Nonclassical minds and indeterminate survival

J. Robert G. Williams*

April 1, 2014

Contents

*This paper grew out of work on indeterminate survival I worked on while on an AHRC research leave award 'Intrinsic survival, Multiple survival, Indeterminate survival'. Material relevant to this was presented in several places: in Oxford, Cambridge, Leeds, Barcelona, MIT, and the UNC metaphysics workshop—thanks to my commentator on that occasion, Dana Goswick. This work was carried out while supported by a British Academy Research Development Award (BARDA: 53286) and the AHRC Metaphysical Indeterminacy early career research grant. I'm grateful to all with whom I've discussed these issues, and especially to reviewers for *Philosophical Review*. My greatest debts, as ever, are to Elizabeth Barnes, Ross Cameron and Jason Turner.

Beliefs aim at truth, and should be logically consistent. A consequence is that if we revise our views on truth and logical consistency by shifting from a familiar classical setting to a nonclassical one, our theory of belief will have to change too. This paper looks at life in a nonclassical world, through a case study of attitudes to *indeterminate survival*.

The paper is divided into three parts. Part I is an overview of the implications of nonclassicism for our philosophical account of rational belief, desire and decision-making—what I here call a 'theory of mind'. Parts II and III consider what nonclassical theory of mind is appropriate to cases of indeterminate personal identity, with Part II focused on puzzles arising from *degrees* of psychological connectedness between person stages, and Part III focused on puzzles of Fission and Longevity.

I Nonclassicism and the mind

Nonclassicists argue for frameworks that diverge from orthodox classical logic and semantics. Some argue for divergence with respect to logical consequence: which arguments are valid, which collections of sentences inconsistent, or which sentences are logically true. Others diverge from classicism over the nature and distribution of alethic statuses. If Patchy is a borderline red/orange colour swatch, the first kind of nonclassicist might deny that the following is logically true:

(LEM) Patchy is red or Patchy is not red.

whereas the second kind of nonclassicist can agree that (LEM) is logically true, but may argue that the following instance of bivalence fails:

(BIV) Either 'Patchy is red' is true or 'Patchy is red' is false.

We may distinguish two senses in which a philosophical position can be revisionary. It may be revisionary of commonsense. A paradigm is the nihilist position that denies that tables and rocks exist.[1] Or it may be revisionary of theory by undercutting what we previously took to be the leading account of some phenomenon. Nonclassicism isn't necessarily revisionary in the first sense, but it is characteristically revisionary of theory. It begins by challenging the orthodox account of logic and semantics, and in doing so it undercuts theoretical work in the sciences and philosophy which has presupposed a classical backdrop. Since the default is to make free use of the "standard" rules of classical reasoning when working out what theories predict, the body of theory thereby threatened may be quite extensive.

If your philosophical position is revisionary of common opinion it's hard to do anything to change that fact, short of embarking on a programme of public information films. But by developing better theories of the Fs one can make a difference to whether or not one's philosophical position is compatible with *current best* theory of Fs, and it is this kind of project in which I am engaged here. One theory that is threatened by nonclassicism is our theory of mind. In the remainder of this section I describe the alterations required to that theory, to make it compatible with nonclassicism.

I.1 The centrality of nonclassical theory of mind

The *theory of mind*, as I'll use the phrase here, picks out a cluster of views about rational belief, belief change, desires, and decision-making. One *classical* version of this theory includes the following claims:[2]

[1](van Inwagen, 1990)

[2]For discussion and defense, see (Howson & Urbach, 1993; Jeffrey, 1965).

Probabilism: Ideally rational beliefs come in degrees, and meet the constraints of probability theory.

Conditionalizing: The rational way for an agent with beliefs b to respond to receiving total evidence E, is by moving to the belief state b' that results from conditionalizing b on E.

Utilities: The rational belief-desire states for agents to be in are those that are represented by probability-utility pairs meeting the constraints of Jeffrey's decision theory.

Maximizing: The rational act to choose, out of given set of options, is the one that has maximum utility.

The model is committed to classicism. This is apparent from the first, most foundational element—probabilism. In standard probability theory, every classical tautology has probability 1, and so for the probabilist, ideally rational agents should assign full confidence in classical tautologies such as (LEM). But that is not something that the nonclassicist doubter of (LEM) can accept! Such examples can be multiplied.

(A housekeeping point. The objects of belief are usually taken to be propositions, whereas logical consequence and the theory of probability alluded to above take sentences as objects. This apparent mismatch makes it hard to interpret the claim that rational degrees of belief are probabilistic. However, standard theories of logic, truth and probability can be easily generalized to take fine-grained *propositions* or Fregean *thoughts* as objects, and so bring the two into line. Conversely some have held that beliefs do consist in a subject's relation to a (mentalese) sentence.[3] I will assume that there is a match between the objects of belief and objects of logic, truth and probability, without making further assumptions about what these are.[4])

There are good reasons for the nonclassicist to start by seeking to generalize the classical theory of mind, leaving consideration of other classically based theories (whether in chemistry, geography, ecology or even philosophy) for later. The first reason for this is that *every* nonclassicism undercuts the classical theory of mind, whereas it does not automatically undercut a classically-based theory of ecology. The core principles of the theory of mind themselves concern truth and logic *explicitly* (as the aim of belief, or norm of coherence), so revisions to our account of the latter have an immediate impact on our account of the former. A theory of ecology, by contrast, needn't mention truth-statuses at all, and draws on logic only implicitly, when we derive consequences of its principles. Logical or semantic revisionism *may* undercut a classically-based ecology, but equally it may turn out to leave it untouched: that (LEM) fails to be a logical truth does *not* mean that the law of excluded middle has exceptions *expressible in the language of ecological theory*; and an invalid argument form may preserve truth in all ecological applications.[5]

The second reason why nonclassicists are especially obligated to develop a nonclassical theory of mind is that *in accepting the nonclassical theory in the first place* they already take on commitments in this area. To communicate their theory to others, they need to convey what these commitments are. And since those commitments should be open to critical scrutiny, the best way to do this is to formulate (relevant parts) of the nonclassical theory of mind. Let me take these points in turn.

[3]See (Field, 1978)

[4]One position on the objects of attitudes that would make a considerable difference is the 'coarse-grained' framework of (Stalnaker, 1984), on which the objects of belief (and probability) are sets of possible worlds. I set these aside here, but discuss their relation to the present framework in (**?**).

[5]Compare and contrast (Beall, 2009) on restricting nonclassicism to liar-like phenomena and (Field, 2008) on effective classicism across set theory. For discussion of whether revisionism about truth statuses (e.g. in supervaluationism) induces logical revisionism of a damaging kind, see (Williamson, 1994) and (Williams, 2008).

The nonclassicist is proposing a certain theory T, of future facts, conditionals, semantic predications, colour patches, or whatever. This involves attributing nonclassical truth statuses to various propositions about the future, conditionals, semantics, and colour patches respectively. In accepting T, you must of course believe the claims that appear in T explicitly. But more than this is rationally required of one who accepts the theory. Suppose T entails that *it is indeterminate whether if I flip this fair coin, it will land heads*, where indeterminacy is some third status incompatible with truth or falsity. It would be *prima facie* surprising if I could rationally persist in a belief that *if I flip the fair coin, it will land heads*, given that I'd just agreed that this proposition was indeterminate. Part of the job of the nonclassical theory of mind is exactly to tell us about the rational constraints here. According to one nonclassical theory of mind, accepting that p is indeterminate rationally requires you to disbelieve p. A rival view says that accepting it rationally requires you to suspend judgement about p. On a third view, it rationally requires you become half-confident in p. (The part of the nonclassical theory of mind at issue I call 'the cognitive role of indeterminacy', covering the rational relation between *belief* and indeterminacy.)

The theorist herself, accepting T, adopts some stance or other towards the propositions that T describes as indeterminate, whether that stance is belief, disbelief, half-confidence, suspension of judgement or something else entirely. In doing so, she accords with some accounts of the cognitive role of indeterminacy, and violates others. That is why I claimed above that in accepting T, the theorist already implicitly takes on commitments in the nonclassical theory of mind.

In communicating one's theory to others, one needs to convey the commitments just identified. The thought here is simple: the theorist wants her audience to come to believe T. But unless she conveys the intended cognitive role of indeterminacy one way or another, then she will leave the audience at a loss about how they're expected to integrate T with their wider beliefs. (When someone presents a nonclassical logico-semantical treatment of some area, I for one am often in exactly this position: I don't immediately know *what* I'm being asked to accept.)

The theorist *could* convey her intentions without explicitly formulating a nonclassical theory of mind. She can leave clues to the intended cognitive role of the nonclassical statuses in the overtones of the labels she assigns to them, or trust her audience to guess correctly based on what makes best sense of the illustrative examples given. But it is theoretical good practice to set such things out explicitly. The intended cognitive role for nonclassical statuses can make a huge difference to the plausibility of the theory being proffered, and in such cases one shouldn't tolerate ambiguity about the intended reading of one's theory.

A first illustration of the difference this can make: suppose you thought that accepting that *it is indeterminate whether p* rationally requires you to utterly reject p and utterly reject $\neg p$ (Field, 2003b). Then accepting that future contingents are indeterminate (as an articulation of the 'open future' thesis) would be drastically revisionary of common opinion. Future contingents include propositions such as *that the coin now spinning in the air will land heads*, and *that there will be blue skies overhead on the way home*. These are paradigmatically things in which we have middling confidence, and *do not* utterly reject. So either we radically revise our pretheoretic opinions in such matters, or we must reject the theory that tells us that they are indeterminate *insofar as this is paired with the 'rejectionist' cognitive-role mentioned*. Clearly that last clause is crucial: assign a different cognitive-role to indeterminacy and the open future thesis may be perfectly compatible with pretheoretic opinion.

A second illustration: suppose that indeterminacy requires a middling level of confidence. Take two borderline red/orange colour patches with distinct but indiscriminable shades. On this theory, accepting that *it is indeterminate whether the first is red and the second not red* rationally requires a middling confidence in whether the red/not-red cut-off occurs between the two shades. But the basic phenomenon that generates the sorites is that such cut-off claims are utterly wild and must be rejected! So again, the theory here requires revision of pretheoretic opinion.

The point in both cases is that we're not in a position to measure how much of a shift in our prior beliefs a nonclassical theory requires, nor the plausibility of that shift, unless and until an account of the cognitive role of the nonclassical statuses is provided.

In sum, a nonclassical generalization of the theory of *mind* has a peculiar relevance to our ability to accept, recommend and evaluate nonclassicism in the first place. It should be the first item on a nonclassicist's agenda.[6]

I.2 A space of nonclassical theories of mind

Two projects now suggest themselves. One is global: to scope out the space of possible views on indeterminacy and associated rational mental states. The second is more local: to pinpoint where in that space particular phenomena lie. That is: assuming nonclassicism is the right view to take of conditionals, or the future, or borderline predications, what is the right nonclassical theory of mind to pair it with?

The arguments of Parts II and III are a contribution to the second enterprise. In the remainder of this section, I will briefly discuss the first, map-making exercise, reporting on results discussed in detail elsewhere. In sum: moving to a nonclassical setting means that the doctrine that degrees of belief should be representable as (classical) probability functions needs to be given up, and with it the classical theory of mind based upon it. But for a vast range of nonclassical theories, generalized versions can be provided, in each case delivering a baseline nonclassical theory of mind. (I do not claim that *every* nonclassical theory in the marketplace can be fitted into the space I will describe; but very many of them can, including all those relevant for later parts of this paper).

We start from three assumptions about truth, belief and logic that will be held fixed throughout. First, rational beliefs should be broadly *coherent*, and logical consistency is one important articulation of that coherence requirement. Second, beliefs in p are correct if and only if they correspond to p's truth value (and so in that sense 'aim at' truth). Third, truth and logic relate: logical consequence is (necessary) *truth* preservation.

The classical theory of mind set out earlier regiments these connections in a classical setting. We assume that the way to describe the belief state of an agent is by a function b that assigns to each sentence or proposition the credence (degree of belief) that the agent invests in it. We write $|A|_w$ for the truth value of sentence or proposition A at world $w \in W$. We take the truth value to be a *number*—1 if it is True, and 0 if it is False. These numbers reflect the cognitive roles of classical Truth and Falsity. God should have credence 1 in each Truth, and credence 0 in each Falsity.

The three claims above can then be formalized as follows:

Truth-Logic: B is a consequence of A iff in all situations, the truth value of B is no less than that of A.[7]

[6] The literature on the nonclassical theory of mind is small in comparison to the vast literature on nonclassical logic and semantics. Some highlights include (Field, 2000, 2003a,b, 2008), (Wright, 2001), (Paris, 2001), (Schiffer, 2003), (Hawthorne, 2005), (Priest, 2006), Smith (2008, 2010), (Barnett, 2009), (MacFarlane, 2010), (Caie, 2012). See also: (Williams, 2012c, 2014).

[7] That is, $A \models B$ iff in all worlds w, $|A|_w \leq |B|_w$

Belief-Truth: A rational belief state b is *probabilistic*$_1$, where that condition is met iff for each A, $b(A)$ is a suitable weighted average of the truth values A takes at worlds.[8]

Belief-Logic: A rational belief state b is *probabilistic*$_2$, where that condition is met iff $b(A)$ is one for A a tautology; zero for A a contradiction; b never assigns higher credence to A over B when B follows from A; and treats disjunctions and conjunctions appropriately.[9]

The overall shape is what is crucial here. We have a characterization of rational belief directly in terms of truth values, a characterization of rational belief via logical constraints; and an interconnection between truth and logic. Standard results reassure us that these will play nicely together—for example, that the conditions probabilistic$_1$ and probabilistic$_2$ pick out the same set of functions from propositions to numbers—the classical probability functions. Given that equivalence, (Belief-Logic) and (Belief-Truth) are two interchangeable ways of articulating *probabilism*—the doctrine that rational belief states are describable as probability functions.[10]

As we know, this classical theory of mind will be unattractive to a nonclassicist. Given that logic is classical, (Belief-Logic), for example, says that the rational agent shouldn't have credence less than 1 in any classical tautology. But the nonclassicist thinks that we *should* doubt the classical tautology (LEM). Naturally enough, the nonclassicist will diagnose the source of the problem as the appeal to classical logic and truth within the accounts.

