Supervaluationism and good reasoning

Timothy WILLIAMSON

ABSTRACT: This paper is a tribute to Delia Graff Fara. It extends her work on failures of meta-rules (conditional proof, RAA, contraposition, disjunction elimination) for validity as truth-preservation under a supervaluationist identification of truth with supertruth. She showed that such failures occur even in languages without special vagueness-related operators, for standards of deductive reasoning as materially rather than purely logically good, depending on a context-dependent background. This paper extends her argument to: quantifier meta-rules like existential elimination; ambiguity; deliberately vague standard mathematical notation. Supervaluationist attempts to qualify the meta-rules impose unreasonable cognitive demands on reasoning and underestimate her challenge.

Keywords: Delia Graff Fara, supervaluationism, supertruth, global validity, vagueness, sorites, ambiguity, meta-rule.

RESUMEN: Este artículo es un tributo a Delia Graff Fara. Extiende su trabajo acerca del colapso de las metareglas (prueba condicional, reductio ad absurdum, contraposición, eliminación de la disyunción) para la validez entendida como preservación de la verdad bajo una identificación de la verdad con superverdad. Graff Fara demostró que dicho colapso se da incluso en lenguajes que no poseen operadores especiales relacionados con la vaguedad, para estándares de razonamiento deductivo tanto material como lógicamente correctos, dependiendo de factores sensibles al contexto. En este artículo se extiende su argumento a: meta-reglas cuantificacionales, como la eliminación existencial; ambigüedad; notación matemática estándar deliberadamente vaga. Los intentos supervaluacionistas de cualificar las meta-reglas imponen exigencias cognitivas poco cabales sobre el razonamiento y subestiman sus retos.

Palabras clave: Delia Graff Fara, supervaluacionismo, superverdad, validez global, vaguedad, sorites, ambigüedad, meta-regla.

The word I most associate with Delia Graff Fara is 'sharp'—sharp philosophy, sharp wit, sharp Delia. I first met her in early 1994, as a visiting professor at MIT. She came to my graduate class on vagueness, based on the proofs of my forthcoming book (Williamson 1994). The other students included Jason Stanley (my TA), Zoltán Szabó, Kathrin Koslicki, and Steven Gross; Judith Jarvis Thomson and Paul Horwich came from the faculty. I can, I hope, take some of the credit for sparking Delia's interest in vagueness. She was cautiously quiet in class, but towards the end of the semester presented me with rough notes for what later developed into 'Phenomenal Continua and the Sorites' (Graff 2001). I was struck by her ruthless application of precise mathematics to what might look on the surface like hopelessly recalcitrant material. Of course, the brilliantly creative and ingenious ideas

about vagueness for which she is known are categorically hers. Still, I was gratified when I once suggested to her that her distinctively interest-relative theory of vagueness in 'Shifting Sands' (Graff 2000) has a layer of epistemicism in its foundations; she immediately agreed.[1]

Delia and her husband Michael became my good friends, as well as my co-authors (Graff and Williamson 2002; Fara and Williamson 2005). My paper on higher-order vagueness (Williamson 1999) emerged out of extended conversation and email exchanges with her; together, we worked through the drawbacks of one proposal after another about how to define nth-order vagueness. We had a constructive exchange on the mathematical consequences of epistemicist models of vagueness (Graff 2002; Williamson 2002). The more technical a problem got, the more Delia relished it, while never losing sight of the bigger philosophical picture. When I was invited to edit an anthology on vagueness, I asked her to co-edit it with me, because I trusted her judgment; also, her knowledge of the recent literature was better than mine. We wrote the introduction together (Graff and Williamson 2002).

In discussion, Delia was in the best way a stickler for accuracy and precision. If her point, typically a steely-sharp point, was not properly answered, she would politely but relentlessly insist. She would accept no diplomatic fudge. But she was the opposite of grim, the first to see the funny side. She was exactly the kind of person philosophy can least spare: a temptation fate would not resist.

The present paper takes up one of Delia's lesser-known contributions to the theory of vagueness, an insight whose significance has, I believe, been underestimated. I will explain her idea, illustrate extensions of it, and argue that attempts to deny it vastly underestimate the difficulty of meeting her challenge. It concerns *supervaluationist* theories of vagueness.

