans [1982].

[17]The relevant technical treatment of the 'actually' operator is essentially due to Davies and Humberstone [1980]. See Soames [2007b] for a development and defense of the claim that this treatment at least approximates the everyday use of 'actually'. See Yalcin [2015] for resistance to this last claim.

comprehensively attacked.[18] On the other hand, for any vaguely plausible explication of 'acquaintance', it is hard to deny that a speaker is acquainted with the referents of indexicals such as 'I' and 'now' and 'actuality'. Hence, examples of the indexical type seem impervious to worries about acquaintance. Further, it has been argued (for instance, by Hawthorne and Manley [2012]) that one can accept examples of the contingent a priori without endorsing implausible epistemic leaps, so long as one's theory of knowledge accommodates a category of *vacuous knowledge* under which contingent a priori knowledge can plausibly be classified. One might label this the *deflationary approach.*

The cheap trick is emblematic of this revived case for the contingent a priori, and will prove a fruitful stalking horse.

4.2.4 Remark: Gibbard's diagnosis

What moral does Gibbard draw from the cheap trick? Gibbard uses the cheap trick as a reductio in support of his main contention: that the fact expressed by a claim φ is *not* in general the object of the knowledge attitude reported by "a knows that φ" (cf. [Gibbard, 2012, pp.252-256]). For instance, for Gibbard, a's utterance "I know that I am here now" does *not* express that a knows that she is in Stanford library at noon (the time and place of her utterance). Rather, it expresses that a stands in the knowledge relation to what Gibbard calls the *import* of her utterance: in this case, the trivial proposition. Gibbard's attitude to the cheap trick is thus similar in spirit to the reductio strategy we consider (and reject) in 4.4.3, modulo an alignment of terminology.[19]

Some of Gibbard's remarks complicate this assessment, suggesting that he takes the expressed fact to be a *component* of the speaker's object of thought.[20] If so, his view is more in line with the bullet-biting strategy considered in 4.4.4.[21]

[18] See Hawthorne and Manley [2012].

[19] Gibbard, for instance, insists on a terminology according to which: if $\langle P, I\rangle$ is the two-dimensional meaning of φ, then P is a proposition, while I is not best thought of a proposition (since I might essentially involve an agent's *perspective* in a way that proposition's do not). We can afford to ignore such a subtleties in the present discussion.

[20] Consider: "Speaker and hearer thus do end up thinking the same structured proposition [i.e. the fact expressed by the speaker's utterance] from their respective standpoints. That's a requirement if communication is to be successful ..." [Gibbard, 2012, pg. 271].

[21] A last remark concerning the relationship between our discussion and Gibbard's: Gibbard intends the cheap trick to refute, in particular, that the *unstructured* proposition expressed by φ is the object of the knowledge attitude reported by "a knows that φ". He offers alternative

It is beyond the scope of this chapter to settle Gibbard's complex position (what he commits to; what he remains neutral on; to what extent he is consistent throughout) once and for all. At worst, the position in section 4.4.3 may be regarded as an educational caricature.

4.3 The cheap trick as a paradox

In this section, I aim to boost the cheap trick's credentials as a paradox. I consider three possible objections to the cheap trick, and offer replies. The objections are chosen to represent standard lines of attack against the existence of the contingent a priori or the validity of unrestricted epistemic closure. The moral is that the objections fail in interesting ways.

4.3.1 Attacking P2 by denying epistemic closure

Proposal Consider *epistemic closure*, the principle that knowability is closed under deductive consequence. As discussed in Chapter 3, denying closure extricates us from a variety of skeptical paradoxes. It seems that I am in a position to know that the wall before me is white, on the basis of it *appearing* white. Yet, intuitively, I am in no position, on the basis of the same information, to know that the appearance of the wall is not a result of clever lighting.[22] Further, it seems that I am in a position to know that the water in my glass is safe to drink on the basis of my ordinary information. Yet it seems doubtful that I am in a position to know that the water supply has not been poisoned by a highly unusual happening (a deranged sanitation worker; a CIA plot; or a terrorist attack etc.) on the basis of just ordinary experience. In short, closure denial explains why our everyday and scientific knowledge is not hamstrung by paranoid fantasies that pathologically resist incompatibility with our empirical information.

Reply To reject closure, one must defend a restricted version of the principle, that ensures that deductive reasoning has an intuitively satisfying epistemic

"cheap tricks" to refute the claim that the *structured* proposition associated with φ is the object of reported knowledge. However, our present discussion need not to depart from consideration of (mere) unstructured propositions: since structured propositions with the same truth set are a priori equivalent (I propose), the distinction between such objects washes out at the level of knowability.

[22] This is the basis for the "problem of easy knowledge" of Cohen [2002].

scope. One plausible desideratum for a restricted closure principle is that it secure mathematical reasoning. The skeptical paradoxes that motivate closure rejection involve contingent, empirical claims. On the other hand, there is little motivation for undermining the sanctity of mathematical reasoning. It is therefore notable that rejection of **P2** seems to be a rejection of the core epistemic principle that grounds mathematical reasoning.

4.3.2 Attacking P1 on cognitive grounds

Proposal Consider this objection: (i) on reflection, every purported instance of the contingent a priori is an illusion resting on a dubious epistemic leap from knowing that sentence φ must be true (whatever fact it expresses) to knowing the particular fact P expressed by φ, with (ii) the failure, by ordinary agents, to clear this gap best explained by *cognitive limitations*, due to a lack of acquaintance with the objects that P concerns.[23] In other words: though one knows that φ must express a truth, one fails to *grasp* the expressed proposition, and therefore do not know it.

Reply It is hard to deny that one who utters the sentence "*this* is the way things are" is acquainted with the referent of 'this'.[24] On the causal front, it cannot be denied that the speaker is involved in immediate causal relationships with the actual world. On the epistemic front, it cannot be denied that the speaker can establish the existence of an actual world (and presumably has a good deal more knowledge about that world).

Of course, it is true that there are many things that one does not know about the actual way things are; true that one is presumably not confronted by the actual way things are in an *unmediated* way; and true that one can only ever perceive a *part* of the way things are when indicating it. However, none of these failures indicate the failure to meet a plausible criterion of acquaintance. One can presumably think (or even know) "that ship is large" when viewing a mere part of the ship through a window; and one can presumably think (or even know) "that ship is far away" after spotting the hazy outline of a ship on the horizon. In neither case, do we confront the ship without the mediation of perception. In

[23] Cf. Donnellan [1977].

[24] Compare a more benign purported instance of contingent a priori knowability, due to Kaplan [1989]: "I am here now".

the first case, only a part of the ship is confronted. In the second, very little is known or knowable about the object, without more information.

In summary: the peculiar trickiness of the cheap trick is largely due to the fact that if we are acquainted with *anything* (for the purposes of reference and singular thought), surely we are acquainted with the actual way things are.

4.3.3 Attacking P1 on informational grounds

Proposal Assume that it is knowable a priori that an utterance of "*this* is the way things actually are" expresses a fact. Instead of positing cognitive limitations that impede an agent from knowing (a priori) the fact so expressed, suppose that we deny a priori knowability of that fact on purely *informational* grounds. Thus, as in the previous objection, one knows something a priori about sentence φ, the vehicle for expressing P, without knowing a priori that P. However, the charge in this case is *not* that one fails to grasp the content of φ - merely that one fails to have sufficient information to know that content.[25] Intuitively, to have sufficient information to know P is to be in a position to *discriminate* P being the case from not. Now, it is clear enough that an ordinary agent who utters "*this* is the way things actually are" is in no position to discriminate between things being *this* way and things being any other way that shares with actuality those (generally meager) facts that are known to the agent. Suppose that it is raining in London, but the speaker lacks empirical information relevant to this fact (they are not standing outside in London at time t; cannot access a weather report etc.). Intuitively, this speaker cannot discriminate between the way things actually are and another way things could be at which it *is* raining in London, and so cannot know that things are exactly as they actually are.

Reply It is difficult to present a theory of 'knowability as discriminability' that is both free of objectionable consequences *and* fails to support **P1**. (Since I am ultimately going to offer a theory that denies the cheap trick on discriminability grounds, note that it will pay off later to dig into details in this reply.)

Start with a simple account of discrimination: information E discriminates fact P from $\neg P$ just in case E is inconsistent with $\neg P$. In this case, we say

[25] This might be construed as an interesting instance of *closure failure*. In what follows, we use $\ulcorner \varphi \urcorner$ and φ to distinguish mentioning φ from using it. The current proposal is that one might be in a position to know $\ulcorner \varphi \urcorner$ is true, yet not be in a position to know φ. If it is admitted that $\ulcorner \varphi \urcorner$ being true *entails* φ, then we have closure failure.

that E *comprehensively rules* $\neg P$ *out*. In terms of our abstract framework, this is the suggestion that $\mathbf{A}$ be constrained so that: if $A \in \mathbf{A}_P$ then $\neg P \subseteq A$. On this account, no ordinary agent has information that discriminates $\{@\}$ from $\mathcal{W}-\{@\}$.

However, this places stringent demands on the information needed to discriminate P from $\neg P$, inviting Cartesian skeptical worries. The account lacks explanatory power in other ways that will prove relevant to our discussion. Assume that "*That* is the president" and "Barack Obama is the president" express the same fact in context. It is a familiar observation that one claim may be harder to know than the other,[26] yet the current account renders their conditions of knowability identical.

Turn instead to a *belief-tracking* account of discriminability.[27] Start with a *safety* account:[28] P is knowable to agent a on the basis of E exactly when it is true that if a were to rationally believe P on the basis of E, then P would be the case. That is: at every nearby world w to @ where a rationally believes P on the basis of E, P holds at w. Now, far from supporting the current objection to the cheap trick, this account renders "*this* is the way things are" as knowable a priori. On the current view, a natural proposal is that P is knowable a priori exactly when: if a were to rationally believe P without basis in any empirical information, then P would be the case. Now, suppose that a comes to believe $\{@\}$ as follows: she utters "*this* is the way things are", notes a priori that the expressed fact must be true and on this basis forms a belief in that fact (recall that we are not here denying that she can grasp that fact). a's belief both seems rational and is assured to track $\{@\}$, as she is guaranteed to believe $\{w\}$ in whatever world w she finds herself.

Similar remarks may be made with respect to a *sensitivity* account.[29]

Finally, consider a *Dretskean information-tracking account of discriminability*: information E discriminates P from $\neg P$ just in case if P were not the case, then E would be false. That is: at the nearest worlds where $\neg P$ holds, $\neg E$ also holds. The resulting account of knowability is not an obvious victim of skeptical pressures.

[26]Cf. Perry [1979].

[27]Cf. Goldman [1976].

[28]Cf. Sosa [1999], Williamson [2000]. Note, however, that the account above is a blend of the approach of Sosa (which does not explicitly mention the information available to the agent) and that of Goldman [1976].

[29]That is: P is knowable to agent a on the basis of E exactly when it is true that if P were not the case, then a would not believe P on the basis of E. Cf. Goldman [1976], Nozick [1981].

It also succeeds in rejecting **P1**: clearly, given ordinary empirical information E, there are distinct nearby worlds to @ at which E holds, and so E does not discriminate things being *this* way from things being every other (nearby) way.

However, this account of knowability appears subject to devastating counter-examples, one of which received attention in Chapter 1. (We return to the *broader* efficacy of these counter-examples in section 4.6).

Counter-example 1 (fake barns mismatch):[30] Suppose that Max is driving through fake barn country. He glances out the window at a barn in the countryside. Max might say: (1) "there is a barn on *this particular field*" (gesturing at the field before him), or (2) "there is a barn on *the field before me*". A key difference between (1) and (2) is that (1) makes use of a demonstrative to rigidly designate the field; in (2) he makes use of a non-rigid definition description. Intuitively, Max is in a position to know (1) just in case he is in a position to know (2), for he *certainly* knows "this particular field is the field before me". (Further, we'll agree with Goldman [1976] that being in fake barn country means that Max is *not* in a position to know (1), and so not (2).) But Dretskean tracking does not deliver this result, if we embellish the story with what seems like an irrelevant detail (irrelevant to what Max is intuitively in a position to know, at any rate). Suppose that if there were no barn where Max is actually looking, then there would be no building (the owner is dead-set against erecting a fake barn on his property). Thus, in the nearest worlds where there is no barn on that field, Max does not have the same empirical information as the actual world (for it would not appear as if there is a barn). However, there are nearby worlds where the thing in the field before Max (in that world) is not a barn. In *some* such worlds, Max *would* have the same empirical information as the actual world, for there are nearby possibilities where he observes a visually indistinguishable barn facade (we assume that this is so in virtue of his being in fake barn country. Insisting on total visual indistinguishability is unrealistic, but separates the force of the current counter-example from the next).

Counter-example 2 (missed clue):[31] Suppose that Jane is browsing through the Italian edition of The Bird Almanac. Jane does not read Italian, but is enjoying the book's photos. On page 300, she comes across a photo of a bird with

[30] I adapt this example from a counter-example to Nozick's sensitivity theory of knowledge due to Kripke [2011].

[31] I adapt a counter-example to Lewis's contextualism due to Schaffer [2002].

red plumage (the clue) that otherwise seems (to Jane) similar in appearance to a wild canary. Jane is somewhat familiar with canaries, but knows no general facts about the plumage of wild canaries (in particular, she does not know that all wild canaries have *yellowish-green* plumage. All she knows is that the plumage of *domestic* canaries is diverse - and sometimes red). Thus, Jane is in no position to know that the depiction on page 300 is not of a wild canary. But a Dretskean tracking account does not furnish this answer. For if the depiction were of a wild canary, then Jane's evidence would be different: in the nearest worlds where Jane is looking at a photo of a wild canary, the depicted bird would have yellow plumage (worlds which contain mutant wild canaries with red plumage are presumably remote).

In total: if we wish to claim that our ordinary information does not discriminate things being *this* way from other ways things could be, it appears that we do not have an account of discrimination at hand that is both defensible and delivers the desired result.

4.4 Biting the bullet via two-dimensionalism

We now examine the cheap trick through the lens of (what I'll call) *Chalmersian epistemic neo-Fregean two-dimensionalism* (shorthand: '2Dism'). This is based, with some qualifications, on Chalmers [2011], which I consider the most sophisticated account of epistemic two-dimensionalism to date. 2Dism affords us consideration of a further line of response to the cheap trick: that of biting the bullet, and accepting its premises and conclusion. The case that emerges is roughly this: 2Dism is an elegant theory that convincingly handles a variety of otherwise puzzling linguistic/epistemic data (section 4.4.1); 2Dism is committed to the premises of the cheap trick, and therefore its conclusion (section 4.4.2); at the same time, 2Dism provides conceptual tools for deflating the significance of this conclusion, stripping it of its shock value (section 4.4.4). Hence, every reason to accept 2Dism is a reason to accept the cheap trick's conclusion; and any impulse to treat the cheap trick as a *reductio* against 2Dism is likely unwarranted. (Along the way (section 4.4.3), we critique a second, and less convincing, line of response to the cheap trick - inspired by Gibbard [2012] - that emerges in the 2D setting: that of treating it as a reductio against a naive view on attitude reports.)

Be this as it may, I argue in section 4.4.5 that we *nevertheless* require a theory of knowability that *evades* the cheap trick. Namely, this is a requirement on a theory of *substantive* knowability. Hence, at best, the appeal to 2Dism only serves to refine the challenge presented by the cheap trick.

4.4.1 Epistemic neo-Fregean two-dimensionalism

I introduce 2Dism by outlining its six core features: two-dimensionalism; epistemicism; neo-Fregeanism; integration by identification; logical conservativeness; and 2D attitude contents. As I will note, my account incorporates two simplifying assumptions which Chalmers [2011] does not endorse. But since the issues in question don't bear on our arguments, trading sophistication for increased simplicity is sensible in our context.

Two-dimensionalism: A two-dimensional semantics posits two dimensions of truth, and so assigns a *pair* of unstructured propositions $\langle P_1, P_2 \rangle$ to every sentence φ. Call such a pair a *2D proposition*.

Simplifying Assumption: I ignore any motivation for developing 2D propositions as pairs of *structured* (Russellian) propositions (cf. Chalmers [2011]).

Epistemicism: an epistemic approach to two-dimensionalism interprets the second component of a 2D proposition $\langle P_1, P_2 \rangle$ along *epistemic* lines: to be in a position to know φ - where this sentence means $\langle P_1, P_2 \rangle$ - is to have information that rules out $\neg P_2$. In contrast, P_1 is understood as the (alethic) *modal* dimension of meaning: if φ means $\langle P_1, P_2 \rangle$, then P_1 is the set of worlds at which φ is considered true when evaluating claims with alethic modal operators (such as "it could be that φ", "it could be that both φ and ψ", or "it must be that φ is not the case").

Neo-Fregeanism: a neo-Fregean approach interprets the components of a 2D proposition using the Fregean distinction between reference (content; what is said; what is asserted) and sense (mode of presentation; way of thinking; guise).[32] In this context, we refer to P_1 as the *content* and P_2 as *the guise* of $\langle P_1, P_2 \rangle$, or *a guise* for P_1. It is instructive, if unrealistic, to think of this in simple terms. Consider a sentence that contains referring terms, such as "Clark Kent is wearing a grey suit": take its *content* to be the singular proposition that Kal-El is wearing a grey suit (where 'Kal-El' is a name in our meta-language). On the other hand,

[32] Compare, in particular, the discussion in [Kaplan, 1989, sect. XVII].

think of referential terms as associated with not only a referent, but a *role*. For instance, 'Clark Kent' may associate with Kal-El the role of the mild-mannered reporter (as opposed to the role of, say, the super-powered hero). Then, think of the guise of "Clark Kent is wearing a grey suit" as the proposition that the mild-mannered reporter is wearing a grey suit. Using technical vocabulary: "Clark Kent is wearing a grey suit" expresses that Kal-El *qua* mild-mannered reporter is wearing a grey suit.

Integration by identification: 2Dism *combines* the epistemic and neo-Fregean perspectives: the second component of $\langle P1, P2\rangle$ is interpreted as *both* an epistemic component and a guise. Thus, epistemic component and guise are *identified*. On this view, then, to be in a position to know that Clark Kent is wearing a grey suit is to have sufficient empirical information to rule out that the mild-mannered reporter is not wearing a grey suit.

Logical conservativeness: 2Dism holds that content and guise agree in their logical structure. If sentence φ means $\langle P_1, P_2\rangle$ and sentence ψ means $\langle Q_1, Q_2\rangle$, then "φ and ψ" means $\langle P_1 \wedge Q_1, P_2 \wedge P_3\rangle$; "$\varphi$ or ψ" means $\langle P_1 \vee Q_1, P_2 \vee P_3\rangle$; and "it is not that φ" means $\langle \neg P_1, \neg P_2\rangle$.

2D attitude contents: 2Dism holds that it is not only the meaning of sentences that are 2D propositions, but such objects also serve as the contents of *propositional attitudes*.

