Episteme

http://journals.cambridge.org/EPI

Additional services for *Episteme:*

Email alerts: Click here
Subscriptions: Click here
Commercial reprints: Click here
Terms of use : Click here

KNOWING OUR DEGREES OF BELIEF

Sinan Dogramaci

Episteme / *FirstView* Article / October 2015, pp 1 - 19
DOI: 10.1017/epi.2015.48, Published online: 22 October 2015

Link to this article: http://journals.cambridge.org/abstract_S1742360015000489

How to cite this article:
Sinan Dogramaci KNOWING OUR DEGREES OF BELIEF. Episteme, Available on CJO 2015 doi:10.1017/epi.2015.48

Request Permissions : Click here

CAMBRIDGE JOURNALS

Downloaded from http://journals.cambridge.org/EPI, IP address: 128.83.63.20 on 23 Oct 2015

Episteme (2015) page 1 of 19 © Cambridge University Press
doi:10.1017/epi.2015.48

KNOWING OUR DEGREES OF BELIEF

SINAN DOGRAMACI
sinan.dogramaci@gmail.com

ABSTRACT

The main question of this paper is: how do we manage to know what our own degrees of belief are? Section 1 briefly reviews and criticizes the traditional functionalist view, a view notably associated with David Lewis and sometimes called the theory-theory. I use this criticism to motivate the approach I want to promote. Section 2, the bulk of the paper, examines and begins to develop the view that we have a special kind of introspective access to our degrees of belief. I give an initial assessment of the view by examining its compatibility with leading theories of introspection. And I identify a challenge for the view, and explain why I'm optimistic that the view can overcome it.

How do you or I know anything about what any of our own degrees of beliefs are? My question is about the first-personal case: how does a person know what their *own* degrees of beliefs are?

Of course, no one can honestly report exact numbers they really think are their actual degrees of belief, at least for most ordinary matters outside of a casino. But I'm not at a complete loss for what my degrees of belief are. I can tell you things like: this gets a fairly low degree of belief, that gets a very high degree of belief, and whether or not this other claim is true is roughly fifty-fifty. We're able to say something informative about our degrees of belief in various things.

Degree of belief, as it concerns me here, is not something that was first discovered by scientists or a technical notion first introduced by definition. It is supposed to be an ordinary (or "folk") notion that ordinary people have at least some vague or implicit understanding of (even if that understanding is later improved and refined by scientific or philosophical investigation). Ordinary language is varied and messy and there is no single word for this notion, but common terms include: degree of belief, partial belief, (degree of) confidence, strength of opinion (or belief), credence, subjective or personal probability, and subjective chance. Degree of desire is an ordinary notion too, and has various names too, such as: utility, partial desire, strength of desire and (subjective) value.

Now, the traditional view among Bayesian formal epistemologists has long been that, in fact, there is nothing special about the first-personal case: our access to our own degrees of belief is the same kind of access we have to other people's degrees of belief. Tradition has it that degree of belief is a functional property that we have no special kind of introspective access to.

§ 1 briefly reviews this traditional view, and the answer it gives to my main question. This is a functionalist view, notably associated with David Lewis and sometimes called the *theory-theory*, which aims to provide both the metaphysics of our partial attitudes together with an accompanying epistemology for them. While I won't refute this

approach, I'll say enough negative things to hopefully motivate giving some attention to the approach I want to promote.

§ 2, the bulk of the paper, examines and begins to develop the view that is unpopular but also – it seems to me – more commonsensical. This is the view that we have a special kind of introspective access to our degrees of belief. I give an initial assessment of the view by examining its compatibility with leading theories of introspection. And I identify a big challenge for the view, and explain why I'm optimistic that the view can overcome it.

1. THE THEORY-THEORY

The most well developed answer to my main question is a certain type of functionalism. It's the sort of functionalism that addresses not only the metaphysics of our attitudes, but the epistemology as well. The view says (i) a person's causal relations to her surroundings – relations the view characterizes in non-mental terms – determine, i.e. necessitate, the facts about what attitudes she has, and (ii) it is possible to know a person's attitudes by an inference premised on knowledge of just those same non-mentally characterized causal relations. Though answering my epistemological main question doesn't require providing a *non-mental* basis for our knowledge, this functionalist approach does offer such a basis.

David Lewis's view is the paradigm of this sort of functionalism, a view addressed to both the metaphysics and epistemology specifically of the partial attitudes.[1] Lewis's view is representative of so-called analytic or commonsense functionalism, a variety of functionalism that says the metaphysics of the mind is a priori. Lewis's view is also called the *theory-theory*, because Lewis proposes an epistemological *theory* according to which the folk implicitly know and rely on a metaphysical *theory* in order to attribute attitudes to themselves and to others.[2]

If the epistemological theory of how we know our minds is to be found in a metaphysical theory, this is the right variety of functionalism to consider. It promises to explain our epistemic access to our minds on the basis of our presumably less mysterious epistemic access to the causal relations we bear to our environment. (Some of these causal relations are dispositional, so this approach will need to explain how we have knowledge of any relevant counterfactuals, where that knowledge is not based on prior knowledge of mental states.)

An alternative variety of commonsense functionalism says that differences in the hidden, internal causal structure of a person's mind can make a difference to what she believes.[3] However plausible this might be as a metaphysical view, it will not yield a plausible answer to our epistemological question here, since ordinary people do not have access to (non-mental characterizations of) the actual inner mechanics of their brains, yet we all somehow know what we are thinking. Fans of a metaphysics on which internal structure is important should favor an epistemology like the one I will offer in the positive part of the paper, § 2.

A rival to commonsense functionalism is so-called empirical functionalism, also called psychofunctionalism. Such views claim that the metaphysics of the mind is not a priori;

1 See especially Lewis (1974; 1994/1999). Schwarz (2014) is also a valuable exposition.

2 See, e.g., Bermúdez (2005: 185, 192, 343), Goldman (2006: 8), Ramsey (2013: §2).

3 See, e.g., Braddon-Mitchell and Jackson (1996/2007: 114–22).

that is, an armchair philosopher, given only folk knowledge, cannot discover what non-mental conditions determine what attitudes a person has.[4] Fans of this view should also favor the proposal I'll give in § 2.

It is not the aim of this paper to thoroughly criticize alternatives to the positive proposal I want to offer. But I will briefly report the problems, developed in other literature, that lead me to believe the theory-theory cannot explain the epistemology of degrees of belief, and that motivate my search for a new proposal.

The standard way of putting flesh on this functionalist approach is to use a so-called *representation theorem* to characterize how partial attitudes, both degrees of belief and utilities together, are determined. Such a theorem aims to derive a unique set of degrees of belief and (unique up to linear transformation) utilities, given only certain facts about a person's (actual and counterfactual) behavior. But the use of representation theorems to characterize partial attitudes has by now been devastatingly criticized. Powerful objections can be found in Hampton (1994), Hampton (1998: chs 7–8), Howson and Urbach (2005: 57) and Eriksson and Hájek (2007: see especially the problem raised on pp. 196–7, for which they credit Joyce). Meacham and Weisberg (2011) provide a general critique, drawing on decades of discrediting research and adding original objections. (I'll describe one of the major objections these authors all raised in just a moment.)