It turns out, however, that the nonclassicist can preserve (Belief-Logic), (Belief-Truth) and (Truth-Logic), so long as she understands the appeal to 'truth values' and 'logic' nonclassically. Crucially, to get this started we need to generalize the classical treatment of alethic statuses. Just as in the classical case, we need to associate each alethic status with a 'cognitive load'—a numerical representation of the level of confidence God would have in A, when A has the status in question.[11] We call this the truth value of A (at the relevant world w), and just as before write it as $|A|_w$. These values needn't be restricted to 1's and 0's—intermediate cognitive loads (truth-values between 0 and 1) are perfectly possible, if the nonclassicist wishes to claim that the Godlike credence to have in a proposition with a given alethic status is intermediate. The

[8] By "suitable weighted average" I mean a convex combination: that is, there are some indices λ_w, taking values in $[0,1]$ that sum to 1, such that for each A,

$$b(A) = \sum_{w \in W} \lambda_w |A|_w.$$

Note that λ_w can then be thought of as the degree of belief invested in world w; though the issues of interpretation here are somewhat subtle, and no such interpretation is required to state the constraint.

[9] More precisely:

Tautology If $\models T$, then $b(T) = 1$

Zero If $F \models$, then $b(F) = 0$;

Monotone If $A \models B$, then $b(A) \leq b(B)$;

Inclusion-Exclusion $\forall A, B$: $b(A \vee B) + b(A \wedge B) = b(A) + b(B)$.

In fact, the last step can be formulated more generally, in terms of a (generalization of) the logical notion of a *partition*—see (Williams, forthcoming) for details.

[10] For simplicity, I'm working here with a finite set of worlds in formulating the last two constraints. But if the worlds are to cover all logical possibilities, then it is very likely we'll want infinitely many. There are of course generalizations of the formulations, but we won't need to enter into the intricacies here.

[11] A note on terminology. I use 'cognitive role' for the general issue of the interaction of belief and indeterminacy, including describing the rational constraints between beliefs whose contents attribute indeterminacy to p, and belief in p itself. 'Cognitive load' instead describes what the *ideal* credence to have in p is, given the *facts* about its alethic status. 'Cognitive load' is thus interchangeable with 'truth value'—the latter terminology reminds us of its role within logic and semantic theory; the former terminology reminding us of its significance in the deployments of logic and semantics in the wider theory of mind.

nonclassical system will tell us about how these are distributed.[12] If a nonclassicist won't or can't tell us what the cognitive load of her envisaged nonclassical statuses are, we can't offer her a corresponding nonclassical theory of mind, and she will have to find some other way to discharge her obligations. But a vast range of different nonclassicists have the resources to spell out their position in terms of truth values, so characterized.[13]

With the nonclassical truth values specified, (Belief-Truth) now says rational beliefs are required to be suitable weighted averages of the *nonclassical* truth values of propositions. (Truth-Logic) should similarly be read in terms of nonclassical truth values. If the nonclassicist's consequence relation satisfies (Truth-Logic) then we immediately recover much of (Belief-Logic), including the principle that *nonclassical* tautologies should be probability 1, contradictions probability 0, and that probability shouldn't be lower in C than A when C follows from A (Williams, forthcoming, 2012a,b). What more we can get—and in particular whether we can completely characterize these nonclassical probabilities (i.e. those that are probabilistic$_1$) by means of the nonclassical logic—requires more detailed case-by-case investigation.[14] In sum: the first element of the classical picture of mind—probabilism—has a natural generalization to the nonclassical case. This is not a purely formal exercise: influential *justifications* for probabilism—dutch book and accuracy domination arguments—can be adapted to support the generalized probabilistic claims in the nonclassical setting.[15]

Probabilism is but one element of the theory of mind that we offer in the classical framework. So the next step is to offer accounts and defences of other aspects of that theory of mind that dovetail with the generalized probabilism just discussed. Here we again find a neat network of connections in the classical case, between categorical belief, conditional belief, and rational combinations of desire and belief as characterized in decision theory:

Categorical-Conditional If one's belief state is rational, then the conditional credence invested

[12]Note that more than one nonclassical status can have the same truth *value*. This corresponds to Dummett's distinction (1959) between ingredient and assertoric sense. Truth values determine the right attitude to have to a sentence. But more fine-grained distinctions may be required to spell out how the status of complex sentences are determined by the statuses of their parts.

[13]Logics characterized by the preservation of 'designated' truth statuses can often be handled by giving designated statuses cognitive load 1, and undesignated statuses cognitive load 0, for example. Certain kinds of substructural logics are hard to interpret in this framework. Since the truth values will be real-valued, (Truth-Logic) effectively ensures that consequence will be transitive. (On the other hand, in (Williams, 2011) I discuss a logic within this system which fails to vindicate the structural rule $A \vdash B, A, B \vdash C \therefore A \vdash C$. So it's not that substructural logics as a class are beyond our reach).

[14]Results due to Jeff Paris (2001) and others give a central, elegant result, albeit one that does not apply to the specific cases we will look at in the next section. Suppose (i) that truth *values* are just 1 or 0 (though perhaps distributed very differently from the classical case) and (ii) their distribution meets the following pair of conditions:

$$\begin{array}{lccc} (\mathcal{T}a) & V(A)=1 \wedge V(B)=1 & \iff & V(A\wedge B)=1 \\ (\mathcal{T}b) & V(A)=0 \wedge V(B)=0 & \iff & V(A\vee B)=0. \end{array}$$

Then the generalized probabilities (which we argued were the nonclassically rational belief states) are exactly those functions that meet the constraints indicated in (Belief-Logic)—simply replacing appeal to classical logic with the nonclassical logic characterized by (Truth-Logic). So we *uncover* a certain normative role for logic, so characterized. If either (i) or (ii) fail, then the picture is more complex, and axioms may need to be tweaked or ambitions scaled back. But across a huge range of theories, analogues of the three basic interconnections can be spelled out. One important shift here is to move away from familiar axiomatizations of logic that make explicit appeal to disjunctions and conjunctions, and replace this key axiom with one that appeals to *partitions*. The former is sensitive to the exact behaviour of the logical connectives, and may have to be tweaked in nonclassical settings; the latter can apply invariantly (Williams, forthcoming).

[15](Paris, 2001) adapts the dutch book argument to a nonclassical setting. I generalize the arguments of (Joyce, 1998, 2009) to a nonclassical setting in (Williams, 2012b), and discuss the relation between these two in (Williams, 2012a).

in A given C is the ratio between the credence given to $A\&B$ and the credence given to C.[16]

Belief-Desire For an arbitrary partition of propositions Γ, the desirability of A is the average desirability of $A\&S$ for $S \in \Gamma$, weighted by the conditional credence in S given A as characterized in (Categorical-Conditional).[17]

Desire-Action Faced with a range of options for action, one should opt for the one with maximimum utility/desirability, as characterized in (Belief-Desire).

In the first case, the ratio formula connects conditional credences—which go on to have an important theoretical role in standard stories about belief update and decision—to unconditional credence. The (Belief-Desire) connection, together with the (Desire-Action) link, allows us evaluate the desirability of a given course of action (A) so long as we know how desirable its consequences would be in each of a range of possible background states of the world. The credence we attach to these background states (conditionally on A being the case) allows us to aggregate these assessments.

Once more there is a motivated generalization of this classical package to our nonclassical settings. In the classical case, the conjunction & that figures in the principles can be identified with the classical truth-function $\wedge$. What we need for the nonclassical generalization is a connective & which satisfies one of the features of classical $\wedge$: that the truth value of $A\&B$ is the product of the truth values of A and B. We also need to extend the notion of a 'partition' of propositions so it makes sense when those propositions can take nonclassical truth values (here is the appropriate generalization: Γ form a partition if at each world, the sum of the truth values of the elements of Γ is 1). With an appropriate connective & and the generalized notion of partition to hand, the principles above can be underpinned in the nonclassical setting by synchronic and diachronic dutch book arguments.[18]

Thus, while nonclassicism is incompatible with the classically-based theory of mind, for a vast range of nonclassical theories, generalized versions can be provided that preserve a familiar network of connections between truth, logic and rational belief. In each case these deliver a baseline nonclassical theory of mind.

I.3 The supervaluational family

Section 1.2 outlined a space of nonclassical theories of mind at a high level of generality. In this section, I argue that it allows us to pin down four different versions of 'supervaluational' nonclassical semantics for indeterminacy.[19] Despite a common semantico-logical skeleton, the

[16]Assuming probabilism, rational categorical beliefs are representable as a probability function, and this link requires that rational conditional beliefs match the associated conditional probabilities under the standard definition: $P(A|C) = \frac{P(A\&C)}{P(C)}$, for $P(C) \neq 0$

[17]That is, for Γ a partition of possibilities, P rational degrees of belief, rational desires (utilities) should satisfy: $D(A) = \sum_{S\in\Gamma} P(S|A)D(S\&A)$.

[18]See (Williams, 2012a, forthcoming). This is the same notion of partition we need to formulate a neat nonclassical analogue of the Inclusion-Exclusion principle in classical probabilism. Note that & won't even be well-defined if the space of truth values for our nonclassical setting isn't closed under multiplication (consider a three valued setting with values $\{1, 0.5, 0\}$. In such a setting, these principles would not make sense as stated. In (Williams, forthcoming) I argue (following Milne, 2007, 2008) that we can still theorize about conditional credence and desirability in the standard ways in those settings, by introducing an 'expanded' set of truth values for the purely instrumental purpose of characterizing conditional credence and desirability of propositions whose truth-values always remain confined to the original (non-instrumental) space of truth values.

[19]For supervaluationism, see (Keefe, 2008). (Keefe, 2000) contains a canonical presentation and defence of what I will call 'standard' supervaluationism.

members of the supervaluational family are very different accounts of indeterminacy. We will illustrate this variation, and in doing so, set the scene for the arguments of parts II and III.

A classical interpretation of a language (relative to a possible world) induces a distinctive truth-status distribution; each sentence is either true or false (and not both) and the usual truth-functional rules for connectives are respected. In the supervaluational family, a language is associated at a world not with one but with several such classical truth-status distributions. We still assign classical truth-statuses to sentences in the language, but only relative to parameters we call *sharpenings*. The 'supertruths' are then those sentences true at every sharpening at the actual world; the 'superfalsehoods' those false at every sharpening at the actual world; the indeterminacies are those true at some sharpenings and false at others at the actual world.

Consider the predicate 'red'. Looking at a red-orange sorites series, there are some clearly red patches, and some clearly non-red patches—and many places in the middle which are borderline. There are many classical interpretations of 'red' that are consistent with all the clear verdicts, and general principles like *anything redder than something that counts as red, should also count as red*. We can think of these rival classical interpretations of 'red' as varying in where they place the cut-off between the red and non-red patches along the red-orange sorites. Relative to the choice of a cut-off one has a definite answer to questions such as 'is patch N red?'. But the trouble is that any such choice seems arbitrary. While classical semantics would say that one among these classical interpretations (and so choices of cut-off) is uniquely correct, the supervaluationist captures the arbitrariness by allowing the cut-off to vary from sharpening to sharpening—the rival classical interpretations are then just the truth-status assignments at rival sharpenings.

While classical logic and semantics allows a scattering of just two statuses (truth, falsity) across sentences like 'Patchy is red' or 'Patchy is red, and some patch indiscriminable in colour from Patchy is not red', supervaluationism allows for a scattering of at least three statuses—supertruth, superfalsity, and supertruth-gap.[20] In the area where cut-offs might be drawn, simple predications such as 'Patch number N is red' will count as red on some sharpenings, and non-red on others, making them supertruth-gaps overall. Indeed, more statuses can be discerned, for we can make sense of gradations of supervaluational indeterminacy: we say a sentence is determinate to degree d iff $d\%$ of sharpenings make it true.[21] Thus, if Patchy is in the middle of the borderline region, perhaps half the sharpenings make 'Patchy is red' true, and half make it false; so it will have 0.5 degree of determinacy. If Patchy were redder, the degree of determinacy would increase; and if Patchy were less red, the degree of determinacy would decrease. Our main concern, later, will be with a more complicated application of the supervaluational framework relevant to personal identity. But the underlying structure will be the same as in this paradigm case.

One of the striking features of the supervaluational setup, however elaborated, is that classical tautologies are always supertrue. After all, they are true on all classical interpretations, so a fortiori are guaranteed to be true on all sharpenings.[22]

[20] A reminder that 'statuses' are being distinguished from 'truth values'. Let *basic statuses* of a sentence be whatever is used in the semantics as the compositional determinate of semantic values of more complex sentences. In the supervaluational setting, a good candidate for basic statuses are functions from sharpenings to truth/falsity. Supertruth, etc are definable from such notions, as we have seen, and so statuses in a secondary sense. As we'll see, they are the statuses which many supervaluationists would use to interpret ordinary talk of truth and falsity.

[21] In a general setting, we would appeal to a normalized *measure* over the sharpenings, and then let the degree of determinacy of a claim be the measure of the sharpenings on which it is true. Exactly how the measure is to be determined is an interesting question—but I will suppose that all supervaluationists can appeal to it.