Simplifying assumption: I assume that the ascription "a knows that φ" expresses that the subject a holds a knowledge attitude towards the content $\langle P_1, P_2\rangle$, where this 2D proposition is the *meaning* of φ (in context). Cf. Section 4 of Chalmers [2011].[33]

The 2D theory of knowability

2Dism draws a tight connection between meaning and knowability, generating a theory of knowability - the *2D theory* - as follows. According to 2Dism, $\langle P, Q\rangle$ is

[33]Instead, Chalmers [2011] holds that "a knows that φ" expresses that the subject a holds a knowledge attitude towards a content $\langle P_1, Q\rangle$, where $\langle P_1, P_2\rangle$ is the meaning of φ and Q is *coordinated* with P_2 in a sense that is left unspecified. The purpose of this maneuver is to acknowledge the difficulty of cases of the following type: suppose that a exclaims "Fred knows that I am hungry". Then, as Chalmers [2011] puts it: "To satisfy [this] ascription, Fred need not have a belief with the same primary intension [i.e. guise] as 'I am hungry'. If he did, he would believe that he is hungry. Rather, Fred can satisfy the ascription with a belief that picks the ascriber out via a quite different primary intension" [pg. 605].

a meaning just in case $\neg Q$ is way of knowing P. That is: if one is in a position to rule out $\neg Q$, then one is in a position to know P (under a certain guise, at least).

We capture this with our abstract model as follows: the alternatives function $\mathbf{A}_P$ is understood to yield the complements of possible guises for the content P. Thus, $A \in \mathbf{A}_P$ indicates that $\neg A$ is a guise under which P can be thought (and so $\langle P, \neg A \rangle$ is a potential meaning for a sentence). With this mind, various constraints (and an enrichment) on the class of epistemic scenarios are needed to capture the 2D theory. For one, a distinguished subset $\mathbf{G}$ of propositions - those fit to serve as a *guise* - must be singled out. Then, $\mathbf{A}$ must be constrained so that: for all P, if $A \in \mathbf{P}$, then $A \in \mathbf{G}$; and if $G \in \mathbf{G}$ then $\mathbf{A}_G = \{\neg G\}$. Finally, we constrain the model to capture the logical aspects of 2Dism: if $A_1 \in \mathbf{A}_P$ and $A_2 \in \mathbf{A}_Q$ then $A_1 \vee A_2 \in \mathbf{A}_{P \wedge Q}$; if $A_1 \in \mathbf{A}_P$ and $A_2 \in \mathbf{A}_Q$ then $A_1 \wedge A_2 \in \mathbf{A}_{P \vee Q}$; and if $A \in \mathbf{A}_P$ then $\neg A \in \mathbf{A}_{\neg P}$.

Motivation for accepting 2Dism

2Dism elegantly and uniformly accounts for various puzzling phenomena.

Frege's puzzle: 2Dism accommodates the examples that fuel Frege's puzzle. It can be true to say "*a* knows that Clark Kent is wearing a grey suit" but not true to say "*a* knows that Superman is wearing a grey suit". A Fregean solution: though both 'Clark Kent' and 'Superman' refer to Kal-El, each name is associated with a different role for Kal-El. 2Dism elaborates as follows: the respective knowledge reports attribute knowledge of 2D propositions that share the *same content*, but *differ in guise*. Formally: 2Dism allows that $\mathbf{A}_P$ need not be a singleton.

Semantic externalism: 2Dism accommodates examples that motivate semantic externalism.[34] Ed says "I know that water is refreshing". Twin-Ed says "I know that water is refreshing". By stipulation, they make use of the same concept for 'water', but this term refers to H_2O for Ed, and XYZ for Twin-Ed. Intuitively, Ed and Twin-Ed report different knowledge.[35] 2Dism can explain the contrast: the respective knowledge reports attribute knowledge of 2D propositions that share the *same guise*, but *differ in content*. Formally: 2Dism allows that $P \neq Q$ yet $\mathbf{A}_P \cap \mathbf{A}_Q \neq \emptyset$.

Necessary a posteriori truth: 2Dism accommodates examples that motivate a

[34] Cf. Putnam [1973].

[35] For another example, see the case of Castor and Pollux in [Kaplan, 1989, pg. 531].

commitment to necessary a posteriori truths. That water is H_2O is a necessary truth, and so no body of empirical information can contradict it. Yet, intuitively, that water is H_2O can only be established a posteriori. The epistemic two-dimensionalist can make sense of this. Roughly, the guise for "water is H_2O" is something like "the clear, potable, ubiquitous liquid is H_2O". But the latter claim is no necessity, and its negation issues the alternatives that must be ruled out to know that water is H_2O. Formally: 2Dism allows for the existence of $A \in \mathbf{A}_P$ such that $A \not\subseteq \neg P$.

Guise-relative hardness of knowing: Intuitively, a fact can be easier or harder to know depending on the choice of its expression: it is easy for the amnesiac Ortcutt to know the truth of his utterance "I am here now", but harder to know the truth of "Ortcutt is in the Stanford library on 1 May 1985". 2Dism captures this point, since different guises can generate more and less strict ways of knowing. Formally: 2Dism allows for $A_1, A_2 \in \mathbf{A}_P$ such that $A_1 \subset A_2$.

4.4.2 Two-dimensionalism is committed to the cheap trick

2D theory is committed to the premises of the cheap trick.

Consider $\varphi =$ "things are exactly *this* way", meaning $\langle\{@\}, P_2\rangle$. In line with our neo-Fregean thinking, the guise P_2 is expressed by "things are exactly the way things are". Since this is true at every world, P_2 is $\mathcal{W}$, the trivial proposition. This establishes:

T1. The meaning of "things are exactly *this* way" is $\langle\{@\}, \mathcal{W}\rangle$.

Combining this with the following characteristic commitment of 2Dism delivers **P1**:

T2. If Q is a guise for P, then $\neg Q \in \mathbf{A}_P$.

Next, the following commitment of 2Dism delivers **P2**:

T3. If $A_1 \in \mathbf{A}_P$ and $A_2 \in \mathbf{A}_Q$ then $A_1 \wedge A_2 \in \mathbf{A}_{P \vee Q}$.

Suppose that $\emptyset \in \mathbf{A}_P$ (i.e. $\mathbb{K}(P)$), that $A \in \mathbf{A}_Q$ and that $P \Rightarrow Q$. Thus, $\emptyset \in \mathbf{A}_{P \vee Q}$. But since $P \Rightarrow Q$, we have that $P \vee Q = Q$. Thus, $\mathbb{K}(Q)$.

Thus, 2Dism offers new directions for probing the cheap trick: to question **P1**, on this perspective, one must challenge either **T1** or **T2**; to question **P2**, one must challenge **T3**.

4.4.3 Biting the bullet without biting the bullet: against naivety regarding attitude reports

Proposal (cf. [Gibbard, 2012, appendix 1].) What to make of 2Dism's commitment to the ubiquity of a priority? The two-dimensionalist had better offer an interpretation that renders this commitment benign. One possible approach is to embrace an *non-Chalmersian* epistemic neo-Fregean two-dimensionalism: one that rejects **T2** and *2D attitude contents*. On this approach, the meaning of a sentence remains a 2D proposition $\langle P, Q\rangle$ (a content and a guise), but the propositional attitudes are understood to have single propositions as their contents. Further, *naive infallibilism* is maintained: for all propositions P, $\mathbf{A}_P = \{\neg P\}$. Hence, to be in a position to know proposition P is, exactly, to have sufficient information to rule out $\neg P$. Then, the advantages of 2Dism can be largely preserved by rejecting a naive view of attitude *reports*: instead of claiming that "a knows that φ" expresses that a stands in the knowledge relation to the 2D proposition $\langle P, Q\rangle$ (i.e. the meaning of φ), or even the single proposition P, one proposes that it expresses that a stands in the knowledge relation to Q, the guise of φ. Thus, "it is knowable that φ given information E" indicates that $\mathbb{K}(Q, E)$ holds, as opposed to $\mathbb{K}(P, E)$.

This proposal has appeal for a neo-Fregean. For she is inclined to take the *cognitive significance* of φ as captured by its guise Q rather than its content P. Why not conclude that only that part of meaning that is cognitively significant is relevant to cognitive attitude reports?[36] Someone says "Lois Lane knows that Clark Kent is wearing a grey suit". What does this report? Perhaps only that Lois stands in the knowledge relation to the fact that the mild-mannered reporter is wearing a grey suit. Someone says "I know that I am here now". Perhaps this reports only that the speaker stands in the knowledge relation to a trivial necessary truth.

Here is the pay-off: the current view can be understood as, in one sense, *rejecting* the cheap trick. Consider the following argument: **(P1)** $\mathbb{K}(\{@\})$; **(P2)** $\mathbb{K}(\{@\}) \Rightarrow \mathbb{K}(Q)$ for any fact Q; therefore: **(C)** $\mathbb{K}(Q)$. Now, according to 2Dism, the content of "the way things are is exactly *this* way" (φ) is $\{@\}$, while the guise is $\mathcal{W}$. According to the current variant of 2Dism, that φ is contingent and

[36] Gibbard [2012] seems sympathetic to this way of thinking. Further, Soames [2007a] takes this line as the *standard one* among two-dimensionalists.

knowable a priori amounts to a benign pair of facts: $\{@\}$ is contingent, while $\mathcal{W}$ is necessary. Since neither entails $\mathbb{K}(\{@\})$, **P1** can be rejected. The argument in section 4.4.2 is unsound.

In another sense, the current view *bites the bullet*. Consider the following argument, in natural language: **(P1)** It is knowable a priori that things are *this* way; **(P2)** A priori knowability is closed under entailment; therefore: **(C)** it is knowable a priori that there are right now exactly 3, 894, 411, 561 eggs in China. The current view can agree that this argument is sound, but interprets these natural language claims as reporting (more precisely): $\mathbb{K}(\mathcal{W})$; if $\mathbb{K}(P)$ and $P \Rightarrow Q$ then $\mathbb{K}(Q)$; therefore: $\mathbb{K}(\mathcal{W})$. This is a *deflationary* account: though the conclusion of the cheap trick is embraced, it is interpreted in a manner that drains it of import.

Reply The above view has the drawback that it sacrifices some explanatory power of 2Dism, with the result that it is an ill-fit with certain linguistic data. In particular, this theory cannot account for semantic externalism. Ed says, staring at some H_2O: "I know that is water". Twin-Ed says, staring at the same substance: "I know that is water". Intuitively, what Ed says is true, and what Twin-Ed says is false. But on the current view, they say the same thing: they each attribute to themselves (something like) knowledge that the substance before them is a transparent, potable, ubiquitous liquid.

Consider some further examples along the same theme.[37]

Erroneous identification: Jon says, in Dubai: "I know that I am here now". Mary says, in New Jersey: "I know that I am here now". The current theory has the consequence that Jon and Mary self-attribute identical knowledge. But that is intuitively incorrect.

Erroneous distinction/communication failure: Jon reports his knowledge to Mary: "I am here now". Mary accepts Jon's assertion and intuitively thereby comes to know something new: namely, the same knowledge Jon self-attributes by saying "I know that I am here now". However, according to the current account, if Mary comes to know Jon's current location, then the second point is false (*erroneous distinction*). And if Mary comes to know the trivial proposition (the knowledge Jon self-attributes, according to the current account), then she does *not* come to know anything new (*communication failure*). Either way, the

[37] Adapted from [Stalnaker, 2008, pp.48-52], originally directed by him towards the theory of *de se* content of Lewis [1979]. It is notable that they transfer easily to the current setting.

current account defies our intuitions.

4.4.4 Biting the bullet redux: cheap knowledge is vacuous knowledge

Proposal: We now consider a deflationary, bullet-biting response to the cheap trick that is available to (Chalmersian) 2Dism.

Again, suppose that "the way things are is *this* way" means $\langle\{@\}, \mathcal{W}\rangle$, and therefore $\emptyset \in \mathbf{A}_{\{@\}}$. Neo-Fregeans generally take the guise of a 2D proposition as encapsulating its cognitive significance. Hence, a neo-Fregean might conclude that though $\langle\{@\}, \mathcal{W}\rangle$ is knowable, this would be *cheap* but *vacuous* knowledge of little significance, in light of the *trivial guise* under which the content is represented. This relies on the view that, in general, only some ways of knowing a content P render that knowledge *significant*, and that a 2D proposition $\langle P, Q\rangle$ is insignificant at least in the case where P is contingent and Q is trivial i.e. $Q = \mathcal{W}$.

This proposal shares commonalities with the variant of 2Dism considered above (in 4.4.3): in particular, the idea that the guise of a claim carries its cognitive significance, while its content is cognitively inert. However, the current incarnation of these ideas avoids the errors of identification, distinction and communication identified above. Jon says, in Dubai: "I know that I am here now". Mary says, in New Jersey: "I know that I am here now". On the current account, both know claims of no cognitive significance. Nevertheless, what they know *differs* (along the dimension of content). Jon reports his knowledge to Mary with "I am here now". Mary accepts Jon's assertion and intuitively thereby comes to know something new: namely, the same knowledge Jon self-attributes by saying "I know that I am here now". The current account accommodates this: Mary comes to know the *content* of Jon's assertion (under an appropriate guise, different to Jon's), an aspect of her knowledge state that is shared with Jon's knowledge state. (Though note that the current account has the odd consequence that Jon's knowledge is cognitively insignificant in his mouth but somehow *attains* significance - for Mary - through communication).

Altogether, we have a rationale for saying the cheap trick is sound: every fact is knowable a priori, *but* cheap knowledge is vacuous knowledge.

4.4.5 Towards a theory of substantive knowability

This seems to me an intriguing case for biting the bullet. Nevertheless, even if we accept it, there remains significant interest in producing a theory of knowability that *resists* the cheap trick. By the lights of the above account, 2D theory is a theory of knowability *per se*, covering both vacuous (cognitively insignificant) knowledge and substantive (cognitively significant) knowledge. Presumably, what we care about *most* is what can be known substantively. For instance, a key purpose in acquiring knowledge is to guide action in a rational and effective manner. But, clearly, vacuous knowledge does not supply this steering function. If one knows merely vacuously that things are exactly *this* way, one is hardly positioned to choose one's actions as if one knows *exactly* what the world is like. In practice, vacuous knowledge is no better than ignorance.

At a high level of abstraction, it is easy to specify what a theory of substantive knowability looks like in our precise framework: $A \neq \emptyset$ for all $A \in \mathbf{A}_P$. However, saying more than this is *not* trivial (even if 2Dism is assumed). To see this, first note the deficiency of a naive proposal: suppose we say that P is substantively knowable exactly when one's information E rules out *at least some* $\neg P$ worlds (in contrast to P being vacuously knowable exactly when P can be known without ruling out *any* $\neg P$ worlds). We might add: the *more* $\neg P$ worlds that can be ruled out, the *more substantive* the knowledge. However, counter-examples immediately spring to mind. There are many cases where one is in a position to know P substantively without having information that eliminates any $\neg P$ worlds: think of mathematical propostions (the *substantive* a priori) and cases of the necessary a posteriori (one need not rule out any worlds where water is not water to know, substantively, that water is H_2O). There are also many cases where one might have ruled out numerous $\neg P$ worlds, and yet it seems that one is no closer to knowing P in any real sense. For instance, consider Q: there are exactly 3, 894, 411, 561 eggs in China. Suppose that I investigate the question of the number of eggs in China as follows: I open the phone book, and record which US citizens have a phone number ending in 6. Every time I verify "X has a phone number that ends in 6", I rule out some $\neg Q$ worlds: those where both $\neg Q$ holds and X's number does not end in 6. Now, obviously, I could rule out any number of $\neg Q$ worlds in this way, and never get any closer to knowing Q.

Furthermore, a satisfying account of substantive knowability must carry out some requisite explanatory work. For one, our theory must explain why the conclusion of the cheap trick can be denied. Thus, the theory must deny one of the cheap trick's premises, escaping the quagmire of section 4.3. Further, our theory must account for all of the data that motivates 2Dism (section 4.4.1). For the key motivating examples all apparently involve *substantive* knowledge claims (even the data showcasing guise-relative hardness: depending on the circumstances, it might be easier or harder for John Perry to substantively know "I am in Paris" rather than "John Perry is in Paris"). In short, our theory of substantive knowability is subject to the same explanatory constraints as 2D theory, and more.

We therefore proceed to investigate a novel theory of knowability that evades the cheap trick. We need not here rule once and for all on the soundness of the above case for biting the bullet. This is not to say that nothing can be said to challenge that case - I offer some tentative considerations in an appendix.

4.5 Evading the cheap trick: resolution theory

We have cast doubt on various strategies for attacking the cheap trick (section 4.3), including that of Gibbard [2012] (section 4.4.3). With the conclusions of the last section firmly in mind, we now develop, in two stages, a theory that evades the cheap trick (by rejecting **P1**) while preserving the explanatory power of 2Dism. In fact, we would like a theory that disallows the contingent a priori in general. (If the reader wishes, they may read 'knowability' in the present section as shorthand for 'substantive knowability'.) Our strategy for attacking **P1** is a novel variant of that in section 4.3.3: an ordinary agent is in no position to know that things are exactly *this* way because their information cannot discriminate this way from other relevant possible ways things could be. (One may nevertheless accept that ordinary agents are positioned to know that the sentence "things are exactly *this* way" is always true in context, in the spirit of Donnellan [1977].) As in the case of 2Dism, our approach is to develop a semantical theory (resolution semantics) that *generates* a theory of knowability.

4.5.1 Stage 1: two-dimensionalism without identification

In this section, I present a minimally altered two-dimensionalist theory, as a stepping stone to a superior solution in the next section.

First, I indicate the form of the solution, with the aim of filling out its motivation in the preceding sections. Then I cast doubt on orthodox 2Dism's commitment to the identification of guise and the epistemic component of meaning. Then I present my variation. The basic idea is simple: while the theory continues to maintain that the meaning of φ can be associated with both a 2D proposition $\langle P, Q\rangle_e$ capturing the content and epistemic component of φ and a 2D proposition $\langle P, R\rangle_g$ capturing the content and guise (mode of presentation) of φ, it merely posits that Q is a function of R (and P), not that Q is *identical* with R.

Finally, I raise a worry: the resulting view makes knowability implausibly demanding and, in particular, makes it implausibly difficult to know "I am here now".

Preliminary: guise as subject matter

First, a preliminary: we will make use of a notion of *subject matter* in the development of our theory, drawn from recent work on this topic.[38] I introduce this here, and indicate that the notion of guise is naturally understood as part of subject matter, in our sense. Since 2Dism is intended to be a semantic theory, this is in general a useful way to avoid thinking of a content's guise in overly psychological or meta-semantic terms.