Can a functionalist metaphysics of degrees of belief be developed without relying on a representation theorem? Lewis himself tried to do this, and he is the only theorist I know of who attempted this. Schwarz (2014: 21) reports Lewis's position like this: "So-called representation theorems in decision theory seem to show that if an agent's choice dispositions satisfy certain qualitative constraints, then there is a unique system of beliefs and desires that matches her dispositions [Savage (1954/1972)]. Lewis did not trust these results and argued that, on the contrary, our choice dispositions leave some aspects of our attitudes radically underdetermined [Lewis (1983: 50–2)]." Hájek and Smithson (2012: §3) agree that if Lewis's view is right, then our degrees of belief are not determinate.

What was Lewis's view? Lewis claimed that two principles together determine a person's partial attitudes: *Rationalization* says a person has expected utility maximizing attitudes given her behavior (her actual and some counterfactual behavior); and *Charity* says a person has rational attitudes given her total evidence (or rather, given the total stimulation of her senses by her environment, all non-mentally characterized). Lewis said there may be indeterminacy in a person's attitudes *when the principles conflict*. However, Lewis thought that if the principles *don't* conflict, they should assign determinate attitudes. He admitted that if multiple assignments of attitudes are perfectly compatible with both principles, then his theory falls short (Lewis 1974: 343).

Here's why Lewis's view is no more plausible than views that rely on representation theorems. The representation theorems only apply to a person whose behavior conforms to many substantial constraints, and, even then, Hampton (1998: 238–9), Eriksson and Hájek (2007: 196–7) and Meacham and Weisberg (2011: 658) show that there fails to be a unique set of expected utility maximizing attitudes that can be assigned to her. What the theorems show is only that, for such a person, there is a unique set of degrees of belief *that obeys the probability axioms*, and a unique-up-to-linear-transformation

4 See, e.g., Braddon-Mitchell and Jackson (1996/2007: 85).

set of utilities, that together make the person an expected utility maximizer. These critics complain that we have been given no reason to think these people have degrees of belief that obey the axioms, so they conclude that representation theorems don't reveal our degrees of belief. But if the representation theorems fall short here, then Lewis's *Rationalization* principle must fare even worse. For, Lewis doesn't impose all those substantial constraints on a person's behavior; *Rationalization* merely instructs us to assign to any person expected utility maximizing attitudes. So, *Rationalization* clearly won't get us farther than the representation theorems, which fail even after making lots of substantial extra assumptions about the person's behavior.

The representation theorems can assign a unique set of degrees of belief if we *assume* our person's beliefs obey the probability axioms. But the critics of representation theorems express doubts: why should we think that a person who *can* be assigned degrees of belief that obey the axioms of probability really *does* have those degrees of belief? (Eriksson and Hájek 2007: 196–7; Hájek 2008: §3; Meacham and Weisberg 2011: §3) Lewis would reply that *Charity* instructs us to assign rational degrees of belief that obey the axioms. But saying only that much is unlikely to ease the doubters' worries. Why think that, just because someone *might* be living up to the ideal of obeying the axioms of probability, then she really *must* be? Anyone with serious doubts about the use of representation theorems is unlikely to find Lewis's approach to be any better.

Finally, functionalists cannot answer these criticisms just by switching to a model on which people's degrees of belief are mushy rather than sharp. Mushy degrees of belief are usually understood as a set of values for each proposition, rather than a single (sharp) value. The critics think a person could have some degree of belief d though Lewis and the representation theorems cannot pin this on her. Suppose we say this person has not a single (sharp) degree of belief d in the proposition in question, but a (mushy) set. If d is not in the set, then the introduction of sets doesn't offer any new help. But if d is to go in the set, Lewis and the representation theorems still have no resources for explaining how it belongs there – the credence was one their principles couldn't pin on the person. So, Lewis and the representation theorems can't answer the critics just by saying our credences are mushy.

2. INTROSPECTION AS FOUNDATION FOR KNOWLEDGE OF OUR DEGREES OF BELIEF

Lewis, as emphasized, aims to provide both the metaphysics and epistemology of our attitudes at once. There is a functionalist theory, and not only does this theory play the metaphysical role of specifying what determines that we have those attitudes, but implicit knowledge of the theory serves the epistemological role of allowing us to infer a person's attitudes from some non-mental facts.

I sympathize with the contrary view that the metaphysics of our partial attitudes is *not* implicitly known; the non-mental facts that determine what attitudes a person has – whatever they may be – are *not* discoverable a priori, *not* discoverable from the armchair; it is aposteriori what these facts are, if it is knowable at all. We can call any such view a *no-theory theory*. For all this position says, one or another functionalist theory has the metaphysics right; this position only denies that the epistemology is as the theory-theory describes.

If I think a no-theory theory is right, though, then how can I explain how we know our degrees of belief? Somehow we manage to know a good deal about our degrees of belief. An answer to the epistemological question need not identify *non-mental* facts from which we infer any person's attitudes. But some account is needed, and all the more urgently for those who reject Lewis's view that we can make an a priori inference to attitudes from the non-mental facts that determine them.

An alternative view is that we have *introspective* access to our degrees of belief. I want to use the rest of this paper to sympathetically consider the prospects for the introspection-based approach, in one or another form.

In doing this, I am motivated by my doubts about the theory-theory, but let me note that the theory-theory is technically consistent with our having introspective access to our degrees of belief. It could be a priori which non-mental facts determine our degrees of belief, and thus the form of access afforded by the theory-theory is available to us (at least in principle), while it is also true that our degrees of belief can be, and presumably more normally would be, introspected. So, even someone certain of the theory-theory should hear out the case for the introspection-based approach.

I am also motivated by the unjustly bad reputation that introspection carries among theorists who study degrees of belief. It got its bad reputation right from the start, with one of modern Bayesianism's founding figures, Frank Ramsey. Here is what he had to say about introspection:

> We can, in the first place, suppose that the degree of a belief is something perceptible by its owner; for instance that beliefs differ in the intensity of a feeling by which they are accompanied, which might be called a belief-feeling or feeling of conviction, and that by the degree of belief we mean the intensity of this feeling. This view would be very inconvenient, for it is not easy to ascribe numbers to the intensities of feelings; but apart from this it seems to me observably false, for the beliefs which we hold most strongly are often accompanied by practically no feeling at all; no one feels strongly about things he takes for granted. (Ramsey 1926/1931: 169)

(Ramsey then went on to endorse an early version of a functionalist view of degree of belief, which he elaborates in the rest of the paper, even providing the first representation theorem.)

Ramsey doubts that we can introspect our degrees of belief because some of our very strongest beliefs are associated with no strong feelings at all.[5] Well, I completely agree. I concede to Ramsey that knowledge of my own degrees of belief with their various strengths is not based on detection of a measurable quantity of some kind of cognitive phenomenology. But I want to still deny that I cannot introspect my degrees of belief. The introspection view need not be developed in a way that has us introspecting special feelings in order to know our degrees of belief.

How, then, should the introspection view be developed? I propose to make a start by briefly considering four currently popular theories of introspection of full belief, and see what promise each has for being extended to cover partial belief (§ 2.1). Then I will identify what I think is a big challenge to the introspection based approach, and I will begin to address it (§ 2.2). I'll end with some brief comments on how the picture I'll recommend might be extended to explain our knowledge of other people's degrees of belief (§ 2.3).

5 See Joyce (2005: 176) for an example of contemporary Bayesians endorsing Ramsey on this point.

2.1 *Theories of introspection: seeking special access to degrees of belief*

If there is introspection of degrees of belief in any non-trivial sense of 'introspection', then there is a kind of epistemic access we have to our own degrees of belief that is different from the kinds of access we have to others' attitudes. In other words, anything worthy of the distinctive label 'introspection' plausibly must involve, as it's called, *special access*.[6] So, our question is whether there is special access to partial attitudes, in particular to degrees of belief.