[22] Although they coincide with classical logic over what is a tautology, it is often claimed that supervaluational systems fail to vindicate classical logic as a whole. (Williamson, 1994) influentially argued that the counterexamples to *reductio*, *reasoning by cases*, *conditional proof* and the like, meant that adopting supervaluationism was in

(I've talked of supervaluationism as a theory of sharpening-relative truth-conditions to sentences. But—to pick up again on a point made earlier—the same pattern could be equally be extended to sharpening-relative truth-conditions to fine-grained *propositions* or Fregean *thoughts*. The loci of this indeterminacy appropriate to our setting will depend on what one's views on the objects of attitudes are, a matter on which I'm not taking a stand.)

Despite this shared basis, supervaluationism is not a single view, but a family. I will differentiate members of the family in the first instance by identifying what each says about truth and falsity (note that above we relied on the neologism 'supertruth', which can be taken as stipulatively defined). I will explain why I think these differ substantively and not merely verbally in a moment.

Standard supervaluationism. (cf Fine, 1975; Keefe, 2000) Truth is supertruth. Falsity is superfalsity. (BIV) fails, as supertruth gaps are truth-value gaps.

Subvaluationism. (cf. Hyde, 1997) Truth is lack of superfalsity. Falsity is lack of supertruth. (BIV) holds; but contradictory claims can be simultaneously true.

Degree supervaluationism. (cf. Kamp, 1975; Lewis, 1970) Truth comes in degrees, and the degree of truth of a claim is its degree of determinacy.

Classical supervaluationism. (cf McGee & McLaughlin, 1994; Barnes, 2007) It is indeterminate which of the sharpenings is the right one; but one of them is. Truth is truth-on-the-correct-sharpening. It can be indeterminate whether or not something is true; the determinate truths are the supertruths and the determinate falsehoods are the superfalsehoods.

If we identify logical consequence with guaranteed truth preservation (or in the case of degree supervaluationism, a guarantee against drops in truth value) then we have correspondingly four different definitions of logical consequence.

I say that the four members of the supervaluational family described above are distinct positions. But as I've presented them, they share an underlying semantic machinery, differing mostly on which constructs are picked out as 'truth' or 'logic'. What difference does this make? I think it is exactly that they indicate differences in the cognitive role associated with the various statuses made available by the bare supervaluational framework. Recall that the crucial issue in characterizing the cognitive role is to settle what the truth-value of each semantic status is to be (glossed as the credence God would have in a proposition with that status). Accordingly, here is my proposal for four distinct ways of supplementing the core supervaluational account:[23]

Standard supervaluationism-TV. $|S|_w = 1$ iff it is supertrue at w; otherwise $|S|_w = 0$.

Subvaluationism-TV. $|S|_w = 1$ iff it is not superfalse at w. If it is superfalse at w, $|S|_w = 0$.

Degree supervaluationism-TV. $|S|_w = d$ iff it is true on proportion d of the sharpenings at w.

the most important sense, logically revisionary. If so, that would be a disappointment for those (like myself) who are attracted to the supervaluational family exactly because it promised to minimize the extent to which we have to reconstruct extant theory on a nonclassical basis. I critically examine Williamson's charge in (Williams, 2008), and argue that all *instances* of reductio et al that we had previously been relying on, are in fact vindicated by the supervaluationist—and so, pace Williamson, supervaluationism is not revisionary of inferential practice.

[23]The focus of Part II is on Degree-supervaluationism, and that of Part III on Subvaluationism. In appendix B, I consider cases which mix the phenomena of these two sections, and give a characterization of the truth values of indeterminate sentences/propositions which generalize both. Such combined positions provide yet further members of the supervaluational family.

Classical supervaluationism-TV. $|S|_w = 1$ iff it is true; otherwise $|S|_w = 0$. If S is indeterminate, it's indeterminate whether $|S|_w$ takes value 1 or 0.

Compare these to the earlier four proposals for which property in the generic supervaluational framework gets to be called 'truth'. I am interpreting the claim that truth is supertruth, as meaning that for one's beliefs to be fully accurate (Godlike), one should have credence 1 in all the supertruths, and no positive credence in anything that lacks supertruth. I suggest that this is a reasonable way of regimenting that slogan, given that truth is something at which belief aims. Likewise, I am interpreting the claim that truth comes in degrees, as saying that to be fully accurate your credence in A should match its degree of truth. Again, if truth is that at which belief aims, and comes in degrees, I think this is a reasonable regimentation of the earlier slogans. Mutatis mutandis for the others. My core claim is that *so interpreted*, we can see that the four slogans we started with embody four substantively distinct ways of elaborating the generic supervaluational framework. (To be clear: I'm not proposing the interpretation as an exegesis of how these terms have previously been understood in the literature, but as a proposal about how we can usefully understand them going forward. I find it useful to take what slogans a given theorist endorses as a defeasible clue about the intended cognitive load associated with a given supervaluational framework. But sometimes such clues point in conflicting directions. Lewis, for example, uses both the classical supervaluational and the degree supervaluational terminology at various points. As I emphasized earlier, we can avoid the need simply laying out the intended nonclassical theory of mind.)

These four options for treating the 'cognitive load of indeterminacy' pick out points in the space of nonclassical theories of mind discussed earlier.[24] The (Belief-Truth) link, for example, will characterize in each case an appropriate range of 'generalized probabilities'.[25] A distinctive nonclassical treatment of conditional belief and desire follows on.

Despite the formal similarities in the way that the respective nonclassical theories of mind are generated, the actual implications can vary dramatically depending on which member of the supervaluational family one buys into. Consider the implications for rational belief. If I know that it's indeterminate whether Patchy is red, the first member of the supervaluational family tells me to utterly reject the proposition that Patchy is red; the second to accept it; the third to give it intermediate credence; the fourth tells me (inter alia) that none of the preceding advice is determinately correct. And this matters for the plausibility of theories. As already noted, future contingents are an area where the plausibility of a nonclassical theory turned on what theory of mind it is paired with. Supervaluational techniques have been used to model the indeterminacy of future contingents (Thomason, 1970). A standard supervaluational interpretation of that formalism commits one to exactly the 'rejectionism' that makes such a theory implausibly revisionary, whereas a degree supervaluational or classical supervaluational (Barnes & Cameron, 2009) interpretation is not subject to the same objection.

[24] Classical supervaluationism, however, just gives us back a purely classical theory of mind. As an interpretation of supervaluationism, it is rather different from the others. I discuss a suitable framework in (Williams, 2014).

[25] As it happens, none of the first three proposals falls under the general result of Paris, quoted in an earlier footnote. Standard and subvaluationism work with truth values 1 and 0, but do not meet the distributional requirements. Degree supervaluationism meets the distributional requirements but involves intermediate truth values. One can check that degree-supervaluationist probabilities vindicate (Additivity). But unlike the classical setting, satisfying the four classical axioms is not sufficient to ensure a given belief state a probability. The standard and subvaluationist can only endorse weakened forms of Additivity. The standard supervaluationist requires the following

$$(\mathcal{DS}) \quad P(\bigvee_{i=1}^{m} A_i) \geq \sum_S (-1)^{|S|-1} P(\bigwedge_{i\in S} A_i)$$

This makes the generalized probabilities 'Dempster-Shafer functions'. See (Williams, 2012b, forthcoming), (Paris, 2001) and compare the discussion in (Field, 2000).

The different implications extend to the treatment of rational desires. Consider the desirability of eating an apple. According to the classical belief-desire connection, this is given by averaging the utility of worlds where one does in fact eat an apple, weighted by how likely one takes each scenario to *be* the one where one eats the apple. Now shift to the nonclassical setting. To what extent does the desirability of eating something that counts as a borderline case of an apple (a pepple, say—half pear, half apple) factor in to your pro-attitude to eating apples? The belief-desire link gives different answers in different cases. For the standard supervaluationist, your views on eating pepples are quite independent from the desirability of eating apples. (The standard supervaluationist has credence zero that a pepple is an apple; hence has zero conditional credence that a pepple-eating-possibility will obtain given that the fruit eaten is an apple. But that conditional credence is the weight given by (Belief-Desire) to pepple-eating possibilities in determining the desirability of eating an apple. The desirability of pepple-eating-possibilties is therefore multiplied by zero when calculating the desirability of the apple, and so the desirability of pepple-eating can vary entirely independently of the desirability of apple-eating. Hence the claimed independence.)

By contrast, for the degree supervaluationist, the desirability of pepple-eating scenarios contributes to the desirability of apple-eating, weighted by the degree to which pepples count as apples. For the subvaluationist, the contribution of desirability of pepple-eating scenarios to the overall desirability of apple-eating is no less than the contribution of the desirability of scenarios where one eats something that is clearly an apple. And for the classical supervaluationist, it's indeterminate whether or not the desirability of pepples does or does not contribute to the overall desirability of apple-eating.

What we have seen is that issues of cognitive load of indeterminacy, and the implications for rational belief and desire, fail to supervene on even a quite richly described semantico-logic skeleton. One can quite happily endorse the background apparatus of sharpenings, supertruth, degrees of determinacy, and subsequently take the theory in very different directions. This isn't just a hypothetical possibility. Edgington (1997) advocates an appeal to 'verities' in analyzing vagueness, thinking of them as 'degrees of closeness to clear truth', with a structure isomorphic to the supervaluational degrees of determinacy just discussed. Her background framework, for present purposes, can be taken to be identical to the one that we'll see Lewis appealing to in Part II. But she explicitly advocates an *independence* claim—the one associated with standard supervaluationism above—as the correct treatment of vague desire.[26] As we'll see in the next section, Lewis himself is committed to a different role for indeterminacy. The availability of these alternative nonclassical theories of mind illustrates that the argument that is to come, identifying reasons to endorse one particular member of the family, targets a substantive conclusion.

Recap of Part I

This first part of the paper has argued that nonclassicists have a distinctive obligation to tell us about their theory of mind. Theories sharing the same semantico-logical resources (like the members of the supervaluational family) can have radically different commitments on this front.

But I have also described a systematic space of theories of mind that nonclassicists can take 'off the shelf' to flesh out their views. These are nonclassical generalizations over a certain kind of probabilistic theory of mind familiar from the classical case, and include stories about the rational belief states (generalized probabilism), rational updating (generalized conditionalization)

[26]The apples/pepples example is Edgington's. As another example: in (Williams, 2014) I develop an account of the cognitive role of indeterminacy that makes heavy use of supervaluational degrees of determinacy, but which does not take the form ascribed here—instead of corresponding to ideal degrees of belief, degrees of determinacy of p fix the weight of a disposition to flat-out *believe* that p.

and rational desires and choice (generalized decision theory). This gives in each case a candidate nonclassical theory of mind to work with, and a starting point for future development—and provides a prototype against which rival approaches to the nonclassical theories of mind can be compared and contrasted, and a basis for further extensions.

But describing a space of possible theories is one thing; it is another to locate ourselves within that space, to give arguments for one nonclassical theory of mind against another. It is the burden of Part II to exhibit one such argument.

II Indeterminate survival: by degrees

What kind of cognitive role does indeterminacy have? What nonclassical theory of mind is right? The task of Part II is to give an argument for one particular framework (at least in application to a specific kind of indeterminacy): the degree supervaluational setting of section 1.3.

The argument will take the following form. I will present a familiar puzzle from the literature on personal identity. I will endorse (at least for the sake of argument) one particular way of resolving that puzzle—also familiar. But I will argue that this solution is (implicitly) committed to a certain thesis in the nonclassical theory of mind—a thesis concerning the interaction of desire and indeterminacy. And I argue that (on principles that are invariant across the space of theories identified in Part I) this commits one to a distinctive cognitive role for indeterminacy, and indeed, exactly the degree supervaluational setting.

It's important to bear in mind the ambitions of this part of the paper: what is being assumed, and what argued for. The central assumptions are squarely about the value of personal identity. I will assume, following Parfit, that degrees of psychological connectedness matter in survival. And I will also assume, against Parfit, that this is truly a *self* interested point of view—that identity matters in survival. Note that we haven't mentioned indeterminacy at all here; and so it's interesting and somewhat surprising that we can argue for a particular nonclassical theory of mind on this basis. The connection is forged principally by Lewis's metaphysics of personal identity, as elaborated in response to Parfit. However, Lewis never addresses the cognitive role of indeterminacy directly. I'll explain why not, and pinpoint the progress made here, in wrapping up Part II. The same assumptions will be in play in Part III, but with the focus on Fission and Longevity puzzles rather than puzzles of degrees of connection.

II.1 Parfittian therapy

In *Material Beings*, van Inwagen (1990) reminds us that indeterminacy can infect our own continued existence, and sets up a particular example that we will use in what follows:

> If, at the extremes of a spectrum along the length of which are arranged more and more radical disruptions of lives, we can find definite cases of the end of a life and definite cases of the continuation of a life, then it seems reasonable to suppose that somewhere between the extremes will be found disruptions of which it is not definitely true or definitely false that they constitute the end of a life. And if this is so, then there are possible adventures of which it is not definitely true or definitely false that one would survive them. Let us call such episodes 'indeterminate adventures'.
>
> . . . Suppose that a person, Alpha, enters a certain infernal philosophical engine called the Cabinet. Suppose that a person later emerges from the Cabinet and we immediately name him 'Omega'. *Is* Alpha Omega? . . . Let us suppose the dials

> on the Cabinet have been set to provide its inmates with indeterminate adventures. (We need not agree on what would constitute an indeterminate adventure to suppose this. Let each philosopher fill in for himself the part of the story that tells how the dials are set). Alpha has entered and Omega has left. It is, therefore, not definitely true or definitely false that Alpha is Omega. (van Inwagen, 1990, p.243-4)

Suppose a psychological theory of personal identity is correct. After encabination, it is indeterminate whether Alpha is Omega because, although there exist psychological connections between them, it is indeterminate whether the connections are sufficiently strong and extensive.