Following Lewis [1988b], subject matter may be considered a set of *distinctions* in logical space i.e. a set of unstructured propositions. Thus, metaphorically, a subject matter imposes a *resolution* on logical space, placing the focus on certain distinctions and back-grounding others. By these lights, the subject matter of a *conversation* is the set of distinctions the interlocutors concern themselves with.[39] For instance, if the discourse topic is Julia's profession, then the subject matter is the set {Julia is a lawyer, Julia is an accountant, ...}. Similarly, it is natural to say that both the proposition *that Julia is a lawyer* and the proposition *that Julia is a fisherman* are part of the subject matter of the complex *sentence* "Julia is a lawyer and a fisherman".

Can atomic predications have complex subject matter, opening up the possibility that two such sentences express the same content against different backdrops of salient distinctions? The 2Dist can answer 'yes'. Consider the atomic

[38] Cf. Yablo [2014], Yalcin [2016], Roberts [2012].

[39] The subject matter, in this sense, is also helpfully thought of as the set of *issues* which that conversation 'seeks' to resolve. See Roberts [2012].

claim "Clark Kent is wearing a grey suit", understood in context as the claim that Kal-El *qua mild-mannered reporter* is wearing a grey suit. By 2D lights, it is natural to think of this claim as involving at least two distinctions: whether Kal-El wears a grey suit or not, and whether the mild-mannered reporter (read opaquely) wears a grey suit or not.

We need *not* then say, however, that "Clark Kent is wearing a grey suit" expresses that Kal-El is wearing a grey suit *and* that the mild-mannered reporter is wearing a grey suit. It is helpful, in this regard, to introduce the terminology of semantic *truth-makers* and *false-makers* (Cf. Yablo [2014]). Given a sentence φ and its subject matter $\mathsf{T}_\varphi = \{P_1, P_2\}$, consider the *ways things can be with respect to* T: the propositions $P_1 \wedge P_2$, $\neg P_1 \wedge P_2$, $P_1 \wedge \neg P_2$ and $\neg P_1 \wedge \neg P_2$ (i.e. every proposition that decides all and only the distinctions in T_φ). Some of these ways serve to make φ true, the rest serve to make φ false, with the former together determining what is *said* by uttering φ. Now, if the 2D proposition $\langle P_1, P_2\rangle$ assigned to atomic sentence φ captures φ's subject matter, it is natural to add that the content P_1 determines the truthmakers for φ. That is: the truth-makers for φ are those ways that entail that P_1 is true: namely, the propositions $P_1 \wedge P_2$ and $P_1 \wedge \neg P_2$. Thus our theory of subject matter accommodates the idea that the content of φ is the proposition that is *said* or *asserted* by φ, while the guise for φ registers distinctions capturing how that proposition is *represented* in context. Though not determining truth value, sensitivity to these distinctions can play other roles: for instance, in determining when that claim is *on-topic* relative to a discourse topic, and determining what possibilities need to be ruled out in order for that claim to be known.

The form of a solution

Recall the notion of an epistemic scenario $\langle \mathcal{W}, @, \mathbf{A}\rangle$. Here is an abstract (and natural) way to constrain the class of epistemic scenarios so that contingent a priori truth is ruled out: for all A, if $A \in \mathbf{A}_P$ then $\neg P \subseteq A$ (cf. infallibilism). That is, A represents a way of knowing P only if eliminating every world in A ensures that every world in $\neg P$ is eliminated. Now, if P is contingent, then $\neg P$ is non-empty, and so is every way A of knowing P.

Let us translate the proposal into two-dimensionalist terms. Suppose we stick to the claim that a meaning $\langle P, Q\rangle$ can be divided into a modal component (I continue to call this *content*) and an epistemic component (I will now refrain from calling this a guise, for reasons to become clear momentarily). Suppose, further, that we stick with the idea that standing in the knowledge relation to $\langle P, Q\rangle$ requires empirical information that rules out $\neg Q$ (thus: $\neg Q \in \mathbf{A}_P$). Then the constraint we impose is this: if $\langle P, Q\rangle$ represents a meaning, then $Q \Rightarrow P$.

Notice, immediately, that our proposal has the attractive feature that we can essentially recreate the motivations from 4.4.1. That it preserves the two-dimensional approach to Frege's puzzle and semantic externalism is obvious. Further, it allows for the possibility of a posteriori necessary truth. For it allows for meanings of the form $\langle \mathcal{W}, Q\rangle$, where Q is non-empty. Finally, though we have jettisoned contingent a priori truth, the approach allows that some ways of knowing can be *more demanding* than others: $\langle P, Q\rangle$ and $\langle P, R\rangle$, where $Q \subset R$.

This is a neat picture. Can it be motivated philosophically?

Factorization of knowability into parts

I now challenge an aspect of orthodox 2Dism that can be replaced at little cost: *integration through identification* (cf. section 4.4.1).

Consider meaningful sentence φ. The orthodox epistemic two-dimensionalist holds that φ can be associated with a pair $\langle P, G\rangle_g$ consisting of its content and a guise (the latter, minimally, a proposition generated by the roles associated with referring terms in context). She also holds that φ can be associated with a pair $\langle P, \neg A\rangle_e$ consisting of a content and an associated epistemic component (the latter a proposition generated in some sense by the guise G, such that ruling out A is a way of knowing P). Finally, she identifies G and $\neg A$.

I propose that this last commitment be rejected (i.e. **T2** be rejected, as in 4.4.3), on the basis of intuition (and at little theoretical cost). Consider the claim "Perry White holds one of the most demanding jobs in Metropolis", in a context where this means: X *qua the editor of the Daily Planet* holds one of the most demanding jobs in Metropolis (X is a logically proper name for the person called 'Perry White'). The orthodox two-dimensionalist proposes a necessary condition for knowing this claim: that one have sufficient empirical information to rule out that the editor of the Daily Planet (read opaquely) does not hold one of the most demanding jobs in Metropolis. This sounds correct. However, the sufficiency of this condition for knowability may be challenged, for there seem at least two conditions which, if not met, intuitively undermine a claim to knowability:

- (A) the agent must have sufficient information to rule out that X is not the editor of the Daily Planet.

- (B) the agent must have sufficient information to rule out that X does not hold one of the most demanding jobs in Metropolis.

To appreciate the apparent necessity, consider an example: Lex Luthor has sufficient information for knowing that the editor of the Daily Planet has one of the most demanding jobs in Metropolis (for he knows that the Daily Planet has a reputation for exceptional quality). However, Lex's belief is clearly *de dicto*: intuitively, he does not believe anything about X in particular. Indeed, Lex's information cannot discriminate between X and Kal-El being the editor of the Daily Planet (that is: if Kal-El *were* the editor, Lex's information would be the same). I propose, under the circumstances, that we (speakers who associate the role of 'editor of the Daily Planet' with the name 'Perry White') would not say that Lex is in a position to know that Perry White holds a demanding job. Rather, he *merely* knows that the editor of the Daily Planet holds a demanding job. Nor would we say he is in a position to know that Perry White is the editor of the Daily Planet (though he is in a position to know that the editor of the Daily Planet is the editor of the Daily Planet).

It is further notable that if Lex's information establishes both that X is the editor *and* that the editor holds a demanding job, then this is sufficient for Lex to be in a position to know that Perry White holds a demanding job. Establishing the former two facts together is a *way of knowing* the latter, and plausibly exactly the way that is relevant in context.

Resolution semantics: two-dimensionalism without identification

I sketch a semantical theory suggested by the previous observations.

Let φ be a meaningful sentence, containing denoting terms $\tau_1, \ldots, \tau_n$ (denoting, respectively, $o_1, \ldots, o_n$), associated (in context) with the role predicates $\rho_1, \ldots, \rho_n$ (denoting, respectively, $R_1, \ldots, R_n$).

Call the *guise* of φ the proposition G_0 expressed by the sentence φ^g, obtained by replacing τ_i with "the ρ_i" for each i. Call the *enriched guise* $\mathbf{G}$ the set containing G_0 along with the facts $G_1, \ldots, G_n$:

- o_1 is the R_1
- ...
- o_n is the R_n

(One might call these *the acquaintance conditions*. Note that their knowability is *not* offered as a requirement on entertaining the thought expressed by φ, however.)

Recalling our identification with guise and subject matter in 4.5.1: note that $\mathbf{G}$ is a potentially richer subject matter than $\{G_0\}$.

Now, on the current picture, the meaning of φ can be associated with two related pairs: first, $\langle P, \mathbf{G}\rangle_g$ and, second, $\langle P, \wedge\mathbf{G}\rangle_e$, where $\wedge\mathbf{G}$ stands for $G_0 \wedge G_1 \wedge \ldots \wedge G_n$. In our current setting, we continue to call $\wedge\mathbf{G}$ the *epistemic component* for $\langle P, \mathbf{G}\rangle_g$ (or φ, in context). We say that $\wedge\mathbf{G}$ is a way of knowing P that *factorizes* into $\{G_0, G_1, \ldots, G_n\}$.

On the assumption that context does not supply extra subject matter, we say that the *subject matter* of φ (in context) is $\mathsf{T}_\varphi = \{P, G_0, G_1, \ldots, G_n\}$, with the *semantic truthmakers* being those conjuctive combinations of literals (formed from $P, G_0, \ldots, G_n$) according to which P is positive (e.g. $P \wedge G_0 \wedge \neg G_1 \wedge \ldots \wedge G_n$).

Knowability

Then, our theory of knowability is as follows: in order for φ to be knowable (in context) given information E, it must be that $E \wedge \neg \wedge \mathbf{G} = \emptyset$. That is, we adopt a comprehensive notion of *discrimination* (cf. section 4.3.3), and require for the knowability of φ that the agent's information can discriminate between the actual truth of $G_0 \wedge G_1 \wedge \ldots \wedge G_n$ (which entails P) and the possible truth of $\neg G_0 \vee \neg G_1 \vee \ldots \vee \neg G_n$.

The cheap trick defused

Consider the claim "things are exactly *this* way", where *this* is associated with the role 'the way things are'. Thus, the guise in question is $\mathcal{W}$, the fact that things are exactly as things are. However, G_1 is the fact that @ is the way things are, and so the epistemic component for φ is $\{@\}$. Thus, to know "things are exactly *this* way" requires sufficient information to rule out that @ is *not* the way things are. But, presumably, no ordinary agent has evidence that is inconsistent with $\mathcal{W} - \{@\}$. Hence, the cheap trick is defused.

A problem: the stage 1 view is too demanding

We have proposed a theory with appealing features: the advantages of the 2Dism without opening the door to the cheap trick.

However, the current theory makes knowability implausibly demanding. On the current picture, to be in a position to know fact P, one must *at least* have empirical information that is inconsistent with $\neg P$. For most guises for P, one requires *stronger* empirical information. This opens the way both to traditional Cartesian worries, and to skeptical hypotheses that are peculiar to our current setting.

Consider "I am here now" as uttered by Ortcutt, waking up with amnesia in the Stanford Library at noon. Suppose that the guise of this claim is $\mathcal{W}$, expressed by something like "The speaker is at the location of utterance, at the time of utterance". In this case, G_1 is the fact that the speaker is Ortcutt; G_2 is the fact that the location of utterance is the Stanford library; G_3 is the fact that the time of utterance is noon. In that case, the epistemic component $\wedge\mathbf{G}$ is the fact that the speaker is Ortcutt and the location of utterance is the Stanford library and the time of utterance is noon. Intuitively, Ortcutt is in a position to know this claim on minimal empirical information, simply by noting his surroundings (whatever they are). However, things are not so simple for Ortcutt on the current theory, which insists that he needs sufficient information to eliminate the world, say, that Ortcutt is at a library that looks identical to his surroundings, but is located at Harvard. This is too demanding.

Thus our current theory does *not* replicate one important advantage of 2Dism: explaining the relative ease with which "I am here now" can be known.

4.5.2 Stage 2: resolution theory

I now present a refinement of the previous solution that preserves the advantages of 2Dism and is not implausibly demanding i.e. resolution theory.

In particular, we weaken the demands of the previous theory of knowability weakening the account of *discriminability* built into it (though we retain the semantic picture). Recall that the previous account held that ruling out a proposition Q required having information E *inconsistent with* Q i.e. $E \cap Q = \emptyset$. Instead, I propose that a *Dretskean account of discriminability* be merged with the semantics presented in the previous section to yield a theory of knowability (cf. 4.3.3).

In 4.3.3, we noted some counter-examples to treating Dretskean information-tracking theory as a theory of knowability. An intriguing by-product of merging Dretskean information-tracking with the previous neo-Fregean framework will be that these counter-examples lose effect (section 4.6).

The proposal

Consider a Drestkean construal of what it is for information E to *rule out* proposition P: if P were not the case, then E would not be the case. That is: the nearest $\neg P$ worlds are also $\neg E$ worlds.

Recall also our sketched theory of *subject matter* in 4.5.1. Given subject matter T, one can consider the *ways things can be for that subject matter:* the set of propositions that decides (true or false) every proposition in T. To illustrate: suppose $\mathsf{T} = \{P, Q, R\}$. Then the way things can be for the subject matter is: $P \wedge Q \wedge R$, $P \wedge \neg Q \wedge R$, $P \wedge \neg Q \wedge \neg R$, and so on. Notice that these ways together form a *partition*, precisely representing the 'resolution' of that subject matter on logical space. We denote the set of ways with π_{T}.

Every meaningful sentence φ is assumed to have an associated subject matter T_φ, where $P \in \mathsf{T}_\varphi$ if P is the fact expressed by φ. Denote the associated set of ways things can be for T_φ by π_φ. The *truthmakers* for φ are those ways that entail that P is true. To illustrate: if $\mathsf{T}_\varphi = \{P, Q, R\}$ where P is the expressed fact, then the truthmakers for φ are: $P \wedge Q \wedge R$, $P \wedge \neg Q \wedge R$, $P \wedge Q \wedge \neg R$ and $P \wedge \neg Q \wedge \neg R$. Denote the set of truthmakers with $\mathbf{T}_\varphi$, and denote the complement of *falsemakers* $\pi_\varphi - \mathbf{T}_\varphi$ by $\mathbf{F}_\varphi$.

Now, since we retain resolution semantics, assume that the subject matter of φ, in context, is the set $\{P, G_0, G_1, \ldots, G_n\}$ where $\{G_0, G_1, \ldots, G_n\}$ is the *enriched guise* introduced in the previous section. Denote by $\mathbf{R}_\varphi$ the set of ways that entail the falsity of the epistemic component $\wedge\mathbf{G}$. Call these the *relevant alternative propositions* to φ. To illustrate: if $\mathsf{T}_\varphi = \{P, G_0, G_1\}$, then

$$\mathbf{F}_\varphi = \{\neg P \wedge G_0 \wedge G_1, \neg P \wedge \neg G_0 \wedge G_1, \neg P \wedge G_0 \wedge \neg G_1, \neg P \wedge \neg G_0 \wedge \neg G_1\}$$

and

$$\mathbf{R}_\varphi = \{P \wedge \neg G_0 \wedge G_1, P \wedge G_0 \wedge \neg G_1, P \wedge \neg G_0 \wedge \neg G_1,$$
$$\neg P \wedge G_0 \wedge G_1, \neg P \wedge \neg G_0 \wedge G_1, \neg P \wedge G_0 \wedge \neg G_1, \neg P \wedge \neg G_0 \wedge \neg G_1\}$$

Note that: $\mathbf{F}_\varphi \subseteq \mathbf{R}_\varphi$. Thus, to rule out every member of the latter is to rule out every member of the former.

Here then is our final theory of knowability: to be in a position to know φ is to have sufficient information E to rule out (in the Dretskean sense) every relevant alternative to φ. Or rather: to be in a position to know the content of φ, in the way determined by the guise of φ, is to have sufficient information E so as to eliminate the nearest worlds (to @) in every proposition in $\mathbf{R}_\varphi$.

Notice that, since the current account maintains that meaning is structured into content and guise, the motivating factors met by 2Dism in 4.4.1 are also essentially met by resolution theory.

Resolution theory via our abstract epistemic model

In terms of our abstract account of epistemic scenario, the key constraints associated with resolution theory are as follows: if $A \in \mathbf{A}_P$ and $P \neq \mathcal{W}$ then $A \cap \neg P \neq \emptyset$ (with each way of knowing A understood as associated with an enriched guise). That is, for $\mathbb{K}(P, E)$ to hold for contingent P, it must be that some $\neg P$ worlds are inconsistent with the available evidence. This is enough to ensure that there is no contingent a priori truth, and thus no cheap tricks. However, unlike 2Dism without identification, it is not required that $\neg P \subseteq A$ (nor is it required that $A \subseteq \neg P$). Thus, the current theory is less demanding on $\mathbb{K}(P, E)$. Finally, for every contingent P, the members of $\mathbf{A}_P$ have a *non-empty* intersection, containing the set of nearest $\neg P$ worlds to @.

The cheap trick defused

To see that resolution theory evades the cheap trick, consider the following example.

Things are exactly this way: The subject matter of φ is composed of the distinctions: @ is the way things are (P, G_1); and the way things are is the way things are (G_0). Thus, $\mathbf{R}_\varphi$ contains exactly one (non-empty) alternative proposition. To rule out $\neg G_1$, however, requires information that is inconsistent with the nearest worlds in which @ is not the way things are. Of course, there are countless such nearby worlds that are compatible with ordinary information. Thus, as desired, the cheap trick is defused.

The problem of demand resolved

To see that resolution theory does not render knowability excessively demanding (unlike the stage 1 theory), consider the following example.

I am here now: The subject matter of φ is composed of the distinctions: Ortcutt is in Stanford library at noon (P); the speaker is at the location of utterance at the time of utterance (G_0); Ortcutt is the speaker (G_1); the location of utterance is Stanford library (G_2); the time of utterance is noon (G_3). For Ortcutt to know "I am here now", he must therefore rule out relevant alternatives such as: that Ortcutt is in Stanford library at noon, but the speaker is not Ortcutt, the location of utterance is not Stanford library and the time of utterance is not noon. But such propositions are relatively easy to rule out with the minimal information Ortcutt receives by scanning his surroundings: if the speaker were not Ortcutt, for instance, then Ortcutt would not have the same perceptual experiences (presumably, he would not have the experience of uttering "I am here now", for a start); if Ortcutt were not at the Stanford library, then his surroundings would appear different (for worlds in which he is in a different location that appears identical are 'far-fetched' i.e. do not occur in relatively nearby worlds). And so on.

4.6 Fake barns and missed clues

Objection In section 4.3.3, we encountered counter-examples to treating Dretskean tracking theory as a theory of knowability. These seemed compelling. Since resolution theory incorporates Dretskean tracking, is it not also subject to these counter-examples?

Reply The answer is 'no': resolution theory provides sufficient resources for repelling the counter-examples. The *key* resource is that, on this view, the richer the guise for a content, the more distinctions need to be discriminated for that content to be knowable under that guise. To evade the counter-examples, we take the subject matter of the claims in question to reflect an intuitive *way of thinking* about the relevant content.