Many contemporary theories of introspection credit us with special access to our own minds; indeed, it is standardly taken as a primary constraint guiding the construction of a plausible theory.[7] However, in the enormous literature on introspection, it is almost impossible to find any discussion at all of introspection of partial attitudes. So, let's start looking at this for ourselves, and see what initial assessments we can make. Let's briefly bring up the leading contemporary theories of introspection to see if they immediately foreclose on, or leave room open for, special introspective access to degrees of belief.

Consider, first, **the self-scanning theory** of introspection.[8] This theory makes special access a contingent truth about us. The theory says the human brain just happens to be equipped with mechanisms for reliably detecting its own mental states; we simply form these introspective beliefs non-inferentially, as a reliable causal effect of our being in those mental states, similar to how we happen to be built to non-inferentially and reliably perceptually detect features of our environment. (The introspective belief might be a full belief, or perhaps just a high degree of belief. A similar choice for what introspection delivers, full belief or partial belief, needs to be made for each theory of introspection we'll consider.)

This theory seems to be perfectly compatible with the variety of special access we're seeking. On this theory, we simply might or might not be built with a mechanism for detecting degrees of belief too. (Or we might be built to at least detect orderings among our degrees of belief, and perhaps our utilities too.) It would just depend on whether we're built with the deluxe version of the self-scanner. (The theory does not say a mind *must have* special access, as some philosophers think a plausible theory of introspection must say, but this was not our particular concern.)

Next, consider the **constitutivist theory** of introspection.[9] This view says that to (fully) believe that p is to be in a state partly constituted by a disposition to have an introspective belief that you believe that p. It is possible to elaborate and extend this view such that a degree of belief that p is partly constituted by a disposition to introspect that very degree of belief that p. (Or, the view might say that to believe p more strongly than q is to be in a state partly constituted by a disposition to introspect this state, the state of believing-p-more-strongly-than-q.) Again, there appears to be no structural incompatibility with the required sort of special access to degrees of belief.

So, the first pair of leading theories of introspection just considered *appear* to be favorable to, or at least allow room for, the sort of no-theory theory I am motivated by. The

6 See, e.g., Shoemaker (1993), Byrne (2005: §2), Gertler (2011: ch. 3). Byrne uses "peculiar" instead of "special".

7 See Gertler (2011: §3.2.2 and p. 85).

8 See Armstrong (1963), Lycan (1996: ch. 2), Nichols and Stich (2003: §4.3), Goldman (2006: §§9.1, 9.7–9.8).

9 See Shoemaker (1996, 2009), Boyle (2011), Schwitzgebel (2011).

other pair of leading theories that we are now going to consider, in contrast, *appear* to raise obstacles to the possibility of our having special access to our own degrees of belief. These final two views are Peter Carruthers' **ISA (interpretive sensory-access) theory**,[10] and the so-called **transparency theory**, varieties of which have recently become popular with many authors.[11] I'll argue, however, that there are plausible elaborations of these two views that can overcome the obstacles. Then, in § 2.2, I'll say why they may even be the best bet.

Carruthers thinks that philosophers have been over-generous in crediting people with special access to their own propositional attitudes. He develops his ISA theory intending it as a less generous alternative theory. Here is a statement of the view:

> [T]he system that is employed when one identifies and attributes mental states to oneself is none other than the mindreading system that underlies one's capacity to attribute mental states to other people. Moreover, this system only receives sensory input (including visual, auditory, and motor imagery as well as perceptions of the world and of one's own body). It follows, then, that one's mindreading system lacks direct access to one's underlying attitudes. The latter operate entirely in the background, competing with one another to help influence the contents of consciousness, but remaining inaccessible to the mindreading faculty. Yet there is no other system or mechanism that gives one access to one's own propositional attitudes. In order to attribute thoughts to oneself, then, the mindreading faculty is forced to interpret the available sensory evidence. This can concern one's physical circumstances and overt behavior, or it can involve one's own visual imagery, affective feelings, and **inner speech**. The result is that all access to one's own propositional attitudes is sensory-based and interpretive in nature. (Carruthers 2013*b*: 145–6)

I added the bold highlighting above to draw attention to an important mental feature involved here, *inner speech*. Carruthers allows that we *do* have special access to inner speech, though his own view is that this aspect of self-interpretation has mixed epistemic significance:

> There is also evidence that people's speech actions do not directly express their underlying thoughts but rather (like all other actions) are subject to a variety of competing motivational influences. So when people say what they think (either aloud or in inner speech) this provides some *evidence* of their thoughts without by any means providing direct and reliable access to those thoughts, either to others or to themselves. (This remains true even when people's statements are acknowledged to be sincere.) (Carruthers 2013*b*: 147, italics in original)

In a moment, I'll come back to the ISA theory and say why it may not be as unfavorable as it initially looks to an introspection based view of our knowledge of our degrees of belief. First, let me introduce the transparency theory.

Most fans of the transparency theory today are inspired by the old famous example of Gareth Evans:

> If someone asks me 'Do you think there is going to be a third world war?', I must attend, in answering him, to precisely the same outward phenomena as I would attend to if I were answering

10 See Carruthers (2009, 2010, 2011, 2013*b*).

11 See Dretske (1995: ch. 2), Gordon (1995, 2007), Gallois (1996), Peacocke (1998), Tye (2000: ch. 3), Byrne (2005, 2011, 2012), Setiya (2011), Silins (2012).

> the question 'Will there be a third world war?' I get myself in a position to answer the question whether I believe that *p* by putting into operation whatever procedure I have for answering the question whether *p*. (Evans 1982: 225)

What exactly is the transparency theory? The transparency theory of introspection comes in many varieties, but the basic common idea is this: I can have introspective knowledge *that I believe that* ***p*** on the basis of the very introspected belief *that* ***p***.[12]

This view has some notable virtues. As Byrne (2005: §§7.2–7.3) emphasizes, any time I manage to form the belief *that I believe that* ***p*** on the basis of my belief *that* ***p***, then my newly formed belief will be true. Byrne suggests this explains introspection's high reliability. The view also faces a number of challenges. How, for example, do we use anything like transparency to know our own desires? Advocates of transparency have recently said a lot about how the view should handle this and other challenges.[13] But since there's a big debate over how to extend the transparency theory into these contentious areas, let's take this as another good reason (in addition to space limitations) to focus our attention here on how we can introspect our beliefs.

How, now, can the ISA theory or the transparency view credit us with introspection of our degrees of belief? Consider the transparency theory: I introspect whether I believe there will be a third world war by asking myself whether there will be a third world war. How can such a process deliver the information that I have a certain degree of belief that there will be a third world war? Or consider the ISA theory: what sensory imagery, or what inner speech, gives us good evidence of my having a specific degree of belief in this or that? At first glance, these theories look less well suited to explaining our access to our degrees of belief than a self-scanning or a constitutivist view is.

I want to argue that the transparency theorist and the ISA theorist have resources here. They should follow some guiding ideas provided by a view of how we express our degrees of belief in language. I have in mind a view recently developed and promoted by Eric Swanson and Seth Yalcin, which Yalcin has named *credal expressivism*.[14] Let me explain.

