Veritities and truth-values*

Nicholas K. Jones

Non-citable draft of: 24–01–14
Final version forthcoming in 2014 in Lee Walters and John Hawthorne, eds., *Conditionals, probability, and paradox: themes from the philosophy of Dorothy Edgington*, OUP.

Studying for a PhD under Dorothy Edgington's supervision was a great intellectual pleasure. She was a patient, dedicated, rigorous and inspiring teacher. Both in person and in print, Dorothy provides a model for me of how philosophy should be done. It is an honour to acknowledge my debt with this study of her thinking about vagueness.

A powerful motivation for degree-theoretic semantics is to semantically capture one of the most striking features of vagueness: the gradual transition from uncontroversial truth to uncontroversial falsity along an appropriate series. The devil, as ever, is in the details. Standard many-valued semantic theories cannot capture penumbral connections, deliver unattractive logics, and violate such seeming truisms as the T-biconditionals. Edgington's theory promises a solution to these difficulties. Edgington uses an analogy between the sorites and lottery paradoxes to motivate a probabilistic version of degree theory that builds penumbral connection in at the ground-floor, does not invalidate classical logic, and in which the T-biconditionals hold without restriction. This paper examines Edgington's view. My primary focus will be the relationship between Edgington's degrees and truth, in light of her claim that the former do not disturb or displace classical bivalent truth.

§1 introduces Edgington's view. I describe her probabilistic treament of the connectives in §1.1, her account of validity in §1.2, and her solution to the sorites in §1.3. The rest of the paper concerns the relationship between Edgington's framework and classical semantics. One of the central questions is whether, and in what sense, the many degrees assigned to sentences in Edgington's framework are new truth-values to replace the classical two. §2 examines and rejects Edgington's intriguing suggestion that her framework and the classical framework provide non-competing descriptions of a single underlying semantic phenomenon. §2.1 introduces this idea. §2.2 articulates a notion of semantics that's neutral between Edgington's framework and the classical view. §2.3 then argues that Edgington's framework and the classical one cannot be non-competing descriptions of semantics understood in this sense. That leaves two options open at the start of §3: (a) classical truth and falsity are semantic notions, whereas Edgington's degrees are not; (b) Edgington's degrees are semantic notions, whereas classical truth and falsity are not. In recent work, Edging-

*Thanks to John Hawthorne, Rosanna Keefe, Peter Sutton, the SLOBs, the members of the Hossack-Textor work-in-progress group, and the audience at the *Conditionals and paradox* conference in London for comments and discussion. I'm especially grateful to Dorothy Edgington for discussion of this material.

ton endorses a primitivist view about vagueness that fits well with option (a). She argues for this view by arguing that bivalence holds without restriction, even in borderline cases, which appears to tell against option (b). §3.1 argues that Edgington's arguments for bivalence presuppose a substantive thesis about truth. §3.2 closes the paper by arguing that this thesis is incompatible with option (a). I conclude by suggesting that the best version of Edgington's proposal will be a version of option (b).

A note on terminology before I begin. Clarity will be my theoretically neutral concept for distinguishing borderline cases from the rest. It's clear that p iff p and it isn't borderline whether p; it's clear that not-p iff not-p and it isn't borderline whether p; it's borderline whether p iff it's neither clear that p nor clear that not-p. I'll say that x is clearly F when it's clear that x is F. So, for example, scarlet is clearly a shade of red, pure orange clearly isn't a shade of red, and somewhere between them in the spectrum lie shades that are only borderline shades of red (and borderline shades of orange too).

1 Edgington's proposal

This section introduces Edgington's view: §1.1 outlines her account of the connectives; §1.2 turns to validity; §1.3 looks at the sorites. My primary source is "Vagueness by degrees".[1]

1.1 The framework

Edgington's view has four key features. Firstly, her approach is degree-theoretic: sentences are assigned degrees, represented by real numbers in $[0, 1]$. Edgington calls these degrees *verities*. Throughout, the number assigned to a sentence A will be $v(A)$.[2] Secondly, verities are not degrees of truth, or new truth-values in competition with the classical two; verities are degrees of *closeness to clear truth*.[3] §§2–3 examine the relation between verity and truth-value in more detail. Thirdly, greater numbers represent closer proximity to clear truth, with 1 as clear truth and 0 as clear falsity. Fourthly, the connectives and verities interact via the probabilistic rules:

- $v(\neg A) = 1 - v(A)$.
- $v(A \vee B) = v(A) + v(B) - v(A \wedge B)$.
- $v(A \wedge B) = v(A) \times v(B|A)$.

Here, $v(B|A)$ is the conditional verity of B given A:

> "Conditional verity may be explained thus. Borderline cases of a vague term, e.g. "red", with verities between 1 and 0, are cases which it would not be definitely wrong to count as red. We can hypothetically decide to count such a case

[1] (Edgington, 1997), henceforth "*VbD*".

[2] I use 'A', 'B', 'C' as metalinguistic variables ranging over sentences, and 'p', 'q', 'r' as object-language propositional constants (though I'll also sometimes use them as placeholders for sentences in the body-text).

[3] (*VbD*, p299), (Edgington, MS, §§2, 4), (Edgington, 2010, p105–6).

> as above the borderline—as definitely red—and see what consequences this hypothetical decision has for other propositions. $v(B \text{ given } A)$, the conditional verity of B given A, is the value assigned to B on the hypothetical decision to count A as definitely true."[4]

One important feature of the way we use vague language is a kind of classificatory discretion about borderline cases: one may choose to count borderline cases either way, as cases or as non-cases, provided one makes it clear to one's conversational partners that that's what one is doing and leaves oneself open to future revision as the conversation evolves.[5] Edgington's rules for the connectives and gloss on conditional verity place classificatory discretion at the heart of her semantics.

Conditional verity induces failures of verity-functionality: the verity of a complex sentence is not a function of the verities of its components.[6] Verity-functionality fails because $v(p) = v(q)$ is compatible with both $v(p|r) \neq v(q|r)$ and $v(r|p) \neq v(r|q)$; so substitution of 'p' for 'q' within a complex sentence can preserve the verities of its components without preserving the conditional verities on which the verity of the whole complex depends. An example. Suppose a and b are both borderline red, though a is only just slightly redder than b, so that:

- $v(Ra) = v(\neg Ra) = 0.5$.
- $v(Rb|Ra) = 0.9$.

  (b counts as almost clearly red, given the hypothetical decision to count something, a, just redder than it as clearly red.)
- $v(Rb|\neg Ra) = 0$.

  (b counts as clearly not red, given the hypothetical decision to count something, a, just redder than it as clearly not red.)

Hence:

- $v(Ra \wedge Rb) = v(Ra) \times v(Rb|Ra) = 0.5 \times 0.9 = 0.45$.
- $v(\neg Ra \wedge Rb) = v(\neg Ra) \times v(Rb|\neg Ra) = 0.5 \times 0 = 0$.

These conjunctions have different verities, though they differ only by verity-preserving substitution of conjuncts. That is a good thing. Since a is redder than b, it is natural to regard one who said that a is not red and b is red as clearly mistaken, unlike one who said that both are red. Edgington's account of conditional verity thus allows her to accommodate the conceptual relations between the borderline regions of vague predicates, known as penumbral connections:[7] hypothetical classificatory decisions about borderline cases have implications for how other borderline cases are classified under the relevant hypotheses.

It is not entirely clear how the classificatory decisions Edgington uses to explicate conditional verity should be understood. The quote above suggests that they are decisions to

[4] (*VbD*, p306).
[5] (Shapiro, 2006) emphasises this phenomenon.
[6] This differentiates Edgington from the more traditional degree theories criticised in (Williamson, 1994, ch4).
[7] (Fine, 1975, pp123–125)

count a sentence as clearly true, or to count a predicate as clearly applying to an object. That is also required by her solution to the sorites, which I'll describe §1.3. Central to that solution is an account of the verities of certain material conditionals, derived from Edgington's rules for the connectives. The derivation relies on the following procedure for calculating $v(Rb|Ra)$ when b is further from clearly red than a:[8]

- First, assign maximum verity to '*Ra*'. Then rescale the degrees to which things further from clearly red than a count as clearly red, holding the clear non-red cases fixed. The conditional verity $v(Rb|Ra)$ is $v(Rb)$ following rescaling.