Fake barn mismatch: Max is once again in fake barn country (unbeknownst to him). Is Max in a position to know "there is a barn *in this particular field*" and is he in a position to know "there is a barn in *the field before me*"? Notice, first, that on the current account these knowability claims are intertwined (as, intuitively, they should be): for, presumably, the guise of "there is a barn in this particular field" is something like "there is a barn in the field before me" (with the role of 'this particular field' understood as 'the field before me'). Thus, "there is a barn in this particular field" has the following subject matter (where X is a meta-language name for the field actually before Max): there is a barn in X (P); there is a barn in the field before Max (G_0); and X is the field before Max (G_1). Thus, according to resolution theory, Max is *not* in a position to know "there is a barn in this particular field" without being in a position to know "there is a barn in the field before me". Further, Max's empirical information does not rule out all of the associated relevant alternatives for the former. For instance, the nearest worlds in which there is no barn in the field before Max (in that world), are worlds in which he is having an indistinguishable experience before a field with an fake barn (as per our 'fake barn country' hypothesis). Thus, resolution theory delivers the intuitively correct verdict: Max can neither know that there is a barn in this particular field, nor that there is a barn in the field before him.

Missed clue I: Jane is browsing through The Bird Almanac and (in her ignorance) wonders if the red-plumed bird depicted on page 300 is a wild canary. Is Jane in a position to know "this particular bird is not a wild canary"? On the current account, the answer is 'no'. First, let's focus on ways of thinking that are naturally associated with *a theorist* evaluating Jane's epistemic position. Using X as a name for the bird in question, take the subject matter of φ as: X is not a wild canary (P); X is not a kind of bird with yellowish-green plumage

(G_0); and wild canaries are a kind of bird with yellowish-green plumage (G_1). Here, we (unrealistically but illustratively) understand the role associated with 'wild canary' as 'kind of bird with yellowish-green plumage'. After all, one who is evaluating Jane's position relative to the 'missed clue' of red plumage is thinking of wild canaries as birds with a characteristic plumage. Now, Jane's empirical information is not sufficient to rule out all of the relevant alternatives in this context. Consider the relevant alternative: X is not a wild canary; X is not a bird of the kind with yellowish-green plumage; wild canaries are not a kind of bird with yellowish-green plumage. Plausibly, the nearest world in which this relevant alternative is realized is one in which Jane is looking at the same picture (of the same bird) but an alternative form of wild canary has evolved that does not, typically, have yellowish-green plumage. Since Jane's information is not sufficient for knowing that wild canaries have yellowish-green plumage, we conclude that Jane is in no position to discriminate actuality from nearby worlds in which wild canaries have different plumage. Thus, as required, Jane is in no position to know that the pictured bird is not a wild canary.

Missed clue II: Consider Jane, this time focusing on ways of thinking that are naturally associated with Jane herself, as she evaluates her own epistemic position (presumably, in her ignorance, she is not thinking of wild canaries as a kind of bird with yellowish-green plumage). Name the bird X, and take Jane's way of thinking about X as a *visual* one: 'the bird with visible features V'. In this context, the subject matter of φ is: X is not a wild canary (P); the bird with visible features V is not a wild canary (G_0); and X is the bird with visible features V (G_1). It follows that Jane's information is not sufficient to rule out all relevant alternatives. For consider the following relevant alternative: X is a wild canary, X is the bird with features V and the bird with features V is a wild canary. Presumably, Jane's information (her glance at the picture in the book) is perfectly compatible with one of the nearest worlds in which this possibility is realized.

4.7 Conclusion

We have surveyed a diverse array of strategies for responding to the cheap trick: including appeals to standard lines of objection against the contingent a priori and a bullet-biting neo-Fregean strategy. The cheap trick proves resilient against these lines of attack. On a more promising note, we tweaked epistemic neo-Fregean two-dimensionalism in a way that forbids the possibility of contingent a priori truth (an outcome I argued to be no great theoretical loss). As a counter to objections that the resulting theory is too demanding to be a realistic theory of knowability, and thereby undermines important theoretical desiderata that gives epistemic two-dimensionalism its appeal, we further refined the theory so as to weaken the notion of discriminability at play. The final result was the neo-Fregean resolution theory, which combines a Dretskean tracking theory of discriminability, a subject matter based framework for relevant alternatives theory and a neo-Fregean view on the nature of a sentence's subject matter. We note again its promising features: the theory forbids the contingent a priori and so is impervious to cheap tricks; the theory meets the most important motivating desiderata for an epistemic two-dimensionalist theory (including resources for dealing with Frege's puzzle and the necessary a posteriori); and, incidentally, provides resources to escape counter-examples that seem to undermine treating the Dretskean tracking theory in a naive way as a theory of knowability.

4.A Evaluating the case for biting the bullet

In section 4.4, we constructed a case for biting the bullet and accepting the deflated premises and conclusion of the cheap trick. Taken most ambitiously, this case is based on three lines of support:

B1. *Only game in town:* every obvious strategy for attacking the premises of the cheap trick fails. This, one might think, is the moral of section 4.3.

B2. *A consequence of two-dimensionalism:* a uniquely attractive (and independently motivated) theory of knowability - namely, neo-Fregean epistemic two-dimensionalism - entails the premises of the cheap trick, and thus its conclusion. (See 4.4.2.)

B3. *Vacuous knowledge is respectable:* the conclusion of the cheap trick is properly interpreted as saying that vacuous knowledge is amply available. The possibility of cheap but vacuous knowledge of empirical facts is a respectable neo-Fregean position, that follows from three key theoretical positions: (i) that the meaning of φ is a 2D proposition $\langle P, Q\rangle$; (ii) that Q, the guise of φ, represents the cognitive significance of accepting φ; (iii) if Q is trivial, then knowledge of φ is easily acquired, but cognitively insignificant.

Here are some tentative considerations that weigh against this case.

First, note that **B1** and **B2** are undermined significantly if resolution theory is accepted as a sensible semantics for knowability ascriptions. For then we have a sensible proposal that *does succeed* in denying a premise in the cheap trick, and does so without relinquishing the key advantages of 2Dism.

Second, suppose we accept that if one believes the content of "things are exactly *this* way" (on the basis of noticing that this sentence *must* be true), then one cognizes this content under a guise that renders it of little cognitive significance. However, unlike the above bullet-biter, one might hesitate to call this belief an instance of *knowledge*. Does the ordinary knowledge concept really allow for fundamentally useless knowledge? Knowledge is ordinarily taken to be a valuable (sometimes scarce) resource.

Perhaps debating whether to include vacuous knowledge in one's theory is merely a *verbal* dispute, with an outcome under-determined by ordinary linguistic data. The bullet-biter is inclined to draw a distinction between two kinds of knowledge: vacuous (cognitively insignificant) knowledge, and proper (cognitively significant) knowledge. Others might be inclined to rule that cognitively insignificant beliefs cannot constitute knowledge. Perhaps we should not expect the ordinary knowledge concept to decide every obscure case that has arisen in the debate on the contingent a priori.

However, it seems more *in the spirit* of the ordinary knowledge concept to declare cognitive significance a necessary requirement for a knowledge state. Again, knowledge regulates rational action: to know fact P is to be in a position to act, appropriately, as if P is true. But if an object of thought is cognitively insignificant, presumably it fails to guide action in any meaningful way. Certainly, this seems to be the case when one entertains (and even accepts) "things are exactly *this* way".

Third, a theory that accommodates the contingent a priori (or, at least does so in the manner of 2Dism) might be at odds with ordinary linguistic data. Consider a standard candidate for the contingent a priori, of the indexical kind (due to [Kaplan, 1989, pp. 508-509]): "I am here now", uttered in context. Suppose I make the following two claims, in the same context:

- (1) I am in a position to know a priori that I am here now.
- (2) Gideon is not in a position to know that I am here now.

(2) is clearly true. But I do not see how this could follow if one is committed to (1) on the basis of the meaning of "I am here now" (in context).[40] For, surely the embedded 'that' clauses in the above two sentences contribute the same semantic value to the sentences as a whole, and presumably the relation expressed by 'knows' remains constant whether talking about myself or Gideon.

(None of this is to deny the obvious fact that it is *easy* for a speaker to know "I am here now" (and less easy for someone else to acquire the same knowledge, and less easy for the speaker to know another claim "Ortcutt is in the Stanford library on 26 June 1982" that expresses the same fact). However, this is easily explained as follows: being in the speaker's position immediately furnishes empirical information that suffices for knowledge of "I am here now". One need not suppose that *no* empirical information is required for this knowledge.)

It is also possible to offer a *principled* argument against accommodating the contingent a priori. Consider the following principles (as in [Chalmers and Rabern, 2014, pg. 4]):

A1. If it is knowable a priori that φ is true then it is necessary that it is knowable a priori that φ is true.

A2. It is necessary that: if it is knowable a priori that φ is true then φ is true.

It follows straightforwardly from these principles that: if it is knowable a priori that φ is true then it is necessary that φ is true. (Note that I assume that the distribution axiom for modal logic is valid.) This has been identified as a puzzle for 2Dism,[41] leading Chalmers and Rabern [2014] to reject **A1** and Fritz [2013] to reject **A2**. However, if the sanctity of contingent a priori knowledge is in doubt, there seems a much simpler response: **A1** and **A2** are intuitive principles that together entail the rejection of the contingent a priori.

[40] Cf. [Soames, 2007a, pg. 284].
[41] See Forbes [2011].

Chapter 5

Conclusion and Further Directions

5.1 Overall conclusion

Over the course of the last three chapters, we have answered the challenges to resolution theory outlined in chapter 1, section 1.4.

The first challenge was to render the closure denial of resolution theory a respectable position. In chapter 3, I concluded that resolution theory endorses a *restricted closure principle* that, apparently, avoids both egregious violations and non-violations of closure, relative to an approach that takes the generation of a skeptical paradox as grounds for discarding an instance of closure.

The second challenge was to respond to forceful arguments for the conclusion that easy knowledge is rampant (and so there is nothing objectionable about a theory of knowability that preserves closure at the cost of positing knowledge ascriptions that express vacuous knowledge). In chapter 4, I argued that this is best viewed as a *deflationary* rationale for rampant easy knowledge. This leaves the question open as to how best to theorize about *substantive* knowability, and I proposed that resolution theory exhibits key advantages on this front. For one, it explains why Gibbard's cheap trick fails when applied to substantive knowability. For another, it incorporates the key explanatory advantages of epistemic two-dimensionalism.

The third challenge was to explain why resolution theory does not deliver the wrong verdict when confronted by the *Jane's missed clue* case. At the end of chapter 4 (section 4.6), I proposed that resolution theory delivers the right result when a natural Fregean guise is associated with the content that is being evaluated for knowability (by Jane).

The fourth challenge was to ensure that resolution theory does not lean on an account of relevance that is ill-motivated. In chapter 2, I defend the notion of relevance at play in resolution semantics. In particular, I argue that the underlying theory of subject matter uniformly meets an extensive list of intuitive desiderata for such a theory, in contrast to various extant theories in the literature.

I conclude that resolution theory is a promising candidate for the correct theory of knowability. Not only does it incorporate compelling insights from various quarters, but it refines and combines them in a subtle and explanatory manner.

5.2 Further directions

Various avenues for further study present themselves.

Incorporation into a theory of knowledge. As I have emphasized throughout, resolution theory is a theory of knowability, thereby exploring a *necessary condition* on having knowledge. We have said virtually nothing about how to extend this account to a full-fledged theory of knowledge. In particular, how should our theory of knowability interact with belief, and with justification, in a complete theory?

Various proposals could be explored. Here is a simple one: to know P is to (i) be in a position to know P given one's empirical information E and (ii) believe P on the (ultimate) *basis of* one's empirical information. Note that this account does *not* say anything about justification. In terms of semantics, one might propose: "a believes that φ" is true just in case: "a is positioned to know that φ" is true, a believes the content of φ and a disbelieves every defeater for φ.

What should the relationship between knowability and justification be? Here is one (recursive) suggestion worth exploring: an agent a has *propositional justification* for P just in case either (i) a is positioned to know P relative to her information or (ii) P is supported by the content of one of a's propositionally justified beliefs. This is a *knowledge-first* approach to the interaction between justification and knowledge (cf. Williamson [2000]).

Relationship to contextualism. We have avoided engaging with the complex debate between contextualists, subject-sensitive invariantists and relativists. Resolution semantics, as a topic-sensitive account, is primed to introduce a contextualist dimension for knowability ascriptions. For there are two basic sources of subject matter in a conversational context: the topic of the discourse and the topics associated with individual claims (that these come apart is obvious when we think about the possibility of stating a claim that is *off-topic* relative to the discourse). We have concentrated on the latter, understood as a semantical fact. However, one might think that epistemic relevance is determined by *both* the discourse topic and the topic of the claim being evaluated for knowability. There is a complex network of motivations for and objections to contextualism, however, that such a proposal must contend with.[1]

Epistemic logic. Over the course of our discussion, we have presented aspects of resolution theory with precision. This culminates with the presentation in appendix B. Besides gesturing at various validities for this framework in chapters 3 and 4, we have not embarked on a serious logical study of its features. The scope here for further work is large: up for grabs are completeness results; expressivity results; and complexity results.

Understanding similarity to actuality. Resolution theory operates at two levels. At the level of the theory of the nature of knowledge, it posits a notion of relevance in terms of *similarity to the actual world*. At the level of the theory of knowledge ascriptions, it posits a notion of relevance in terms of *subject matter* and *Fregean guise*. We have concentrated on motivating the latter. When it comes to an account of similarity, we have relied on the thought that such an account is independently of interest, however we finally make sense of it, since it grounds the orthodox semantics for counter-factual conditionals. Nevertheless, it remains an open question how best to understand similarity (if this can be accomplished at all) and such a study is of utmost import for grounding resolution theory.[2]

[1] See Preyer and Peter [2005], Stanley [2005], DeRose [2009] and MacFarlane [2005].

[2] Cf. the suggestions in Kripke [2011] that the notion of 'relevant alternative' in epistemology does not neatly align with that of 'nearest/nearby possibility'.

Appendix A
Relevance in Epistemology and Logic

The current appendix offers a more detailed introduction to the relevant alternatives (RA) approach to the theory of knowledge and knowability, and provides some indication of the complex landscape such theories inhabit. I aim to emphasize the breadth and versatility of the RA approach - at least in the very general form we expound and develop it here.[1]

Another important theme is that RA theory, in its many guises, is typically amenable to being studied with precise formal methods. The RA approach is, therefore, not only a unifying framework for diverse, nuanced and intriguing philosophical theories of knowledge (encompassing a significant bulk of major recent developments), but is also a notable site for the interaction between epistemology and logic. This interaction extends fruitfully in both directions [Holliday, 2012, 2013, 2015a,b]. In one direction, logical techniques allow the RA theorist to operate at an unusual level of technical precision when framing rival positions and their consequences. In the other, the RA approach is a source of novel, sophisticated variants of epistemic logic, worthy of detailed logical study in their own right.

In the next section, I review the main motivations for an RA approach, citing, for instance, some striking linguistic considerations and the idea that RA theory captures the 'common man' response to Cartesian skepticism. In addition, we briefly draw out connections and contrasts between the RA approach and similarly themed discussions in the logic, epistemology and scientific methodology

[1] This appendix refines and significantly compresses Hawke [2016b].

literature. In the third section, I propose a series of basic 'choice points' for the RA theorist. The leading claim here is that a particular, 'concrete' RA theory of knowability is essentially the product of settling each choice point. Hence, our list of choice points offers a basic schema for *classification* of RA theories and provides a tool for studying RA theory at different levels of abstraction (where a higher level of abstraction corresponds to leaving more choice points open). We discuss each choice point in turn and briefly mention techniques for formalizing the options at each choice point. With that, we conclude.

A.1 The motivation for an RA approach

Why be an RA theorist? We outline a number of important motivations (many of which are related, though still worth separating out). Here's an overview.

- *Striking linguistic data*, concerning ordinary epistemic claims, seem to support an RA approach.
- The RA approach provides a unique and compelling reply to *Cartesian skepticism.*
- More generally, RA theory provides a universal strategy for dealing with *underdetermination problems.*
- RA theory is suggested by our intuitive reaction to *Goldman-Ginet barn cases.*
- The RA framework has theoretical value (for contextualists and others) as a tool for modeling *epistemic standards.*

A.1.1 Suggestive linguistic data

Two kinds of purported linguistic data have been used (in concert) to support the RA approach. First, linguistic data seems to indicate a *fallibilist* aspect to our ordinary knowledge concept. That is, it seems that ordinary agents will sometimes happily attribute knowledge to themselves (or others), but, *if pressed*, will concede that certain possibilities for error are compatible with the available evidence. Second, linguistic data seems to indicate an *infallibilist* aspect to our ordinary knowledge concept. That is, ordinary agents seem uncomfortable to state the conjunction of a knowledge claim with an explicit acknowledgement of live possibilities of error.[2] These points are emphasized by both Dretske [1981] and Lewis [1996], following the observations of Unger [1975].

Clearly, there is a tension between these tendencies of ordinary knowledge ascriptions. The RA approach is alleged to have the capacity to resolve this tension.

To illustrate the fallibilist tendency, we (ab)use an influential case. Dretske [1970] points out that, under ordinary circumstances (using ordinary visual evidence), one is happy to say that one knows that the zebra-looking animal one sees at the zoo - in the zebra enclosure - is a zebra. But one will be *less* happy to say that one knows that the animal is not a mule painted to appear exactly like a zebra. The immediate visual evidence does not to settle the latter issue.

The legitimacy, exact diagnosis and consequences of this purported linguistic data are controversial [Vogel, 1990, Luper, 2016]. For our purposes, however, the description of the example can be altered in a telling way, by weakening the proposed judgement concerning the "painted mule" possibility: under ordinary circumstances - it is plausibly suggested - one is happy to say one *knows* the enclosed animal is a zebra, yet, if pressed, one will hesitate to add that one's evidence *rules out* that the animal is a painted mule. Thus, it appears we have everyday linguistic data to the effect that we are often willing to ascribe knowledge of φ, yet will quickly concede the limitations of the available evidence when it comes to ruling out certain alternatives to φ.

As has been pointed out by critics of the RA approach [Vogel, 1990, 1999], this modest reading of the zebra case supports fallibilism in general, rather than the RA approach in particular. Consider, for instance, a Bayesian that holds that knowledge of φ is a matter of not-φ being *sufficiently improbable* on the evidence. Such a Bayesian is a rival to the RA theorist that sees no need for evidence (in its role as a constraint on the space of 'serious' possibilities) to be supplemented with an independent notion of "relevance".