According to credal expressivism, we express a degree of belief when we use an *epistemic modal* like 'probably' to make an assertion like "Ed is probably late", or when we make similar assertions using other epistemic modals such as 'certainly', 'likely', 'maybe', (the epistemic use of) 'possibly', 'might', and others. (Chisholm (1989: 16) catalogues 13 of these terms, giving them special significance in his epistemological theory.) The degree of belief here has for its *content* just the proposition that Ed is late.

The notion of *expresses* here is not self-description. It is meant to be the notion familiarly used to characterize the relation between most ordinary assertions and the speaker's full beliefs. (When I utter "Birds fly", I am describing birds, not myself; I don't say that I believe birds fly, but I express that belief.) On a metaethical expressivist view like that of

---

12 See footnote 11 for advocates of the transparency theory. Silins would insist that we replace, in the above formulation of the view, "on the basis of the very introspected belief *that* ***p***" with "on the basis of the very introspected judgement *that* ***p***". I intend to include such a variation under the umbrella of the transparency theory as it will be discussed here.

13 See, e.g., Byrne (2011, 2012), Setiya (2011).

14 See Swanson (2006: 38; 2011: 251), Yalcin (2007: 1020–1; 2012: 125). See also Moss (2013: 5) and Rothschild (2012: 102) for further development and defense of the expressivist approach to probabilistic language. See van Fraassen (1984: §V; 1989: §§7.2–3) for an under-appreciated early articulation of the view. Christensen (2004: 39) talks as though he accepts the view.

Gibbard (2003), an assertion that applies a normative predicate expresses a planning state, though Gibbard says these planning states can still be considered beliefs or judgments, just ones of a special kind; they are normative judgments. So, like Gibbard, credal expressivists say that a special kind of assertion (one that uses an epistemic modal) expresses a special type of belief (a partial belief).

I presuppose no particular definition of expression.[15] What I will presuppose about expression is just that an utterance is some defeasible evidence that the utterer is in the mental state expressed by that utterance.

Now, credal expressivists tell us what kind of mental state is expressed by a certain use of language, by a certain kind of assertion. Our main question in this paper is about how we have epistemic access to that kind of mental state, degree of belief. Let's link up credal expressivism and our main question.

Consider first the ISA theory, which can help itself to credal expressivism in a fairly direct way. If we have access to assertions made using epistemic modals, and the assertions are at least some evidence of what credal expressivism says they express, then the ISA theorist has the outlines of an answer to our main question. While Carruthers advertises himself as highly skeptical of special access to propositional attitudes, I emphasized earlier that he credits us with special access to *inner speech*. Inner speech is an externally unobservable mental episode that has content, and whose content Carruthers says the subject can immediately understand (Carruthers 2013*b*: p.146). It is produced in the same way ordinary speech is, only it is silent, and it expresses the same mental states that ordinary speech expresses (Carruthers 2009: §2.1; 2010: 79). I noted that Carruthers himself questions the strength of the evidence provided by our speech acts, including inner speech, for the mental states they express. But so long as we are not complete skeptics, so long as inner speech bears some supportive evidential relation to the mental states that our speech acts serve to express, then the ISA theorist is now in a position to give an account of our epistemic access to our own degrees of belief. Thus, when we find that our inner speech is using epistemic modals to assert something like "There probably will be a third world war", we can take this to provide evidence of a degree of belief that such speech expresses. Since this evidence is immediately available only to the subject herself, we have here a form of epistemic access to our degrees of belief that qualifies as special access.[16] This sort of

15 Yalcin, like me, wishes to remain non-committal on the details of what expressing involves; see Yalcin (2012: 139). He does suggest that an utterer must be in the mental state their utterance expresses, and the mental state must cause the utterance. I won't assume any of this, and am personally inclined to deny it. (I am inclined to take it that, in asserting that p, a speaker still expresses the belief that p even if the assertion is insincere and she doesn't believe that p. But this is a somewhat conventional choice about how to understand the relation of expressing.)

Definitions proposed in the literature often follow a Gricean approach. For example, Bach and Harnish (1979: 15) say something roughy like this: a speech act expresses the mental state that the speaker intends their audience to, on the basis of that speech act, take them to be in. See van Fraassen (1984: §V) for a different, but I believe compatible, sketch. Again, I'm not taking on such a definition or any other. For some more discussion of expressing, including its place in the Gricean model, see Pagin (2014: esp. §6); Green (2014: esp. §§2.1, 3.3).

16 Some authors follow a usage on which the described epistemic access does not qualify as "special access" because it is not (in some relevant sense) *direct* access to the introspected attitude. (It involves *direct* access to inner speech, but a further step is required to reach the degree of belief.) And some furthermore say that this sort of epistemic access therefore does not qualify as "introspection". Carruthers himself is an example; he states his view by saying that we cannot introspect our attitudes.

access can be far from comprehensive; we only get access to attitudes that receive verbal expression. But, I don't think this is an implausible consequence; we may have many beliefs that are very hard to uncover.

Again, as indicated by the last block quote from him above, Carruthers has concerns about the reliability of inner speech as evidence of a person's real attitudes, but he still grants them some evidential value (Carruthers 2009: §2.1). My proposal here is, in any case, qualified: as long as it has some positive evidential status, there is room for a view that allows us to ultimately base our knowledge of (or at least justified beliefs about) our degrees of belief (or at least some of them) on such introspection.

Consider, next, how the more popular transparency theory could utilize the guiding ideas of credal expressivism to explain our access to our degrees of belief. Whereas the ISA theory offers us a way to base introspective beliefs on specially perceived silent *speech*, the transparency theory tells us to base introspective beliefs on *other beliefs* (about the external world, typically, as in Evans' world war example). The right strategy here for the transparency theorist is to somehow extend credal expressivism. Credal expressivism concerns assertions of the form *probably* ***p***, where a sentence is embedded in an epistemic modal that is itself unembedded. We should now consider an extension of the view that applies to more complex assertions, specifically to assertions that embed epistemic modals within belief reports, that is, assertions of the form *S believes probably* ***p***. The extension is actually one that Yalcin has said that he suspects is tacitly accepted by Bayesians in their thinking about beliefs that we report using such embedded epistemic modals:

> Credal expressivism is, I think, already tacit in the way that many Bayesians tend to informally describe epistemic modal beliefs. The usual way of modeling, within a Bayesian framework, someone describable as believing (for example) that it is probably raining would be to let the credence function characterizing their credal state map the proposition that it is raining to some highish value. Whether someone accepts what an epistemic modal clause says is thus generally taken to be a matter of their credence in the proposition expressed by the sentence embedded under the modal – not a matter of their credence in a proposition about their credence. (Yalcin 2007: 1021)

Yalcin is right. Some earlier authors, for example Bas van Fraassen and Stephen Schiffer, even explicitly endorsed the picture. Here is how Schiffer described his understanding of degrees of belief (see also van Fraassen 1989: §7.3):

> SPBs [standard partial beliefs] generate corresponding likelihood beliefs. Thus, if Sally s-believes to degree .5 that she left her glasses in her office, then she thinks it is just as likely that she left them there as that she didn't; she thinks, as she would put it, that there's a fifty-fifty chance that her glasses are in her office. If she s-believes to degree .98 that it will rain tonight, then she believes that it will almost certainly rain tonight. If she s-believes to degree .32 that she'll pass her course in number theory, then she thinks it is somewhat unlikely that she'll pass. (Note that these beliefs aren't about any kind of 'objective probability'. They are really just *redescriptions* of a particular kind of partial belief. In the relevant sense, to say that Sally thinks there's a fifty-fifty chance that

See, e.g., Carruthers (2010). As for whether "special access" to our attitudes must be direct, or need only be a route of access that is not equally available to others, I find the former usage in Shoemaker (1993: 79), and the latter, which I have said I am following, in Byrne (2005: 81), though Byrne prefers "peculiar" to "special". Gertler (2011: §3.2.2) clarifies that direct access is one natural way in which access may be special, but directness is not part of her definition of special access.