1. Background

Supervaluationists treat vagueness as a kind of semantic underdetermination. The community's use of its language fails to determine unique semantic values for its expressions. Many different assignments are in some (under-explained) sense compatible with use—even in a given context (this relativity to context will henceforth be left tacit). Call those assignments compatible with use *reasonable*. There is nothing vague in any one such assignment, considered in itself. A given sentence is either true on a given assignment or false on that assignment, and not both.[2] The vagueness consists in the multiplicity of reasonable assignments. In particular, a sentence is borderline if it is true on some reasonable assignments and false on others: its use and the world together underdetermine its truth-value. By contrast, a sentence is *supertrue* if and only if it is true on every reasonable assignment, and *superfalse* if and only if it is false on every reasonable assignment. The borderline sentences are those which are neither supertrue nor superfalse.

---

[1] Delia characterizes herself as an epistemicist at Graff 2003, 216. Elsewhere she comments that boundary-shifting theories of vagueness, whether contextualist or her own interest-relative invariantist theory, 'are not in competition with any particular account of what it is to be a borderline case of a vague predicate' (Graff Fara 2010, 379n13).

[2] For simplicity, the object-language is assumed to contain only declarative sentences.

Traditional supervaluationism identifies truth with supertruth and falsity with super-falsity, not with truth and falsity respectively on the intended assignment, because there is no intended assignment. Use determines only supertruth and superfalsity. Thus every theorem of classical logic is true, because supertrue, because true on every reasonable assignment. That includes all instances of the law of excluded middle. However, not every sentence of a vague language is either true or false, because borderline sentences are neither supertrue nor superfalse. Consequently, standard disquotational biconditionals for truth are held to fail.

A variant form of supervaluationism identifies truth with disquotational truth rather than supertruth (McGee and McLaughlin 1994). An ascription of disquotational truth to a sentence α is true on an assignment if and only if α itself is true on that assignment. This variant has the advantage over the traditional version of making standard disquotational principles for truth come out true, indeed supertrue. It also has the result that every sentence is either true or false. However, as a consequence, although its proponents are officially anti-epistemicist, they have the greatest difficulty in distinguishing their view from a form of epistemicism (Williamson 1994, 162-4; 2004). The present paper concerns only the more standard form of supervaluationism, which identifies truth with supertruth, and falsity with superfalsity. This view goes back to the 1950s (Williamson 1994, 143-6). A classic exposition is by Kit Fine (1975), though he no longer holds the view. A contemporary defender of standard-form supervaluationism is Rosanna Keefe (2000).

Speakers may wish to talk in the object-language about borderline status. The natural way to do that is by adding an operator D, read 'definitely', corresponding to supertruth, to the object-language. A borderline case is neither definitely on one side nor definitely on the other. The idea is roughly that, for any sentence α, $D\alpha$ is true on an assignment if and only if α is true on every reasonable assignment.

We can easily turn these ideas into a simple model theory for a formal propositional language with a D operator and a standard set of truth-functors. A model is an ordered pair <W, V>, where W is a nonempty set, whose members are known as *points* (rather than the *worlds* of modal logic); informally, we think of them as the reasonable assignments; V is a function mapping each atomic formula of the language to a subset of W, the set of points at which it is true. In a given model, if a formula α is true at every point, then $D\alpha$ is also true at every point in W; otherwise, $D\alpha$ is false at every point in W. The semantics for the truth-functors is as usual: $\neg\alpha$ is true at a point w if and only if α is not true at w, $\alpha \vee \beta$ is true at w if and only if α is true at w or β is true at w, and so on. Correspondingly, α is supertrue in the model if and only if α is true at every point, in which case so is $D\alpha$; α is superfalse in the model if and only if α is false at every point, in which case so is $D\neg\alpha$. Whereas models for modal logic often have one designated world, informally conceived as the 'actual' world, supervaluationist models have no designated point, because there is not meant to be an 'intended' world.