Maximum verity is clear truth. So the first step of the procedure makes '*Ra*' clearly true. Edgington wants to explicate conditional verity using classificatory decisions. So such decisions must be capable of making '*Ra*' clearly true, when it's borderline before the decision. This raises some unanswered questions for Edgington.[9]

Suppose '*Ra*' is borderline. Suppose I decide to count it as clearly true. What happens? There are two options: (a) '*Ra*' remains borderline ; (b) '*Ra*' becomes clearly true. Option (a) undermines the above procedure for calculating $v(Rb|Ra)$. So some classificatory decisions cannot be of type (a).

According to option (b), my decision changed '*Ra*' from borderline to clearly true. So borderline status for sentences must be relative to time, or to the classificatory decisions in force at a conversational context.[10] Now, Edgington endorses unrestricted bivalence, even for borderline cases. (§2.1 and §3 discuss bivalence further.) So before my decision: '*Ra*' was either true or false. Suppose it was true. Then my decision changed '*Ra*' from borderline to clearly true, without affecting whether its truth-value. How did it do so? My decision presumably affected '*R*'s meaning in some other way. We need to know more; a story is owed of exactly how this works.

One might attempt to avoid this commitment by maintaining that '*Ra*' was false before my decision, and likewise for any other borderline sentence. The attempt will fail. Suppose there's another vague predicate '*O*' whose extension (within the relevant domain D) comprises exactly the anti-extension (within D) of '*R*'. Then if '*Ra*' is false before my decision, '*Oa*' is true. The previous paragraph's argument can now be used to show that classificatory decisions can change '*Oa*' from borderline to clearly true without affecting its truth-value. Plausibly, 'red' and 'orange' are like this: their semantic values are related to ensure that they partition a certain region of colour space. An account of how classificatory decisions affect borderline status without affecting truth-value is still required.

What about if '*Ra*' was false before my decision? Since clear truth implies truth, my decision changes '*Ra*' from false to true. So my decision changed the truth-conditional content of '*Ra*' in some way. How? Again, a story is owed.

We've seen that Edgington is committed to three ideas. Firstly, classificatory decisions can change a sentence from borderline to clearly true. Secondly, some such decisions don't

[8] Edgington employs this procedure in the central paragraph of (*VbD*, p308).

[9] For a different criticism see (Keefe, 2000, pp98–100).

[10] (Shapiro, 2006, pp31–36, 73–75) develops this kind of view, using extensional variation to analyse variation in clarity. Because Shapiro's extensions and anti-extensions are not exhaustive, however, Edgington cannot adopt his approach. Note also that Shapiro distinguishes context-relative *internal* clarity from context-invariant *external* clarity (though he uses 'determinacy' rather than 'clarity').

affect the sentence's truth-value. Thirdly, others such decisions change the sentence's truth-conditional content. I am not claiming that these ideas are problematic. I am claiming only that a central component of Edgington's view remains under-described until an account of how this all works is supplied. Rather than getting bogged down in these issues, let us move on.

Edgington supplements the standard connectives with an operator D, intended to formalise 'It is clear that…'. Since verities are degrees of closeness to clear truth and anything with verity less than 1 is at some distance from clear truth, the rule for D is:

- $v(\mathsf{D}A) = 1$ iff $v(A) = 1$; otherwise $v(\mathsf{D}A) = 0$.

Using D, we can say in the object-language that it's borderline whether p:

$$\neg \mathsf{D}p \wedge \neg \mathsf{D}\neg p$$

Without D, the logic and semantics of borderline status can be studied in Edgington's framework. With D in place, the logic and semantics of claims about borderline status can also be examined.

1.2 Validity

This section presents Edgington's approach to validity.

Edgington introduces the *verity constraining property* (VCP) thus:[11]

- The argument from $A_1, \ldots, A_n$ to C has the VCP iff it is not possible that: $[1 - v(C)] > [(1 - v(A_1)) + \ldots + (1 - v(A_n))]$.

Call $1 - v(A)$ the *unverity* of A. In this terminology, an argument has the VCP iff the unverity of the conclusion cannot exceed the sum of the unverities of the premisses; that is, iff there cannot be an increase in total unverity between premisses and conclusion. Edgington argues that exactly the valid arguments possess the VCP:

Valid≡VCP An argument is valid iff it possesses the VCP.

On Edgington's view, the valid arguments are those in which total unverity cannot increase between premisses and conclusion. In order to understand this, we need to know what 'possible' means in the definition of the VCP. Edgington does not consider this directly, though her remarks about validity do suggest an answer.

Edgington uses modality to introduce validity:

> "By a valid argument…I mean an argument such that it is impossible that all its premisses are true and its conclusion false."[12]

Edgington does not explicate the relevant variety of possibility further. We may, however, reasonably assume that it's the same notion as she uses to characterise the VCP. On this view **Valid≡VCP** says that, for an unspecified background notion of necessity, necessary

[11] (*VbD*, pp302, 307)
[12] (*VbD*, pp299–300)

possession of one of the following properties is equivalent to necessary possession of the other: (a) truth-preservation from premisses to conclusion; (b) non-increase in total unverity between premisses and conclusion. A question remains: equivalent in what sense?

It is tempting to interpret Edgington as offering a verity-theoretic analysis of validity. Temptation should be resisted. Since **Valid≡VCP** follows trivially from the claim that its right-hand-side analyses its left, that interpretation renders it mysterious why Edgington offers a *proof* of **Valid≡VCP**.[13] One might respond that the proof shows that the VCP is coextensive with a particular notion of validity, e.g., classical validity. But the proof requires only minimal assumptions about the formal characteristics of validity, never mind distinctively classical characteristics.[14] So it cannot show any such thing.

A better interpretation takes **Valid≡VCP** to capture a necessary coextensiveness between validity and the VCP, without attributing conceptual priority to one or the other. On this view, Edgington's proof of **Valid≡VCP** is intended to establish that this relationship holds between the verity-theoretic properties of arguments and validity, without commitment to an explication of one in terms of the other. This allows the logical properties of arguments to be studied via their verity-theoretic properties. Edgington uses that approach to resolve the sorites. Now, although Edgington left the formal properties of modal validity unspecified, it is plausibly identified with classical validity; for that is what follows from the orthodox conception of possible worlds as maximal and consistent.[15] Edgington thus offers a verity-theoretic solution to the sorites that, unlike typical degree-theoretic approaches, brings no revision in classical logic. It is to this I now turn.

1.3 The sorites

How does **Valid≡VCP** help with the sorites? For brevity, I consider only the long-form sorites here.[16] Let $a_1, \ldots, a_n$ be a series of coloured patches with a_1 clearly red, a_n clearly orange (hence clearly not red), and successive patches pairwise visually indiscriminable to unaided normal human observers under ideal conditions. The sorites paradox arises because:

- 'Ra_1' is clearly true.
- 'Ra_n' is clearly false.
- Each material conditional of the following form (where $1 \leq i \leq n$) is extremely plausible, and its negation extremely implausible:

[13] (*VbD*, pp307–308)

[14] The right-left direction requires: (1) if A entails C, then $v(C|A) = 1$; (2) if $A_1, \ldots, A_n$ entail C, then $\neg C$ entails $\neg A_1 \vee \ldots \vee \neg A_n$. The left-right direction requires: (3) if it's possible that $A_1, \ldots, A_n$ are true and C false, then it's possible that $v(A_1) = 1, \ldots, v(A_n) = 1$ and $v(C) = 0$. Now, (3) is pretty close to what Edgington's trying to show: it's an instance of the contrapositive of the left-right direction of **Valid≡VCP**. Representing verities as proportions of sharpenings in a supervaluationist setting may, however, allow us to justify (3); it's verified by the one sharpening model in which $A_1, \ldots, A_n$ are all true and C false; that model exists whenever $A_1, \ldots, A_n$ can be true with C false.

[15] If the modality is metaphysical modality, then modal validity plausibly extends classical validity. Given standard views about metaphysical modality, '2+4=6' and hence also '$\exists x, y(x \neq y)$' are modal logical truths but not classical logical truths. Likewise, the argument from 'Dorothy exists' to 'Dorothy is a human being' is modally but not classically valid. The slightly modified modal notion of validity in §2.2 does not have this feature.

[16] For more detail see (*VbD*, pp307–312).