> she left her glasses in her office is just another way of saying she s-believes to degree .5 that she left her glasses in her office.) (Schiffer 2003: 200, italics in original)

Here Yalcin and Schiffer are not just saying that our assertions that use simple, unembedded epistemic modals bear a link with our degrees of belief (the link of expressing). They are saying that certain beliefs, ones that are reportable by embedding an epistemic modal within the belief report, bear an even closer link with our degrees of belief: these are the same thing. One belief admits of two different belief reports, reports that are simply redescriptions of it. I can report the same belief twice by saying (i) my degree of belief that ***p*** is high, and (ii) I believe it's very likely, or it's highly probable, that ***p***. Weisberg (2013: §3.1) also lends some more support for this view.

If the transparency theorist borrows this view, she can now suggest a reliable rule for introspecting your own degrees of belief. Her old reliable rule told you to believe *I believe that* ***p*** on the basis of your (full) belief that ***p***. Suppose, now, that the content ***p*** here is one we would express using an epistemic modal. So, to illustrate, imagine that the result of the procedures that Evans put into operation when answering whether there will be a third world war was a mental state he would express by saying, as he would put it, *there probably will be*. (He may merely make a judgment and say nothing; this is only what he would say.) From this, the transparency theorist's old rule gets Evans to the belief, as he would report it, *I believe there probably will be a third world war*. Now, from that, we can invite Evans to draw a conclusion that is explicitly about his degrees of belief; he may conclude, *I have a high degree of belief that there will be a third world war*. In general, the transparency theorist can now get you to a belief *I have a high degree of belief that* ***p*** on the basis of a belief that – as you would put it – *probably* ***p***. (All these beliefs' contents, set off in italics, are put as the subject would report them.) And other kinds of degrees of confidence can be introspected on the basis of views expressed by other epistemic modals, like 'might', 'very probably', 'certainly', 'slim chance', and others.

So far, we've indicated how the ISA theory or the transparency theory lets us know at least the rough outlines of our degrees of belief. In the next section, I'll turn to the question of how these views, and the self-scanning and constitutivist views, might grant us more knowledge about the finer structure of our degrees of belief. But before moving on to that, though, let me quickly observe that we're getting the sorts of views I wanted, introspection-based views that don't require the theory-theory. None of these four theories of introspection claim, as the theory-theory did, that our knowledge of our own attitudes is based on beliefs in the non-mental facts that determine that we have those attitudes. Self-scanning views and constitutivist views say our knowledge of our own attitudes is based on no other beliefs. The ISA theory says introspection is based on beliefs in *mental* facts (about sensory data and inner speech). And while the transparency theory does say an introspective belief will typically have a wholly non-mental content in its basis belief (the introspected attitude itself), possession of the introspected attitude is clearly not determined by *that* content. The fact that there will be war won't on its own determine that I have any clue there will be.

2.2 *A challenge for introspection: cardinal structure*

How much can you plausibly learn about your degrees of belief by introspection? One view is that introspection only gives us an ability to introspect comparative intensities

of our degrees of belief: I'm more confident of p than I am of q. For certain theories of introspection, such comparative knowledge might be the best we can plausibly be credited with. From such orderings, we could assign some values with ordinal structure. Any two assignments that shared the same ordinal structure would then have no meaningful difference.[17]

But in order to credit ourselves, and ultimately others, with degrees of belief that have more complex properties like *obeying the axioms of probability*, we would need some way of epistemically accessing features of our degrees of belief that allow assignments of values with some kind of cardinal structure. The particularly challenging property to introspect is your obedience to the so-called *additivity axiom*: your degree of belief in the disjunction of exclusive possibilities must be the *sum* of your degrees of belief in the disjuncts. (The other two axioms say your degrees of belief are non-negative numbers, and that certain trivialities receive the maximum probability. To my mind, these two are somewhat more easily thought of as setting conventions for quantitative measurement of probabilities. In any case, conventional or not, it seems less worrisome how I might introspect whether my degrees of belief obey those two axioms.) It's actually a fairly trivial question whether you or I fully obey the axioms of probability: no ordinary person's credence function, their degrees of belief on the whole, obeys the axioms, since that would require inhuman feats like logical omniscience. But the non-trivial sorts of questions we want to find a way to settle are questions like whether this or that subset of my degrees of belief violates an axiom like the additivity axiom. That's one reason we want access to cardinal structure in our degrees of belief.

Another reason, pointed out by Meacham and Weisberg (2011: §5.2), is that imposing a merely ordinal structure would mean abandoning many useful features of the standard Bayesian model. They point out that it would mean abandoning interpersonal comparisons of degrees of belief. It would mean we could not explain why we struggle to make certain choices by saying that our degrees of belief are very close together. And it would compromise some standard Bayesian solutions to old puzzles: "For example, the standard Bayesian resolution of the raven paradox says that the discovery of a non-black [and non-raven] object should raise our degree of belief in the raven hypothesis, but only by very little" (2011: 660).

So, we may naturally want to impose absolute values with some cardinal structure, so that greater distances between numbers carries significance. But how can introspection deliver this? Can we introspect things like sums of our degrees of belief? Can you introspect that your degree of belief that a die will come up even is *equal* to the *sum* of your degree of belief that the die will come up 2, your degree of belief it will come up 4, and your degree of belief it will come up 6? Let's consider this.

Many theorists have thought it absurd to think we can directly introspect self-standing *numbers* associated with our degrees of belief. (As we saw, Ramsey, and many Bayesians since him, worried about this.) And, you might worry, without access to numbers, what hope is there of introspection telling you whether, say, certain beliefs of yours obey the axioms of probability? This is a big challenge for the introspection based approach.

It might be initially tempting to tackle the problem by reverting to old methods of measuring degrees of belief by appeal to betting preferences. After all, it might be tempting

17 See Zynda (2000) for this view.

to think, if I am allowed to rely on introspection to assure myself that I have no unusual aversion to betting, that I have no diminishing utility for the monetary pay-offs in a given set of bets, and that there is no interference that placing the bets will have on my attitudes, then it might seem I can set aside the classic objections to measuring partial attitudes via betting.[18] However, this tempting line of thought is a dead end. Even if introspection gives me perfect knowledge of my preferences and my utilities, this will not allow me to know whether my degrees of belief obey the probability axioms. Even assuming I maximize expected utility, my preferences and utilities entail only an *ordering* of my degrees of belief.[19] But our present concern is how I could know whether my degrees of belief have cardinal structure, as properties like additivity require.

So, rather than appealing to old techniques involving bets and preferences, it seems to me that the best hope for measuring cardinal structure in our degrees of belief lies, once again, by pairing an introspection-based view of our knowledge of our degrees of belief with the ideas behind credal expressivism. Applied here, the idea would be that our introspectible attitudes can be expressed, or reported, using locutions laden with sophisticated epistemic modals, such as, "[*S* believes] it is *twice as likely* that the die will come up 1–6 as that it will come up even". It's plausible that an assertion like this expresses, and the corresponding belief report describes, cardinal structure in our degrees of belief. (See van Fraassen (1989: §7.2) for endorsement of a similar sketch. Maybe these assertions don't directly express or report such structure, but they are still at least reliably correlated with it; but see Weisberg (2013: §3.1) for some criticism of the mere-correlation position.)