What is the appropriate model-theoretic relation of logical consequence in this setting? One proposal is that a valid argument is one that preserves truth at any given point in any given model, which is known as *local validity*:

Local validity

$\Gamma \vDash_L \alpha =_{def}$ for every model <W, V> and $w \in$ W, if γ is true at w in <W, V> for all $\gamma \in \Gamma$, then α is true at w in <W, V>.

Formally, local validity is a nice classical consequence relation, but it does not suit the spirit of standard supervaluationism. For validity is supposed to be truth-preservation, and standard supervaluationists equate genuine truth with supertruth, not with truth on a given assignment. For them, a genuinely valid argument is one that preserves supertruth in any given model, a different consequence relation known as *global validity*:

Global validity

$\Gamma \vDash_G \alpha =_{def}$ for every model <W, V>, if γ is true at w in <W, V> for all $\gamma \in \Gamma$ and $w \in$ W, then α is true at w in <W, V> for all $w \in$ W.

Kit Fine (1975), Rosanna Keefe (2000, 174) and many others have treated global validity as the appropriate consequence relation for supervaluationism.

As so often with vagueness, there is a catch: higher-order vagueness. 'Reasonable' is itself vague, which makes D vague too, so its semantic value and that of many sentences containing it should themselves vary from one reasonable assignment to another. In the simple semantics just sketched, in any given model, for any given formula α, $D\alpha$ is true at every point or none, so $DD\alpha \lor D\neg D\alpha$ is valid (both globally and locally). In other words, it is always definite whether something is definitely so. Yet 'definitely red' seems just as sorites-susceptible as 'red', and so on for other cases of vagueness. Some supervaluationists have denied that D can be expected to measure its own vagueness (Keefe 2000, 210). However, there is nothing obviously inconsistent or unintelligible in the hypothesis that it is indefinite whether something is definite.

A more natural strategy for the supervaluationist is to follow the example of modal logic, where Kripke showed how to implement model-theoretically its being contingent whether something is necessary, by introducing an accessibility relation R between worlds. Informally, a world x is accessible from a world w if and only if x would be possible if w obtained. Formally, $\Box\alpha$ is true at w if and only if α is true at every world x such that Rwx. Since two worlds may differ in which worlds are accessible from them, $\Box\alpha$ may be true at one and false at the other. The supervaluationist analogue is to implement model-theoretically its being indefinite whether something is definite, by introducing an accessibility relation R between points. Informally, an assignment x is accessible from an assignment w if and only if x is reasonable according to w. Formally, $D\alpha$ is true at w if and only if α is true at every point x such that Rwx. Since two points may differ in which points are accessible from them, $D\alpha$ may be true at one and false at the other. In effect, each assignment makes its own determinations of which assignments are reasonable.

The only formal constraint on the accessibility relation imposed here is reflexivity: every assignment is reasonable according to itself. This validates the plausible principle that whatever is definitely so is indeed so ($D\alpha \rightarrow \alpha$). If accessibility is also symmetric and transitive, then it is an equivalence relation, which validates the analogue for D of the strong modal logic S5; in particular, it validates the problematic theorem $DD\alpha \lor D\neg D\alpha$. However,

that principle can be shown to fail whenever accessibility is either non-symmetric or non-transitive, but still reflexive.[3]

Supervaluationist models are now triples <W, R, V>, where W and V are as before and R is the accessibility relation. Since the point of the D operator is to represent supertruth in the object-language, supertruth must now be defined at a point: in the model, α is supertrue at w if and only if $D\alpha$ is true at w, which in turn is so if and only α is true at all points x such that Rwx. Since validity is meant to be truth-preservation, and standard supervaluationists equate genuine truth with supertruth, in the setting of this more sophisticated model theory they should regard a genuinely valid argument as one that preserves supertruth at any given point in any given model. That is not global validity but what, following Cobreros (2008), we may call *regional validity* (Williamson 1994, 297 n32):[4]

Regional validity
$\Gamma \vDash_R \alpha =_{def}$ for every model <W, R, V> and $w \in$ W, if γ is true at w^* in <W, R, V> for all $\gamma \in \Gamma$ and w^* such that Rww^*, then α is true at w^* in <W, R, V> for all w^* such that Rww^*.