This form of Bayesianism, however, is not as effective as RA theory in accounting for the *infallibilist* tendencies in our ordinary knoweldge ascriptions. Ordinary speakers feel uncomfortable in making or accepting claims along the

[2] This point is best emphasized by DeRose [1995], who labels such conjunctions 'abominable'.

following lines: "I know that φ, though not-φ might well be the case"; "I know that φ, yet my evidence does not vouchsafe certainty that φ"; "I know that φ, though not-φ remains a live possibility". Lewis [1996] sums up the sentiment:

> If you claim that a knows that P, and yet you grant that a cannot eliminate a certain possibility that not-P, it certainly seems as if you have granted that a does not after all know that P. To speak of fallible knowledge, of knowledge despite uneliminated possibilities of error, just *sounds* contradictory [Lewis, 1996, pg.549, his emphasis].

Now, according to the Bayesian, one can know φ when the probability bestowed on φ by the evidence meets an appropriate threshold. But if this threshold is less than 1, then the Bayesian is committed to the possibility that an agent may know that φ and yet not-φ has non-zero probability and, so, is compatible with (if unlikely on) the evidence.

The RA theorist, on the other hand, has a trick to play. She can account for our fallibilist tendencies: if φ is known then an alternative A may well be identifiable as uneliminated by the evidence (for A might be irrelevant). On the other hand, the RA theorist can account for our infallibilist tendencies. She can agree that to know is to leave *no* possibility for error (in some important sense), but add that what is *implicit* in this saying is that the possibilities being quantified over are the *relevant* ones.[3]

A.1.2 The RA strategy against skepticism

Consider the following Cartesian argument for a skeptical conclusion:

P1. To be a handless brain-in-a-vat is an alternative to having hands.

P2. The evidence in my possession is not sufficient to rule out that I am a handless brain-in-a-vat.

P3. In order to know φ, one needs to have evidence that rules out all alternatives to φ.

[3]As we see in the closure debate, our ordinary infallibilist tendencies are used as a weapon in *internal* debates among RA theorists, suggesting the possibility that some versions of RA theory are better suited to account for these tendencies than others. For instance, DeRose [1995] influentially criticizes Dretske's version of RA theory as incorrectly predicting that abominable conjunctions are felicitous in ordinary conversational contexts.

C. Therefore: I do not know that I have hands.

This argument is valid, and **P1** and **P2** might strike one as undeniable (to deny them, it might be said, is to fail to appreciate the nature of the brain-in-vat scenario, or to forgo a sensible epistemic modesty). But there is a way out: deny **P3**. Of course, this is simply to embrace the RA slogan. Thus, the RA approach has theoretical value as a tool for resolving a key philosophical paradox.

Something along these lines is a common response from the layman (i.e. non-philosophers) when skeptical possibilities are raised: the reaction is to deride that possibility as too far-fetched to impact our ordinary epistemic concerns. This reaction seems particularly apt when practical applications of everyday knowledge are afoot. It is in no way adequate to respond to the question "do you know where I left my keys?" with "no, for I cannot rule out that my senses are being deceived by an evil demon". To the extent that she is willing to take the layman as a competent user of the knowledge concept, the RA theorist finds this reaction telling.[4]

A.1.3 RA theory as a response to under-determination problems

The previous motivation can be generalized. Cartesian skepticism, at least in certain forms, is an instance of the class of (what may be called) *under-determination problems*. An under-determination problem has the following form: it is obvious that we know that P, yet, on close inspection, our supposed evidence for P seems compatible with some (maybe odd, but logically possible) alternative Q. Another prominent under-determination problem: Hume's problem, which notes that our sensory evidence of particulars seemingly under-determines the general knowledge held upon its basis. Thus, under-determination has proved a pressing issue in philosophy of science (cf. Stanford [2016]).

The RA approach embodies a universal strategy for dealing with under-determination problems: simply establish that every deviant alternative is properly classified as irrelevant.

[4]Of course, this may be taken as further linguistic data, in the spirit of section A.1.1.

A.1.4 RA theory by way of the Goldman-Ginet barn cases

The Goldman-Ginet barn case from Goldman [1976] has proven a particularly influential example in the contemporary epistemology literature. Suppose subject a clearly observes what is in fact a (genuine) barn out of her car window, as she drives by. Does she know that it is a barn? Our reaction to this question depends, it seems, on whether a is driving through a county in which the only objects that look like barns to the casual observer are, in fact, barns (in which case, she does know), or if she is in the unusual situation where there are as many barn facades ("fake barns") around as real barns (in which case, she does not).

What does the barn case teach us? The RA theorist suggests that it exhibits that an agent a could have exactly the same evidence in states S and S' (not to mention the same beliefs), and yet a knows that φ in S and does not know that φ in S' (where φ is true in both S and S'). This difference, the RA theorist urges, can be accounted for in a natural way: the alternatives to φ that are relevant differ from one case to the other. In particular, a's evidence does not rule out an alternative (that the object is a barn facade) that happens to be irrelevant in S, but is relevant in S'.[5]

A.1.5 RA theory and epistemic standards

Various contemporary authors defend a version of the idea that the epistemic standards that an agent must meet to know φ can vary from context to context. What is chiefly debated, among such authors, is *which* context determines the relevant standards: is it that of the *subject* to whom knowledge is potentially attributed [Stanley, 2005, Hawthorne, 2004], that of the *speaker* who is performing the attribution [Cohen, 1988, DeRose, 1995, Lewis, 1996], or that of an *assessor* potentially different to both speaker and subject [MacFarlane, 2005]?

Whichever view one takes, such perspectives on the semantics of knowledge claims are a natural fit with RA theory. For how are we to understand the idea of an *epistemic standard*? A natural suggestion is that a variation of epistemic

[5] Note that the barn case can be seen to teach a similar lesson to Cartesian skepticism: that one can know something even though one has not ruled out *all* alternatives. However, the barn case potentially teaches us something more: that what counts as a relevant alternative can *vary* with the circumstances: the possibility of fake barns may be properly ignored by knowledge ascribers under one set of circumstances, but is not properly ignored in another.

standards consists in a variation of the space of alternatives that must be ruled out: a higher standard involves more relevant alternatives.

Thus, it is a short path from accepting that epistemic standards vary by context to an acceptance of the RA approach. Arguments for the former may therefore, with the right massaging, be taken as support for the latter.[6]

A.2 Connections and contrasts

In this section, we discuss potential connections and contrasts between the RA approach and other salient developments in the epistemology and logic literature. We will aim for a sense of the theoretical promise of the RA approach (insofar as it can be integrated and unified with similarly motivated concerns in other strands of the literature) while also distinguishing it from only superficially similar projects.

A.2.1 Relevant logic

Begin with *relevant logic*.[7] This area of logic is animated by a desire to build (technically and philosophically sound) logics that avoid validating so-called "fallacies of relevance". In particular, the relevant logician is concerned to avoid two counter-intuitive results of classical logic: that any sentence is a valid consequence of contradictory premises, and that a necessary truth is a valid consequence of any set of premises whatsoever. The difficulty with these results, the relevant logician claims, becomes evident when we consider arguments where the premises and conclusion are *irrelevant to each other*, insofar as they concern disjoint *subject matter*: it does not follow from the claim that the moon is both made of green cheese and not that Barack Obama is president of the USA (nor, for that matter, does this follow from 2+2=5). Further, it does not follow from the fact that Berlin is the capital of Germany that either it is raining in London or it is not (nor, for that matter, that 2+2=4).

Thus, the relevant logician has two chief concerns: (i) to offer an account of when one proposition is "relevant" to another (which at least partly involves those

[6]For a more careful defense of the 'alternatives' approach to capturing the relevant parameter that shifts across contexts, see Schaffer [2005b].

[7]For thorough introductions, see Anderson and Belnap [1975], Burgess [2009] and Mares [2014].

propositions overlapping in subject matter) and (ii) an integration of this account into a logical system, so that only *relevant* conclusions are valid consequences of a set of premises. The concerns of the relevant logician and a RA theorist overlap, therefore, to the extent that (i) and (ii) are pertinent to the RA theorist in question.

Is (i) pertinent to an RA theorist? This will depend on whether the RA theorist and relevant logician mean the same thing by "relevance". Since they are motivated by different starting points (intuitions that point to alternatives that can be properly ignored when evaluating knowledge claims, versus intuitions concerning logical consequence) there is no guarantee that there will be a convergence here. Indeed, there is a quick argument that "relevance" as deployed by a standard RA theorist must have a different import than that deployed by the relevant logician. For: suppose we follow the standard line (we return to this in section 3.5) and say that proposition A is an alternative to P just in case P logically entails $\neg A$. Now, the RA theorist wishes to draw a distinction between "relevant" and "irrelevant" alternatives to P. But this distinction is therefore intended to apply to propositions that are logically related to P. Thus, *both* "relevant' and irrelevant" alternatives, in the RA theorist's sense, are "relevant" to P in the sense of the relevant logician.

Nevertheless, we discuss at various points this dissertation *topic-relative RA theorists* that attempt to account for relevance - in the RA theorist's sense - in terms of subject matter. For such RA theorists, a useful dialogue is to be had with the relevant logician on the nature of subject matter (cf. chapter 2).

Is (ii) pertinent to an RA theorist? On the face of it, the answer is 'yes'. Suppose our RA theorist has settled on an account of relevance. There is then clear theoretical interest for her in tools for building logical systems where relevance is preserved across the proposed logical consequence relation. What is not so clear, however, is that the results of relevant logic provide a general enough framework for carrying out this job for an arbitrary RA theorist, since, again, relevant logic focuses on a notion of relevance closely tied to the preservation of subject matter.

In total: the basic concerns of relevant logic (at the very least) indicate an intriguing notion of relevance tied intimately to that of subject matter, an obvious matter of interest to the RA theorist.

A.2.2 Epistemic relevance between evidence and hypothesis

We turn to a second tradition in philosophy in which the term "relevance" has received prominence. Here, the focus is on when a piece of evidence is relevant to the evaluation of a hypothesis. This is evocative of the concerns of the relevant logician: while the relevant logician is concerned with when a conclusion genuinely follows from its premises, one who investigates "epistemic relevance" in the present tradition is concerned with when evidence is genuinely a *reason* to accept (or reject) a hypothesis. The linchpin in this investigation is a probabilistic account: evidence E is relevant to hypothesis H just in case the conditional probability of H given E is different to the (prior) probability of H. The discussion in the literature - initiated chiefly by Keynes [1921] and Carnap [1950] - traces a series of refinements of this basic idea, as in Floridi [2008].

Analogously to the case of relevant logic, the discussion of "epistemic relevance" hinges on two basic concerns: (i*) what is the correct account of the relevance at issue? (ii*) How is this account to be integrated into a theory of evidential support? Once again, the extent to which the discussion of this sense of epistemic relevance relates to the concerns of the RA theorist depends on the extent to which answers to i* and ii* bear on these concerns.

Our remarks will mirror those those concerning the relationship between the RA approach and relevant logic. With respect to i*: the notion of relevance at work in the discussion of "epistemic relevance" is of interest to the RA theorist insofar as it represents, surely, *one* intriguing candidate for the notion of relevance the RA theorist might identify as at work in the theory of knowledge (namely, a candidate that appeals to notions of *probability* and *independence* as crucial features). It remains to be seen, however, how far such a version of RA theory could be developed with plausibility. With respect to ii*: again, the RA theorist finds interest in any general techniques for integrating an account of relevance into a theory of reasoning or evidential support (perhaps in aid of a *relevant alternatives theory of justification* that underlies the RA theory of knowledge). The focus in the "epistemic relevance" literature on a very specific notion of relevance does not inspire hope, however, that very *general* tools for such integration are to be found there. Once again, however, our remarks are preliminary. Clearly, a deeper investigation is a worthwhile task.

A.2.3 Methodology of science

Various strands in the literature on the epistemology and methodology of science have a notable prima facie affinity to the ideas that animate the RA theorist. We sketch a few such points.

Begin with the plausible idea that the goal of scientific inquiry is knowledge. If we then agree with the RA theorist that knowledge is always relative to a set of relevant alternatives, we conclude that the methods of scientific inquiry - geared towards producing such knowledge - must themselves operate against the backdrop of a set of relevant alternatives. If so, we expect a notion of "relevant alternative" to play a role in both the *context of justification* and *context of discovery* of scientific hypotheses.

Ideas of roughly this ilk have received considerable attention in the literature. For instance, Kuhn [1970] famously proposes that major developments in science are revolutionary upheavals, brought about by a shift in the paradigm for normal science [Kuhn, 1970]. We need not here become engrossed in the substantive or scholarly issues connected to Kuhn's work. But note the potential for an RA approach to offer tools for understanding and investigating such revolutionary shifts: for an RA theorist, a change in paradigm can be understood as a shift in the *relevant alternative hypotheses* that a normal scientist selects between.

More specifically, consider the *context of justification*. The idea that scientific justification is a process of *eliminative induction* has found recent traction (cf. [Earman, 1992, Ch.7]): given a space of hypotheses H_1 through H_n, a particular hypothesis H_i is supported just in case the available evidence rules out every competing hypothesis. This account might at first seem naive, given the unwieldiness of the space of logically possible hypotheses and the inability of our actual evidence to rule out any significant portion of this space. A successful RA theory, however, can defuse this problem by constraining the space of pertinent hypotheses with an appropriate notion of relevance. RA theory seems, therefore, a potential ally to the eliminative inductivist.

Or consider the *context of discovery*. For the RA theorist, a natural way to understand the *discovery* of a new hypothesis is for that hypothesis - through whatever mechanism - to become *relevant* in the context of scientific inquiry. Whether this mechanism is rational or not depends on the exact account of relevance, and how the space of relevant alternatives might evolve. Compare Hintikka

[1999], who posits that a scientific discovery is the posing of a question to a source of information in nature. On this approach, scientific inquiry depends crucially on the *questions* that are (implicitly or explicitly) asked by scientists. To the extent that *background questions* may therefore be understood as a source of *relevant alternatives* (compare our discussion of *question-relative RA theory* in chapter 3), we see another opportunity for convergence between RA theory and the philosophy of scientific methodology.

A.3 Choice points for the RA theorist

In this section, we develop a technical framework for the RA theorist.

First, we present a 'minimal' RA theory that operates at a high level of abstraction yet captures core elements of the approach. The minimal theory is too abstract to engage fully with philosophical debate. Likewise, its abstractness precludes the interesting formal features of more concrete RA theories. To this end, we discuss the potential for considering precise RA theories with more content. We thus list a number 'choice points' for the RA theorist, by which one may divide the family of RA theories into a large number of species. To settle every choice point is to arrive at a 'concrete' RA theory.

Thus, minimal RA theory holds limited theoretical interest in itself. Rather, it is a unifying skeleton upon which to hang the features of more concrete RA theories. Nevertheless, the minimal theory highlights a philosophical point: at its most abstract, the RA approach is *very general*, a point for critics to keep in mind when aiming for blanket objections to the approach.

A.3.1 An epistemic language

Let **At** be a set of atomic proposition letters. We work with the following logical language $\mathcal{L}$:

$$\varphi \ ::= \ p \mid \neg\varphi \mid \varphi \wedge \varphi \mid K\varphi \mid R\varphi \mid I\varphi \mid [\varphi]\varphi$$

where $p \in \mathbf{At}$.

The rest of the connectives are defined as usual. $K\varphi$ is intended to mean "the agent is in an evidential position to know that φ". $R\varphi$ is intended to mean

"φ is relevant". $I\varphi$ is intended to mean "the agent has the information that φ". The intended interpretation of $[\varphi]\psi$ is "after the set of relevant propositions is updated so as to be those relative to φ, ψ is true". This last expression represents a *dynamification* of our logic (cf. van Benthem [2011]).

We may then define a *two-place* relevance operator: $\mathtt{R}(\varphi, \psi) ::= [\varphi]R\psi$. The aim is to express that ψ is relevant relative to φ.

A.3.2 Minimal RA theory

In what follows, $\mathcal{P}(A)$ refers to the power-set of set A.

A.3.1. Definition (Minimal RA model). A *minimal RA model* is a tuple

$$\langle W, \{\mathbf{R}_w\}_{w \in W}, \{E_w\}_{w \in W}, \{*_w\}_{w \in W}, V\rangle$$

where,

- W is a set of *points of evaluation*. The reader may think of these as "possible worlds", subsets of which are "unstructured propositions".
- $\mathbf{R}_w \in \mathcal{P}(\mathcal{P}(W))$ is a set of sets of worlds i.e. a set of propositions. This is the set of *relevant propositions* at world w.
- $E_w \in \mathcal{P}(W)$ is a set of worlds i.e. a proposition. This is the agent's *total information* at world w.
- $*_w$ is an *update operation* accepting a sentence $\varphi \in \mathcal{L}$ and returning an updated model we denote by $\mathcal{M} *_w \varphi$. We stipulate that the only distinction between $\mathcal{M}$ and $\mathcal{M} *_w \varphi$ lies in the relevant propositions.
- V is a valuation that assigns an unstructured proposition to each atom in **At**.

Given minimal RA model $\mathcal{M}$ and world w, define the set $\mathbf{U}_w$ as follows:

$$\mathbf{U}_w = \{A \subseteq W : A \in \mathbf{R}_w \text{ and } A \cap E_w \neq \emptyset\}$$

Call $\mathbf{U}_w$ the set of *uneliminated propositions* at w: the set of propositions that are both relevant and compatible with the agent's evidence at w.

Though we use the ‘worlds’ terminology for our points of evaluation, there is no technical necessity attached to this interpretation. One may equally well talk about ‘scenarios’, ‘centered worlds’, or so forth. (Though the totality of the propositional valuations associated with each world bars thinking of them as mere ‘situations’.)

We turn to semantics. In what follows, we use $|\varphi|_{\mathcal{M}}$ to denote the *truth set*

$$\{w \in W : \mathcal{M}, w \vDash \varphi\}$$

A.3.2. Definition (Minimal RA semantics). Given a minimal RA model $\mathcal{M}$, we define satisfaction at world w as follows:

- $\mathcal{M}, w \vDash p$ iff $w \in V(p)$.
- $\mathcal{M}, w \vDash \neg\varphi$ iff $\mathcal{M}, w \nvDash \varphi$.
- $\mathcal{M}, w \vDash (\varphi \wedge \psi)$ iff $\mathcal{M}, w \vDash \varphi$ and $\mathcal{M}, w \vDash \psi$.
- $\mathcal{M}, w \vDash R\varphi$ iff $|\varphi|_{\mathcal{M}} \in \mathbf{R}_w$.
- $\mathcal{M}, w \vDash I\varphi$ iff $E_w \subseteq |\varphi|_{\mathcal{M}}$.
- $\mathcal{M}, w \vDash K\varphi$ iff $\{A \in \mathbf{U}_w : A \subseteq |\neg\varphi|_{\mathcal{M}}\} = \emptyset$.
- $\mathcal{M}, w \vDash [\varphi]\psi$ iff $\mathcal{M} *_w \varphi, w \vDash \psi$.