Swanson (2011: 252) emphasizes the importance of sophisticated but ordinary uses of epistemic modals in these ways. He points out that it is reasonably ordinary to say something like "However likely it is that ϕ, it's every bit as likely that ϕ and ψ." He makes the important observation that such an assertion plausibly expresses a *conditional degree of belief*; in an assertion of this form, you express a (maximally) strong conditional degree of belief by expressing a comparative relation between your degree of belief in a conjunction and your degree of belief in one of its conjuncts. A weaker position here might only claim that the assertion expresses the cardinal structure between two of your (unconditional) degrees of belief, and this is normally *correlated* with a conditional degree of belief; the correlation here is between (i) the conditional probability of ψ given ϕ and (ii) the ratio of degree of belief in their conjunction to degree of belief in ϕ. The stronger and simpler position is that that ratio *constitutes* the conditional belief, but that is a claim that some reject, and there's no need to dispute it here.[20] (For another (perhaps more controversial) view of how we express conditional belief, see van Fraassen (1989: §7.2), who suggests a special use of 'if' does the job.)

So, my suggested strategy is, in order to gain some insight into the cardinal structure of our degrees of belief, we should look to structure found in the *chances* that we might, in a perfectly ordinary setting, make assertions about, or report ourselves as having beliefs about. I emphasize that, as before, these are not *objective chances*. The present strategy involves no reliance on anything like the Principal Principle (which interrelates subjective

18 See Ramsey (1926/1931: 172), Eriksson and Hájek (2007: 187–92), Eriksson and Rabinowicz (2013) for these objections.

19 See, e.g., Meacham and Weisberg (2011: 658).

20 See, e.g., Hájek (2003) for discussion.

and objective chances), or reliance on the introspector's obeying such a principle.[21] Credal expressivism says that *subjective chances* are expressed by talk of there being "half a chance of this coin coming up heads", even if it is known that the coin is weighted so that one or the other of the sides has a higher objective chance of coming up.

Now we can see how the ISA theory and the transparency theory can use this strategy to explain how we have epistemic access to facts about cardinal structure in our degrees of belief. For the ISA theorist, we use our special access to our inner speech, while for the transparency theorist, we consider beliefs of ours that would be reported using epistemic modals. And what we now look for in these assertions, or in the complement clauses of those belief reports, are locutions like "It's *twice* as likely that the die will come up 1–6 as that it will come up even" or "The die has *half* a chance of coming up even". On the basis of these assertions, or on the basis of the beliefs that would be reported using these locutions, we can gain introspective knowledge, knowledge I express by saying "I believe that there's half a chance the die will come up even". And *that* knowledge can plausibly be equally well expressed (and that known fact can be redescribed) by saying "My degree of belief that the die will come up even *is half* my degree of belief that it will come up 1–6". Paired with knowledge that I'm certain that the die will come up 1–6, this begins to get us close to knowledge that my degrees of beliefs, at least my degrees of belief about these dice, obey all the probability axioms.

Although introspection becomes less clear and no doubt less reliable as we turn our attention toward finer and finer gradations in our attitudes, it's still reasonably plausible that we can sometimes introspect finer cardinal structure than just our being twice as confident of one proposition as another. When I say "the chance of the die coming up 2 is a third the chance of it coming up even", and I say similar things about the die coming up 4 or 6, I'm plausibly expressing my being in a complex mental state in which three of my degrees of belief add up to another degree of belief I have, once again demonstrating my obedience to the additivity axiom. Indeed, Swanson (2011: 250) even mentions the ordinary use of the sentence "Five to one that ϕ" as one of his opening examples telling in favor of credal expressivism, and telling against older semantic theories that he argues fail to "recognize just how finely grained the language of subjective uncertainty can be". (A page later, he even offers "It's between two and three times likelier that it rained than that it snowed" as an example of an ordinary sentence that an ordinary speaker might use to express a mushy degree of belief.) If such assertions are any reliable indication of the mental states that the credal expressivist claims they express, then we can be optimistic about our ability to introspect cardinal structure in our degrees of belief.

Ramsey and others worried that we cannot introspect self-standing numbers in our own heads, and so cannot introspectively measure whether our degrees of belief have sophisticated properties like additivity. The advocate of credal expressivism's guiding ideas together with either an ISA or a transparency theory of introspection is suggesting otherwise: all the fine structure we can see in the right sort of *chances* of various possibilities gives us access to an equally fine structure we can know about in our degrees of belief.

Turning back to the other theories of introspection now, it seems quite unclear whether the self-scanning or constitutivist views of introspection can follow this strategy with the same ease. It will partly depend on the details of the developed views. It is a perfectly

21 See Lewis (1980).

coherent possibility to develop these views such that they claim we can simply non-inferentially know that we are twice as confident of one proposition as another. I'm just unsure how plausible the resulting view would be. Again, since theorists working on introspection almost never discuss partial attitudes, it is hard to interpret existing views in the literature, and we're only making a start on a large project here. But at least at first glance, it seems far more natural to develop this credal expressivism inspired strategy together with a transparency theory or even the ISA theory. This is why I'm inclined to favor adopting these latter theories.

I emphasize the strategy here at best helps me get *close* to knowing when degrees of belief of mine obey the probability axioms. I haven't even outlined how the ideas of credal expressivism could help me to introspectively know I obey additivity in cases where my degrees of belief don't neatly divide. How, for example, do I know that my degree of belief that Home or Visitor will win is equal to my degree of belief that Home wins added to my degree of belief that Visitor wins?

But if we can have introspective knowledge like this of not only the order of the strengths of own degrees of belief, but of the cardinal structure of their strengths, we can at least hope to somehow leverage out knowledge of something as sophisticated as whether a substantial body of degrees of belief obeys the probability axioms, and in particular whether they obey the additivity axiom. More modestly, at least we can see here an avenue toward a good deal of knowledge of our degrees of belief and their cardinal structure.

2.3 Future steps: other people's degrees of belief

I want to end with a few comments gesturing at how we might incorporate the ideas I'm recommending in this paper into a larger, more comprehensive theory of our knowledge of people's degrees of belief. In particular, I want to consider what we could say about how we know about the degrees of belief of people other than ourselves.

There may be some temptation to simply say: just as the ISA theorist says we can look to our own inner speech to discover our own attitudes, we can likewise just listen to the outer speech of others to learn what their attitudes are. But I think this cannot fully explain how we know others' attitudes. The problem is that, while we immediately understand the content of our own inner speech (again, Carruthers 2013*b*: 146), some epistemological basis is required for us to know the contents of others' speech. I need some basis for interpreting your assertions as meaning the same as what I mean when I use my epistemic modals. So, I don't think the right view is as simple as that.

A promising suggestion for how we can have knowledge of another person's degrees of belief is to say that knowledge about others is inferred, in a way, from knowledge of our own case: knowledge about others might be gained by way of an *analogical inference* concerning functional similarities and differences we observe between ourselves and others.

Now, this suggestion is a non-starter if it's taken to imply that any kind of deliberate or self-conscious inference has to take place. Of course, no such thing goes on when we introspect our own attitudes, and likewise it can't be part of any plausible view on which we make analogizing attributions to others. The plausible theory here had better only say that we have a way of making justified attributions to others, where that justification *depends* on *prior* justification for self-attributions and for beliefs about similarities and differences

between an attributee and oneself.[22] You need not *have* the self-attributing beliefs or the analogical beliefs, but you need prior justification *for* these. (In jargon, doxastically justified attributions to others depend on prior propositional justification for self-attributions and analogies.)