The main considerations in this paper are independent of higher-order vagueness. Although we leave room for it by not requiring accessibility to be an equivalence relation, the arguments still work even if that constraint is imposed. Indeed, they work for both regional validity and global validity; for convenience, some will be stated just in terms of regional validity. Although other accounts of validity have been proposed within a supervaluationist setting (Williamson 1994, 148; Varzi 2007; Williams 2008), the standard of regional validity is rooted in a very natural line of thought from a standard supervaluationist perspective. If supervaluationists abandon it, they incur the suspicion that they are not serious about their identification of truth with supertruth.

In the fragment of the language without the D operator, global validity, regional validity, and classical validity all coincide. For if the argument from Γ to α is classically invalid, then on some assignment all formulas in Γ are true while α is false; that assignment can be turned into a supervaluationist model with just one point, at which all formulas in Γ are supertrue while α is superfalse, so the argument is globally and regionally invalid too. Conversely, if the argument is classically valid, then it preserves truth on any given assignment, so it preserves truth at any given point in any given supervaluationist model, since the rest of the model makes no difference to the evaluation of D-less formulas, the argument is globally valid too.

Even in the full language with D, classical truth-functional validity still entails global and regional validity. More precisely, if the argument from Γ to α is classically valid and D-free, then it is also globally and regionally valid, so for any uniform substitution σ into the

---

[3]   One would not expect 'x is reasonable according to w' to be a transitive relation (Williamson 1994, 159). Indeed, arguments against the S4 principle $D\alpha \rightarrow DD\alpha$, corresponding to transitivity, can be modelled on some arguments against the analogous KK principle for knowledge, even within the supervaluationist framework (Williamson 1994, 217-226; 1999). For arguments against the B principle $\alpha \rightarrow D\neg D\neg\alpha$, corresponding to symmetry, see Mahtani 2008 and Dorr 2015. For a theory of vagueness on which transitivity holds but symmetry fails, see Bobzien 2011.

[4]   The distinction between global and regional validity is also involved in the debate between Cobreros (2008) and Graff Fara (2011) on higher-order vagueness.

full language, the argument from σΓ to σα is also globally and regionally valid, because both global validity and regional validity are closed under uniform substitution. In that narrow sense, the logics of global validity and regional validity are classical.

In a broader sense, however, the logics of global validity and regional validity are non-classical. For they violate some standard *meta-rules*, rules connecting the validity of several arguments. In particular, they violate meta-rules involving the discharge of premises (Williamson 1994, 151-2, 297n32). They include conditional proof, otherwise known as the deduction theorem and as the introduction rule for the conditional (if $\Gamma \cup \{\alpha\}$ entails β then Γ entails $\alpha \rightarrow \beta$), reductio ad absurdum (if $\Gamma \cup \{\alpha\}$ entails the contradiction ⊥, then Γ entails ¬α), argument by cases, otherwise known as the elimination rule for disjunction (if $\Gamma \cup \{\alpha\}$ entails γ and $\Delta \cup \{\beta\}$ also entails γ, then $\Gamma \cup \Delta \cup \{\alpha \vee \beta\}$ entails γ), and contraposition (if $\Gamma \cup \{\alpha\}$ entails β then $\Gamma \cup \{\neg\beta\}$ entails ¬α). Such rules are of course frequently used in ordinary deductive reasoning, not least in mathematics.

To see how the failure happens in the case of conditional proof, consider the formula $p \wedge \neg Dp$, where p is atomic. Suppose that, in a model <W, R, V>, $p \wedge \neg Dp$ is supertrue at some $w \in$ W. Thus p is supertrue at w, so p is true at all w^* such that Rww^*. Similarly, $\neg Dp$ is supertrue at w, so $\neg Dp$ is true, and Dp false, at all w^* such that Rww^*. But R is reflexive, by hypothesis, so Rww, so Dp is false at w, so p is false at some w^* such that Rww^*, by the semantic clause for D. Thus p is both true and false at some w^*: a contradiction. So $p \wedge \neg Dp$ is not supertrue at any point in any model. Consequently, the argument from $\{p \wedge \neg Dp\}$ to any conclusion, including ⊥, is vacuously regionally valid: $\{p \wedge \neg Dp\} \vDash_R \bot$. Hence, if $\vDash_R$ obeys conditional proof, $\vDash_R (p \wedge \neg Dp) \rightarrow \bot$, so $\vDash_R p \rightarrow Dp$, since the two formulas are truth-functionally equivalent. But a supervaluationist counter-model to $p \rightarrow Dp$ can be built from any accessibility relation that holds between two points, so not $\vDash_R p \rightarrow Dp$. Thus regional validity violates conditional proof. By a similar argument, global validity also violates conditional proof. Under supervaluationism, restricting the models to validate conditional proof would mean ruling out vagueness.