The clause for $I\varphi$ says: the agent has the information that φ just in case the agent’s total information entails φ. The clause for $K\varphi$ says: the agent is positioned to know φ just in case there is no proposition that entails $\neg\varphi$ that is uneliminated i.e. both relevant and compatible with the agent’s total information. The expression $[\varphi]\psi$ is satisfied if ψ holds after the relevancy sets have been updated using the $*_w$ operation, with φ as input.[8]

Some of our clauses fall within the tradition of *neighbourhood semantics* for modal logic [Chellas, 1980], where the truth clause for $\Box\varphi$ (“it is necessary that

[8]The reader will note that we make no mention of a notion of ‘context’ anywhere in this semantics. We gloss over the role of context, as follows: context may be thought of as settling the valuation V and, potentially, the set of relevant alternatives $\mathbf{R}_w$. Thus, context may be thought of as settling the model in question. We do not explore this thought in any detail here.

φ") is given as: $\Box\varphi$ holds at world w just in case the set of worlds where φ holds is one of a set of "necessary propositions" associated with w.

Presumably, knowability is factive: if P is knowable, then P is true. A flat-footed refinement of our system that ensures this is to add a 'truth condition' to the truth clause for $K\varphi$ i.e. add that φ is true at w. (This is, of course, a structural feature of countless proposed theories of knowledge.) However, such a maneuver might strike some as ad hoc and unsatisfying. Further, an independent 'truth condition' seems to be exactly what allows for the Gettier cases that dog epistemology (cf. Zagzebski [1994]). Thus, an advantage of our RA framework is that it offers alternative routes for ensuring that truth is entailed by knowability. For instance, one could constrain the class of models by stipulating that any true proposition at w is relevant at w and that the proposition E_w is a *true* proposition (i.e. $w \in E_w$).

A.3.3 Choice points

We turn to a series of choice points that the RA theorist must settle in order to fill out the minimal approach.

- What is 'relevance'?
- What is an 'alternative'?
- What is it to 'rule out' an alternative?
- What is the structure of the space of relevant alternatives?
- Does the (ir)relevance of a claim only make sense in contrast to another claim?
- Interaction principles: is relevance a necessary condition on knowing? Is irrelevance a sufficient condition for knowing the denial?

(I do not claim that these choice points are entirely independent of each other: settling one choice in a certain way may well constrain how other choice points must be settled.)

A.3.4 Relevance

What is it for an alternative to be relevant? The literature on RA theory offers a bewildering diversity of suggestions, but little in the way of detailed theories of relevance. Cohen [1999], on page 61, suggests that relevance is a matter of the psychology of the agents in conversation, "determined by some complicated function of speaker intentions, listener expectations, presuppositions of the conversation, salience relations etc". Heller [1999] suggests that relevance is a matter of similarity to the actual world, where the similarity relation is itself settled partially by psychological facts - intentions, salience and so forth - of the speakers in context. Lewis [1996] suggests a complex array of factors that determine relevance, ranging from salience to the speaker to practical stakes. In contrast, Dretske [1981] is somewhat non-committal, but indicates some commitment to the idea that relevance is a purely *objective matter*, independent of any agent's state of mind.

A.3.5 Alternatives

What sort of thing is an alternative? The standard approach treats alternatives as propositions, where A is an alternative to P just in case A entails the negation of P. In our minimal model, we have treated alternatives as *unstructured propositions* i.e. as sets of possible worlds.

This could be refined: an alternative could be understood as a *situation* or *set of situations* [Barwise and Perry, 1981], where formally a situation is associated with a partial valuation; as a structured proposition; as a 'centered' proposition (i.e. as a set of centered worlds); as an interpreted sentence; or as a possible world. On the last approach, a world w is an alternative to proposition P just in case $\neg P$ holds at w. We can capture this within our current framework by adding the restriction that only *singleton sets* can act as relevant propositions.

A.3.6 Ruling out

What is it for an agent's total information to rule out an alternative? In our RA semantics, we utilize a notion of *incompatibility* between unstructured propositions P and E: namely, that the intersection of P and E is empty. This could be supplemented with a richer construal of 'ruling out', in at least two ways. One route is to connect the notion of 'ruling out' with the agent's rational belief: an alternative A is ruled out for the agent just in case it is rational for the agent to find A implausible. Call this the *soft approach* to ruling out. The formal treatment of rational belief (and belief update) is well developed (see, for instance, Ch. 7 of van Benthem [2011]) and could easily be the integrated into our RA framework. More abstractly, one could simply interpret E in a minimal RA model as encompassing not only the agent's total perceptual evidence, but the rational beliefs that are grounded in that evidence.

A second approach gives 'ruling out' a stronger reading: for A to be ruled out for the agent is for the agent to (be positioned to) *know* that A is false. Call this the *hard approach* to ruling out. This account has an interesting consequence: if A being an alternative to P means that P entails $\neg A$, then, on the current view, the RA theorist is committed to the idea that an agent can know P without knowing all of its entailments. For, in this setting, to say that an agent is in positioned to know P but not positioned to rule out irrelevant alternative A, is just to say that P is knowable without $\neg A$ being knowable. Thus, the hard approach and the denial of epistemic closure are intimately related. Formally, they may be construed as equivalent.

A.3.7 The structure of the relevant alternatives

What is the structure of a space of relevant alternatives? Certain developments in the literature show that is worth approaching this question as follows: should we think of the relevance of alternatives as *derived* from another kind of object to which relevance more fundamentally applies?

For instance, Heller [1989] suggests that *possible worlds* should be treated as the primitive objects of relevance. Call this the *worlds-first approach*. How then do we recover the relevance of propositions? One option: we simply take the set of relevant propositions to be the set of singletons $\{\{w\} \mid w \text{ is relevant}\}$. Formally, this leads us down a similar path to treating *alternatives* as singletons (i.e. as individual possible worlds). Another option is to say that A is relevant just in case A holds at *some* relevant world.

On the other hand, as we discuss at length in chapter 3, Jonathan Schaffer proposes in a series of recent papers that knowledge claims can only be evaluated relative to a given *question*. Call this the *question-first approach* [Schaffer, 2004, 2005a, 2007a,b]. To know something, according to this idea, is to know it *rather than* other possible answers to the question, while non-answers and presuppositions to the question are simply ignored. The idea is that in answer to the question "is there a zebra in the cage or nothing at all?" one may know that there is a zebra, but in answer to the question "is there a zebra in the enclosure or a painted mule?" one may not know that there is a zebra. This suggests the following notion of relevance for propositions: A is relevant relative to (relevant) question $\mathbf{Q}$ just in case A is an *answer* to $\mathbf{Q}$. As also discuss in chapter 3, we can draw on an ongoing tradition in the semantics literature for treating questions formally [Hamblin, 1958, 1973, Belnap and Steel, 1976, Ciardelli et al., 2015]. Along this line, we may understand a question $\mathbf{Q}$ as a set of disjoint propositions, representing the set of (least specific) answers to that question. An answer to the question is then any subset of a member of $\mathbf{Q}$. A *partial* answer is is any *union* of subsets of $\mathbf{Q}$. A presupposition to the question is any proposition that contains every member of $\mathbf{Q}$.

To display one last approach, Yablo [2014] proposes, as we discuss in chapter 3, that knowledge claims be evaluated relative to a background *subject matter*. Call this the *topic-first* approach. On this view, one can know that the enclosure contains a zebra so long as the subject of painted mules is suppressed. One can then derive the relevant propositions as follows: A is relevant relative to (relevant) topic $\mathbf{T}$ just in case A 'speaks to' $\mathbf{T}$. Again, fortunately, there are proposals for a formal treatment of subject matters: for instance, following Lewis [1988a], a subject matter can be understood as a *partition* on the space of possible worlds, with two worlds sharing a cell just in case they are exactly the same when it comes to any state of affairs concerning that subject matter (for a more thorough discussion, see chapter 2).

A.3.8 Contrast

Does talk of the 'relevance' of a proposition A makes sense only relative to another claim, to which A must be *contrasted*? Let us say that a theory that answers this question in the affirmative takes the *contrast approach*.

The Dretskean theory of knowability (as in chapter 1) serves as an illustration. On this approach, a relevant alternative to P is a possible world at which P is false, and for which there is no nearer world to actuality at which P is false.

Thus, it makes sense to call w (or $\{w\}$) a relevant alternative only relative to a fixed P, with which it contrasts.

On the other hand, Lewis [1996] does not subscribe to the contrast approach. For Lewis, once the context is fixed, a proposition A is uniformly relevant (or not), no matter the specific claim being evaluated for knowledge.

Subscription to the contrast approach has far-reaching consequences: Holliday [2013] and Holliday [2015a] shows that this is the source of epistemic closure failure in various RA theories.

We can formally capture acceptance of the contrast approach as follows: define $\mathtt{K}\varphi$ ('proper knowability') as follows: $\mathtt{K}\varphi ::= [\varphi]K\varphi$ i.e. proper knowledge of φ is understood as the elimination of the relevant alternatives to φ in the wake of an update that relativizes relevance to φ. On this approach, one treats $\mathtt{K}\varphi$ as the chief object of one's theory of knowability (with K serving a merely technical role). That is, the contrast approach can be incorporated by stipulating that evaluating the knowability of φ requires fixing a relevancy set determined by φ (cf. Holliday [2012]).

A.3.9 Interaction principles between relevance and knowledge

What logical relationship should exist between the relevance of a proposition and knowledge of that proposition? Should there be *no* such relationship? Should the relevance of A act as a necessary condition on knowledge of A (that is, should only *relevant* propositions count as candidates for knowledge)? Should the irrelevance of A be sufficient for $\neg A$ to be known, or for A to be *not* known?

In terms of integration into our framework, stipulating that relevance be a necessary condition on knowledge is at least a simple matter: we simply add the condition $\mathcal{M}, w \vDash R\varphi$ to the clause for $K\varphi$.

A.4 Conclusion

That concludes our whirlwind tour of the landscape of RA theories. We have accomplished the following: we have seen a number of informal philosophical motivations for embracing the RA approach, ranging from ordinary linguistic data to lessons from famous philosophical examples; we have discussed a minimal framework for formalizing RA theory; and have identified various choice points that an RA theorist must decide upon in the construction of her specific theory.

Appendix B

Resolution Logic

We have developed various aspects of resolution semantics with precision (in particular, in chapter 2, section 2.5.3; in chapter 3, section 3.6; and in chapter 4, section 4.5). However, these efforts were piecemeal. In this appendix, I develop (a version of) resolution semantics as a complete logical framework. Doing so leaves us well-positioned to (elsewhere) study the resulting epistemic logic in detail. This investigation should not be regarded as an exercise in mere technicalities, for it raises substantive philosophical and semantic issues related to higher-order knowledge claims: what approach is best for assigning a subject matter and a guise to a knowledge ascription? Relatedly, what approach is best for modeling the relevant alternatives to a knowledge ascription? In what follows, I fill out such details. It is likely that there are alternative proposals that the resolution theorist should give serious consideration. I leave a comprehensive assessment of my choices for elsewhere.

I assume familiarity with the features of resolution theory (and their motivation) that have been discussed in detail in preceding chapters.

B.1 Syntax

We work with a formal language $\mathcal{L}$, described as follows. Our language is composed from predicate letters, terms, the usual connectives, the identity symbol $=$ and the knowability operator K. Let **Pr** be a countable set of predicate letters. To keep things simple, we work only with one-place predicate letters, which we denote by $F, G, H, \ldots$. Let $\mathbf{Rd}$ be a countable set of *name terms* $n, m, \ldots$. Let $\mathbf{Ro}$ be a countable set of *role terms* $r, s, \ldots$. We stipulate that for every $n \in \mathbf{Rd}$ there is a distinguished element in $\mathbf{Ro}$, denoted by n^f. We refer to n^f as the *Fregean guise* of n. We use $\mathbf{Te}$ to denote $\mathbf{Rd} \cup \mathbf{Ro}$, which we refer to as the set of *terms*.

Then, our language is composed recursively as follows:

$$\varphi \ ::= \ Ft \mid t = t \mid \neg\varphi \mid \varphi \wedge \varphi \mid K\varphi$$

where $t \in \mathbf{Te}$ and $F \in \mathbf{Pr}$.

The rest of the connectives are defined as usual. $K\varphi$ is read as "it is knowable for the agent that φ".

We define the set of *atomic claims* $\mathbf{At} \subseteq \mathcal{L}$ as follows:

$$\mathbf{At} ::= \{Ft \mid F \in \mathbf{Pr} \text{ and } t \in \mathbf{Te}\} \cup \{t_1 = t_2 \mid t_1, t_2 \in \mathbf{Te}\} \cup \{K\varphi \mid \varphi \in \mathcal{L}\}$$

We now recursively define two functions: the *Fregean guise function* $\mathfrak{f}$ and the *acquaintance condition function* $\mathfrak{a}$. We write φ^e as shorthand for $\mathfrak{f}(\varphi) \wedge \mathfrak{a}(\varphi)$. We call φ^e the *epistemic guise* of φ.

B.1.1. Definition (Fregean guise). We define the operation $\mathfrak{f} : \mathbf{Te} \cup \mathcal{L} \to \mathbf{Te} \cup \mathcal{L}$ as follows:

1. If $n \in \mathbf{Rd}$ then $\mathfrak{f}(n) = n^f$
2. If $r \in \mathbf{Ro}$ then $\mathfrak{f}(r) = r$
3. $\mathfrak{f}(Ft) = F\mathfrak{f}(t)$
4. $\mathfrak{f}(t_1 = t_2) = (\mathfrak{f}(t_1) = \mathfrak{f}(t_2))$
5. $\mathfrak{f}(K\varphi) = K(\mathfrak{f}(\varphi))$
6. $\mathfrak{f}(\neg\varphi) = \neg\mathfrak{f}(\varphi)$
7. $\mathfrak{f}(\varphi \wedge \psi) = \mathfrak{f}(\varphi) \wedge \mathfrak{f}(\psi)$

We use the shorthand t^f to indicate $\mathfrak{f}(t)$ and φ^f to indicate $\mathfrak{f}(\varphi)$.

B.1.2. Definition (Acquaintance condition). We define the operation $\mathfrak{a} : \mathcal{L} \to \mathcal{L}$ as follows:

1. $\mathfrak{a}(Ft) = (t = t^f)$
2. $\mathfrak{a}(t_1 = t_2) = (t_1 = t_1^f \wedge t_2 = t_2^f)$
3. $\mathfrak{a}(K\varphi) = K(\varphi^f) \leftrightarrow K(\varphi)$
4. $\mathfrak{a}(\neg\varphi) = \mathfrak{a}(\varphi)$
5. $\mathfrak{a}(\varphi \wedge \psi) = \mathfrak{a}(\varphi) \wedge \mathfrak{a}(\psi)$

We adopt the convention of writing φ^a for $\mathfrak{a}(\varphi)$.

Next, we define a crucial class of syntactic objects - the class of *topics* - and assign an object of this type to each formula in $\mathcal{L}$.

B.1.3. Definition (Topic). A topic $\mathbf{T}$ is a finite subset of $\mathbf{At}$. We refer to the class of topics by $\mathbb{T}$.

B.1.4. Definition (Topic assignment). We define the *topic assignment function* $\mathsf{T} : \mathcal{L} \to \mathbb{T}$ as follows:

1. $\mathsf{T}(Ft) = \{Ft, Ft^f, t = t^f\}$
2. $\mathsf{T}(t_1 = t_2) = \{t_1 = t_2, t_1^f = t_2^f, t_1 = t_1^f, t_2 = t_2^f\}$
3. $\mathsf{T}(K\varphi) = \{K\varphi\} \cup \{K\varphi^f\} \cup \mathsf{T}(\varphi)$
4. $\mathsf{T}(\neg\varphi) = \mathsf{T}(\varphi)$
5. $\mathsf{T}(\varphi \wedge \psi) = \mathsf{T}(\varphi) \cup \mathsf{T}(\psi)$

Observation: if two formulae φ and ψ are tautologically equivalent and are composed of the same atomic claims, then: φ^e and ψ^e are tautologically equivalent, and $\mathsf{T}_\varphi = \mathsf{T}_\psi$.

B.1.5. Definition (State Description, Resolution). Given a topic $\mathbf{T}$, a *state description* $\lambda \in \mathcal{L}$ *built from* $\mathbf{T}$ is a conjunction of literals, where every literal is either a member of $\mathbf{T}$ or its negation, and each member of $\mathbf{T}$ appears exactly once in λ. The *resolution* of $\mathbf{T}$ is the set of state descriptions built from $\mathbf{T}$. We say that φ is *at the resolution of topic* $\mathbf{T}$ just in case $\mathsf{T}_\varphi \subseteq \mathbf{T}$.

B.1.6. PROPOSITION.

1. $\mathfrak{f}(\varphi^f) = \varphi^f$

2. $\mathfrak{a}(\varphi^a) = \psi$, *where ψ is tautologically equivalent to φ^a and both are composed of the same atomic claims*

3. $\mathsf{T}_{\varphi^e} \subseteq \mathsf{T}_\varphi$

Proof:
Direct consequences of our recursive definitions, via proof by induction on the complexity of formulae. We highlight one part of the proof.

For the second result, consider the inductive step for $K\varphi$, working with the induction hypothesis that $\mathfrak{a}(\varphi) = \psi$, where ψ is tautologically equivalent to φ^a and $\mathsf{T}_\psi = \mathsf{T}_{\varphi^a}$. From our recursive clauses, it follows easily that $\mathfrak{a}\mathfrak{a}(K\varphi) = \mathfrak{a}(K\varphi^f) \wedge \mathfrak{a}(K\varphi)$. This is equal to $(K(\mathfrak{f}(\varphi^f)) \leftrightarrow K(\varphi^f)) \wedge (K\varphi^f \leftrightarrow K\varphi)$. By item 1, this is in turn equal to $(K\varphi^f \leftrightarrow K\varphi^f) \wedge (K\varphi^f \leftrightarrow K\varphi)$. This last formula is tautologically equivalent to and contains the same atoms as $K\varphi^f \leftrightarrow K\varphi$. $\square$

B.2 Frames

B.2.1. DEFINITION (Ordered RA frame). An *ordered RA frame* $\mathcal{F}$ is a tuple

$$\langle W, D, \{\preceq\}_{w\in W}, \{E_w\}_{w\in W}\rangle$$

where W is a set of worlds; D is a domain of objects; $\preceq_w$ is a similarity ordering on W (i.e. a preorder) with w as the unique minimal element; and $E_w \subseteq W$ - with $w \in E_w$ - is the *empirical information* at w.

Thus, an ordered RA frame captures an evidential position at every world, and a way of measuring similarity between worlds. For simplicity, we think of the domain D as being invariant between worlds.

As usual, a *proposition* is just a subset of W, given frame $\mathcal{F}$.