The view can qualify as a no-theory theory. We can observationally know that certain non-mental, functional properties are contingently associated with the degrees of belief we can introspect in ourselves, and then attribute those attitudes to others we know to have sufficiently similar functional properties. These non-mental functional properties do not determine possession of these degrees of belief, but they can serve as evidence, evidence as strong as the kind that underwrites ordinary inductive knowledge. Some fans of this approach may want to claim that it's only possible to make knowledgeable attributions in this way to people we know to be of the same species as us, and thus whose inner, hidden, non-functional properties (presumably neural properties) are the same as ours. This would allow the view, as a no-theory theory, to say that mental states are determined by inner, hidden states that the armchair philosopher has no knowledge of.

This sort of view was famous in the form of a response to the problem of other minds, the so-called response from analogy.[23] Authors in that debate mostly discussed phenomenally conscious mental sensations, such as pains and itches, though they also discussed full attitudes, but I know of no discussion of *partial attitudes* that advertises itself as a discussion of the skeptical problem of other minds.

In recent decades, the skeptical problem of other minds has been discussed less than it was earlier in the 20th century,[24] but another closely related discussion has risen up concerning, as it's come to be called, *mindreading*.[25] In this literature, philosophers and cognitive scientists examine what cognitive mechanisms we use to form beliefs about the mental states of others. There's emerged a remarkable convergence of views in this literature, one that's especially important for present purposes. The convergence has been toward a position that claims that we attribute attitudes to others always partly on the basis of an initial, default attribution to others of our *own* attitudes, revising and correcting that default attribution according to similarities and differences between our circumstances and our attributee's.[26]

---

22 See Pryor (2000) for some explication of the relevant notion of prior, or antecedent, justification.

23 See, e.g., Russell (1948: 159), or Hyslop and Jackson (1972) and citations therein. Hyslop and Jackson give compelling responses to some traditional objections to the view, for example the objection that the analogical inference rashly generalizes from a single case.

24 As noted by Fodor (1994: 292) and Hyslop (2014: opening paragraph).

25 See, e.g., Davies and Stone (1995*a*, *b*), Nichols and Stich (2003), Goldman (2006).

26 See especially, the work of Nichols and Stich, who once strongly opposed the so-called "simulation theory" (e.g. Stich and Nichols 1992), but later came to endorse the claim that our mechanism for mindreading begins with a default attribution of our own attitudes; see Nichols and Stich (2003: 66–7, 85, 92, 106, 140–1). Goldman (2006) defends a version of the simulation theory that disagrees with Nichols and Stich (2003) on many issues, but agrees on the broad claim noted above; see especially §§2.5 and 7.7.

A note about the view of Carruthers: Carruthers (2011, 2013*a*) is sympathetic to much of the model that Nichols and Stich develop, and Carruthers himself observes that the mindreading literature has converged toward including an essential role for simulation in mindreading (Carruthers 2011: 225, 230). However, Carruthers wishes to insist on a distinction between (i) views where an attribution of a belief that *p* to another person is made merely by "drawing on" one's own belief that *p*, and (ii) views where attributions to others are "based on" or "depend on" introspection of one's own belief

However, our epistemological questions are not directly targeted in the mindreading literature. That literature has almost entirely set aside all normative issues, including the questions of how we have justification for believing, and thus know, facts about others' minds. (As Goldman (2006: 10) says at the start of his book, "The subject matter of mindreading, however, doesn't investigate issues of justification or knowledge; these topics are set aside in the field of mentalizing in general and in this book.") There is thus a program here that is ripe for fresh exploration. I hope, then, that future work will not only examine and develop the ideas we've only begun to sketch about introspection of degrees of belief, but also fit these ideas into a comprehensive epistemological theory of our knowledge of all attitudes.[27]

REFERENCES

Armstrong, D. 1963. 'Is Introspective Knowledge Incorrigible?' *Philosophical Review*, 72: 417–32.

Bach, K. and Harnish, R. M. 1979. *Linguistic Communication and Speech Acts*. Cambridge, MA: MIT Press.

Bermúdez, J. 2005. *Philosophy of Psychology*. New York, NY: Routledge.

Boyle, M. 2011. 'Transparent Self-Knowledge.' *Proceedings of the Aristotelian Society, Supplementary Volume*, 85: 223–41.

Braddon-Mitchell, D. and Jackson, F. 1996/2007. *Philosophy of Mind and Cognition*. 2nd edn. Oxford: Blackwell.

Byrne, A. 2005. 'Introspection.' *Philosophical Topics*, 33: 79–104.

—— 2011. 'Transparency, Belief, Intention.' *Proceedings of the Aristotelian Society, Supplementary Volume*, 85: 201–21.

—— 2012. 'Knowing What I Want.' In J. Liu and J. Perry (eds), *Consciousness and the Self: New Essays*, pp. 165–83. Cambridge: Cambridge University Press.

Carruthers, P. 2009. 'How We Know Our Own Minds: The Relationship between Mindreading and Metacognition.' *Behavioral and Brain Sciences*, 32: 121–38.

—— 2010. 'Introspection: Divided and Partly Eliminated.' *Philosophy and Phenomenological Research* 80: 76–111.

—— 2011. *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford: Oxford University Press.

—— 2013*a*. 'Mindreading in Infancy.' *Mind and Language*, 28: 141–72.

—— 2013*b*. 'On Knowing Your Own Beliefs: A Representationalist Account.' In N. Nottelmann (ed.), *New Essays on Belief: Structure, Constitution and Content*, pp. 145–65. New York, NY: Palgrave MacMillan.

Chisholm, R. 1989. *Theory of Knowledge*. 3rd edn. Englewood Cliffs, NJ: Prentice Hall.

that *p* (Carruthers 2013*a*: 143–4). He says Nichols and Stich (2003) only accept the former, and Carruthers endorses this view, and he says Goldman (2006) accepts the latter, wrongly in Carruthers' view. I don't have space here to interpret and critically examine Carruthers's intended distinction. However, I believe my purposes here are met even by the view that Carruthers does accept, as when he says: "For the standard way of predicting what someone with a given belief will think or do is to assume that belief for oneself, and then to reason on one's own behalf (with suitable adjustments for the context, and for other differences from the target), attributing the result to the other person. (This is the core truth in simulationist models of mindreading; see Nichols and Stich (2003); Carruthers (2011, 2013*a*).)" Carruthers (2013*b*: 160).

27 For help during the preparation of this paper, I must thank a very constructive anonymous referee, Ray Buchanan, Chris Meacham, Adam Pautz, Karl Schafer, Miriam Schoenfield, David Sosa, Jonathan Weisberg, an audience at CUNY's Cognitive Science Speaker Series, and a meeting of the epistemology reading group *officiale* ('ERGo') at the University of Texas at Austin.