Similar arguments show that global validity and regional validity also violate reductio ad absurdum, argument by cases, and contraposition, since conditional proof is derivable from any one of those meta-rules in this setting.[5, 6]

---

[5]   Proof: Suppose that $\Gamma \cup \{\alpha\}$ entails β. Hence $\Gamma \cup \{\alpha \wedge \neg\beta\}$ entails ⊥. Hence, by reductio ad absurdum, Γ entails $\neg(\alpha \wedge \neg\beta)$; equally, by contraposition, $\Gamma \cup \{\neg\bot\}$ entails $\neg(\alpha \wedge \neg\beta)$; in both cases, Γ entails $\alpha \rightarrow \beta$. Also, from the original supposition, $\Gamma \cup \{\alpha\}$ entails $\alpha \rightarrow \beta$. Moreover, $\{\neg\alpha\}$ entails $\alpha \rightarrow \beta$. Hence, by argument by cases, Γ entails $\alpha \rightarrow \beta$. The proofs assume that any set of premises entails its classical truth-functional consequences and that entailment obeys the standard structural rules for logical consequence, in particular the cut rule; these assumptions hold for both global and regional validity.

[6]   One can also consider an alternative consequence relation that holds between premises and conclusion if and only if (i) whenever the conjunction of the premises is supertrue, the conclusion is superfalse and (ii) whenever the conclusion is superfalse then the conjunction of the premises is superfalse. However, since $p \wedge \neg Dp$ is never supertrue, $\neg(p \wedge \neg Dp)$ is never superfalse, so $p \wedge \neg Dp$ entails $\neg(p \wedge \neg Dp)$ vacuously in this new sense; thus $(p \wedge \neg Dp) \rightarrow \neg(p \wedge \neg Dp)$ and so $p \rightarrow Dp$ would be valid if this new sense validated conditional proof, which it consequently does not. Moreover, since $p \wedge \neg Dp$ entails each of $p \wedge \neg Dp$ and $\neg(p \wedge \neg Dp)$ but not their conjunction in the new sense, it does not provide the conjunction introduction rule.

The meta-rules are defined in terms of the given consequence relation for the language, in this case global or regional validity. Thus whether the meta-rules hold for a language supervenes on the consequence relation for that language; no difference in the former without a difference in the latter. For the D-free fragment, they all hold for global and regional validity because they coincide with classical truth-functional validity, for which they hold. They can fail for global and regional validity in the language with D only because there no such coincidence obtains.[7]

Some confusion arose about the status of the meta-rules for global validity because Kit Fine originally illustrated the failure of conditional proof for global validity with the argument from p to Dp (Fine 1975: 290), which is globally valid, even though the formula $p \rightarrow Dp$ is not globally valid. By contrast, Dp is not a regional consequence of p, since the supertruth of p at a point does not imply the supertruth of Dp at that point in a model with a non-transitive accessibility relation. This misled some authors (not Fine) into denying that the meta-rules fail once higher-order vagueness is taken into account.[8] The example used earlier shows that all the meta-rules at issue fail for both global and regional validity even when account is taken of higher-order vagueness (Williamson 1994, 297n32). The proof works when the accessibility relation is (harmlessly) constrained to be reflexive in all models, whether or not it is also constrained to be symmetric and transitive.

As usual, one can complicate the model theory and the models to avoid the inconvenient results, for instance by making the argument from $\{p \land \neg Dp\}$ to $\bot$ invalid in another sense. Such effects are typical artefacts of building too many degrees of freedom into one's models.[9] On a more appropriate model theory for supervaluationism, like that above, the four meta-rules at issue do indeed fail for the relevant standard of validity.