B.2.2. DEFINITION (Nearest P-worlds). Given $\mathcal{F}$, proposition P and world w, we define the set of *nearest P-worlds to w* as:

$$N_P^w := \{u \in P \mid \text{ there exists no } v \in P \text{ such that } v \preceq_w u \text{ but } u \npreceq_w v\}$$

B.2.3. Definition (Ruling out). Given $\mathcal{F}$, propositions P and Q and world w, we say that P *rules out* Q at w if $P \cap N_Q^w = \emptyset$.

That is, P rules out Q at w just in case P is inconsistent with the nearest Q-worlds.

B.3 Models

B.3.1. Definition (Role, rigidity, property). Given worlds W and domain D, a *role* is a function $\mathbf{r} : W \to D$, assigning an object to every world (intuitively, the object that plays the role in question in that world). A role is a *rigid designator* just in case it assigns the same object to every world. A *property* is a function $\mathbf{p} : W \to \mathcal{P}(D)$, assigning a set of objects to every world (intuitively, the objects that have the property in question at that world).

B.3.2. Definition (Resolution model). A *resolution model* $\mathcal{M}$ is an ordered RA frame $\mathcal{F}$ enriched with a valuation function V that assigns a role to every term and a property to every predicate letter in $\mathcal{L}$ (we write this as V_t and V_F). In particular, if $t \in \mathbf{Rd}$, then $V(t)$ is rigid.

B.4 Semantics

B.4.1. Definition (Truth). Relative to $\mathcal{M}$, we define the satisfaction relation $\vDash$ as follows:

- $w \vDash Ft$ iff $V_t(w) \in V_F(w)$
- $w \vDash (t_1 = t_2)$ iff $V_{t_1}(w) = V_{t_2}(w)$
- $w \vDash \neg\varphi$ iff $w \nvDash \varphi$
- $w \vDash \varphi \wedge \psi$ iff $w \vDash \varphi$ and $w \vDash \psi$
- $w \vDash K\varphi$ iff: if D is a defeater for φ, then D is ruled out by E_w

where we define the notion of a *defeater* as follows: proposition D is a defeater for φ just in case (i) for some λ in the resolution of T_φ, D is the set of worlds at which λ is true, and (ii) φ^e is false at every world in D.

Relative to a model $\mathcal{M}$, we define the *truth set of* φ, denoted $|\varphi|$, as:

$$|\varphi| ::= \{w \in W \mid w \vDash \varphi\}$$

B.5 Validities and invalidities

We define the notion of a validity in the usual manner.

B.5.1. Definition (Validity). A formula φ is *valid* just in case $w \vDash \varphi$ holds at every world w in every model $\mathcal{M}$. We denote this by $\vDash \varphi$.

If φ is *not* valid, we call it an *invalidity.*

B.5.2. Proposition. *If φ is at the resolution of* $\mathbf{T}$ *and λ is a state description in the resolution of* $\mathbf{T}$*, then either* $\vDash \lambda \to \varphi$ *or* $\vDash \lambda \to \neg\varphi$*.*

Proof:
By induction on the complexity of formulae, using the recursive definition of T_φ and our truth clauses. □

B.5.3. Theorem (Restricted closure). *If $\varphi^e \to \psi^e$ is valid and $\mathsf{T}_\psi \subseteq \mathsf{T}_\varphi$, then $K\varphi \to K\psi$ is valid.*

Proof:
If $\mathsf{T}_\psi \subseteq \mathsf{T}_\varphi$, then the partition generated by the resolution of T_φ is a refinement of that of T_ψ. Hence, if E_w rules out every cell in the former partition that falsifies ψ^e (i.e. for which it is true that ψ^e is false at every world in that cell), then E_w rules out every defeater for ψ. Next, if $\varphi^e \to \psi^e$ is valid, it follows that every cell in the partition associated with T_φ that falsifies ψ^e also falsifies φ^e, and is therefore a defeater for φ. Putting this together, if $w \vDash K\varphi$ holds, then every defeater for φ is ruled out by E_w, and so every defeater for ψ is also ruled out. Thus, $w \vDash K\psi$. □

B.5.4. Theorem (Some validities). *Given any φ and ψ, the following formulae are valid.*

1. $\mathfrak{f}(\varphi^a)$

2. $\mathfrak{a}(\varphi^f)$

3. $\varphi^e \rightarrow \varphi$

4. $K\varphi \rightarrow \varphi$

5. $K(\varphi \wedge \psi) \rightarrow (K\varphi \wedge K\psi)$

6. $K(\varphi \wedge (\varphi \rightarrow \psi)) \rightarrow K\psi$

7. $K\varphi \leftrightarrow K(\varphi \wedge \varphi^e)$

8. $K\varphi \rightarrow K\varphi^e$

Proof:

1. Induction on the complexity of formulae, using the definitions of acquaintance condition and Fregean guise, and proposition B.1.6.

2. Induction on the complexity of formulae, using the definitions of acquaintance condition and Fregean guise, and proposition B.1.6.

3. We prove the result by induction on the complexity of formulae, with induction hypothesis: for given φ and ψ, it is valid that (i) $\varphi^e \rightarrow \varphi$ and $(\neg\varphi)^e \rightarrow \neg\varphi$ and (ii) $\psi \rightarrow \psi$ and $(\neg\psi)^e \rightarrow \neg\psi$.

   The base cases (basic predications and identity statements) are trivial, as is the case of a negated formula.

   $\wedge$: assume that $w \vDash (\varphi\wedge\psi)^e$ holds. That is: $w \vDash \varphi^e \wedge \psi^e$ holds. Thus, by the induction hypothesis, $w \vDash \varphi \wedge \psi$ holds. Next, assume that $w \vDash (\neg\varphi \vee \neg\psi)^e$ holds. That is: $w \vDash (\varphi^a \wedge \psi^a) \wedge (\neg\varphi^f \vee \neg\psi^f)$ holds. Thus, by the induction hypothesis, $w \vDash \neg\varphi \vee \neg\psi$ holds.

   K: assume that $w \vDash (K\varphi)^e$ holds. That is: $w \vDash (K\varphi^f \leftrightarrow K\varphi) \wedge K\varphi^f$ holds. Thus, $w \vDash K\varphi$ holds. Next, assume that $w \vDash (\neg K\varphi)^e$ holds. That is: $w \vDash (K\varphi^f \leftrightarrow K\varphi) \wedge \neg K\varphi^f$ holds. Thus, $w \vDash \neg K\varphi$ holds.

4. We show that $K\varphi \to \varphi^e$ is valid (combining this with item 3 yields our desired conclusion). We proceed by contraposition. Assume that $w \vDash \neg\varphi^e$ holds, and let λ_w be the state description in the resolution of T_φ such that: $w \vDash \lambda_w$ (since the propositions expressed by the state descriptions in this resolution form a partition of W, λ_w must exist and is unique). By proposition B.5.2, it must be that $|\lambda_w| \subseteq |\neg\varphi^e|$, since φ^e is at the resolution of T_φ. Thus, the proposition D expressed by λ_w is a defeater for φ. However, E_w does not rule out D, for the nearest world to w in D is w itself. Hence, $w \vDash \neg K\varphi$. We conclude that $w \vDash K\varphi \to \varphi^e$.

5. An application of theorem B.5.3.

6. An application of theorem B.5.3.

7. An application of theorem B.5.3 (note that $\varphi \leftrightarrow (\varphi \wedge \varphi^e)$ is a validity, since $\varphi^e \to \varphi$ is valid, and that $\mathsf{T}_\varphi = \mathsf{T}_{\varphi\wedge\varphi^e}$, since $\mathsf{T}_{\varphi^e} \subseteq \mathsf{T}_\varphi$).

8. Follows from items 6 and 7 above.

$\square$

Observation: the third validity indicates that our system does not allow for contingent a priori truth. For it cannot be that the epistemic guise φ^e is a necessity while φ is not.

Observation: the eighth validity indicates that the epistemic guise φ^e does indeed provide one 'way of knowing' the proposition expressed by φ. For if ψ is another sentence that expresses the same proposition as φ, it need not follow that $K\psi$ entails $K\varphi^e$.

B.5.5. THEOREM (Some invalidities). *The following formulae are not valid.*

1. $K(Fn) \to KK(Fn)$

2. $K(Fn) \to K(Fn \vee Gm)$

3. $(K(Fn) \wedge K(Gm)) \to K(Fn \to Gm)$

Proof:

For convenience, in what follows we work with models where the epistemic guise of Fn expresses the same proposition as Fn (i.e. n and n^f are assigned the same role), and the epistemic guise of Gm expresses the same proposition as Gm.

1. Counter-example: consider a model with three worlds w, u, v such that: $w \vDash Fn$; $u \vDash Fn$; $v \vDash \neg Fn$; and $w \prec u \prec v$. Further, set $E_w = \{w, u\}$ and $E_u = \{u, v\}$. Then, $w \vDash K(Fn)$ holds, for the nearest $\neg Fn$ world is ruled out by E_w. But $w \vDash KK(Fn)$ does not hold, for u is the nearest $\neg K(Fn)$ world to w (to see that $u \Vdash K(Fn)$ does not hold, note that v is the nearest $\neg Fn$ world to u, and $v \in E_u$).

2. Counter-example: construct a model where the nearest $\neg Fn$ worlds to w are all Gm worlds, and are incompatible with E_w, but the nearest $\neg Fn \wedge \neg Gm$ worlds are compatible with E_w.

3. Counter-example: construct a model where the nearest $\neg Fn$ worlds to w are $\neg Gm$ worlds, and vice versa, and E_w is incompatible with all such worlds. However, set the nearest $Fn \wedge \neg Gm$ worlds to be compatible with E_w.

$\square$

Observation: the second result indicates that our epistemic logic does not preserve epistemic closure in full generality (cf. chapter 3). The third result indicates that it does not preserve an instance of multi-premise epistemic closure that we related, in section 1.3.4, to the phenomenon of missed clue counter-examples.

Bibliography

A. R. Anderson and N. D. Belnap. *Entailment: The Logic of Relevance and Necessity*, volume 1. Princeton University Press, 1975.

J. L. Austin. Other Minds. In Urmson and Warnock, editors, *Philosophical Papers*, pages 76–116. Oxford University Press, 1946.

J. L. Austin. Truth. In Urmson and Warnock, editors, *Philosophical Papers*, pages 117–133. Oxford University Press, 1950.

Y. Avnur, A. Brueckner, and C. T. Buford. No Closure on Skepticism. *Pacific Philosophical Quarterly*, 92:439–447, 2011.

J. Barwise and J. Etchemendy. *The Liar: An Essay on Truth and Circularity.* Oxford University Press, 1987.

J. Barwise and J. Perry. Situations and Attitudes. *Journal of Philosophy*, 78(11): 668–691, 1981.

N.D. Belnap and T. B. Steel. *The Logic of Questions and Answers.* Yale University Press, 1976.

T. Black. Solving the Problem of Easy Knowledge. *The Philosophical Quarterly*, 58(233):597–617, 2008.

M. Blome-Tillmann. *Knowledge and Presuppositions.* Oxford University Press, 2014.

L. Bonjour. In Defense of the A Priori. In Steup, Turri, and Sosa, editors, *Contemporary Debates in Epistemology*. Wiley Blackwell, second edition, 2014.

D. Bradley. A Relevant Alternatives Solution to the Bootstrapping and Self-Knowledge Problems. *Journal of Philosophy*, 111(7):379–393, 2014.

A. Brueckner. What Missed Clue Cases Show. *Analysis*, 63(4):303–305, 2003.

J. P. Burgess. *Philosophical Logic*. Princeton University Press, 2009.

F. Cariani. Ought and Resolution Semantics. *Noûs*, 47(3):534–558, 2013.

R. Carnap. *Logical Foundations of Probability*. University of Chicago Press, 1950.

D. Chalmers. Propositions and Attitude Ascriptions: A Fregean Account. *Noûs*, 45:595–639, 2011.

D. Chalmers and B. Rabern. Two-dimensional semantics and the nesting problem. *Analysis*, 74(2):210–224, 2014.

B. F. Chellas. *Modal Logic: An Introduction*. Cambridge University Press, 1980.

I. Ciardelli, J. Groenendijk, and F. Roelofsen. On the Semantics and Logic of Declaratives and Interrogatives. *Synthese*, 192(6):1689–1728, 2015.

S. Cohen. How to Be a Fallibilist. *Philosophical Perspectives*, 2:91–123, 1988.

S. Cohen. Contextualism, Skepticism and Reasons. *Philosophical Perspectives*, 13:57–89, 1999.

S. Cohen. Basic Knowledge and the Problem of Easy Knowledge. *Philosophy and Phenomenological Research*, 65(2):309–329, 2002.

A. Coliva. *Mind, Meaning and Knowledge: Themes From the Philosophy of Crispin Wright*. Oxford University Press, 2012.

C. Cross and F. Roelofsen. Questions. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2016 edition, 2016.

M. Davies and L. Humberstone. Two Notions of Necessity. *Philosophical Studies*, 38:1–30, 1980.

K. DeRose. Contextualism and Knowledge Attributions. *Philosophy and Phenomenological Research*, 52(4):913–929, 1992.

K. DeRose. Solving the Skeptical Problem. *Philosophical Review*, 104(1):1–52, 1995.

K. DeRose. *The Case for Contextualism.* Oxford University Press, 2009.

M. Devitt. There is No a Priori. In Steup, Turri, and Sosa, editors, *Contemporary Debates in Epistemology.* Wiley Blackwell, second edition, 2014.

D. Dodd and E. Zardini, editors. *Scepticism and Perceptual Justification.* Oxford University Press, 2014.

K. S. Donnellan. Reference and Definite Descriptions. *Philosophical Review*, 77: 281–304, 1966.

K. S. Donnellan. The Contingent A Priori and Rigid Designators. *Midwest Studies in Philosophy*, 2(1):12–27, 1977.

F. I. Dretske. Epistemic Operators. *Journal of Philosophy*, 67(24):1007–1023, 1970.

F. I. Dretske. Conclusive Reasons. *Australasian Journal of Philosophy*, 49(1): 1–22, 1971.

F. I. Dretske. Contrastive Statements. *Philosophical Review*, 81(4):411–437, 1972.

F. I. Dretske. The Pragmatic Dimension of Knowledge. *Philosophical Studies*, 40:363–378, 1981.

F. I. Dretske. *Knowledge and the Flow of Information.* CLSI Publications, 1999.

F. I. Dretske. The Case Against Closure. In Steup, Turri, and Sosa, editors, *Contemporary Debates in Epistemology.* Wiley Blackwell, first edition, 2005.

J. Earman. *Bayes Or Bust?: A Critical Examination of Bayesian Confirmation Theory.* MIT Press, 1992.

R. L. Epstein. *The Semantic Foundations of Logic.* Oxford University Press, 1994.

G. Evans. *Varieties of Reference*. Oxford University Press, 1982.

J. Fantl and M. McGrath. *Knowledge in an Uncertain World*. Oxford University Press, 2009.

R. Feldman. Contextualism and Skepticism. *Philosophical Perspectives*, 13:91–114, 1999.

K. Fine. Analytic Implication. *Notre Dame Journal of Philosophical Logic*, 27:169–179, 1986.

K. Fine. Truth-maker Semantics for Intuitionistic Logic. *Journal of Philosophical Logic*, 43:549–577, 2014.

K. Fine. Angellic Content. *Journal of Philosophical Logic*, 45(2):199–226, 2016.

K. Fine. Constructing the Impossible. `https://www.academia.edu/11339241/Constructing{_}the{_}Impossible`, 2017a. Accessed: 2017-04-20.

K. Fine. Survey of Truthmaker Semantics. `https://www.academia.edu/10908756/survey{_}of{_}truthmaker{_}semantics`, 2017b. Accessed: 2017-04-20.

L. Floridi. Understanding Epistemic Relevance. *Erkenntnis*, 69:69–92, 2008.

L. Floridi. *Information: A Very Short Introduction*. Oxford University Press, 2010.

G. Forbes. The problem of factives for sense theories. *Analysis*, 71(4):654–622, 2011.

P. Fritz. A logic for epistemic two-dimensional semantics. *Synthese*, 190:1753–1770, 2013.

R. Fumerton. The Challenge of Refuting Skepticism. In Steup, Turri, and Sosa, editors, *Contemporary Debates in Epistemology*. Wiley Blackwell, second edition, 2014.

A. Gibbard. *Meaning and Normativity*. Oxford University Press, 2012.

A. Goldman. Discrimination and Perceptual Knowledge. *Journal of Philosophy*, 73:771–791, 1976.

N. Goodman. About. *Mind*, 70(277):1–24, 1961.

P. Grice. Logic and Conversation. In *The Logic of Grammar*, pages 64–75. Dickenson, 1975.

C. L. Hamblin. Questions. *Australasian Journal of Philosophy*, 36(3):159–168, 1958.

C. L. Hamblin. Questions in Montague English. *Foundations of Language*, 10(1): 41–53, 1973.

P. Hawke. Questions, Topics and Restricted Closure. *Philosophical Studies*, 73 (10):2759–2784, 2016a.

P. Hawke. Relevance in Epistemology and Logic. In Redmond, Pombo Martins, and Nepomuceno Fernàndez, editors, *Epistemology, Knowledge and the Impact of Interaction*, volume 38 of *Logic, Epistemology, and the Unity of Science*, pages 207–235. Springer, 2016b.

P. Hawke. Theories of Aboutness. *Australasian Journal of Philosophy*, 2017. Forthcoming.

J. Hawthorne. *Knowledge and Lotteries*. Oxford University Press, 2004.

J. Hawthorne. The Case for Closure. In Steup, Turri, and Sosa, editors, *Contemporary Debates in Epistemology*. Wiley Blackwell, first edition, 2005.

J. Hawthorne and D. Manley. *The Reference Book*. Oxford University Press, 2012.

M. Heller. Relevant Alternatives. *Philosophical Studies*, 55:23–40, 1989.

M. Heller. Relevant Alternatives and Clousre. *Australasian Journal of Philosophy*, 77(2):196–208, 1999.

J. Hintikka. *Inquiry as Inquiry: A Logic of Scientific Discovery*. Springer, 1999.

J. Hintikka. *Socratic Epistemology: Explorations of Knowledge-Seeking by Questioning.* Cambridge University Press, 2007.

W. H. Holliday. Epistemic Logic, Relevant Alternatives, and the Dynamics of Context. In Lassiter and Slavkovik, editors, *New Directions in Logic, Language and Computation*, Lecture Notes in Computer Science, pages 109–129. Springer-Verlag, 2012.

W. H. Holliday. *Knowing What Follows: Epistemic Closure and Epistemic Logic.* PhD thesis, Stanford University, 2013.

W. H. Holliday. Epistemic Closure and Epistemic Logic I: Relevant Alternatives and Subjunctivism. *Journal of Philosophical Logic*, 44(1):1–62, 2015a.

W. H. Holliday. Fallibilism and Multiple Paths to Knowledge. In T. Gendler and J. Hawthorne, editors, *Oxford Studies in Epistemology*, volume 5, pages 97–144. Oxford University Press, 2015b.