(Si) $$Ra_i \to Ra_{i+1}$$

- The following argument is valid:

(A) $$Ra_1, (\text{S}1), (\text{S}2), \ldots, (\text{S}n); \text{ therefore } Ra_n$$

Edgington's solution to the paradox emerges from the verity theoretic properties of 'Ra_1',..., 'Ra_n'.

As i increases from 1 to n, we gradually leave the clearly red patches, progress through patches increasingly distant (both physically and semantically) from clearly red, and terminate with clearly not red patches. Somewhere around the middle of the series, increasing i gradually decreases $v(Ra_i)$ from 1 to 0. The corresponding (Si)'s have consequents with marginally lower verity than their antecedents; call them *dropping* (Si)'s. On Edgington's semantics, non-dropping (Si)'s have verity 1, and dropping (Si)'s have verity only marginally less than 1. We are therefore right to find each (Si) highly plausible: each is either clearly true or very close to it.

Given that argument (A) is valid, **Valid≡VCP** implies: the unverity of 'Ra_n' cannot exceed the sum of the unverities of 'Ra_1' and the (Si)'s. 'Ra_1' and the non-dropping (Si)'s all have verity 1. What about the dropping (Si)'s? Suppose verity decreases uniformly from 1 to 0 over m steps. Then there are m dropping (Si)'s, each with verity $1 - \frac{1}{m}$. So their total unverity is $m \times \frac{1}{m} = 1$.[17] **Valid≡VCP** thus permits 'Ra_n' to have unverity 1. That is, 'Ra_n' can have verity 0 despite following by valid argument from premisses with high individual verities. The validity of argument (A) is therefore compatible with (a) 'Ra_1' being clearly true, (b) 'Ra_n' being clearly false, and (c) each (Si) being clearly true or very close thereto. The features that generate the paradox are thereby rendered consistent. Moreover, when the premisses of a valid argument have low total unverity, **Valid≡VCP** requires that the conclusion have high verity. So although valid reasoning does permit decrease in verity, belief in the conclusion on the basis of the premisses is not threatened. Provided the premisses have low total unverity, the conclusion will be both close to clearly true and far from clearly false. Adding that conclusion to one's stock of beliefs will not significantly increase the distance from clear truth of one's total doxastic state. Classical reasoning in the presence of vagueness is thereby vindicated.

2 Verities as truth-values

What are verities? In light of their probabilistic structure, one might identify them with credences: "the credence that a person with no relevant ignorance other than about the precise line, would give to a statement like "that's red"."[18] Against this proposal, Edgington argues that verity and credence occupy different theoretical roles; they interact with preferences in

[17] Dropping the uniformity assumption blocks calculation of the verities of individual dropping (Si)'s, though their total unverity remains $m \times \frac{1}{m} = 1$. It is in calculating the unverities of the dropping (Si)'s that Edgington relies on the procedure for calculating conditional verities discussed in §1.1.

[18] (*VbD*, p312)

different ways.[19]

The natural alternative sees verities as a semantic phenomenon: verities are new truth-values. The classical semantic principle of bivalence says that every sentence has exactly one of the two classical truth-values. So if verities are truth-values, bivalence fails. Now, the classically valid Law of Excluded Middle (LEM) licenses assertion of $\ulcorner A \vee \neg A \urcorner$ under any suppositions whatsoever, for any sentence A. Instances of bivalence follow from instances of LEM and the corresponding material conditionals $\ulcorner A \rightarrow \alpha$ is true$\urcorner$ and $\ulcorner \neg A \rightarrow \alpha$ is false$\urcorner$ (where α denotes A). So if verities are truth-values, those conditionals cannot all be true, given Edgington's retention of classical logic. However, Edgington offers an interesting way of avoiding this conclusion, by denying that verities and classical truth-values are competitors. She attempts to reconcile her probabilistic semantics with classical semantics by suggesting that those theories are different descriptions of the same underlying phenomenon. This section examines and rejects that suggestion.

To simplify presentation, I use $\ulcorner \mathsf{B}A \urcorner$ to abbreviate $\ulcorner \neg \mathsf{D}A \wedge \neg \mathsf{D} \neg A \urcorner$ (i.e. the formal counterpart of $\ulcorner$It's borderline whether $A \urcorner$). I also extend the object-language with a truth-predicate 'T' (restricted somehow to avoid contradiction) and quotation-names for its own sentences. In the interests of readability, I omit object-language quotation marks so that, e.g., 'T('$p \wedge q$')' becomes '$\mathsf{T}p \wedge q$'. As usual, falsity is identified with having a true negation. I abbreviate $\ulcorner \mathsf{T}$('$\neg A$')$\urcorner$ as $\ulcorner \mathsf{F}A \urcorner$. I ignore meaningless and non-truth-evaluable sentences. I also abstract away from contextual effects, writing as if all sentences have contents and truth-values *simpliciter*, rather than only declarative sentences as used in particular contexts to make assertions.

2.1 Edgington on degrees of truth

The thesis under examination is:

V=TV Verities are truth-values.

Edgington appears to reject **V=TV**, by arguing that verities are not degrees of truth.[20] This section shows that her argument is no threat to **V=TV**.

Here's how Edgington argues that verities are not degrees of truth:

The degrees of truth argument Suppose it's clearly borderline whether p: $\mathsf{DB}p$. Then it isn't clearly wrong to call 'p' true: $\neg \mathsf{D} \neg \mathsf{T}p$ But it is clearly wrong to say that 'p' has verity 1, that it's clearly true: $\mathsf{D} \neg \mathsf{D}p$. So verity 1 is not truth. So verities are not degrees of truth.

This argument relies on two assumptions that are worth noting before continuing. Firstly, that clear borderline status is a coherent notion. Secondly, that when 'p' is clearly borderline, '$\mathsf{T}p$' is also clearly borderline. The justification is presumably that when 'p' is borderline, so is '$\mathsf{T}p$'. Both assumptions might be questioned, though I won't do so here; the second is examined in more detail later (§3.2).

[19] (*VbD*, pp312–15)
[20] (*VbD*, p299)

The final step in the argument relies on the assumption that if verities are degrees of truth, then maximum verity is truth. On that picture, intermediate verities are degrees of closeness to truth, the greater the degree, the closer to truth. Although that is a version of **V=TV**, it is not the only one.

The degrees of truth argument shows at most that truth is not maximum verity. A parallel argument shows that falsity is not minimum verity. The argument thus threatens one way of locating classical semantic concepts within the framework of verities, by identifying truth and falsity with maximum and minimum verity. Since **V=TV** does not entail any identities between verities and classical truth-values, however, the degrees of truth argument does not threaten **V=TV** itself.

V=TV is in tension with Edgington's claim that "I do not see verity as disturbing or displacing the concept of truth."[21] By truth, Edgington appears to mean classical bivalent truth. **V=TV** replaces the two classical truth-values with many. Surely that counts as displacing the classical picture. However, another of Edgington's remarks suggests a resolution of this tension:

> "[W]e have two ways of classifying the same phenomena, each vague, and only vaguely related to each other. The [classical] way is simple, adequate when vagueness is not at issue, and too basic to be tampered with. The [verity] way is needed when vagueness is in focus: it enables us, we shall see, to solve the sorites paradox, in the context of a general theory of reasoning from vague premisses."[22]

Edgington is endorsing:

Classification Classical semantics and verity semantics are two ways of classifying the same phenomenon.

The following two sections evaluate and reject this interesting proposal. §2.2 offers an account of the phenomenon in question that is neutral between classical and verity semantics. This account will give substance to the notion of a truth-value, thus fleshing out **V=TV**. §2.3 then argues against **Classification**. It follows that the classical theory and verity theory cannot be non-competing descriptions of semantics and truth-values in the sense of §2.2. The goal is to show that at most one of those theories captures the word-world relations that underwrite validity.

2.2 Two ways of classifying what?

What is classical semantics a theory of? This section attempts to answer without prejudice for or against classical semantics. The short answer is: semantics is the theory of truth-value assignments. More detail follows. This detail will be important later, since I'll be distinguishing two conceptions of truth in §3.2, arguing that only one is relevant to semantics. Given a conception of semantics as the study of truth-conditions, the issue would be

[21] (*VbD*, p299)
[22] (*VbD*, p299)

terminological: two notions of truth would yield two notions of semantics. The substantive conception of semantics, truth and the relations between them provided in this section hopefully alleviates this threat of merely verbal dispute.