Supervaluationists have used their semantic framework to provide semantic accounts of comparative and superlative adjectives and of modifiers such as 'rather', 'in a sense', and '-ish', starting with work by David Lewis (1972, 216) and Hans Kamp (1975, 137-45). Such accounts generate further violations of the meta-rules for global and regional validity (Williamson 1994, 295-6n25). Thus the violations are not confined to arguments with a 'definitely'-like operator; they also occur in the presence of more innocuous-looking natural language constructions.

Reasoning to a standard of global or regional validity seems to be a tricky business. Some supervaluationists have responded by proposing non-standard variants of the me-

---

7 Unclarity on which object-language the meta-rules concern vitiates the objection in Hjortland 2017, 639-40 to the account of logical consequence in Williamson 2013, 92-5 and 2017.

8 See McGee and McLaughlin 1997, 224, to which Williamson 2004, 119-20 replies; they later acknowledged the slip (McGee and McLaughlin 2004, 134).

9 An example is Williams 2008, where each model has both an accessibility relation over all delineations (points) and an independently specified subset of delineations, whose members are called *sharpenings*. Supertruth in a model is defined as truth on all sharpenings, and 'global' validity as preservation of supertruth in each model. Key arguments then exploit models where, for each delineation, the set of accessible delineations differs from the set of sharpenings; thus the model's own view of which delineations are 'reasonable' differs from that of any delineation countenanced within the model. This flouts the point of the D operator, which is to project supertruth into the object-language.

ta-rules that avoid such counterexamples. For instance, Rosanna Keefe suggests formulations in which D is inserted at appropriate points into the concluding argument (2000, 179-81). Her modified form of conditional proof states that if $\Gamma \cup \{\alpha\}$ entails β then Γ entails $D\alpha \rightarrow \beta$, which holds for global validity in the simple models with no accessibility relation, and for regional validity in the models with such a relation. However, it fails for global validity in the more complex models, since the argument from p to DDp is globally valid, but $Dp \rightarrow DDp$ is not. By contrast, McGee and McLaughlin avoid inserting D by in effect restricting the reasoning from $\Gamma \cup \{\alpha\}$ to β to 'modes of inference that are known to preserve truth in each acceptable model' (2004, 135; their framework is slightly different from mine). Williams (2008) suggests a version where the argument from $\Gamma \cup \{\alpha\}$ to β is valid by one standard while the argument from Γ to $\alpha \rightarrow \beta$ is valid by another.

Is ordinary deductive reasoning in science—for instance, in drawing consequences from a theory—sensitive to any of the proposed restrictions on the meta-rules? That is far from obvious. Still, supervaluationists might claim that the restrictions are rarely needed, and if we occasionally violate them, that is our fault.

The debate has sometimes focussed on whether the contested meta-rules are 'part of classical logic'. That has not been very fruitful. They are in a broad sense of 'classical logic', but not in a narrower sense. A more interesting question is how much of our deductive practice will be lost if the meta-rules are restricted.

2. Graff Fara's contribution

Graff Fara discusses the more interesting question in the latter half of her seminal paper 'Gap Principles, Penumbral Consequence, and Infinitely Higher-Order Vagueness' (Graff 2003). A component of the supervaluationist defence had been that the problems for the meta-rules arise only when operators that trade on vagueness are in play: for instance, in English, 'definitely', 'rather', 'in a sense', '-ish', and perhaps '-er'. As Rosanna Keefe put it, 'the cases in which these rules fail all involve the D operator (or similar such devices)' (2000: 178). Graff Fara showed that this was not so.

Graff Fara's examples involve arguments in ordinary English, evaluated according to supervaluationist semantics.[10] The standard of validity she applies is not meant to be purely logical; rather, she intends a lower standard, one more relevant to reasoning in at least some contexts. She assumes, very plausibly, that in much of our deductive reasoning we intend the truth of the premises in some sense to guarantee the truth of the conclusion, but we do not care whether the guarantee is purely logical or instead depends on a shared conversational background. For example, the argument from 'It's red' to 'It's coloured' is good enough for the purposes of most contexts, even though its validity is not strictly *logical*. It might be called a standard of *materially good* deduction. In model-theoretic terms, the supervaluationist may naturally conceive such a standard of material goodness as restricting the models, and points within those models, with respect to which regional validity is assessed.