Christensen, D. 2004. *Putting Logic in Its Place: Formal Constraints on Rational Belief*. Oxford: Oxford University Press.
Davies, M. and Stone, T. (eds) 1995*a*. *Folk Psychology: The Theory of Mind Debate*. Oxford: Blackwell.
—— (eds) 1995*b*. *Mental Simulation: Evaluations and Applications*. Oxford: Blackwell.
Dretske, F. 1995. *Naturalizing the Mind*. Cambridge, MA: MIT Press.
Eriksson, L. and Hájek, A. 2007. 'What Are Degrees of Belief?' *Studia Logica*, 86: 183–213.
—— **and Rabinowicz, W.** 2013. 'The Interference Problem for the Betting Interpretation of Degrees of Belief.' *Synthese*, 190: 809–30.
Evans, G. 1982. *The Varieties of Reference*. Ed. J. McDowell. Oxford: Oxford University Press.
Fodor, J. 1994. 'Jerry A. Fodor (Self-Profile).' In S. Guttenplan (ed.), *A Companion to the Philosophy of Mind*, pp. 292–9. Oxford: Blackwell.
Gallois, A. 1996. *The World Without, the Mind Within: An Essay on First-Person Authority*. Cambridge: Cambridge University Press.
Gertler, B. 2011. *Self-Knowledge*. New York, NY: Routledge.
Gibbard, A. 2003. *Thinking How to Live*. Cambridge, MA: Harvard University Press.
Goldman, A. 2006. *Simulating Minds: the Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford: Oxford University Press.
Gordon, R. 1995. 'Simulation without Introspection or Inference from Me to You.' In M. Davies and T. Stone (eds), *Mental Simulation: Evaluations and Applications*, pp. 53–67. Oxford: Blackwell.
—— 2007. 'Ascent Routines for Propositional Attitudes.' *Synthese*, 159: 151–65.
Green, M. 2014. 'Speech Acts.' In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, winter 2014 edition. http://plato.stanford.edu/archives/win2014/entries/speech-acts/.
Hájek, A. 2003. 'What Conditional Probability Could Not Be.' *Synthese*, 137: 272–323.
—— 2008. 'Arguments For – or Against – Probabilism?' *British Journal for the Philosophy of Science*, 59: 793–819.
—— **and Smithson, M.** 2012. 'Rationality and Indeterminate Probabilities.' *Synthese*, 187: 33–48.
Hampton, J. 1994. 'The Failure of Expected-Utility Theory as a Theory of Reason.' *Economics and Philosophy*, 10: 195–242.
—— 1998. *The Authority of Reason*. Ed. R. Healey. Cambridge: Cambridge University Press.
Howson, C. and Urbach, P. 2005. *Scientific Reasoning: The Bayesian Approach*. 3rd edn. Peru, IL: Open Court.
Hyslop, A. 2014. 'Other Minds.' In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, spring 2014 edition. http://plato.stanford.edu/archives/spr2014/entries/other-minds/.
—— **and Jackson, F.** 1972. 'The Analogical Inference to Other Minds.' American Philosophical Quarterly, 9: 168–76.
Joyce, J. M. 2005. 'How Probabilities Reflect Evidence.' *Philosophical Perspectives*, 19: 153–78.
Lewis, D. 1974. 'Radical Interpretation.' *Synthese*, 27: 331–44.
—— 1980. 'A Subjectivist's Guide to Objective Chance.' In R. Jeffrey (ed.), *Studies in Inductive Logic and Probability*, Volume II. Berkeley, CA: University of California Press.
—— 1983. 'New Work for a Theory of Universals.' *Australasian Journal of Philosophy*, 61: 343–77.
—— 1994/1999. 'Reduction of Mind.' In *Papers in Metaphysics and Epistemology*, pp. 291–324. Cambridge: Cambridge University Press.
Lycan, W. 1996. *Consciousness and Experience*. Cambridge, MA: MIT Press.
Meacham, C. and Weisberg, J. 2011. 'Representation Theorems and the Foundations of Decision Theory.' *Australasian Journal of Philosophy*, 89: 641–63.
Moss, S. 2013. 'Epistemology Formalized.' *Philosophical Review*, 122: 1–43.
Nichols, S. and Stich, S. 2003. *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford: Oxford University Press.
Pagin, P. 2014. 'Assertion.' In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, spring 2014 edition. http://plato.stanford.edu/archives/spr2014/entries/assertion/.
Peacocke, C. 1998. 'Conscious Attitudes, Attention and Self-Knowledge.' In C. Wright, B. Smith and C. Macdonald (eds), *Knowing Our Own Minds*, pp. 63–99. Oxford: Oxford University Press.
Pryor, J. 2000. 'The Skeptic and the Dogmatist.' *Noûs*, 34: 517–49.
Ramsey, F. 1926/1931. 'Truth and Probability.' In R. Braithwaite (ed.), *The Foundations of Mathematics*, pp. 156–98. London: Routledge and Kegan Paul.

Ramsey, W. 2013. 'Eliminative Materialism.' In E. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, summer 2013 edition. http://plato.stanford.edu/archives/sum2013/entries/materialism-eliminative/.
Rothschild, D. 2012. 'Expressing Credences.' *Proceedings of the Aristotelian Society*, 112: 99–114.
Russell, B. 1948. *Human Knowledge: Its Scope and Limits*. London: George Allen and Unwin.
Savage, L. 1954/1972. *The Foundations of Statistics*. New York, NY: Dover.
Schiffer, S. 2003. *The Things We Mean*. Oxford: Oxford University Press.
Schwarz, W. 2014. 'Against Magnetism.' *Australasian Journal of Philosophy*, 92: 17–36.
Schwitzgebel, E. 2011. 'Knowing Your Own Beliefs.' *Canadian Journal of Philosophy* 35 (Supplement, Belief and Agency, ed. D. Hunter): 41–62.
Setiya, K. 2011. 'Knowledge of Intention.' In A. Ford, J. Hornsby and F. Stoutland (eds), *Essays on Anscombe's Intention*. Cambridge, MA: Harvard University Press.
Shoemaker, S. 1993. 'Special Access Lies Down with Theory-Theory.' *Behavioral and Brain Sciences*, 16: 78–9.
—— 1996. *The First-Person Perspective and Other Essays*. Cambridge: Cambridge University Press.
—— 2009. 'Self-Intimation and Second Order Belief.' *Erkenntnis*, 71: 35–51.
Silins, N. 2012. 'Judgment as a Guide to Belief.' In D. Smithies and D. Stoljar (eds), *Introspection and Consciousness*, pp. 295–328. Oxford: Oxford University Press.
Stich, S. and Nichols, N. 1992. 'Folk Psychology: Simulation or Tacit Theory?' *Mind and Language*, 7: 35–71.
Swanson, E. 2006. *Interactions with Context*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.
—— 2011. 'How Not to Theorize about the Language of Subjective Uncertainty.' In A. Egan and B. Weatherson (eds), *Epistemic Modality*, pp. 249–69. Oxford: Oxford University Press.
Tye, M. 2000. *Conciousness, Color, and Content*. Cambridge, MA: MIT Press.
van Fraassen, B. 1984. 'Belief and the Will.' *Journal of Philosophy*, 81: 235–56.
—— 1989. *Laws and Symmetry*. Oxford: Oxford University Press.
Weisberg, J. 2013. 'Knowledge in Action.' *Philosophers' Imprint*, 13: 1–23.
Yalcin, S. 2007. 'Epistemic Modals.' *Mind*, 116: 983–1026.
—— 2012. 'Bayesian Expressivism.' *Proceedings of the Aristotelian Society*, 112: 123–60.
Zynda, L. 2000. 'Representation Theorems and Realism about Degrees of Belief.' *Philosophy of Science*, 67: 45–69.

SINAN DOGRAMACI is Assistant Professor of Philosophy at the University of Texas at Austin. His two main areas of research are the epistemology of basic logic and the practical function of epistemic evaluation.