Our next question is: how should Alpha think and feel about Omega's prospects? Is this as bad as death, or as good as ordinary survival, or something in between? Now, Alpha might independently care about Omega—she might love all humanity equally. But set that aside. We're interested in how she should feel from a selfish point of view; and so we take her to be entirely motivated by self-interest. Our question is how that distinctive self-concern should be translated into this setting.

To make the issue concrete, suppose a broker offers her the following deal: pay 100 dollars now, and gain 10,000 later. Future riches would be nice! But it's indeterminate whether Alpha will receive them. By paying the money right now, she will have to give up on a planned party with friends before stepping into the cabinet. Everybody is in a position to describe the prospects of accepting or rejecting the offer on the assumption that Alpha is Omega, or on the contrary supposition:

Broker's offer	$\alpha = \omega$	$\alpha \neq \omega$
Invest	Long life, No party, Riches	Short life, No party, No riches
Reject	Long life, Party, No riches	Short life, Party, No riches

We can assign utilities to each such outcome (cell) in the table in a way that reflects Alpha's self-interest. Let's suppose that the life up to the point of encabination on its own contributes -100 utils; the extra years come with a boost of $+100$ utils on top of this. We'll assume that having riches comes with a boost of $+100$ utils, and a party with a boost of $+10$. Lacking riches and lacking a party don't add or subtract any utils to the life. We're then in a position to figure out the net utility for each cell in the table above. The top left cell includes long life, i.e., a life up to encabination plus extra years ($-100 + 100$) and riches ($+100$), which will give a net utility of $+100$. The bottom right cell includes a short life (-100) and a party ($+10$), and thus has a utility of -90. Similar calculations for the remaining two cells give us:

Broker's offer	$\alpha = \omega$	$\alpha \neq \omega$
Invest	+100	-100
Reject	+10	-90

However, neither $\alpha = \omega$ nor $\alpha \neq \omega$ is determinately correct, and Alpha knows this. Nor do the columns of utilities produced by the two assumptions converge on a single piece of advice—taken individually they point in opposite directions. So no obvious way of making decisions under indeterminacy emerges.

Parfit (1971, 1984) gave a recipe for answering such questions that many find attractive. What matters in survival, he urged, is the underlying facts about psychological continuity and connectedness. Since these are present to a reduced degree between Alpha and Omega, Alpha's self-concern (the intrinsic desire that things go well for *Alpha*) will translate to a scaled concern for Omega (a less strong desire that things good things happen to Omega after she emerges from the cabinet). If we follow Parfit here, we simply set aside the columns of the previous table, which were framed in terms of personal identity, and look to the underlying facts. We reframe the decision table as follows:

Indeterminate cabinet	α is psych-connected to ω to degree 0.7
Invest	+40
Reject	-20

The valuations of the outcomes are determined by discounting good or bad things that happen to people in proportion to their degree of psychological disconnection from Alpha. Thus, the long life of Omega makes the prospects far better than cases of total psychological disconnection (where the second column earlier would be appropriate), and riches are desirable—but slightly less so than in a putative case of full connection (represented by the first column earlier). This time the recommendation is clear: invest.

Henceforth I will suppose with Parfit that this recipe assigns the correct utilities to investing and rejecting the Broker's offer. This immediately raises a further issue—Alpha is supposed to be a *self*-interested agent. But that motivates the original identity-based description of the situation, rather than the Parfittian one. Parfit, of course, recommends that we simply *abandon* framing these issues in terms of identity. What matters in survival, he urges, is psychological connectedness. And insofar as an intrinsic concern for identity induces a different psychological profile of concern, that is simply evidence that personal identity is not what matters in survival. Indeed, the situation we have been dealing with is an instance of one of his central arguments to this effect: that psychological connectedness comes in degrees, but personal identity does not.

But I will be arguing in the opposite direction. I endorse the claim that personal identity *is* what matters in survival, by stipulation in Alpha's case, and as a matter of fact, for the rest of us. And that leaves us with a *Reconciliation* challenge, under our assumption that Parfit nevertheless provides accurate advice. How can a basic desire for one's own future welfare, translate into the kind of evaluation of the scenario in the table above? As the case is, *ex hypothesi*, one where it is indeterminate whether you yourself benefit, this reconciliation will involve taking a stance on the interaction of indeterminacy and desire—which is where it makes contact with broader issues in the nonclassical theory of mind.

In the next section, I review David Lewis's attempt to pursue the Reconciliation project, by endorsing a particular view about the metaphysics of persons, that allows us to bring to bear the supervaluational treatment of indeterminacy described in 1.3 above as a model of the indeterminacy in the case. I will argue in the subsequent section, however, that this metaphysics alone leaves Lewis's Reconciliation project incomplete, exactly because he does not pin down which member of the supervaluational family he is endorsing. Reconciliation of our commitments—to both psychological connectedness and personal identity mattering—requires not just an appropriate metaphysics but also a distinctive nonclassical theory of mind.

II.2 Lewis against Parfit

Parfit famously argued that psychological connectedness and personal identity had to be distinguished because they had different formal properties. Identity is transitive, connectedness is not; connectedness comes in degrees; identity does not. Lewis's (1976) core strategy is to define a certain relation—the *I*-relation—in terms of his favoured theory of what it is for a person to persist through time. Two person-stages are *I*-related iff they are both stages of a single four-dimensional person. He then goes on to argue that the *I*-relation can have the logical properties (degrees, non-transitivity) that Parfit takes as characteristic of psychological connectedness.[27] And then—central to our concerns—Lewis argues that there is a sense in which the *I*-relation comes in degrees, and moreover, degrees that exactly match the degrees of psychological connectedness.

(Some more housekeeping. Lewis emphasizes that the *I*-relation is not itself an *identity* relation among persons. For one thing, it has different relata—it is not a relation among persons, but among *stages* of persons. Lewis' view, apparently, was that we should understand 'personal identity' loosely—as picking out the *I*-relation—when interpreting "personal identity is what matters". It's not immediately clear he has to be this concessive though. Suppose that no stage is part of two distinct persons. Let S be a descriptive name for the person who has s as a part; and T be a descriptive name for the person who has t as a part. Then saying that the welfare of stage t matters to s if and only if $S = T$, is straightforwardly equivalent to saying that t matters to s if and only if t and s are *I*-related. So (at least when there is no stage-sharing) we could reformulate Lewis's discussion in terms of informative strict identities between persons. The puzzle of *degree* that is our focus here doesn't involve stage-sharing, and so in the discussion to follow in Part II 'personal identity' and 'same person as' can be understood in either the loose or the strict sense (I'll flag up any subtleties as they arise). How cases where there *is* stage-sharing (e.g. fission and longevity) relate to our current discussion is something to which I return in part III.[28])

Let's now review how Lewis introduces degrees—and indeterminacy—into personal identity. The background metaphysics identifies persisting objects like persons as four-dimensional objects, fusions of appropriately-related three dimensional *person-stages* (such as the three-dimensional human-shaped object sitting in my chair right now). The person-stages existing at various times fuse to make a single person if they are psychologically interrelated in the appropriate way. But the strength and extent of psychological connections is a matter of degree. Lewis gives his account of personal identity relative to an arbitrary choice of 'boundary number'—we will say that two person stages are *R*-related if the degree of psychological connectedness that obtains between them is greater than that threshold. Persons, for Lewis, are maximal *R*-interrelated fusions. In Figure 1, we represented person stages with circles, and the obtaining of *R*-relations with a line joining them.

Relative to a particular choice of boundary number, we have answers to questions such as 'is the person-stage sitting in this chair part of the same person as the infant thirty-two years ago in location L?'. Perhaps the psychological connections between the person-stage in the chair and the infant are of intermediate degree (memory links have decayed, etc). But given the boundary-number, we can return a yes/no answer to the question. The identification

[27] This happens, recall, when two stages are parts of two distinct persons, who share a stage.

[28] Cases of stage-sharing are the focus of Parfit's (1976) rejoinder to Lewis (though interestingly, Parfit describes the degreed cases as more important, since they arise all the time rather than in recherché reaches of science fiction, and so can't be dismissed as irrelevant to practical concerns). In cases of stage-sharing, there will be no *single* individual who is *the* person containing a shared stage s, to be picked out by a descriptive name S. The idea in Part III is effectively that in such cases, the descriptive name S will be indeterminate in reference between those persons who share stage s.

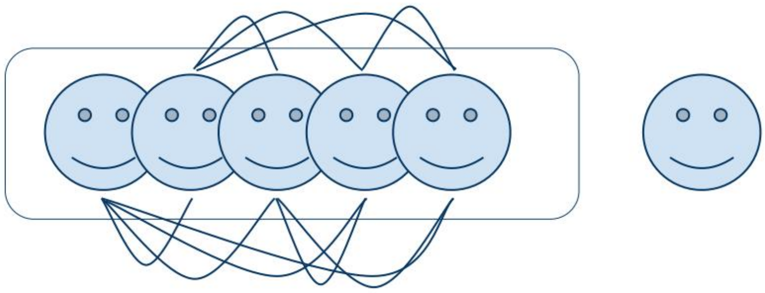

Figure 1: Persons are maximally R-interrelated fusions. The fusion of stages inside the box is the person, since (i) each stage is *R*-related to every other within the box; (ii) there is no larger fusion of stages of which this is true.

of an exact boundary-number would be as unprincipled as the choice of a cut-off in the red-orange sorites, however—and Lewis offers the parallel diagnosis. Whereas the classicist would have it that there is a single privileged boundary number—the one that generates the One True extension of the *same person as* relation—supervaluationists like Lewis allow for a whole range of admissible boundary-numbers, corresponding to different sharpenings. Indeterminacy arises when stages are psychologically connected to intermediate degrees, higher than some boundary numbers and lower than others. This is schematically illustrated in Figure 2, where solid lines represent a high degree of psychological connectedness and dotted lines represent a relatively low degree of psychological connectedness.

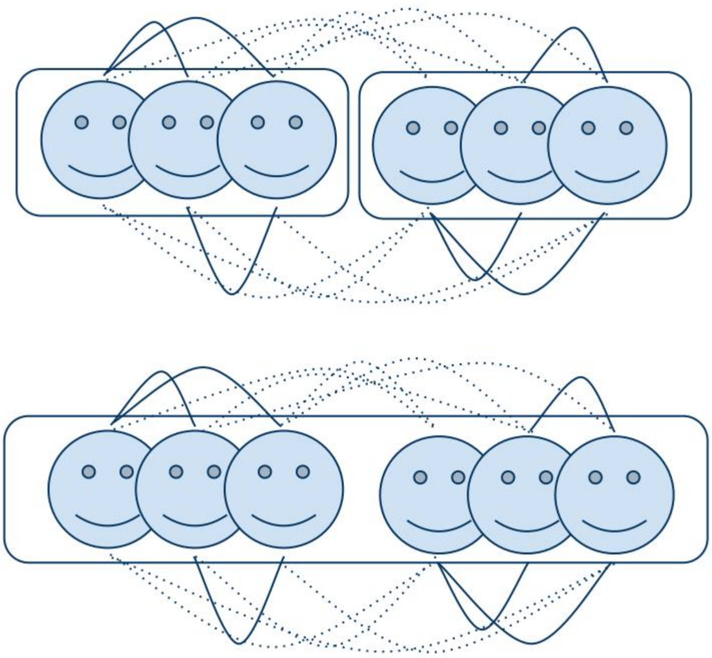

Figure 2: The dotted lines indicate psychological connectedness of borderline degree. If the criterion for R-relatedness is set high (top diagram), then we have two distinct persons—and so no person survives whatever incident took place as time flows left-to-right. If the criterion for R-relatedness is set low (bottom diagram), then we have a single persisting person. It's indeterminate how high the threshold should be, so indeterminate whether we have one person or two. Hence, from the perspective of the leftmost person stage, it is indeterminate whether or not he will survive the upcoming incident.

The standard conceptual repertoire of the supervaluational family can now be brought into play. Thus, all supervaluationists would agree on certain descriptions of figure 2 (assuming that they accept the underlying metaphysics of personal identity). For simplicity, suppose that the two sharpenings shown above are the only ones, and let us suppose that the person-stages illustrated from left-right exist on Monday through Saturday respectively. It will then be supertrue that the Monday stage is part of the same person as the Tuesday stage (and so, that the person on Monday survives till Tuesday at least). But it will be neither supertrue nor superfalse that the Monday stage is part of the same person as the Saturday stage (and so indeterminate whether the person on Monday survives till Saturday). Since it is true on exactly half the sharpenings, its degree of determinacy is 0.5.

It is the last point that Lewis wants to emphasize in response to Parfit. What this comes to is that for Lewis the *I*-relation, i.e. *x being part of the same person as y*,[29] will be indeterminate. On his favoured supervaluational treatment of indeterminacy, the determinacy of the obtaining of the *I*-relation will come in degrees. The degrees of indeterminacy of the *I*-relation exactly coincide with degrees of psychological connectedness (appendix A gives Lewis' argument for this claim).