---

[10] Her model-theoretic framework differs slightly from the present one, but not in ways that matter for present purposes.

Here is one of Graff Fara's examples (Graff 2003, 208). First consider argument IIB (her tag):

> *Argument IIB*
>   *X* is tall.
>   *Y* is not tall.
>   Therefore, *X* is more than a millimetre taller than *Y*.

In some contexts, the supertruth of the premises of IIB guarantees the supertruth of its conclusion. For, given how vague 'tall' is, the borderline zone is definitely more than a millimetre wide (if you doubt that, Graff Fara invites you to work with a smaller distance). Hence, if *X* is definitely tall and *Y* is definitely not tall, then *X* is definitely more than a millimetre taller than *Y*. Thus IIB is regionally valid in the relevant material sense. Now try contraposing the second premise and the conclusion of IIB, and deleting the double negation in the conclusion (the latter being a harmless move for a supervaluationist). The result is IIA:

> *Argument IIA*
>   *X* is tall.
>   *X* is no more than a millimetre taller than *Y*.
>   Therefore, *Y* is tall.

Argument IIA is *not* regionally valid in the relevant material sense. If *X* is definitely tall, and definitely no more than a millimetre taller than *Y*, *Y* may still not be definitely tall. As Graff Fara points out, IIA generates a sorites paradox for tallness, even by supervaluationist standards, while IIB does not.[11] Thus contraposition fails from IIB to IIA.

Strictly speaking, both arguments involve the comparative operators 'more' and '-er', whose semantics may be given in a supervaluationist framework. This is an inessential feature. We could replace the conclusion of IIB and the second premise of IIA with a sentence in mathematical notation about the relative heights of *X* and *Y*. In Graff Fara's other main example, the role of 'is no more than a millimetre taller than' is played by 'looks the same as', which is not a comparative in the relevant sense (Graff 2003, 215-17; 'looks red' plays the role of 'is tall'). Neither 'definitely' nor any similar device occurs in the premises or conclusions of the spruced-up versions of IIA or IIB. Of course, they contain the vague adjective 'tall', but that is quite different from containing an operator, such as 'definitely', 'rather', or '-ish', whose function involves interacting with vagueness in some way. Virtually any argument in natural language contains some vagueness. The putative failure of the meta-rules would be a limited phenomenon if it arose only for arguments containing operators that trade on vagueness; to say that it arises only for arguments containing vague terms is hardly any limitation at all.

----

[11] To generate a sorites paradox from multiple instances of IIA, one needs the structural rule of Cut: if Γ entails α and $\Delta \cup \{\alpha\}$ entails β then $\Gamma \cup \Delta$ entails β. Cut holds even for restricted versions of regional validity, for if all members of $\Gamma \cup \Delta$ are supertrue at a point w, then α is supertrue at w if the argument from Γ to α preserves supertruth, so β is supertrue at w if the argument from $\Delta \cup \{\alpha\}$ to β preserves supertruth; likewise for global validity.

Could a supervaluationist maintain that since the operative standard of validity is supertruth-preservation, and the *D* operator is a device for explicitness about supertruth in the object-language, *D* is implicitly present in all these arguments? As Graff Fara points out, that defence burns supervaluationism with friendly fire, by making *D* far more generally present than it appears to be (Graff 2003, 218).

A more discriminating idea is that to appreciate the material regional validity of IIB one must in effect think *about* vagueness in a way not required for appreciating material regional validity in most arguments containing vague terms. For clearly, if 'tall' were precise, IIB would be materially regionally invalid, even though the conditional probability of the conclusion on the premises might be high. The key to its material regional validity is the width of the borderline zone between the definitely tall and the definitely non-tall.