Semantics is the study of truth-value assignments. Truth-value assignments are ways of distributing truth-values across sentences permitted by the meanings of the connectives. Which connectives? That depends on our interests. Sometimes, our narrow interests may cover only narrowly logical connectives. On other occasions, our more inclusive interests may encompass, say, the '$\Box$' of necessity, or the 'D' of clarity, as when we study modality and vagueness. Our interests may even extend beyond connectives, to predicates like 'T' too. Whatever vocabulary concerns us, the goal is to capture the totality of truth-value assignments that its meanings permit, and to use that totality to characterise validity. Semantic theories will differ over how the meanings of various expressions constrain truth-value assignments, and over exactly how truth-value assignments relate to validity. For present purposes, we can set those differences aside. The difference between classical semantics and verity semantics that now concerns us lies in the truth-values themselves.

Classical semantics admits exactly two truth-values. Given **V=TV**, verity semantics admits many truth-values. What exactly are the truth-values about which these views disagree? They are relational properties of sentences, whose possession is jointly determined by: (i) the sentence's content; (ii) the state of reality.[23] I'll ultimately be arguing (in §3.2) that the best version of Edgington's view will take verities but not classical truth and falsity to be truth-values in this sense. I'll say a few words about content now before returning to truth-values.

The content of a sentence is the way it represents things as being: a sentence's content is that p iff it represents things as being such that p.[24] The analysis of representation is a deep and difficult matter that cannot be addressed here. It is, however, an adequacy condition on theories of content and vagueness that they ultimately mesh. Many-valued semanticists owe an account of representation that comports with their multiplicity of truth-values.

Classical semantics admits two content-determined ways for sentences to stand to reality. First way (truth): things are the way the sentence represents them as being. Second way (falsity): things are not the way the sentence represents them as being. The logical form of these locutions is a delicate matter. Classical semanticists prepared to avail themselves of quantification into sentence position have a ready answer:[25]

- Truth: $\lambda x[\exists P(x$ represents things as being such that P, and $P)]$.
- Falsity: $\lambda x[\exists P(x$ represents things as being such that P, and not-$P)]$.

The variable 'P' occupies sentence-position here; it is not a metalinguistic nominal variable ranging over sentences. Many-valued semanticists owe a similarly detailed account of their multiplicity of truth-values and of the *closer to clearly true than* ordering on them.

The most natural way to explain the many-valued semanticist's multiplicity of truth-values is in terms of the two classical ones, as degrees to which sentences possess the clas-

[23] (Cook, 2009)

[24] Better: iff speakers of the appropriate linguistic community use the sentence to represent things as being such that p.

[25] (Künne, 2003) uses quantification into sentence position in a similar analysis of linguistic truth.

sical truth-values. One sentence's being closer to clearly true than another can then be explained as the former's being true to a greater degree than the latter. Two points are worth noting. Firstly, on this view, intermediate verities are intermediate degrees of truth, and maximum verity is truth. This conflicts with Edgington's degrees of truth argument. As §3.1 and §3.2 will show however, that argument is not decisive. Secondly, this strategy requires the notion of degrees to which things are the way they are represented as being. It is not immediately obvious whether that is intelligible. An adequacy condition on attempts to explicate it is that things can be the way they are represented as being to only an intermediate degree, without thereby being inaccurately represented; for otherwise intermediate truth-values will merely be varieties of falsity. The most developed theory of degrees of truth that I know of, due to Nicholas J. J. Smith, meets this adequacy condition.[26]

On Smith's view, an atomic sentence '*Fa*' represents an object, denoted by '*a*', as having a property, denoted by '*F*'. Vagueness arises from a mismatch between the structures of representation and represented: instantiation comes in degrees, which are absent from the representation. Degrees of instantiation are used to analyse truth-values: '*Fa*' has truth-value n iff the denotation of '*a*' has the denotation of '*F*' to degree n. However, this simply relocates the explanatory burden to the theory of instantiation. Because degrees of instantiation are no more readily understood than degrees of truth, using the former to explicate the latter provides little explanatory advance. The availability of a fully satisfying account of degrees of truth remains an open question.

An important additional question concerns the nature of the differences between truth-value assignments.[27] On one view, different assignments result from differences in meaning. Another view invokes differences in what the world described by the language is like. Hybrids are also available. In line with Edgington's modal conception of validity, however, the second view will be assumed here. On this approach, different truth-value assignments result from different possible states of reality.

I said that semantics is the study of how the meanings of the logical constants—or any other privileged vocabulary—constrain truth-value assignments. Different truth-value assignments result from different possible states of reality. So semantics is the study of how the logical operations expressed by the logical constants—more generally: the semantic values of privileged vocabulary—constrain the possible states of reality, and how those constraints are reflected in the possible assignments of truth-values to sentences. The relevant notion of possibility will be more permissive than metaphysical possibility is typically taken to be. So far as the meanings of the logical constants are concerned, it's possible that $2+2 \neq 4$. The impossibility here lies with the numbers two, four and the *addition* operation, not with *conjunction*, *negation* or any other logical operation. Taking different truth-value assignments to result from different metaphysical possibilities would therefore impose constraints on truth-value assignments beyond those resulting from the meanings of the logical connectives alone, contrary to the present notion of semantics. Isolating the contribution of the meanings of the connectives requires a space of logical possibilities more inclusive than the metaphysical possibilities. Different choices of logical connective will deliver different

[26] (Smith, 2008, pp201–213)
[27] See (Etchemendy, 1990) for discussion.

spaces of logical possibilities.

This section has left many questions open. My goal, however, was not to settle every issue. My goal was to articulate the subject-matter of semantics in a manner amenable to both classical and many-valued semanticists, so that we can understand and evaluate **Classification** and **V=TV**. On the view proposed, semantics is the study of word-world relations, and the constraints imposed on such relations by the meanings of the object language logical constants (or whatever non-logical vocabulary concerns us). This conception of semantics and its connection with truth-value assignments places a substantive constraint on the nature of truth: truth-values are relational properties that sentences possess by virtue of their content and the state of reality. §3.2 will discuss a conception of truth that doesn't meet this constraint. Other notions of semantics may well be available, though they won't concern me here.

2.3 Against Classification

This section argues that **Classification** is false. Central classical semantic principles fail, if classical truth-values and verities provide two ways of classifying the same system of word-world relations. This leaves it open whether verities or classical truth and falsity are truth-values in the sense just outlined. §3 attempts to settle the matter in favour of verities.

How do assignments of verities and classical truth-values interact? Two plausible constraints on the relationship between a language's classical state c and verity state v are:

Upwards closure of truth Necessarily, for any sentences A, B, if $c(A) = \text{True}$ and $v(B) \geq v(A)$, then $c(B) = \text{True}$.

Downwards closure of falsity Necessarily, for any sentences A, B, if $c(A) = \text{False}$ and $v(B) \leq v(A)$, then $c(B) = \text{False}$.

Anything closer to clear truth (falsity) than a truth (falsheood) is also true (false). The classical truth-values are exclusive. So the closure principles define a mapping π from verities to truth-values such that, for all i:

- $\pi(i) = \text{True}$ iff, for some A: $v(A) = i$ and $c(A) = \text{True}$.
- $\pi(i) = \text{False}$ iff, for some A: $v(A) = i$ and $c(A) = \text{False}$.

Consider the *T-set* $\{i : \pi(i) = \text{True}\}$ and the *F-set* $\{i : \pi(i) = \text{False}\}$. The closure principles ensure that every member of the F-set is less than every member of the T-set. So those sets are disjoint. Given the following two assumptions, every verity belongs to the T-set or the F-set. (1) The classical semantic principle of bivalence: every sentence possesses some classical truth-value. (2) Every verity is possessed by some sentence. Since there are fewer sentences than verities, assumption (2) is problematic. The problem can be avoided by considering assignments of verity to open sentences, relative to assignments of values to variables.[28] Assuming that every possible shade between clear red and clear orange is actually instantiated, the open sentence 'Rx' possesses each verity relative to some assignment of value to 'x'.

[28] The right hand side of the first biconditional governing π becomes: iff, for some A and for some variable assignment α, $v(A, \alpha) = i$ and $c(A, \alpha) = \text{True}$. The second biconditional requires a similar modification.

We've just seen that the T-set and the F-set partition the verities, with the members of the former all greater than those of the latter. So the greatest lower bound of the T-set belongs to one or other of those sets. Let's assume it belongs to the F-set.[29] Say that a *truth-threshold* is any verity t that satisfies:

- $c(p) = \text{True}$ iff $v(p) > t$; and $c(p) = \text{False}$ iff $v(p) \leq t$.