II.3 Articulating what matters

Lewis has outlined one sense in which personal identity—understood as the *I*-relation—can 'obtain' to intermediate degrees, and moreover, degrees which perfectly coincide with Parfit's degrees of psychological connectedness. What goes for the *I*-relation also goes for the strict identities between the persons who include the various stages—like Alpha (the person that includes the pre-cabinet stages) and Omega (the person that includes post-cabinet stages). Such identities, too, can be indeterminate to various degrees, and will match the degree of determinacy of the *I* and *R* relations. Appendix B sets out the case for this.[30]

However, that the personal identity relation (however understood) comes in degrees is *not* sufficient to respond to the Parfittian worries I've put forward. What Reconciliation requires is that psychological connectedness and personal identity both matter in survival—that they induce the same pattern of desire over possible outcomes. Lewis's notion of degrees of truth/determinacy is common-ground in the supervaluational family. But the members of this family attribute radically different cognitive roles to it, as we saw in section 1.3. It makes perfect sense to combine the formal skeleton Lewis provides with an account of the cognitive and conative role of indeterminacy that predicts an unParfittian pattern of care given indeterminacy in the *I*-relation. Indeed, people in the literature who work with Lewis-style 'degrees of determinacy' have taken the theory in those directions—Edgington's treatment of 'verities' being one example.

Lewis's metaphysics of personal identity therefore leaves open the question of whether intrinsic concern for personal identity and intrinsic concern for psychological connectedness are compatible. To say that one desires that things go well for *oneself*, simply does not immediately entail *any* result about what one desires in cases where it's indeterminate who one is. In particular, it carries no predictions, yet, for what to say about "degreed" cases like Alpha's. To make it predictive, we need to add assumptions about the conative role of indeterminacy. Specifically, we need:

> (Reconciliation) Personal identity mattering requires that one cares about those individuals who are determinate-to-degree-*d* the same person as oneself, to a reduced

[29] or: the *concept* of *same person as*, or the *predicate* 'same person as'—-how exactly one should regiment this depends on one's preferred views on the loci of indeterminacy. See the beginning of section 1.3.

[30] Evans (1978) famously argued that vague identities were impossible. But as is now familiar, his argument only works on the assumption that the terms flanking the identity be determinate in reference (Lewis, 1988); that will fail in the current case.

but non-zero extent; the scaling factor being d.

(Again, we can read 'same person as' loosely as saying that either that the welfare of t should matter from s's perspective, to an extent scaled by d when it is determinate to degree d that the I-relation holds between s and t; or we can read it 'strictly' as saying the scaling factor is d when it is determinate to degree d that $S = T$, where 'S' is a descriptive name for the person including stage s, and 'T' a descriptive name for the person including stage t.)

In order to draw out the implications of (Reconciliation), it is useful to regiment the situation formally. Suppose we have a rule for assigning values to person-stages in terms of their situation at that very moment—intuitively, this measures how intrinsically desirable being in that particular situation is (that's what I've been calling above a stage's 'welfare'). Set aside potential indeterminacy for a moment. Let the intrinsic value of person-stage y be v^y. A given possible world—w—fixes both the values attaching to each person-stage, and which person-stages are parts of the same person as the stage x whose beliefs and desires are in question. Let the set of such person-stages be A_x. The model has it that the 'self-interested' value of the outcome w for x will be:

Identity: $v(w) = \sum_{y \in A_x} v^y$

We can then use these values over outcomes, as a recipe for determining their *utility* for Alpha. The 'belief-desire links' discussed in Part I of this paper, together with x's credences over the results of acting a particular way, will fix the utility (desirability) of other propositions. (Identity) is a model of self-interested desire which exactly builds in the assumption that personal identity is what matters.[31]

Parfit offers an alternative recipe for fixing the utility of outcomes. We can set this out in a similar fashion. We need to make room for scaling the values attaching to person stages. The adaption is simple: we allow y to range over any person stage whatever, but weight the contribution by a scaling factor, the degree of psychological connection between x and y:

Connectedness: $v(w) = \sum_y \delta_x^y v^y$

Again, if we fix the utilities of outcomes w by their values so characterized, the belief-desire link then allows us to extrapolate utilities for arbitrary propositions. (Connectedness) of course, makes explicit the idea that degrees of psychological connectedness are what matter in survival.[32]

[31]This model of intertemporal utility already builds in some very strong assumptions—but such summative models of inter-temporal utility are well-known in the economics literature (Mas-Colell *et al.* , 1995, ch.x)., so it's a well-motivated starting point. Typically, economic models of inter-temporal utility will include intertemporal *discounting*, so that values of the 'commodity bundles' that your future stages possess contribute less and less as they get further and further into the future. But we'll assume for now that no such 'pure temporal' discounting goes on.

As (Broome, 1994) notes, temporal discounting in a model of intertemporal utility can often be viewed as a fudge, representing factors that aren't explicit in a model. For example, some goods may be worth more in the near future because they could in principle be invested to produce more of the same kind of goods in the future (as an example of 'natural interest rates', think of the wood in a tree this year, vs. the wood from the same tree in ten years time). Likewise, if there's a small chance that you get wiped out every year that isn't represented in the model, the cumulative chance that you'll never get to occupy a given person-stage position will increase over time. 'Pure time-preference' which legitimates care about the near future over the distant future *purely* on the grounds of their temporal situations rather than via one of the rationales just suggested is itself a controversial topic. Indeed, in a recent review article, Parfit's work is cited as one of the few available rationalizations of pure time-preference—see (Frederick *et al.* , 2002).

[32]I assume that x must exist in world w (or have a counterpart there, if individuals are worldbound) in order for w to have self-interested utility for x. δ_x^y is intended to record the degree of psychological connectedness *in w* between x and y.

(Identity) and (Connectedness) are what generate the decision tables that we laid out in an earlier section. With appropriate assumptions about the values of stages, (Identity) generates the assignment of utility to outcomes in the two columned table; and (Connectedness) the assignment in the one-columned Parfittian table.[33]

How should we think of Lewis's intervention in the debate? Here is the tempting but *wrong* view: to view him as proposing to *replace* (Identity) with the following, where I_x^y denotes the degree to which it is determinate that y and x are stages of a single person:

Identity*: $v(w) = \sum_y I_x^y v^y$

I urged earlier that this can't be right. To be sure, (Identity*) reflects a concern for the interests of those *for whom it is determinate to a given degree that they are the same person as Alpha*. But we are not entitled to take this as an articulation of the claim that 'identity is what matters', for its relationship to Alpha caring for *her own* future interests is still obscure. That is so whether we understand 'same person as' in the loose or the strict ways we've been discussing; on either reading, it's a binary relation we're supposed to be caring about, not the degree-relative relation that (Identity*) features.

(Identity), as a substantive claim about the value attached to a world, is our formalization of the thesis that 'personal identity is what matters'. (Connectedness), as a substantive thesis, is our formalization of the thesis that 'psychological connectedness is what matters'. Reconciliation will be achieved if we show that they are equivalent—that they assign the same utility to each proposition. We have seen that Lewis's metaphysics of personal identity ensures that that (Connectedness) is equivalent to (Identity*). It is necessary and sufficient for Reconciliation, therefore, that (Identity) and (Identity*) be equivalent. Since (Identity) talks of the I-relation, and (Identity*) talks of degrees of determinacy of the I-relation, this is a claim about the interaction of indeterminacy with desirability—a thesis in the nonclassical theory of mind.

II.4 From conative to cognitive role

We now draw out the consequences of the Reconciliation thesis (a thesis about the interaction of *desire* and indeterminacy) for the cognitive role of indeterminacy (a thesis about the interaction of rational *belief* and indeterminacy). Consider a situation where Alpha is deciding whether to invest in the broker's offer, and so wishes to evaluate the desirability of this prospect. Alpha is certain that the person who emerges from the cabinet will get the goods, conditionally on her investing. And he's certain that person is Omega, and is certain that it's indeterminate that that person is him.

The first way of calculating the desirability of investing is by appeal to (Connectedness). Suppose the value of the life-segment up to the point of encabination (without a party) is s, and subsequent life of ω (with riches) are worth g to whoever gets them. Then the value to Alpha of this prospect by the Parfittian lights (i.e. by the lights of (Identity*) or (Connectedness)) is $s + I_\alpha^\omega g$.

[33] Here are the assumptions we need. First, the value of investing. Here one assumes that s, the value of the *shorter life* sans party is -100. In the identity-based decision table, that directly vindicates the utility entered in the investing row and $\alpha \neq \omega$ column. Further, assume that the value g of the post-encabination stages is +200. In the identity-based decision table, this vindicates the entry +100 in the investing row under the $\alpha = \omega$ column. In the Parfittian decision table, we are scaling the utility of the future stage by 0.7 for the degree of psychological connectedness stipulated. This gives $-100 + 140 = 40$. And similarly for the Rejecting row. Assume that the value of the shorter life s here is -90: this is the -100 from before, but with a small boost for the extra partying. The value of the post-encabination stages g here is +100, half as good as before, since there are no riches. That delivers the right results in the identity-based decision table; and scaling this again by 0.7 in the Parfittian case, we have $-90 + 70 = -20$.

On the other hand, Alpha may calculate the desirability indirectly. He first appeals to (Identity) to argue that the intrinsic value for Alpha of bringing about the following outcome is simply $s+g$:

> the post-Cabinet person gets the goods, and that person is Alpha

Further, the intrinsic value of the following outcome is just s, since it entails inter alia that Alpha has only the short life without party:

> the post-Cabinet person gets the goods, and that person is not Alpha

Let G be the proposition that the post-Cabinet person gets the goods, $\alpha = \omega$ be the proposition that this person is Alpha, and $\alpha \neq \omega$ be the proposition that it is not Alpha. With U giving the utility of a proposition, C credence in a proposition, and v the values assigned to worlds determined by (Identity), what we have argued is that $U(\alpha = \omega \& G) = v(\alpha = \omega \& G) = s + g$ and $U(\alpha \neq \omega \& G) = v(\alpha \neq \omega \& G) = s$.

These values make it possible to apply the *belief-desire link* to calculate $U(G)$, as follows:

$$
\begin{array}{llll}
1. & U(G) = & U(\alpha = \omega \& G)C(\alpha = \omega|G) + U(\alpha \neq \omega \& G)C(\alpha \neq \omega|G) & \text{(Belief-desire link)} \\
2. & = & (s+g)C(\alpha = \omega|G) + sC(\alpha \neq \omega|G) & \text{(From (Identity))} \\
3. & = & (s+g)C(\alpha = \omega) + sC(\alpha \neq \omega) & \text{(Epistemic Independence)} \\
4. & = & s(C(\alpha = \omega) + C(\alpha \neq \omega)) + gC(\alpha = \omega) & \text{(Rearranging)} \\
5. & = & s + gC(\alpha = \omega) & \text{(Fact about probabilities)}
\end{array}
$$

The first line is just the belief-desire link. The second arises from our discussion above of the utility of the two conjunctions involved. The third relies on the plausible assumption that whether or not Omega gets the goods is evidentially independent of whether or not Alpha is Omega. The fourth line is simple algebra. The final line appeals to a fact about probability—that the probabilities of a sentence and its negation sum to 1. (I will return to the justification of lines (1) and (5) in the next section.)

But we already had from the earlier Parfittian calculation that $U(G) = s + I^{\omega}_{\alpha} g$. Equating the expansions of $U(G)$, we get:

$$s + I^{\omega}_{\alpha} g = s + gC(\alpha = \omega)$$

And if $g \neq 0$, this entails that:

$$C(\alpha = \omega) = I^{\omega}_{\alpha} \qquad (\dagger)$$

That is: the appropriate credence to have in the proposition that Alpha is Omega, matches the degree of determinacy (truth) of that proposition.

Let's step back to see what is happening here. (Identity) and (Connectedness) in effect fix the value (and hence utility) of different *ranges* or *partitions* of propositions, at different levels of grain. (Connectedness) fixes the value of 'outcomes' that are specified in terms of psychological relations between stages of creatures over time. But since facts about personal identity are not settled by such descriptions, it is possible to 'fine grain' or 'sharpen' any proposition describing such an outcome by conjoining with it propositions about who is identical to who. It is at this sharpened level that (Identity) is directly applicable.[34] The two recipes do not conflict directly,

[34]Note that if we thought of 'possible worlds' as corresponding to something like Lewisian space-times, e.g. supervening on facts about particle positions, then it is (Connectedness)/(Identity*) that is most naturally taken to

because they concern different propositions. But they may still implicitly conflict, since they fix (via a belief-desire link) a distribution of utility over all propositions. It becomes a substantive question under what conditions the two recipes can be reconciled. We have seen that the answer is: only if indeterminacy has a specific cognitive role expressed in (†) above.

Starting from assumptions about the *conative* role of indeterminacy, we've argued for a specific *cognitive load* of indeterminacy—the psychological state we should adopt to a proposition (viz. Alpha is Omega) under full information about the indeterminacy-status of that proposition. Amongst the supervaluational family, this unambiguously picks out degree supervaluationalism as the nonclassical theory of mind required to achieve Lewis's reconciliatory goals. To see this, suppose that we have a situation in which it is neither supertrue, nor superfalse, that Alpha is Omega, this proposition being true on some sharpenings and false on others (we can suppose it is true on 60% of them). Then the standard supervaluationist would say that one should have credence zero in Alpha being Omega (regarding it as not true, albeit also not false); the subvaluationist would say one should have credence one in it (regarding it both true and false). The classical supervaluationist thinks it indeterminate whether the proposition is true or false, hence claiming it is indeterminate what credence it is appropriate to invest. But in the situation mentioned $I^{\omega}_{\alpha} = 0.6$, hence none of the above options can be combined with (†), which tells us that in this case the appropriate credence to have in Alpha being Omega is 0.6 rather than any of the options hitherto mentioned. On the other hand, degree-supervaluationism tells us to match our credence in Alpha being Omega to that proposition's degree of determinacy, which exactly matches the degree of determinacy of the I-relation, I^{ω}_{α}, just as (†) requires.