F. Jackson. *From Metaphysics to Ethics.* Oxford University Press, 1998.

P. Jacob. Intentionality. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, winter 2014 edition, 2014.

M. Jago. *The Impossible: An Essay on Hyperintensionality.* Oxford University Press, 2014.

D. Kaplan. Demonstratives. In J. Almog, J. Perry, and H. Wettstein, editors, *Themes From Kaplan*, pages 481–614. Oxford University Press, 1989.

J. M. Keynes. *A Treatise on Probability.* Macmillan, 1921.

A. Kratzer. *Modals and Conditionals.* Oxford University Press, 2012.

S. Kripke. *Naming and Necessity.* Harvard University Press, 1980.

S. Kripke. Nozick on Knowledge. In *Philosophical Troubles: Collected Papers, Volume 1.* Oxford University Press, 2011.

T.S. Kuhn. *The Structure of Scientific Revolutions.* The University of Chicago Press, second edition, 1970.

J. Kvanvig. *The Knowability Paradox.* Oxford University Press, 2006.

T. Lando. Conclusive Reasons and Epistemic Luck. *Australasian Journal of Philosophy*, 94(2):378–395, 2016.

K. Lawlor. Living Without Closure. In M. Blaauw, editor, *Epistemological Contextualism*, volume 69 of *Grazer Philosophische Studien*, pages 25–49. Rodopi, 2005.

K. Lawlor. *Assurance: An Austinian View of Knowledge Claims.* Oxford University Press, 2013.

A. Leite. Austin, Dreams and Skepticism. In Gustafsson and Sorli, editors, *The Philosophy of J. L. Austin*, pages 78–113. Oxford University Press, 2012.

D. Lewis. Attitudes De Dicto and De Se. *Philosophical Review*, 88(4):513–543, 1979.

D. Lewis. Statements Partly About Observation. In *Papers in Philosophical Logic*, pages 125–155. Cambridge University Press, 1988a.

D. Lewis. Relevant Implication. *Theoria*, 54(3):161–174, 1988b.

D. Lewis. Elusive Knowledge. *Australasian Journal of Philosophy*, 74(4):549–567, 1996.

S. Luper. The Epistemic Predicament: Knowledge, Nozickian Track, and Skepticism. *The Australasian Journal of Philosophy*, 62:26–50, 1984.

S. Luper. Epistemic closure. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, spring 2016 edition, 2016.

J. MacFarlane. The Assessment Sensitivity of Knowledge Attributions. In Gendler and Hawthorne, editors, *Oxford Studies in Epistemology*, pages 197–234. Oxford University Press, 2005.

D. Makinson. The Paradox of the Preface. *Analysis*, 25(6):205–207, 1965.

Edwin Mares. Relevance logic. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2014 edition, 2014.

D. Nolan. Impossible Worlds: A Modest Approach. *Notre Dame Journal of Formal Logic*, 38(4), 1997.

R. Nozick. *Philosophical Explanations*. Harvard University Press, 1981.

W. T. Parry. The Logic of C. I. Lewis. In Schlipp, editor, *The Philosophy of C. I. Lewis*, pages 115–154. Cambridge University Press, 1968.

W. T. Parry. Analytic Implication: its History, Justification and Varieties. In Norman and Sylvan, editors, *Directions in Relevant Logic*, pages 101–118. Kluwer Academic Publishers, 1989.

J. Perry. The Problem of the Essential Indexical. *Noûs*, 13:3–21, 1979.

J. Perry. Possible Worlds and Subject Matter. In *The Problem of the Essential Indexical and Other Essays*, pages 145–160. CSLI publications, first edition, 1989.

J. Perry. *The Problem of the Essential Indexical and Other Essays*. CSLI publications, expanded edition, 2000.

G. Preyer and G. Peter, editors. *Contextualism in Philosophy: Knowledge, Meaning and Truth*. Oxford University Press, 2005.

D. Pritchard. Relevant Alternatives, Perceptual Knowledge and Discrimination. *Noûs*, 44(2):245–268, 2010.

D. Pritchard. *Epistemological Disjunctivism*. Oxford University Press, 2012.

H. Putnam. Formalization of the Concept "About". *Philosophy of Science*, 25 (2):15–130, 1958.

H. Putnam. Meaning and Reference. *Journal of Philosophy*, 70:699–711, 1973.

S. Read. *Relevant Logic: A Philosophical Examination of Inference*. Basil Blackwell, 1988.

M. Reimer and E. Michaelson. Reference. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University, spring 2017 edition, 2017.

C. Roberts. Topics. In Maienborn Heusinger and Portner, editors, *Semantics: An International Handbook of Natural Language Meaning*, volume 2, pages 1908–1934. De Gruyter Mouton, 2011.

C. Roberts. Information Structure: Towards an Integrated Formal Theory of Pragmatics. *Semantics and Pragmatics*, 5:1–69, 2012.

S. Roush. Closure on Skepticism. *Journal of Philosophy*, 107(5):243–256, 2010.

B. Russell. Knowledge by Acquaintance and Knowledge by Description. *Proceedings of the Aristotelian Society*, 11:108–128, 1910.

G. Ryle. About. *Analysis*, 1(1):10–12, 1933.

P. Rysiew. Motivating the Relevant Alternatives Approach. *Canadian Journal of Philosophy*, 36(2):259–279, 2006.

J. Schaffer. Knowledge, Relevant Alternatives and Missed Clues. *Analysis*, 61(3): 202–208, 2002.

J. Schaffer. From Contextualism to Contrastivism. *Philosophical Studies*, 119(1): 73–103, 2004.

J. Schaffer. Contrastive Knowledge. In Gendler and Hawthorne, editors, *Oxford Studies in Epistemology 1*, pages 235–271. Oxford University Press, 2005a.

J. Schaffer. What Shifts? Thresholds, Standards or Alternatives? In Preyer and Peter, editors, *Contextualism in Philosophy: Knowledge, Meaning and Truth*, pages 115–130. Oxford University Press, 2005b.

J. Schaffer. Closure, Contrast and Answer. *Philosophical Studies*, 133(2):233–255, 2007a.

J. Schaffer. Knowing the Answer. *Philosophy and Phenomenological Research*, 75(2):383–403, 2007b.

S. Soames. *Reference and Description: The Case Against Two-Dimensionalism.* Princeton University Press, 2007a.

S.. Soames. Actually. *Aristotelian Society Supplementary Volume*, 81(1):251–277, 2007b.

E. Sosa. How to Defeat Opposition to Moore. *Philosophical Perspectives*, 13: 141–153, 1999.

E. Sosa. Relevant Alternatives, Contexualism Included. *Philosophical Studies*, 119:35–65, 2004.

R. Stalnaker. Pragmatics. In Davidson and Harman, editors, *Semantics of Natural Language*, pages 380–397. Reidel, 1972.

R. Stalnaker. Assertion. In P. Cole, editor, *Syntax and Semantics 9: Pragmatics*, pages 315–332. New York Academic Press, 1978.

R. Stalnaker. *Inquiry.* MIT Press, 1987.

R. Stalnaker. *Our Knowledge of the Internal World.* Oxford University Press, 2008.

K. Stanford. Underdetermination of scientific theory. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy.* Metaphysics Research Lab, Stanford University, spring 2016 edition, 2016.

J. Stanley. *Knowledge and Practical Interests.* Oxford University Press, 2005.

J. Stanley. *Know How.* Oxford University Press, 2011.

G. Stine. Skepticism, Relevant Alternatives and Deductive Closure. *Philosophical Studies*, 29:249–261, 1976.

P. Unger. *Ignorance: A Case for Skepticism.* Oxford University Press, 1975.

J. F. A. K. van Benthem. What One May Come To Know. *Analysis*, 64:95–105, 2004.

J. F. A. K. van Benthem. *Logical Dynamics of Information and Interaction.* Cambridge University Press, 2011.

B. van Fraassen. Facts and Tautological Entailment. *Journal of Philosophy*, 66 (15):477–487, 1969.

J. Vogel. Are There Counterexamples to the Closure Principle? In Roth and Ross, editors, *Doubting*, pages 13–27. Kluwer Academic Publishers, 1990.

J. Vogel. The New Relevant Alternatives Theory. *Philosophical Perspectives*, 13: 155–180, 1999.

J. Vogel. The Refutation of Skepticism. In Steup, Turri, and Sosa, editors, *Contemporary Debates in Epistemology*. Wiley Blackwell, second edition, 2014.

T. Williamson. *Knowledge and its Limits*. Oxford University Press, 2000.

T. Williamson. How Deep is the Distinction Between A Priori and A Posteriori Knowledge? In A. Cusollo and J. C. Thurow, editors, *The A Priori in Philosophy*, pages 291–312. Oxford University Press, 2013.

C. Wright. Some Reflections on the Acquisition of Warrant by Inference. In Nuccetelli, editor, *New Essays on Semantic Externalism and Self-Knowledge*, pages 57–78. MIT Press, 2003.

C. Wright. On Epistemic Entitlement: Warrant for Nothing (and Foundations for Free). *The Aristotelian Society Supplementary Volume*, 78(1):167–212, 2004.

C. Wright. On Epistemic Entitlement (II): Welfare State Epistemology. In Dodd and Zardini, editors, *Scepticism and Perceptual Justification*, pages 213–247. Oxford University Press, 2014.

S. Yablo. Knowing About Things. `http://www.mit.edu/~yablo/home/Papers_files/Knowing%20About%20Things.pdf`, 2012. Accessed: 2017-04-20.

S. Yablo. *Aboutness*. Princeton University Press, 2014.

S. Yalcin. Nonfactualism about Epistemic Modality. In Egan and Weatherson, editors, *Epistemic Modality*, pages 295–332. Oxford University Press, 2011.

S. Yalcin. *Actually*, Actually. *Analysis*, 75(2):185–191, 2015.

S. Yalcin. Belief as Question-Sensitive. *Philososophy and Phenomenological Research*, 2016. Forthcoming.

L. Zagzebski. The Inescapability of Gettier Problems. *The Philosophical Quartely*, 44(174):65–73, 1994.

Titles in the ILLC Dissertation Series:

ILLC DS-2009-01: **Jakub Szymanik**
Quantifiers in TIME and SPACE. Computational Complexity of Generalized Quantifiers in Natural Language

ILLC DS-2009-02: **Hartmut Fitz**
Neural Syntax

ILLC DS-2009-03: **Brian Thomas Semmes**
A Game for the Borel Functions

ILLC DS-2009-04: **Sara L. Uckelman**
Modalities in Medieval Logic

ILLC DS-2009-05: **Andreas Witzel**
Knowledge and Games: Theory and Implementation

ILLC DS-2009-06: **Chantal Bax**
Subjectivity after Wittgenstein. Wittgenstein's embodied and embedded subject and the debate about the death of man.

ILLC DS-2009-07: **Kata Balogh**
Theme with Variations. A Context-based Analysis of Focus

ILLC DS-2009-08: **Tomohiro Hoshi**
Epistemic Dynamics and Protocol Information

ILLC DS-2009-09: **Olivia Ladinig**
Temporal expectations and their violations

ILLC DS-2009-10: **Tikitu de Jager**
"Now that you mention it, I wonder...": Awareness, Attention, Assumption

ILLC DS-2009-11: **Michael Franke**
Signal to Act: Game Theory in Pragmatics

ILLC DS-2009-12: **Joel Uckelman**
More Than the Sum of Its Parts: Compact Preference Representation Over Combinatorial Domains

ILLC DS-2009-13: **Stefan Bold**
Cardinals as Ultrapowers. A Canonical Measure Analysis under the Axiom of Determinacy.

ILLC DS-2010-01: **Reut Tsarfaty**
Relational-Realizational Parsing

ILLC DS-2010-02: **Jonathan Zvesper**
Playing with Information

ILLC DS-2010-03: **Cédric Dégremont**
The Temporal Mind. Observations on the logic of belief change in interactive systems

ILLC DS-2010-04: **Daisuke Ikegami**
Games in Set Theory and Logic

ILLC DS-2010-05: **Jarmo Kontinen**
Coherence and Complexity in Fragments of Dependence Logic

ILLC DS-2010-06: **Yanjing Wang**
Epistemic Modelling and Protocol Dynamics

ILLC DS-2010-07: **Marc Staudacher**
Use theories of meaning between conventions and social norms

ILLC DS-2010-08: **Amélie Gheerbrant**
Fixed-Point Logics on Trees

ILLC DS-2010-09: **Gaëlle Fontaine**
Modal Fixpoint Logic: Some Model Theoretic Questions

ILLC DS-2010-10: **Jacob Vosmaer**
Logic, Algebra and Topology. Investigations into canonical extensions, duality theory and point-free topology.

ILLC DS-2010-11: **Nina Gierasimczuk**
Knowing One's Limits. Logical Analysis of Inductive Inference

ILLC DS-2010-12: **Martin Mose Bentzen**
Stit, Iit, and Deontic Logic for Action Types

ILLC DS-2011-01: **Wouter M. Koolen**
Combining Strategies Efficiently: High-Quality Decisions from Conflicting Advice

ILLC DS-2011-02: **Fernando Raymundo Velazquez-Quesada**
Small steps in dynamics of information

ILLC DS-2011-03: **Marijn Koolen**
The Meaning of Structure: the Value of Link Evidence for Information Retrieval

ILLC DS-2011-04: **Junte Zhang**
System Evaluation of Archival Description and Access

ILLC DS-2011-05: **Lauri Keskinen**
Characterizing All Models in Infinite Cardinalities

ILLC DS-2011-06: **Rianne Kaptein**
Effective Focused Retrieval by Exploiting Query Context and Document Structure

ILLC DS-2011-07: **Jop Briët**
Grothendieck Inequalities, Nonlocal Games and Optimization

ILLC DS-2011-08: **Stefan Minica**
Dynamic Logic of Questions

ILLC DS-2011-09: **Raul Andres Leal**
Modalities Through the Looking Glass: A study on coalgebraic modal logic and their applications

ILLC DS-2011-10: **Lena Kurzen**
Complexity in Interaction

ILLC DS-2011-11: **Gideon Borensztajn**
The neural basis of structure in language

ILLC DS-2012-01: **Federico Sangati**
Decomposing and Regenerating Syntactic Trees

ILLC DS-2012-02: **Markos Mylonakis**
Learning the Latent Structure of Translation

ILLC DS-2012-03: **Edgar José Andrade Lotero**
Models of Language: Towards a practice-based account of information in natural language

ILLC DS-2012-04: **Yurii Khomskii**
Regularity Properties and Definability in the Real Number Continuum: idealized forcing, polarized partitions, Hausdorff gaps and mad families in the projective hierarchy.

ILLC DS-2012-05: **David García Soriano**
Query-Efficient Computation in Property Testing and Learning Theory

ILLC DS-2012-06: **Dimitris Gakis**
Contextual Metaphilosophy - The Case of Wittgenstein

ILLC DS-2012-07: **Pietro Galliani**
The Dynamics of Imperfect Information

ILLC DS-2012-08: **Umberto Grandi**
Binary Aggregation with Integrity Constraints

ILLC DS-2012-09: **Wesley Halcrow Holliday**
Knowing What Follows: Epistemic Closure and Epistemic Logic

ILLC DS-2012-10: **Jeremy Meyers**
Locations, Bodies, and Sets: A model theoretic investigation into nominalistic mereologies

ILLC DS-2012-11: **Floor Sietsma**
Logics of Communication and Knowledge

ILLC DS-2012-12: **Joris Dormans**
Engineering emergence: applied theory for game design

ILLC DS-2013-01: **Simon Pauw**
Size Matters: Grounding Quantifiers in Spatial Perception

ILLC DS-2013-02: **Virginie Fiutek**
Playing with Knowledge and Belief

ILLC DS-2013-03: **Giannicola Scarpa**
Quantum entanglement in non-local games, graph parameters and zero-error information theory

ILLC DS-2014-01: **Machiel Keestra**
Sculpting the Space of Actions. Explaining Human Action by Integrating Intentions and Mechanisms

ILLC DS-2014-02: **Thomas Icard**
The Algorithmic Mind: A Study of Inference in Action

ILLC DS-2014-03: **Harald A. Bastiaanse**
Very, Many, Small, Penguins

ILLC DS-2014-04: **Ben Rodenhäuser**
A Matter of Trust: Dynamic Attitudes in Epistemic Logic

ILLC DS-2015-01: **María Inés Crespo**
Affecting Meaning. Subjectivity and evaluativity in gradable adjectives.

ILLC DS-2015-02: **Mathias Winther Madsen**
The Kid, the Clerk, and the Gambler - Critical Studies in Statistics and Cognitive Science

ILLC DS-2015-03: **Shengyang Zhong**
Orthogonality and Quantum Geometry: Towards a Relational Reconstruction of Quantum Theory

ILLC DS-2015-04: **Sumit Sourabh**
Correspondence and Canonicity in Non-Classical Logic

ILLC DS-2015-05: **Facundo Carreiro**
Fragments of Fixpoint Logics: Automata and Expressiveness

ILLC DS-2016-01: **Ivano A. Ciardelli**
Questions in Logic

ILLC DS-2016-02: **Zoé Christoff**
Dynamic Logics of Networks: Information Flow and the Spread of Opinion

ILLC DS-2016-03: **Fleur Leonie Bouwer**
What do we need to hear a beat? The influence of attention, musical abilities, and accents on the perception of metrical rhythm

ILLC DS-2016-04: **Johannes Marti**
Interpreting Linguistic Behavior with Possible World Models

ILLC DS-2016-05: **Phong Lê**
Learning Vector Representations for Sentences - The Recursive Deep Learning Approach

ILLC DS-2016-06: **Gideon Maillette de Buy Wenniger**
Aligning the Foundations of Hierarchical Statistical Machine Translation

ILLC DS-2016-07: **Andreas van Cranenburgh**
Rich Statistical Parsing and Literary Language

ILLC DS-2016-08: **Florian Speelman**
Position-based Quantum Cryptography and Catalytic Computation

ILLC DS-2016-09: **Teresa Piovesan**
Quantum entanglement: insights via graph parameters and conic optimization

ILLC DS-2016-10: **Paula Henk**
Nonstandard Provability for Peano Arithmetic. A Modal Perspective

ILLC DS-2017-01: **Paolo Galeazzi**
Play Without Regret

ILLC DS-2017-02: **Riccardo Pinosio**
The Logic of Kant's Temporal Continuum

ILLC DS-2017-03: **Matthijs Westera**
Exhaustivity and intonation: a unified theory

ILLC DS-2017-04: **Giovanni Cinà**
Categories for the working modal logician

ILLC DS-2017-05: **Shane Noah Steinert-Threlkeld**
Communication and Computation: New Questions About Compositionality

ILLC DS-2017-06: **Peter Hawke**
The Problem of Epistemic Relevance

ILLC DS-2017-07: **Aybüke Özgün**
Evidence in Epistemic Logic: A Topological Perspective