Truth is verity greater than the truth-threshold t. The greatest lower bound of the T-set is a truth-threshold. So the closure principles imply:

Threshold There is a truth threshold.

The argument for **Threshold** relies only on the exclusivity of classical truth and falsity, the closure principles, and assumptions (1) and (2) above.[30] Only (1) is controversial in the present context. But (1) is bivalence, hence part of classical semantics. So **Classification** implies **Threshold**. I'll now argue that **Threshold** is incompatible with classical semantics, and hence that **Classification** is false.

As far as the argument for **Threshold** goes, each verity assignment may be compatible with many classical assignments. Bivalence and the closure principles imply that a truth threshold exists without settling its location. None of the arguments below assume that a language's verity state uniquely settles its classical state.

One may interpret Edgington as rejecting **Threshold** when she says:

> "A sufficient condition for the truth of A is that $v(A) = 1$; a sufficient condition for the falsity of A is that $v(A) = 0$. Beyond that, there is no precise mapping."[31]

However, there is a complication with this interpretation of Edgington. She only explicitly rejects the existence of a precise mapping from verities to classical truth-values. Taking 'precise' to mean non-vague, this is consistent with the existence of a vague such mapping: for some verities, it's vague whether π maps them to True or to False. This fits Edgington's earlier quoted suggestion that verity assignments and classical assignments are both "vague, and only vaguely related to each other."[32] Properly evaluating this suggestion would require a theory of vague mappings. Fortunately, we needn't provide one here. The arguments below assume only that a truth threshold exists, not that its location is precise. Any (total) mapping from verities to truth-values that respects the closure principles verifies **Threshold**. So the following arguments show that Edgington should deny the existence of any mapping whatsoever, whether vague or precise.

The closure argument

Threshold implies that truth isn't closed under conjunction and falsity isn't closed under disjunction. In classical semantics, truth is closed under conjunction and falsity is closed under disjunction. So **Threshold** is incompatible with classical semantics.

[29] Nothing turns on this: if the assumption is false, redefine 'truth-threshold' by substituting '$\geq$' and '$<$' for '$>$' and '$\leq$' respectively.

[30] The argument also requires the verity-theoretic analogue of bivalence: every sentence possesses some verity. Otherwise, the truths (or falsehoods, or both) may outstrip the sentences with any given kind of verity. I take this assumption to be uncontroversial in the present context.

[31] (*VbD*, p299)

[32] (*VbD*, p299)

To see why **Threshold** implies that truth isn't closed under conjunction, suppose 'p' and 'q' are independent. Then $v(q|p) = v(q)$. So:

- $v(p \wedge q) = v(p) \times v(q|p) = v(p) \times v(q)$.

Suppose:

- Truth threshold $t = 0.7$.
- $v(p) = v(q) = 0.75$.

Then:

- $v(p \wedge q) = 0.75 \times 0.75 = 0.56$.

Since $v(p)$ and $v(q)$ are both above the threshold, 'p' and 'q' are both true. Since $v(p \wedge q)$ is below the threshold, '$p \wedge q$' is not true but false. So truth isn't closed under conjunction, and a false conjunction has true conjuncts. This counterexample to the modal validity of conjunction introduction also undermines Edgington's claim to preserve classical logic.

This problem doesn't arise if $t = 1$. However, that generates three other problems. Firstly, it conflicts with Edgington's degrees of truth argument in §2.1. Secondly, all verities other than 1 become varieties of falsity. Thirdly, falsity cannot be identified with having a true negation: when $0 < v(p) < 1 = t$, neither 'p' nor '$\neg p$' exceeds the truth-threshold; so if bivalence holds, 'p' is false despite '$\neg p$' being untrue.

To see why falsity isn't closed under disjunction, suppose 'p' and 'q' are independent. Then:

- $v(p \vee q) = v(p) + v(q) - v(p \wedge q) = v(p) + v(q) - (v(p) \times v(q))$.

Suppose:

- Truth threshold $t = 0.7$.
- $v(p) = v(q) = 0.65$.

Then:

- $v(p \vee q) = 0.65 + 0.65 - 0.65^2 = 1.3 - 0.42 = 0.88$.

Since $v(p)$ and $v(q)$ are both below the threshold, 'p' and 'q' are both false. Since $v(p \vee q)$ is above the threshold '$p \vee q$' is not false but true. So falsity isn't closed under disjunction, and a true disjunction has false disjuncts.

The T-implication argument

Threshold invalidates the argument from 'p' to '$\mathsf{T}p$'. Classical semantics validates that argument. So if **Threshold** is true, classical semantics is false.

To show that **Threshold** invalidates the argument from 'p' to '$\mathsf{T}p$', we need a semantic clause for 'T'. The natural candidate is:

- $v(\mathsf{T}p) = 1$ iff $v(p) > t$; $v(\mathsf{T}p) = 0$ otherwise.

This ensures that 'T' serves as an object-language device for reporting the presence of what the meta-theory regards as truth: 'Tp' is clearly true when 'p' exceeds the truth threshold, and clearly false otherwise. Now, suppose $0 < v(p) < t$. By the clause just given: $v(\mathsf{T}p) = 0$. So 'Tp' has maximum unverity: $1 - v(\mathsf{T}p) = 1$. But 'p' does not have maximum unverity: $1 - v(p) < 1$. So the unverity of 'Tp' is greater than the unverity of 'p'. So the argument from 'p' to 'Tp' lacks the VCP, and, by **Valid≡VCP**, is therefore invalid on Edgington's semantics.

One might object to this clause for 'T' on the grounds that the boundary between truth and untruth is vague, and hence that $v(\mathsf{T}p)$ shouldn't sharply shift from 0 to 1 when $v(p)$ exceeds t. The natural implementation of this idea allows $v(\mathsf{T}p)$ to increase from 0 to 1 gradually, as $v(p)$ approaches and exceeds t. So suppose $v(\mathsf{T}p)$ begins to increase from 0 when $v(p)$ exceeds $t - \varepsilon$. Then the argument goes through as before, if we vary the initial supposition to: $0 < v(p) < t - \varepsilon$. Nothing has been gained.

Why does classical semantics validate the argument from 'p' to 'Tp'? Modal validity, recall, is necessary truth-preservation. The point of 'T' is to report the presence of that which valid arguments preserve. So 'Tp' will have that feature whenever 'p' does; that is, the argument from 'p' to 'Tp' is truth-preserving. Since that rests on no contingent assumptions, the argument from 'p' to 'Tp' is necessarily truth-preseving.

Note how radical a view this is, even by the standard of non-classical semantics. Supervaluation semantics has familiarised us with the idea that a material conditional '$p \to \mathsf{T}p$' might not be true.[33] However, that's not because supervaluation allows 'p' to be true without 'Tp' being true. Rather, supervaluation makes the whole conditional neither true nor false when 'p' is neither true nor false. Supervaluation thus validates the argument from 'p' to 'Tp'. Supervaluationists may therefore retain the biconditional 'p iff Tp' by interpreting 'iff' as mutual entailment. The T-implication argument shows that defenders of **Threshold** cannot. I'll discuss this kind of issue in more detail later (§3.2).

These arguments expose the radical consequences of **Threshold**. The closure argument shows that central classical semantic principles fail. The T-implication argument shows that the logical properties of truth are radically non-classical. So **Threshold** is incompatible with classical semantics, hence also with **Classification**. Defenders of **Classification** must reject the closure principles from which **Threshold** follows. That, I submit, is a deeply unattractive view. Although it allows one to regard bivalent truth and verity as ways of classifying the same phenomenon, it is obscure why one would wish to do so, given the non-classical conceptions of truth and semantics that result.

The option remains of regarding classical semantics as a limiting case of verity semantics, in the sense of being perfectly adequate when borderline cases are ignored. The second and third sentences in the quote from Edgington on p9 might be interpreted as suggesting this kind of view. Yet this is an unsatisfying defence of classical semantics and **Classification**. It amounts to an admission that borderline cases are counterexamples to bivalence. This is not a way for the classical theory and the verity theory to be non-competing descriptions of a single semantic phenomenon.