Further confirmation can be gathered by generalizing the above argument: we can directly derive the result that credence in a proposition should be the expectation of its degree of determinacy/truth—essentially the (Belief-Truth) thesis characteristic of degree supervaluationism. Suppose that Alpha is uncertain over the exact degree of determinacy of his being Omega. We can represent this as his being uncertain which of the worlds $w_1, \ldots, w_n$ will be realized, with respective credences $\lambda_1, \ldots, \lambda_n$ summing to 1, where the respective degree of determinacy of the obtaining of the I-relation between Alpha and Omega in each is $I_1, \ldots, I_n$.

We can rerun the (Identity)-based argument just as before, with the same conclusion: $U(G) = s + gC(\alpha = \omega)$.

We cannot however simply identify the utility of G by the Parfittian desirability of a specific outcome, since now the agent is uncertain what that outcome would be. But we can use the Belief-desire link to fix this. By (Connectedness)/(Identity*), the utility of $w_i \& G$ is $(s + I_i g)$. The utility of the *uncertain* prospect G is therefore:

1′.	$U(G)$	$=$	$\sum_i U(w_i \& G) C(w_i \mid G)$	(Belief-desire link)
2′.		$=$	$\sum_i (s + I_i g) C(w_i \mid G)$	(Utility of $w_i \& G$)
3′.		$=$	$\sum_i (s + I_i g) \lambda_i$	(Epistemic Independence, $C(w_i) = \lambda_i$)
4′.		$=$	$s \sum_i \lambda_i + \sum_i I_i g \lambda_i$	(Rearranging)
5′.		$=$	$s + g \sum_i I_i \lambda_i$	(λ_i sum to 1)

This replicates the form of the argument above, but with a partition of worlds w_i rather than $A, \neg A$. Equating the two characterizations of $U(G)$, we derive

$$C(\alpha = \omega) = \sum_j \lambda_j I_j.$$

assign utilities to worlds. But where we have utility, there are *propositions* around that carve things up *more finely than any possible world does*. These are still objects of attitudes, and hence (Identity) can coherently describe the value of propositions at this finer level of grain.

That is, the credence appropriate to 'Alpha is Omega' when there is uncertainty over what degree of truth attaches to this proposition, is the *expectation* of the degree of truth of the proposition.[35]

II.5 An objection and a response

The central argument of the previous section purported to derive a commitment to the way that indeterminacy interacts with belief from a thesis about the interaction of indeterminacy and desire. The argument would be less interesting if we assumed along the way *substantive* theses about the cognitive role of indeterminacy.

The argument does (of course!) make assumptions about how beliefs and desires interact under indeterminacy: the belief-desire link, in step (1). It also requires one assumption pertaining to rational beliefs alone, in step (5). Now, a version of the belief-desire link of step 1 is endorsed by *every* nonclassical decision theory of the kind described in Part I. However, there is a significant assumption hidden in our appeal to it. The generalized formulation I mentioned earlier supports the version used in (1) only if A and $\neg A$ form a 'partition'—in the specific sense that (necessarily) their truth values sum to 1. This is so in many nonclassical settings, but not all. The other assumption we needed is that the probability of A and that of $\neg A$ sum to 1. A sufficient condition for this to be the case, in the space of nonclassical theories of mind described earlier, is again that A and $\neg A$ form a partition.

It might seem, therefore, that our argument for the conclusion that the 'cognitive loading' of an indeterminate proposition matches its degree of determinacy, requires a substantive presupposition about how the cognitive loads (truth values) shape up. What makes this *dialectically* worrying is that two *rival* stories about the cognitive loads of indeterminacy in a broadly supervaluational setting are among those that do *not* verify the partition assumption.[36]

What we need to address this is a slightly stronger deployment of our original assumptions. Think about what *would* make the argument go through *without* extra assumptions about how negation behaves. Well, suppose that we have a proposition $N(\alpha = \omega)$ such that $\{\alpha = \omega, N(\alpha = \omega)\}$ form a partition. Then if we substitute $N(\alpha = \omega)$ for $\alpha \neq \omega$ in the above argument, it becomes valid.[37]

What of the premises of this revised argument? We need to assume that the utility of what follows is simply s, i.e. the utility of a short life, without riches or party:

> the post-Cabinet person gets the goods, and N(that person is Alpha)

But I argue we are entitled to this by the original description of the case. Recall the initial motivation: that Alpha is to be entirely selfish or self-interested. That is, she *only* attributes value to states if he himself benefits. An outcome whose canonical description is *inconsistent* with Alpha benefitting should not receive any of Alpha's 'self-interested' utility. It is *because*

[35] Here I am in step with how Nick Smith (2008, 2010) recommends we think of degrees of truth and credence as relating—though I am working within a degree-supervaluational rather than fuzzy treatment of degrees of truth. Not every degree theorist would be happy with this idea of truth-values as the aim of credence—see MacFarlane (2010) for a very different proposal. Among supervaluationists who make prominent appeals to degrees of determinacy that I'm aware of, Lewis and Kamp (1975) say nothing about this issue, and as already mentioned, Dorothy Edgington is committed to a rival view. The idea of degrees of truth as aims of credence is explicit in the discussions of degree supervaluationism in Williams (2011, 2012a,b).

[36] These are the subvaluational and supervaluational settings. I think we can argue independently that the subvaluational pattern won't deliver the right results; but the case of supervaluationalism is more worrying, and requires the sort of considerations given in the text.

[37] Indeed, there is an easy extension of the argument where all we need would be *some partition or other* that includes $\alpha = \omega$ as one cell.

we standardly assume that *the post-Cabinet person is not Alpha* is inconsistent with *the post-Cabinet person is Alpha*, that we take outcomes meeting the former description not to be assigned any utility on account of the post-Cabinet person getting the goods.[38] But exactly the same rationale can be given for denying any self-interested utility (at least on account of the post-Cabinet person getting the goods) to outcomes where N(the post-cabinet person is Alpha) obtains, since by construction that is inconsistent with *Alpha* getting the goods.

Note that this revised version of the argument does not assume that there is a generally applicable operator N, such that for any p, $\{p, Np\}$ form a partition (i.e. an operator N that acts as an exclusion-negation). That would certainly *suffice* for my purposes, but it is not required. What the generalized argument relies on is the existence of some partition or other, not any general recipe for constructing them. Still, that leaves a loophole for a nonclassicist who is prepared to deny that any such partition can be built around $\alpha = \omega$. But it's important to remember that in order to be relevant to the current dialectic, they have to offer this *in the context of Lewis's reconciliation of self-interest and Parfittian concern*, for my argument has been that *in order to achieve this*, he is committed to treating indeterminacy in the degree-supervaluationist way. The salient rivals to the degree-supervaluational treatment that I discussed in the first section all provide a suitable $N(\alpha = \omega)$, and I'm not aware of any *dialectically relevant* ways of exploiting the remaining loophole.[39]

Recap of Part II

The discussion in Part I gave us an abstract characterization of a whole space of nonclassical theories of mind. Even narrowing things down to members of the supervaluational family leaves open radically different conceptions of indeterminacy. The abstract logico-semantical framework appropriate to indeterminacy is neutral in important respects (in respects I argued were absolutely crucial if we're even to know *what it is* to accept that theory of indeterminacy).

Part II has shown how this lacuna can be filled. The materials we need are two substantive theses about intrinsic desire—the thesis that identity matters in survival (Identity), and that psychological connectedness does too (Connectedness). From this, ultimately, can we argue for a specific degree-supervaluational nonclassical theory of mind.

At the start of part II, I emphasized that it's a *prima facie* surprise that these first-order theses about personal identity result in such a conclusion. One might feel, however, that it would be less news to those already familiar with Lewis's response to Parfit. After all, he already emphasizes degree-indeterminacy as the key to reconciling identity and connectedness. I have not proposed altering his core metaphysics of personal identity; I build upon it. So what does our discussion add?

There are three respects in which I think we have made progress. First, the way I've set up the challenge is not the way that Lewis does. I emphasize the problem that *caring about survival* and *caring about psychological connectedness* look at first glance like they do not assign the same utilities to outcomes and actions. So I see the challenge as first and foremost a puzzle in normative psychology. But Lewis formulates the challenge differently: he presents Parfit's puzzle as just that psychological connectedness and identity are relations with a different formal character—that one is transitive, the other not; one comes in degrees, the other not. So for him, the puzzle is most directly one in philosophical logic. That difference is significant, since unless

[38] In fact, in a subvaluational setting, the negation of a proposition will not be inconsistent with the original.

[39] Thanks here to a referee who pressed me to be clearer on the status of $N(\alpha = \omega)$ in this argument. For the record, the relevant proposition for standard supervaluationism is that it is not determinate that $\alpha = \omega$; for subvaluationism is that it is determinate that $\alpha \neq \omega$; and for degree and classical supervaluationism is simply the ordinary negation.

desirability enters the picture somewhere, there's no chance of extracting the kind of theses that I've been discussing. And indeed, I take it that his different conception of the challenge to be addressed is exactly why Lewis doesn't explicitly address such questions in his paper.

But, even granted the above, you might think that degree-supervaluationism (understood as a position in the nonclassical theory in mind, as in Part I) was really implicit in Lewis's proposal all along; and so at best we've been offering a rational reconstruction of his reasons for endorsing it. I would of course be pleased to have such support for my conclusions. However, I see no independent evidence that Lewis was (implicitly) thinking of indeterminacy in the way I recommend. I have emphasized that there are many alternate, reasonable ways to pair the supervaluational framework with a nonclassical theory of mind, and that simply introducing the formal constructs 'degrees of determinacy' is common property among supervaluationists, not a clever gambit available to degree supervaluationists alone. Indeed, on the matter of Lewis-exegesis I'm not at all confident that degree-supervaluationalism fits with the things he says about indeterminacy and vagueness elsewhere—we're about to see one source of pressure in Part III. So while I argue that he *needs* to endorse degree-supervaluationism to have an answer to Parfit, I don't suggest that he *does* so. It might be a quite unwelcome commitment for him.

Third, I don't think we should underplay the significance of having a detailed *argument* taking us from premises about personal identity to conclusions about the nonclassical theory of mind—even if those conclusions supported an already-familiar account. For while I don't think it's shocking that degree-supervaluationism (in the technical sense in which I'm understanding that term) is one way of understanding Lewis's proposal, and one way of reconciling identity and connectedness mattering, that it's the *only* way to do so is a far stronger claim. And that's what this section, if cogent, achieves. To see that this is a live issue, consider one who would like to endorse the Lewis formalism under a standard-supervaluationist, rather than a degree-supervaluationist, reading. Is it *obvious* that such a person won't be able to endorse both (Identity) and (Connectedness)? Not at all. In effect, this was the issue in section II.5, where I argued that in order to rule out the standard supervaluationist view without begging the question, we needed to start from stronger premises than we had been working with up to that point. Of course, I think that we are entitled to the strengthened premises. But that isn't beyond dispute, and neither was it obvious *ex ante* that it would be required. One virtue of working patiently through the nitty gritty details of an argument for a particular thesis is that we get insight into the nature of paths not taken. Here we learn just how close a completely different cognitive interpretation of Lewis's machinery comes to fitting with the degree-ish desires that Parfit posits.

III Fission and Longevity

Parfit had two basic forms of argument against taking identity to be what matters in survival. Part II examined one of these forms, based around degrees of psychological connectedness. But Parfit also raised puzzles based on *non-transitivity* in psychological connectedness. Two famous examples of this are:

Fission Scotty steps into the teletransporter, and a malfunction occurs. Scotty's matter and psychological states are recorded, transferred, replicated and two copies of Scotty appear on the bridge. A moment after the accident, there are two post-accident person stages which are psychologically connected, to the fullest extent, to the pre-accident Scotty stage. The post-accident stages are not directly psychologically connected to each other.

Longevity Methuselah lives 969 years. His psychological connections, however, are time-limited. At any point, he can remember 150 years past; and effectively plan for 150 years to come, but no further. His 500th year stage is thus psychologically connected to his 400th year stage; and his 400th year stage is psychologically connected to his 300th year stage. But the 500th year stage is not psychologically connected to his 300th year stage.

Lewis's metaphysics of persons (as maximally psychological-continuity-interrelated fusions of person-stages) gives him a distinctive take on both scenarios. His metaphysics predicts that in Fission, there are two persisting persons—Scotty$_1$ and Scotty$_2$. Both Scotties are collocated (overlap) before the accident–the same person stages are part of both. Post-accident, they diverge. Longevity is an even more dramatic example of the same phenomenon. The 500-year stage of Methuselah is a part of many distinct persisting people, one for each maximally psychological-continuity interrelated fusion of stages.

I-relatedness between person stages, for Lewis, consists in there being *some* person containing both stages as parts. The pre-accident Scotty-stage is thus *I*-related to post-accident stages of both Scotty$_1$ and Scotty$_2$, though they are not *I*-related to each other. The 500-year Methuselah-stage is *I*-related to those Methuselah-stages within 150 years. Since there is no person containing both the 300th and 500th year stages, they are not *I*-related, though the 400th year stage is *I*-related to both. By contrast to the case of degrees, in these scenarios it is an all-or-nothing matter whether or not stages are *I*-related. No indeterminacy need enter the picture.

III.1 What's the puzzle?

As we've seen, Lewis interprets Parfit's challenges as puzzles in philosophical logic. On this reading, the problem was that psychological continuity had formal properties (degrees, non-transitivity) that identity does not share. Lewis points out that the *I*-relation that emerges from his metaphysics of personal identity shares these properties (for degrees we needed to appeal to the interaction of Lewis's metaphysics with indeterminacy, for non-transitivity, we read it straight off the metaphysics).

But my concern in this paper has been with Parfittian cases read as puzzles for normative psychology. Adapting our earlier case can bring out the need to say more. Suppose that pre-accident Scotty is told that he will Fission later in the day, and is offered a deal benefitting only Scotty$_1$, for a small present sacrifice, exactly paralleling the deal the broker offers Alpha in the Cabinet scenario. Supposing Scotty is entirely self-interested. What should he do?

(I focus on Fission for simplicity, but Longevity raises the same issues. Suppose the broker offers an investment for present sacrifice at year 500, with benefits accruing to the year 600 Methuselah-stage. Some of the persons of whom the year-500 stage is a part would benefit from such investment (e.g. the one whose life runs from year 475 to 625), others do not benefit at all (e.g. the one whose life runs from year 400 to 550). Again, how should *Methuselah* act, if he is *self*-interested?)

Now, Parfit would give clear advice to the pre-accident Scotty-stage. Since psychological connectedness is what matters in survival, and this obtains to the highest degree between the pre-accident Scotty stage and the post-accident Scotty$_1$ stage, Scotty should take the deal. It benefits someone he cares about for little cost.

Parfit can stop there. But someone who wants to reconcile *identity being what matters* needs to add something extra. Once self-interested desires and decisions are spelled out in terms of self-identity, they might give different advice. I suspect, however, *Lewis* wouldn't see any new puzzle. As we saw, his metaphysics of personal identity ensures that the *I*-relation between person stages will exactly track psychological connectedness (its non-transitivities, as well as its varying strengths). Lewis seems happy to unpack 'caring about identity' as *caring about stages to whom one's current stage is I-related*. So Lewis like Parfit might argue that pre-accident Scotty should invest in the scenario just given, and that no new Reconciliation challenge has been identified.

I argued in Part II that a stricter and more principled reading of 'personal identity being what matters' could be combined with Lewis's treatment of degrees of indeterminacy. Let us therefore look and see what happens, under the hypothesis that Scotty cares about future individuals iff they are strictly the same person as himself.

III.2 The case for subvaluational indeterminacy

The broker offered a deal on which for a small sacrifice, later stages of Scotty$_1$ benefit; and later stages of Scotty$_2$ gain nothing. A natural thought is that if they are each self-interested, Scotty$_1$ must take the deal; and Scotty$_2$ should not. The trouble is that pre-accident Scotty is full-well aware that his current person-stage (who is the one who is going to have to make the choice!) is part of both Scotty$_1$ and Scotty$_2$. So the natural thought above just generates apparently conflicting advice.

But this way of formulating the issue does suggest a way to present the decision-situation in a way that parallels the one that faced Alpha in Part II:

Broker's offer	$S = S_1$	$S = S_2$
Invest	Long life, No party, Riches	Short life, No party, No riches
Reject	Long life, Party, No riches	Short life, Party, No riches

S_1 and S_2 are names for the persons Scotty$_1$ and Scotty$_2$ respectively. S, on the other hand, is the name that pre-accident Scotty gives himself. In asking whether $S = S_1$, Scotty is asking himself "am *I* Scotty$_1$?". As the parallels suggest, *indeterminacy* re-enters the scene. For there are two salient persons who could be the referent of the 'I'-thoughts in the pre-accident-Scotty-stage's head, and they have completely symmetrical claims to be the referent. Assertion: in such a scenario, it is indeterminate which of *I am Scotty*$_1$ or *I am Scotty*$_2$ is true, as thought by the pre-accident-Scotty-stage. Thus, in asking whether pre-accident Scotty should take the broker's deal, we have another situation of decision-making under indeterminacy. I will argue that we can achieve Lewisian Reconciliation of Parfittian concern with caring about strict self-identity in the fission cases, but that this requires a *different* cognitive role for indeterminacy to that argued for in Part II.

First: the degree-supervaluational cognitive role will not work. Based on the symmetry of the situation, we may assume pre-accident Scotty is Scotty$_1$ on half of the sharpenings; and is Scotty$_2$ on the other half. If the cognitive role of indeterminacy was as described in previous sections, then this would mean that Scotty would value benefits accruing to future Scotty$_1$ stages in a scaled down way, weighted by the degree of determinacy. The prediction would be that pre-accident Scotty should only give up the funds that buy a party right now if the benefits accruing to Scotty$_1$ were intrinsically more than *twice* as good. However, that conflicts with Parfit's predictions (and Lewis's, on the reading on which the *I*-relation is what matters). For the pre-accident Scotty stage is fully psychologically connected (and *I*-related) to the Scotty$_1$ stage who receives the returns on the investment. As these don't get scaled down, the prediction is that Scotty should give up funds whenever the intrinsic benefits accruing to Scotty$_1$ *exceed* that of the party pre-accident Scotty has to give up. If the broker offers returns whose utility is strictly between that of a party and twice that amount, then the Parfittian advice is to accept but the degree-supervaluationist advice is to reject.

Second: a *subvaluational* cognitive role for the indeterminacy involved in Fission and Longevity will achieve Reconciliation. I emphasized in the introductory sections that there are many candidate cognitive roles for indeterminacy, even within the supervaluational family. Subvaluationism tells us that sentence/proposition is true iff it is true on at least one sharpening. Applied to Scotty's scenario, this will tell us that it's true that Scotty is Scotty_1 (since this is so on at least one sharpening) but likewise that it's true that Scotty is Scotty_2 (for parallel reasons). The conjunction of these claims, however, is false, since the conjunction is never true on any sharpening. Subvaluationism does not allow for true contradictions, but it is a theory that attributes truth value gluts—a proposition and its negation can both be true/have truth value 1.[40]

The interaction with beliefs and desires is what is crucial for present purposes. The characteristic claim will be that Scotty should *fully* believe that he is Scotty_1, and equally fully believe the claim that he is Scotty_2. And insofar as he is interested in securing benefits for himself, Scotty will desire that future Scotty_1 stages get benefits. There will be no 'scaling down' predicted as there was on the degree supervaluation view. If Lewis wants to reconcile Parfittian patterns of concern with care about identity in Fission, he needs to endorse a subvaluational conception of the relevant indeterminacy.

(Exactly the same diagnosis applies to Longevity. The 500-year Methuselah stage is psychologically connected to all stages between 350-years and 650-years. By Parfittian lights, he should care fully and in an unscaled-down way for what happens to all these stages (and not at all for those to which he's not so connected). By Lewis's lights, there are many—perhaps infinitely many—persons who share the 500-year old Methuselah stage. I argue it is indeterminate which of these persons *is* 500-year Methuselah. A subvaluationist treatment of the cognitive role of this indeterminacy is necessary and sufficient for concern for strict self-interest to coincide with Parfittian concern in such cases.)

III.3 Full Reconciliation requires pluralism about cognitive role

In Part II I argued that a degree-supervaluational cognitive role was required for the Lewisian Reconciliation project, in application to cases where psychological connectedness comes in degrees. I've just argued that a rival *subvaluational* cognitive role is required for other cases that Parfit raises. One hostile to reconciliation might take the moral to be that there is no coherent account of the cognitive role of indeterminacy that simultaneously reconciles all of Parfit's puzzles. Going consistently degree-supervaluationist would give the wrong results in Fission and Longevity; going consistently subvaluationist would mess up the cases of degrees from Part II.

As I read him, Lewis treats the puzzles of degree and non-transitivity differently. In the former case, he needs to appeal to indeterminacy in the I-relation; but in the latter case, the distribution of I-relatedness is perfectly determinate. Since Lewis seems content to say that it is *caring for the I-relation* that we need to reconcile with Parfittian concern for psychological connectedness, indeterminacy only enters in the first case. Indeterminacy enters the second case too, however, if we adopt the stricter interpretation of *personal identity being what matters*. As we saw, how Scotty should react to a broker's offer that benefits one fission-product but not the other turns on what attitude he takes to another indeterminate question: *which person he is*.

[40]Various treatments of fission predict similar patterns, though subvaluationism is distinctive in tracing this to indeterminacy. Compare stage theories (Hawley, 2001; Sider, 1996); occasional identities (Gallois, 1990) and paraconsistent theories (Tanaka, 1998) as well as the many-one identity account of (Baxter, 1988).

Those who want to achieve full reconciliation of care for personal identity with Parfittian concern need to face up to the challenge that they seem committed to giving inconsistent answers to the cognitive role question in the two cases.

However, the arguments of Parts II and III only deliver *incompatible* conclusions if we presuppose that there is a *single* cognitive role for indeterminacy. If indeterminacy in the extent of psychological connectedness required for identity between stages has a *different* cognitive role from the indeterminacy induced when a stage is a common part of multiple persons, then there is no principled problem in giving distinct treatments of each. (An interesting secondary issue arises however, concerning the proper treatment of *mixtures* of the two kinds of indeterminacy—for example, cases of Longevity where psychological connectedness is both intransitive, and 'fades out'. I propose a treatment of this in Appendix B that delivers degree-supervaluations and subvaluations as special cases).

I have argued against the general assumption that there is a *single* cognitive role associated with indeterminacy elsewhere (Williams, 2012c). It is striking that the Reconciliation project places such tight constraints on the cognitive role of indeterminacy, including the demand for *variation* for different *types* of indeterminate survival. But this is not an objection to the project; it is a result we can embrace.

Conclusion

I argued in Part I that developing a nonclassical theory of mind was a mandatory item on the to-do list of anybody who wants to diverge from standard classical logic and semantics. That is so even if one's favoured nonclassicism involves only minimal departures from classicality, minimal enough so as to vindicate the classical *reasoning* that is the usual way that theory-building manifests classical presuppositions. Members of the supervaluational family illustrate this. Despite being very close to classical logic, they require substantive changes to our theory of rational belief, desire and doxastic modality—and different members of the family point in different directions. The good news is that we can indeed generalize a baseline classical account of rational states of mind, generating a zoo of nonclassical theories of mind.

In Parts II and III, I presented arguments that can pin down commitments to a distinctive theory of mind. As it turns out, they point in different directions in the two kinds of case I consider. The arguments are based, most fundamentally, on a Parfittian treatment of our attitudes to survival. Independently of whether one agrees with its starting point (and I am personally agnostic) the connections here established afford a model for how these issues can be discussed and resolved.

A The *I*-relation

Lewis argues that degree of determinacy of the *I*-relation can exactly coincide with degree of psychological connectedness. His argument for is based on his metaphysics of personal identity: a person is a fusion of a maximally *R*-interrelated collection of stages; and the *I* relation is then defined as obtaining between stages whenever there is some person of whom they are both parts. This appendix reviews Lewis's argument for the coincide of degrees of psychological connectedness with degrees of determinacy of the *I*-relation.

The argument involves two moves (only the second is explicit in (Lewis, 1976)):

(A) whenever the *R* relation holds between *a* and *b* on a sharpening, then on that sharpening *a* and *b* will be parts of some person.

(B) the measure over the sharpenings that fixes degrees of determinateness makes degrees of determinateness coincide with degrees of psychological relatedness.

On (A): we need to show that whenever *Rab* holds (on a sharpening), then (on that sharpening) there's a maximal *R*-interrelated fusion (i.e. a person) that includes *a* and *b*. To see this, assume (plausibly) that there are only set-many person-stages in existence. Appealing to the axiom of choice, well-order the set of person-stages, giving it order-type isomorphic to ordinal σ, so that each person-stages can be written c_α for some $\alpha \in \sigma$. Now let $A_0 := \{a,b\}$, and then in general set $A_{\alpha+1} := A_\alpha \cup \{c_\alpha\}$ iff $\forall x \in A_\alpha : Rxc_\alpha$; and $A_{\alpha+1} := A_\alpha$ otherwise (we set A_λ for limit ordinals be the union of $A_\alpha, \alpha < \lambda$). Each set is by construction *R*-interrelated, and lower-indexed sets are included in those with higher index. A_σ is then by construction a maximal *R*-interrelated set of person stages—after all, any person stage c_α not included must, by construction, fail to be *R*-related to some stage in A_α, which is a subset of A_σ. So we couldn't add that stage into the set and preserve *R*-interrelatedness. The fusion of the stages in A_σ counts as a person by Lewis's lights (relative to the operative sharpening of the *R*-relation). And it includes *a* and *b* as parts.

On (B): Parfit's idea was psychological connectedness comes in degrees, which scale our concern to others so-related to us. Let's assign the degrees numbers, where 1 indicates a level of connectedness sufficient for full care, and 0 no care at all. We can get the degree-of-determinateness of the *I*-relation, (i.e. 'both parts of the same person') to coincide with these numbers so long as the measure over these sharpenings induces the uniform measure over this set of boundary-numbers $(0,1)$; that is, the measure of sharpenings which set the boundary number within $(0,k]$ is exactly k. Suppose that x and y are psychologically connected to degree r. Under those assumptions, the measure of sharpenings on which there is a person with x and y as parts, will be exactly the measure of the set Z of sharpenings on which they count as *R*-related (by the paragraph above). But Z is exactly those sharpenings who set the boundary number within $(0,r]$, and so (by the uniformity assumption) the measure of this set will be r.

So degrees of connectedness and degrees of determinacy of the *I*-relation *can* coincide—though only if the measure over sharpenings that underpins the latter takes the right form. Lewis himself notes the importance of the 'scale and measure' to securing this result.

B Personal identity

The *I*-relation, as Lewis emphasizes, is not the relation of personal identity—it is a relation that holds between (often distinct) person-stages. Can we argue further that degree of determinacy of *personal identity* strictly construed coincides with the degree of determinacy of the *I*-relation?

Because of the non-transitivities that Fission and Longevity illustrate, this cannot be the case in general if degrees of determinacy are understood in the degree-supervaluational manner. This claim was argued for in Part III. In this appendix I provide a treatment of the truth values of indeterminate sentences which ensures that the *truth values* of identity claims and the *truth values* of I-relatedness claims will exactly match.