Edgington is guided by a parallel between credence and the lottery paradox on the

[33] The *locus classicus* is (Fine, 1975).

one hand, and vagueness and the sorites on the other. An analogue of **Classification** for credence and belief is available: credence and belief co-classify the same doxastic phenomenon.[34] Analogues of **Threshold** and the closure argument fit quite nicely with that view. Arguably, lotteries show that rational belief isn't closed under conjunction: one can rationally believe, of each ticket-holder, that he won't win, without believing the conjunction of those beliefs. The lesson is that Edgington's guiding parallel can only take us so far.[35]

3 Vagueness as *sui generis*?

I've argued that **Classification** is false: the classical theory and verity theory are not non-competing descriptions of the same semantic phenomenon. At most one of those theories captures the semantics of vague language. In more recent work, Edgington also rejects **Classification**. On her new view, however, the theory of verities is not a semantic theory at all.[36] This section examines and rejects this proposal.

Drawing on a proposal of David Barnett's, Edgington suggests that vagueness is *sui generis*, something that cannot "be illuminatingly understood as a species of some more general phenomenon".[37] Call this view primitivism. Primitivism is incompatible with **V=TV**, the thesis that verities are truth-values in the sense of §2.1; for **V=TV** makes borderline status a species of the more general phenomenon of semantic evaluation.

Edgington retains classical logic. So borderline status is compatible with LEM: if it's borderline whether *p*, then *p* or not-*p*, it's just not clear which. Edgington retains bivalence in a similar manner: if it's borderline whether *p*, then '*p*' is true or false, it's just not clear which. This is just what primitivism should lead us to expect; for if vagueness isn't a semantic phenomenon, then borderline status shouldn't interfere with classical semantic evaluation, or with the classical logic it validates. Primitivism thus provides a way to make classical logic and semantics compatible with borderlines status.

Despite preserving classical semantics and logic, primitivism is methodologically unattractive. We should be content to take *X* as primitive only when two conditions are satisfied. Firstly, we can use *X* to provide satisfying accounts of other phenomena. That, for example, is why knowledge is plausibly taken as primitive.[38] Nothing similar applies here; clarity's explanatory connections beyond the theory of vagueness are thin at best. Secondly, no satisfying analysis of *X* should be available. Since **V=TV** promises one such analysis, we should be content with primitivism only if **V=TV** is problematic.

Edgington and Barnett present two arguments that appear to threaten **V=TV**. Both arguments purport to show that bivalence should hold without restriction, even in borderline cases. It appears to follow that we should reject **V=TV** because it makes the multitude of verities incompatible with bivalence. This section shows that these arguments against **V=TV**

[34] (Edgington, 1997, p299) endorses that thesis.
[35] Rosanna Keefe's contribution to this volume discusses the putative parallel between prefaces, lotteries and vagueness in more detail.
[36] (Edgington, MS, §§1-4), (Edgington, 2010, §5.6).
[37] (Edgington, MS, §1, footnote 1, MS p2). (Barnett, 2009).
[38] (Williamson, 2000), (Hossack, 2007).

fail. The problem is that bivalent truth and falsity needn't be truth-values in the sense of §2.2; they might play some other, non-semantic, theoretical role. Indeed, I will argue that Edgington and Barnett's arguments for bivalence presuppose that they do.

§3.1 outlines Edgington and Barnett's arguments for bivalence. The first argument is invalid unless supplemented by a substantive principle about truth. That principle also provides the best justification for one of the second argument's key inferences. §3.2 argues that this principle is incompatible with the conceptions of semantics and truth outlined in §2.2. Rather than presenting a problem for **V=TV**, Edgington and Barnett's arguments for bivalence require a non-semantic conception of bivalent truth. An appropriate such conception is outlined that satisfies the principle underwriting the arguments for bivalence.

Primitivism and the view I'll be defending are structurally similar: bivalent truth and vagueness are fundamentally different phenomena. This renders borderline status compatible with bivalence. The views differ over the theoretical roles of verities and classical truth. The primitivist says: vagueness and verities are *sui generis*, to be understood in their own terms or not at all; semantic evaluation involves classical truth.[39] The view I'll be defending says: verities are truth-values, i.e. relational properties possessed by virtue of content and the state of reality; classical truth plays a merely expressive, non-semantic role.[40]

3.1 Bivalence at the border

Although Edgington does not discuss **V=TV** directly, she does argue against views on which borderline cases present counterexamples to bivalence. **V=TV** replaces the two classical truth-values with many, assigning intermediate values to borderline cases. So on **V=TV**, borderline cases are counterexamples to bivalence. Edgington's arguments thus appear to threaten **V=TV**. These arguments also appear to motivate primitivism, by showing that borderline status shouldn't be understood as a semantic phenomenon that interferes with bivalence. I'll argue that these appearances are illusory. Edgington's arguments do show something interesting, but not something that conflicts with **V=TV**.[41]

First argument for borderline bivalence '$p \vee \neg p$' is classically valid, hence valid on Edgington's view. So if the following material conditionals are true, so is the corresponding instance of bivalence:

$(\mathsf{T} \rightarrow)$ $\qquad p \rightarrow \mathsf{T}p \qquad \neg p \rightarrow \mathsf{F}p$

So if 'p' is neither true nor false, one of those conditionals is untrue. Edgington rejects this because "it follows that either [p] but it's not true that [p], or [not-p] but it's not

[39] Primitivism could in principle be combined with the non-semantic, expressive conception of classical truth outlined in §3.2. An account is then required of the truth-values in the sense of §2.1. However, the end of §3.2 claims that classical semanticists are in the privileged position of allowing the expressive and semantic roles to coincide.

[40] My own preference is for a different option, that there isn's space to discuss here. This view treats vagueness as a metasemantic phenomenon, residing in the relation between meaning-determining linguistic behaviour and content, rather than interfering with classical content and alethic evaluation. Verities measure how well assignments of content fit with meaning determining behaviour.

[41] The arguments considered in the text are from (Edgington, 2010, p104) and (Edgington, MS, §2); versions are also found in (Barnett, 2009, pp14-16).

false that [p]."[42] But "what more could it take to make it true that [p] than [p]?"[43]

Second argument for borderline bivalence Suppose it's clearly borderline whether p: $\mathsf{DB}p$. If borderline cases are counterexamples to bivalence, then 'p' clearly isn't true: $\mathsf{D}\neg\mathsf{T}p$. But "this is in tension with our ambivalence about the question whether ['p'] is true, our temptation to see it as 'sort of' true."[44] When it's clearly borderline whether p, it should be clearly borderline whether 'p' is true, not clear that it isn't. So when it's borderline whether p, 'p' should be borderline true, rather than untrue. Likewise for not-p and falsity. So borderline cases shouldn't be counterexamples to bivalence.

In both arguments, the choice of 'p' was arbitrary. So we can generalise to bivalence. What excatly do these arguments show?

The first argument is invalid in some many-valued settings. In classical semantics, untrue material conditionals are false; their antecedents are true and their consequents false. Yet in a many-valued setting, an untrue material conditional may not be false. Many-valued semanticists can allow untrue '$p \to \mathsf{T}p$' without true '$p \wedge \neg\mathsf{T}p$', by denying that untruth of a material conditional requires true antecedent plus false consequent; untrue antecedent with false consequent may suffice. This is what happens in standard supervaluationist semantics, where (super)truth-valueless antecedent plus (super)false consequent suffices for (super)truth-valuelessness of the whole conditional. Although '$p \vee \neg p$' is (globally) valid, the corresponding instance of bivalence can fail because '$p \to \mathsf{T}p$' is neither (super)true nor (super)false when 'p' is neither (super)true nor (super)false.[45] As Edgington and others have noted, verities can be modelled by proportions of sharpenings.[46] So this problem with Edgington's first argument for borderline bivalence afflicts her verity-theoretic setting too.

Can we fix the argument? Not simply by replacing the material conditionals ($\mathsf{T}\to$) with logical implications. Supervaluation semantics validates the arguments from 'p' to '$\mathsf{T}p$' and from '$\neg p$' to '$\mathsf{F}p$', though bivalence still fails. The reason is that (global) validity is necessary preservation of (super)truth, and '$p \vee \neg p$' can be (super)true when neither disjunct is; both '$\mathsf{T}p$' and '$\mathsf{F}p$' are (super)false in such cases, and so ($\mathsf{T}\to$) lack (super)truth-value. This is consistent with '$\mathsf{T}p$' being (super)true whenever 'p' is, and likewise for '$\mathsf{F}p$' and '$\neg p$'. On this view, no more is required for 'p' to be true than that p.