First of all, what personal *identity* claim is it that we are targeting? I continue with the model outlined in Parts II and III. Let a and b be person stages, A a descriptive name for that person which contains a as a stage, and B a descriptive name for that person which contains b as a stage. There is then a well-formed question of whether persons A and B are strictly and literally identical.

First case: assume a and b are each parts of exactly one person. Then, determinately, $A = B$ iff a and b are parts of the same person (since that person is then the referent of both capitalized terms). And so the measure of sharpenings on which the identity holds is exactly the measure of sharpenings on which the I-relation holds between the stages.

Second case: we allow that a and b can be parts of several people (that will fall out of Lewis's metaphysics of personal identity as formalized in the first Appendix, in scenarios such as Fission and Longevity). I will assume that in a scenario where a is a stage of several persons, a term like 'A' will be indeterminate in reference between the persons. Thus in addition to that aspect of indeterminacy that turns on the choice of a boundary number for the R-relation, there is now also an aspect that has to resolve which among the various persons overlapping at a, gets to count as 'the' person A.

The sharpenings on which the I-relation obtains between a and b now *subdivide* according to how the indeterminacy of reference of A and B is resolved. So on some sharpenings where a and b are I-related, $A = B$ is false (of course, $A = B$ is never true unless a, b are I-related). As discussed in Part III, this means we cannot simply identify the truth value of the identity claim with the measure of sharpenings on which it is true (its degree of determinacy) if we want to get the match of truth values between I-relatedness and identity claims that is our target.

Now for the account of truth values that will do the job. Model sharpenings as having internal structure reflecting the two aspects of indeterminacy in the cases we are considering: a boundary number b (telling us what degree of psychological connectedness suffices for being R-related, and hence what objects count as persons) and a selection function f (telling us which of the various persons sharing stage a terms like 'A' refer to).Whether a sharpening $\langle b, f\rangle$ classifies a and b as I-related depends solely on b—since whether or not there is some person or other containing both is fixed solely by which objects count as persons. Within the set of sharpenings that *do* classify a and b as I-related, whether $\langle b, f\rangle$ classifies $A = B$ as true turns on f alone. Given that there is some person that contains both a and b as parts, f tells us whether 'A' and 'B' both name that person.

I propose we define the truth value of p as *the measure of boundary numbers b such that there is some sharpening $\langle b, f\rangle$ such that p is true at $\langle b, f\rangle$*. This has the result that the truth value of $A = B$ exactly matches the degree of determinacy of the I-relation obtaining between a and b.

The definition just proposed has two limiting cases. Suppose there is no stage-sharing, so that the parameter f is redundant (as we presupposed throughout section II). Then the definition tells us that degrees of truth match degrees of determinacy, just as in the first case above. Suppose on the other hand that psychological connectedness is all-or-nothing, as in the cases of Fission and Longevity in the final section. Then the parameter b is redundant (the measure of boundary numbers meeting the stated condition will always be 1 or 0), and the characterization tells us that the truth value of $A = B$ is 1 iff there is some person containing both a and b as stages; and otherwise 0. This is the 'subvaluationist' model of Part III.

The combined characterization allows us to account for cases that combine intermediate degrees of connectedness and stage-sharing. To illustrate this, I adapt the Longevity case described earlier. Suppose the psychological connectedness between Methuselah's stages 'fades out' in a uniform way when they are between 100 and 150 years apart. If m_{500} is Methuselah's 500th-year-stage, and m_{625} is a successor stage 125 years later, then the degree of psychological connectedness between the two stages will be of strength 0.5. Lewisian reconciliation requires that in the mindset in which strict and literal identity is what matters in survival, what happens to m_{625} matters to m_{500} to a scaled down extent—the scaling being 0.5. The account here achieves this, by predicting that the truth value of '$M_{500} = M_{625}$' is 0.5, where the capitalization indicates descriptive names based on the respective person-stages, as before. All boundary numbers < 0.5 will be such that the two stages are not common parts of any person. Boundary number ≥ 0.5 will be such that *many* persons containing m_{500} will also contain m_{625}, but equally, that there are many persons containing m_{500} that do not contain the later stage. Thus in order to achieve Reconciliation, it's vital that we define the truth value as we have done: 0.5 is not the measure of sharpenings which make the identity claim true, but rather the measure of b such that there is some sharpening $\langle b, f \rangle$ such that p is true at b, f.

The combined treatment gives a nonclassical theory of mind within the supervaluational family that generalizes both degree and subvaluational settings.

References

BARNES, ELIZABETH. 2007. *'Conceptual room for Ontic Vagueness'*. Ph.D. thesis, University of St Andrews.

BARNES, ELIZABETH, & CAMERON, ROSS P. 2009. 'The Open Future: Bivalence, Determinism and Ontology'. *Philosophical Studies*, **146**, 291–309.

BARNETT, DAVID. 2009. 'Is vagueness sui generis?'. *Australasian Journal of Philosophy*, **87**, 5–34.

BAXTER, DONALD L. M. 1988. Many-One Identity. *Philosophical papers*, **XVII**(3).

BEALL, J. C. 2009. *Spandrels of Truth*. Oxford University Press.

BROOME, J. 1994. 'Discounting the future'. *Philosophy and Public Affairs*, **23**(2), 128–156.

CAIE, MICHAEL. 2012. 'Belief and Indeterminacy'. *Philosophical Review*, **121**(1), 1–54.

DUMMETT, MICHAEL A. E. 1959. 'Truth'. *Pages 1–24 of: Truth and Other Enigmas*. London: Duckworth. First published in *Proceedings of the Aristotelian Society* (59):141–62.

EDGINGTON, DOROTHY. 1997. 'Vagueness by Degrees'. *Pages 294–316 of:* KEEFE, ROSANNA, & SMITH, PETER (eds), *Vagueness: A reader*. Cambridge, MA: MIT Press.

EVANS, GARETH. 1978. 'Can there be vague objects?'. *Analysis*, **38**(13), 208. Reprinted in Keefe and Smith (eds) *Vagueness: A reader* (Cambridge, MA: MIT Press, 1997).

FIELD, HARTRY H. 1978. 'Mental Representation'. *Erkenntnis*, **13**, 9–61. Reprinted in Field, *Truth and the Absence of Fact* (Oxford University Press, Oxford, 2001) pp. 30–67.

FIELD, HARTRY H. 2000. 'Indeterminacy, Degree of Belief, and excluded middle'. *Nous*, **34**, 1–30. Reprinted in Field, *Truth and the Absence of Fact* (Oxford: Oxford University Press, 2001) pp. 278-311.

FIELD, HARTRY H. 2003a. 'No fact of the matter'. *Australasian Journal of Philosophy*, **81**, 457–480.

FIELD, HARTRY H. 2003b. 'Semantic paradoxes and the paradoxes of vagueness'. *Pages 262–311 of:* BEALL, J. C. (ed), *Liars and Heaps*. Oxford: Oxford University Press.

FIELD, HARTRY H. 2008. *Saving Truth from Paradox*. Oxford: Oxford University Press.

FINE, KIT. 1975. 'Vagueness, truth and logic'. *Synthese*, **30**, 265–300. Reprinted with corrections in Keefe and Smith (eds) *Vagueness: A reader* (MIT Press, Cambridge MA: 1997, pp.119-50).

FREDERICK, SHANE, LOEWENSTEIN, GEORGE, & O'DONOGHUE, TED. 2002. Time Discounting and Time Preference: A Critical Review. *Journal of Economic Literature*, **40**(2), 351–401.

GALLOIS, ANDRE. 1990. 'Occasional Identity'. *Philosophical Studies*, **58**(3), 203–24.

HAWLEY, K. 2001. *How Things Persist*. Oxford: Oxford University Press.

HAWTHORNE, J. 2005. Vagueness and the Mind of God. *Philosophical studies*, **122**(1), 1–25.

HOWSON, C, & URBACH, P. 1993. *Scientific Reasoning: the Bayesian approach*. Michigan: Open Court.

HYDE, DOMINIC. 1997. 'From Heaps and Gaps to Heaps of Gluts'. *Mind*, **106**, 641–60.

JEFFREY, RICHARD C. 1965. *The Logic of Decision*. 2nd edn. Chicago and London: University of Chicago Press. Second edition published 1983.

JOYCE, JAMES M. 1998. 'A non-pragmatic vindication of probabilism'. *Philosophy of Science*, **65**, 575–603.

JOYCE, JAMES M. 2009. 'Accuracy and coherence: prospects for an alethic epistemology of partial belief'. *Pages 263–297 of:* HUBER, FRANZ, & SCHMIDT-PETRI, CHRISTOPH (eds), *Degrees of belief*. Springer.

KAMP, J. A. W. 1975. 'Two theories about adjectives'. *Pages 123–55 of:* KEENAN, EDWARD L. (ed), *Formal Semantics of Natural Language*. Cambridge: Cambridge University Press. Reprinted in Davis and Gillon (eds) *Semantics: A reader* (Oxford University Press, Oxford, 2004) pp.541–62.

KEEFE, R. 2008. Vagueness: supervaluationism. *Philosophy Compass*, **3**(2), 315–324.

KEEFE, ROSANNA. 2000. *Theories of Vagueness*. Cambridge: Cambridge University Press.

LEWIS, DAVID K. 1970. 'General Semantics'. *Synthese*, **22**, 18–67. Reprinted with postscript in Lewis, *Philosophical Papers I* (Oxford University Press, 1983) 189–229.

LEWIS, DAVID K. 1976. 'Survival and Identity'. *In:* RORTY, AMÉLIE OKSENBERG (ed), *The Identities of Persons*. University of California Press. Reprinted in Lewis, *Philosophical Papers vol. I* (Oxford University Press, 1983, pp. 55–72.).

LEWIS, DAVID K. 1988. 'Vague Identity: Evans misunderstood'. *Analysis*, **48**, 128–30. Reprinted in Keefe and Smith (eds) *Vagueness: A reader* (Cambridge MA: MIT Press 1997).

MACFARLANE, JOHN G. 2010. 'Fuzzy epistemicism'. *In:* DIETZ, RICHARD, & MORUZZI, SEBASTIANO (eds), *Cuts and Clouds*. Oxford: Oxford University Press.

MAS-COLELL, A., WHINSTON, M.D., & GREEN, J.R. 1995. *Microeconomic theory*. Oxford University Press.

MCGEE, VANN, & MCLAUGHLIN, BRIAN. 1994. 'Distinctions without a difference'. *The Southern Journal of Philosophy*, **supp XXXIII**, 203–51.

MILNE, PETER. 2007. 'Betting on fuzzy and many-valued propositions (long version)'. *In: ms.*

MILNE, PETER. 2008. 'Betting on fuzzy and many-valued propositions'. *In: The logica yearbook*.

PARFIT, DEREK. 1971. 'Personal Identity'. *The Philosophical Review*, **80**, 3–27.

PARFIT, DEREK. 1976. 'Lewis, Perry and what matters'. *In:* RORTY, AMÉLIE O. (ed), *The identities of persons*. Berkeley: University of California Press.

PARFIT, DEREK. 1984. *Reasons and Persons*. Oxford: OUP.

PARIS, J.B. 2001. 'A note on the Dutch Book method'. *Pages 301–306 of: Proceedings of the Second International Symposium on Imprecise Probabilities and their Applications, ISIPTA*. Ithaca, NY: Shaker.

PRIEST, G. 2006. *In contradiction: a study of the transconsistent*. USA:: Oxford University Press.

SCHIFFER, STEPHEN. 2003. *The things we mean*. Oxford: Oxford University Press.

SIDER, TED. 1996. 'All the world's a stage'. *Australian Journal of Philosophy*, **74**, 433–453.

SMITH, NICHOLAS J. J. 2008. *Vagueness and degrees of truth*. Oxford: Oxford University Press.

SMITH, NICHOLAS J. J. 2010. 'Degree of Belief is Expected Truth Value'. *In:* DIETZ, RICHARD, & MORUZZI, SEBASTIANO (eds), *Cuts and Clouds*. Oxford: Oxford University Press.

STALNAKER, ROBERT. 1984. *Inquiry*. Cambridge, MA: MIT Press.

TANAKA, KOJI. 1998. 'To be or to be something else: Dialethic Tense Logic'. *Logique et Analyse*, **161-162-163**, 189–202.

THOMASON, RICHMOND H. 1970. 'Indeterminist time and truth-value gaps'. *Theoria*, **3**, 264–281.

VAN INWAGEN, PETER. 1990. *Material Beings*. Ithaca and London: Cornell University Press.

WILLIAMS, J. ROBERT G. 2008. 'Supervaluations and logical revisionism'. *The Journal of Philosophy*, **CV**(4).

WILLIAMS, J. ROBERT G. 2011. 'Degree supervaluational logic'. *Review of Symbolic Logic*, **4**(1), 130–149.

WILLIAMS, J. ROBERT G. 2012a. 'Generalized Probabilism: dutch books and accuracy domination'. *Journal of Philosophical Logic*, **41**(5).

WILLIAMS, J. ROBERT G. 2012b. 'Gradational accuracy and non-classical semantics'. *Review of Symbolic Logic*.

WILLIAMS, J. ROBERT G. 2012c. 'Indeterminacy and normative silence'. *Analysis*, **72**, 217–25.

WILLIAMS, J. ROBERT G. 2014. 'Descision making under indeterminacy'. *Philosopher's imprint*.

WILLIAMS, J. ROBERT G. forthcoming. 'Nonclassical logic and probability'. *In:* HAJEK, ALAN, & HITCHCOCK, CHRISTOPHER (eds), *A companion to the philosophy of probability*.

WILLIAMSON, TIMOTHY. 1994. *Vagueness*. London: Routledge.

WRIGHT, CRISPIN. 2001. 'On being in a quandary'. *Mind*, **110**, 45–98.