Of course, one might simply insist on the material conditionals ($\mathsf{T}\to$). That amounts to treating them as logical truths in the following sense: the meanings of the logical connectives and 'T' (and quotation names) preclude truth-value assignments on which ($\mathsf{T}\to$) are untrue. Since '$p \vee \neg p$' is logically true, the corresponding instance of bivalence is too. Generalising: every instance of bivalence is a logical truth. Indeed, instances of bivalence are inter-derivable with (corresponding sentences of the form) ($\mathsf{T}\to$) given LEM and the uncontroversial converses of (those sentences of the form) ($\mathsf{T}\to$).

[42] (Edgington, 2010, p104)
[43] (Barnett, 2009, p15)
[44] (Edgington, 2010, p104)
[45] I'm assuming that '$\mathsf{T}p$' is true on a sharpening iff 'p' is supertrue—i.e. true on all sharpenings—and hence that '$\mathsf{T}p$' is superfalse when 'p' is anything other than supertrue.
[46] (Edgington, 1997, pp315-6), (McGee & McLaughlin, 1994, §5). Note that Edgington sees the parallel with supervaluation semantics as formally but not philosophically/metaphysically helpful.

Treating every sentence of the form ($\mathsf{T} \rightarrow$) as valid embodies a substantive commitment about truth. The converse conditionals are uncontroversially valid. So every instance of these material biconditionals is valid too:

($\mathsf{T} \leftrightarrow$) $\qquad A \leftrightarrow \mathsf{T}A \qquad \neg A \leftrightarrow \mathsf{F}A$

The material biconditional is a test for identity of truth-value. If $\ulcorner A \leftrightarrow B \urcorner$ is logically true, it's logically impossible for A and B to receive different truth-values; the meanings of the logical constants alone preclude states of reality to which the contents of A and B stand in different ways. Now, if 'T' isn't counted a logical constant, instances of ($\mathsf{T} \leftrightarrow$) are not logically true: so far as *conjunction*, *negation* etc. are concerned, the atomic predication '$\mathsf{T}p$' may have a different truth-value from 'p'. On the present proposal, the meaning of 'T' (and quotation names) suffices to exclude such states. The meaning of 'T' is therefore captured by some function f from sentences to sentence-contents such that, for any sentence A and possibility w: $f(A)$ and the content of A stand to w in the same way. That's the minimum required for the meaning of 'T' to preclude truth-value assignments that differentiate between A and $\ulcorner \mathsf{T}A \urcorner$. What function could f be? There appears to be only one candidate: $f(A) =$ the content of A. Unless that identity holds, it's mysterious what prevents states of reality from inducing truth-value assignments that falsifies instances of ($\mathsf{T} \leftrightarrow$). So on the present proposal, that treats the instances of ($\mathsf{T} \leftrightarrow$) as logical truths, the meaning of 'T' guarantees:

Identity For any sentence A, the content of $\ulcorner \mathsf{T}A \urcorner$ is the same as that of A.

What we have just seen is that a fixed up version of the first argument for borderline bivalence, one that's not invalid in the semantic setting Edgington's attacking, presupposes **Identity**.[47]

Consider the second argument for borderline bivalence. It assumes that clear borderline cases are possible. This is not trivial, but I won't question it here. The argument also assumes that when it's borderline whether p, we are (or perhaps should be) ambivalent about whether 'p' is true. Why should we grant that? I see two candidate reasons.

The first takes it as part of the data to be accommodated, part of our ordinary linguistic behaviour that semantic theory should capture. That may be questioned. When we take it to be borderline whether p, we certainly are ambivalent, confused and conflicted about whether p. This confusion about whether p might lead ordinary speakers to express similar attitudes about whether 'p' is true. We should not assume without argument, however, that ordinary speakers are right to do so, or that this is relevant to truth. Our interest is in the interaction of borderline status with truth, not with ordinary speakers' usage of 'true'. Ordinary uses of 'true' may not always express truth. Ordinary speakers are also often mistaken, about even the most commonplace of concepts. And truth is not the most commonplace of concepts. Truth has a complex theoretical role to play, one aspect of which is its place in semantic theory. The fundamental concept of semantic theory should not be held hostage to ordinary usage of 'true'.

[47] For an interesting response to a similar argument for **Identity**, see (Asher et al., 2009, §3.2), (Dever, 2009, esp. §§9, 12).

A better reason to think that it being borderline whether p entails borderline status in '$\mathsf{T}p$' is this: 'p' and '$\mathsf{T}p$' are always intersubstitutable *salve veritate*. Given this assumption, '$\mathsf{B}p$' implies '$\mathsf{BT}p$', and hence that the confusion and ambivalence characteristic of borderline status is appropriately directed towards '$\mathsf{T}p$' whenever appropriately directed towards 'p'.

Unrestricted substitutability entails **Identity**. One way to see this is by substituting '$\mathsf{T}p$' into one side of the logically true material biconditional '$p \leftrightarrow p$', to obtain a logically true instance '$p \leftrightarrow \mathsf{T}p$' of ($\mathsf{T} \leftrightarrow$). Since '$p$' was arbitrary, we can generalise to the logical truth of every instance of ($\mathsf{T} \leftrightarrow$). As argued above, that implies **Identity**. Alternatively, let $\ominus$ be a content-identity connective: '$p \ominus q$' is true iff 'p' has the same content as 'q'. Since '$p \ominus p$', intersubstitutability implies '$p \ominus \mathsf{T}p$'. Since 'p' was arbitrary, we can generalise to **Identity**.

In sum, the first argument for borderline bivalence is unpersuasive because invalid in the many-valued settings it is supposed to refute. Fixing the argument leads to **Identity**. **Identity** also underwrites the second argument for borderline bivalence. Consider too the degrees of truth argument in §2.1. The key inference there is from '$\mathsf{B}p$' (= '$\neg\mathsf{D}p \wedge \neg\mathsf{D}\neg p$') to '$\neg\mathsf{D}\neg\mathsf{T}p$'. Appeal to **Identity** is the most natural way to justify that. Rejecting **Identity** therefore allows identification of truth and falsity with verities 1 and 0 respectively, and of intermediate verities with intermediate degrees of truth.

This diagnosis of **Identity**'s place in Edgington's thought is reinforced by:

> "There is no reason to deny the equivalence of 'It is true that A' and 'A', or of 'It is false that A' and '$\neg A$'. If $v(A) = 0.5$, $v(\text{It is true that } A) = 0.5$. As $v(A \text{ or not } A) = 1$, so $v(\text{It is true that } A \text{ or it is false that } A) = 1$, even if each disjunct gets 0.5.…A principle which is stronger than bivalence is rejected: the principle that every proposition is either [clearly] true or [clearly] false, that every proposition has verity 1 or 0. We saw earlier: a disjunction can be [clearly] true, without either disjunct being [clearly] true."[48]

Commitment to **Identity** manifests here in the equivalence of 'It is true that A' with A. What should many-valued semanticists, and defenders of **V=TV** in particular, make of Edgington's use of **Identity**? The next section argues that they should take it to show that, on Edgington's view, classical truth and falsity don't concern the word-world relations that underwrite validity; their theoretical role differs from that of truth-values in the sense of §2.2.

3.2 Two roles for truth

Edgington's arguments for the compatibility of borderline status with bivalence presuppose **Identity**. It might therefore appear that many-valued semanticists should reject **Identity**. This section argues that the appearance is misleading. A different moral can be drawn instead: two theoretical roles associated with truth come apart.[49] I'll begin by describing these roles.

A truth-predicate allows linguistic expression of contents it would otherwise be impossible or impractical to express. Suppose I know that Keith has uncannily good judgement,

48 (*VbD*, p310)
49 (McGee & McLaughlin, 1994), (McGee & McLaughlin, 2000), (McGee, 2005) advocate a similar view.

so good in fact that he's incapable of error. Without a truth-predicate (or equivalent), the closest I can come to expressing this is to begin uttering an infinite conjunction: if Keith says that whales are fish, then whales are fish, and if Keith says that grass is blue, then grass is blue, and if Keith says that necessity is *a priority*, then necessity is *a priority*,.... Since the conjunction is infinite, however, I cannot finish expressing it. So I cannot communicate just how good Keith's judgement is. But it may be important for me to do so. If Dorothy is going to the races with Keith tomorrow, where she will be gambling with my money, it matters greatly to me whether she will bet the same way as Keith. A truth-predicate allows me to overcome this difficulty by wrapping up the infinitely many conjuncts into a finite universal generalisation:

- For any sentence x that Keith uses tomorrow to make an assertion, x as then used by Keith will be true.

A truth-predicate thus serves an expressive function; it allows linguistic communication of otherwise inexpressible contents. Call this first theoretical role the *expressive role* for truth.[50] A predicate F cannot occupy the expressive role unless: for any content-bearer A, 'Fy' is guaranteed to have the same content as A under an assignment of A to 'y'. So the expressive role for truth requires **Identity**.

Truth's second theoretical role lies in empirical semantics. It is a non-trivial matter what kind of worldly circumstances suffice for the truth of 'the sky is blue'. That expression could have been used in many different ways, to express many different contents. The way a concrete (or syntactically individuated) string represents things as being is contingent, determined largely by its use within a linguistic community. One aspect of empirical semantics is the association of expressions with worldly items and circumstances—i.e. with contents—on the basis of contingent features of use. One central goal of this enterprise is to delineate the worldly circumstances under which sentences are true (as they are used within the relevant linguistic community) on the basis of the meanings of their sub-sentential components plus syntactic structure. Truth is thus involved in a minimal standard of correctness for empirical semantics.[51] Call this the *semantic role* for truth.

It is not *a priori* that a single notion occupies both the expressive and semantic roles. The expressive role requires a certain type of predicate and arises from a practical need. The semantic role concerns the underlying structure of the language to which the expressive truth-predicate belongs, and the word-world relations involving expressions of the language, as determined by use. It should not be assumed without argument that an expressive truth-predicate fills the needs of empirical semantics. Because the arguments for borderline bivalence presuppose **Identity**, they target expressive truth most directly. We can therefore retain **V=TV** without faulting those arguments, by taking **V=TV** to concern semantic truth. Indeed, we should take it to do so anyway. The point of an expressive truth-predicate is not to classify sentences on the basis of how their content stands to reality (or a possible state thereof), but to form a sentence with the same content as another sen-

[50] For discussion of the expressive utility of truth, see (Azzouni, 2004, ch 1). Speakers of languages that permit quantification into sentence position have no need for a truth-predicate. Even if such quantification is intelligible, however, it does not appear to be present in natural language: *we* need an expressive truth-predicate.

[51] Another truth-involving standard of correctness is that the semantics get entailment relations correct.

tence A from an expression that denotes A. To the extent that an expressive truth-predicate does effect such a classification, it does so parasitically on the word-world relations arising from the contents of the sentences classified. Explicating those relations is the purpose of semantic truth. And that way of classifying sentences—on the basis of how their content stands to reality—is what I used in §2.2 to characterise the truth-value assignments that underwrite validity. Differentiating these roles for truth allows bivalent expressive truth to cohabit with many-valued semantic truth.

Why connect validity with semantic truth? Because an argument is valid when a certain relation holds between its premisses and conclusion at every possibility. To make sense of that, we need relations between sentences and such possibilities. These sentence-possibility relations are, in effect, what §2.2 used to explicate truth-values. They result from the association of strings with contents that underwrites semantic truth. One might reply that an expressive truth-predicate allows us to introduce appropriate sentence-possibility relations, by evaluating '$\mathsf{T}p$' relative to such possibilities (holding fixed the actual meaning of 'p'). However, insofar as an expressive truth-predicate allows us to simulate sentence-possibility relations, it does so parasitically on the contents of the sentences themselves, and the relations between sentences and possibilities induced by those contents. When 'T' is an expressive truth-predicate, the evaluability of '$\mathsf{T}p$' relative to a possibility w depends on the evaluability of the concrete string 'p' relative to w, which itself depends on (a) 'p' representing things as being a certain way, and (b) w adjudicating whether things are that way. The evaluability of '$\mathsf{T}p$' relative to w thus relies on a semantic notion of truth.

Another way in which this parasitism emerges is via the fact that an expressive truth-predicate does not require its satisfiers to be any distinctive way. An expressive truth predication '$\mathsf{T}p$' is (typically) not even about 'p'; it is about whatever 'p' itself is about. The evaluability of '$\mathsf{T}p$' relative to w therefore rests on the evaluability of 'p' relative to w. Since 'p' is an intrinsically meaningless concrete string, its evaluability relative to w relies on its being used in a contentful manner, which is just what semantic truth captures. Since validity involves evaluating sentences relative to possibilities, validity requires semantic truth.

An independent argument shows that predicates for the many-valued semanticist's truth-values will not serve the expressive role. Let 'F' be an object-language predicate for reporting the presence of verity n. Then 'Fp' should have maximum verity when $v(p) = n$ and minimum verity otherwise. But with an expressive truth-predicate 'T': $v(\mathsf{T}p) = v(p)$. We should therefore expect divergence between an expressive truth-predicate and the linguistic resources in which many-valued semantics is couched. Only in a two-valued setting should we expect these to coincide, where bivalent semantic truth-conditions exhaust content. The separation of expressive and semantic truth is no threat to many-valued semantics.

On this approach, classical bivalent truth plays no role in empirical semantics, or in the analysis of validity. It exists only to fill an expressive need. If LEM holds without restriction, then so does one form of bivalence: expressive bivalence. Instances of expressive bivalence are mere linguistic variants on instances of LEM: their content is the same. Expressive bivalence thus places no constraints on the semantics of vagueness, beyond validation of LEM. In constructing such a semantics, one may use as many truth-values as one likes.

Edgington's arguments for borderline bivalence do not refute **V=TV**. They do, however, force advocates of **V=TV** to distinguish the semantic and expressive roles for truth. The question remains whether verities or classical truth-values should occupy the semantic role. Given the unattractiveness of primitivism about vagueness, I conclude by suggesting that the best version of Edgington's view will treat verities as semantic truth-values, hence concerned with word-world relations, whereas classical bivalent truth serves a merely expressive role.

References

Asher, N., Dever, J. & Pappas, C. (2009), "Supervaluations debugged," *Mind*, 118(472): 901–33.

Azzouni, J. (2004), *Deflating existential consequence: a case for nominalism*, OUP.

Barnett, D. (2009), "Is vagueness sui generis?" *Australasian Journal of Philosophy*, 87: 5–34.

Cook, R.T. (2009), "What is a truth value and how many are there?" *Studia Logica*, 92(2): 183–201.

Dever, J. (2009), "The disunity of truth," in R.J. Stainton & C. viger, eds., *Compositionality, Context and Semantic Values: Essays in Honour of Ernie Lepore*, chapter 7, Springer Netherlands.

Edgington, D. (1997), "Vagueness by degrees," in R. Keefe & P. Smith, eds., *Vagueness: a reader*, chapter 16, MIT Press.

— (2010), "Sorensen on vagueness and contradiction," in R. Dietz & S. Moruzzi, eds., *Cuts and Clouds: Vagueness, its Nature and its Logic*, chapter 5, OUP.

— (MS), "Vagueness: filling the gaps," *unpublished*.

Etchemendy, J. (1990), *The concept of logical consequence*, Cambridge Mass., London England: Harvard University Press.

Fine, K. (1975), "Vagueness, truth and logic," *Synthese 30*, pp. 265–300. Reprinted in (Keefe & Smith, 1997, ch9/pp119–150).

Hossack, K. (2007), *The Metaphysics of Knowledge*, Oxford: OUP.

Keefe, R. (2000), *Theories of Vagueness*, Cambridge: Cambridge University Press.

Keefe, R. & Smith, P., eds. (1997), *Vagueness: a reader*, MIT Press.

Künne, W. (2003), *Conceptions of Truth*, OUP.

McGee, V. (2005), "Two conceptions of truth? – comment," *Philosophical Studies*, 124(1): 71–104.

McGee, V. & McLaughlin, B. (1994), "Distinctions without a difference," *The Southern Journal of Philosophy*, 33: 203–261.

— (2000), "The lessons of the many," *Philosophical Topics*, 28: 129–51.

Shapiro, S. (2006), *Vagueness in Context*, Clarendon Press: OUP.

Smith, N.J.J. (2008), *Vagueness and Degrees of Truth*, OUP.

Williamson, T. (1994), *Vagueness*, London, New York: Routledge.

— (2000), *Knowledge and its Limits*, OUP.