In response, the supervaluationist might point out that we rarely accept arguments like IIB, so the problem with contraposing them rarely arises. However, drawing attention to the unnaturalness of IIB is a dangerous move for a supervaluationist to make. For IIB itself feels far less compelling than does the step of contraposition from IIB to IIA, even if one reduces the distance at issue from a millimetre to something much less (but not zero). If one wants to avoid being drawn into the soritical argument IIA, the pre-theoretically natural way to do it is by not accepting IIB in the first place, rather than by accepting IIB and then refusing to contrapose. However materially restricted, regional validity does not come naturally as a standard for judging the material goodness of vague deductive arguments in one's native language. That is awkward for the supervaluationist who proffers regional validity as the appropriate standard.

Of course, native speakers may be confused about the logic of their own language. Nevertheless, supervaluationists have typically defended their account of the logic of vague languages by arguing that it is not seriously revisionary of inferential practice. But if we avoid reliance on those numerous instances of the meta-rules condemned by supervaluationists only because we are unwilling to rely on numerous object-language arguments (like IIA) *endorsed* by supervaluationists, their account is still highly revisionary.

The case of IIA and IIB is far from an isolated example. Similar pairs of arguments can be constructed for any sorites-susceptible term. Graff Fara has presented a deadly serious challenge to the supervaluationist account of deductive validity for vague languages.[12]

3. Extensions

In this section, I present some variations on Graff Fara's theme. In the final section, I will comment on the prospects for restricting the meta-rules in the ways supervaluationists have proposed.

---

[12] In Graff 2003, Graff Fara draws together her argument about supervaluationist consequence relations with considerations about higher-order vagueness from the first half of the paper to complete an argument against supervaluationist treatments of higher-order vagueness. In Graff Fara 2010, she extends her critique of supervaluationist accounts of the logic of vague languages. There is no space here to discuss these powerful further arguments.

3.1. Ambiguity

In his classic presentation of a supervaluationist treatment of vagueness, Kit Fine describes vagueness as 'ambiguity on a grand and systematic scale', and gives it a supervaluationist treatment: 'an ambiguous sentence is true if true for all disambiguations' (Fine 1975, 282; Keefe 2000, 156-8 rejects the analogy). It is therefore significant that Graff Fara's point applies even to the very simplest kind of ambiguity, under the supervaluationist treatment, in a language with no D operator or similar devices.

We start with a standard propositional language L, assumed free of vagueness, ambiguity, and additional operators. The semantics is the standard bivalent one, for which classical propositional logic is sound and complete. Let p and q be two atomic sentences in L; thus the semantics treats them as logically independent. We now expand L to a language L^+ by adding one new atomic sentence r. The idea is to make r ambiguous between just two meanings, one that of p, the other that of q. For simplicity, we can identify points in the models for L^+ with ordered pairs $<a, j>$, where a is an assignment of truth-values to atomic sentences of L and j is either 0 (flagging that r is disambiguated as p) or 1 (flagging that r is disambiguated as q). Thus any atomic sentence other than r is true at $<a, j>$ if and only if it is true on a; r is true at $<a, 0>$ if and only if p is true on a; r is true at $<a, 1>$ if and only if q is true on a. The models for L^+ correspond one-one with the original assignments a, where the set of points in the corresponding model is just $\{<a, 0>, <a, 1>\}$. The two points in any model are accessible from each other and from themselves; thus accessibility is always an equivalence relation. This implements the idea that supertruth is just truth on both disambiguations of r, the only source of ambiguity or vagueness in L^+. Of course, when the models are restricted like this, the resulting regional consequence relation $\vDash_R$ is not purely logical, since a non-logical constraint on the truth-value of r has been built in, but that is just what we want for Graff Fara's point.

Some simple observations can now be made.

[a] $r \vDash_R p$, for if r is supertrue at $<a, j>$ in a model, then r is true at $<a, 0>$ in the model, so p is true on a, so p is true at both $<a, 0>$ and $<a, 1>$, so p is supertrue at $<a, j>$.
[b] $r \vDash_R q$, by an argument like that for [a].
[c] $r \vDash_R p \leftrightarrow q$, from [a] and [b] because $p \leftrightarrow q$ follows classically from p and q.
[d] $\neg r \vDash_R \neg p$, by an argument like that for [a].
[e] $\neg r \vDash_R \neg q$, by an argument like that for [a].
[f] $\neg r \vDash_R p \leftrightarrow q$, from [d] and [e], as for [c].

Consequently, if $\vDash_R$ obeys argument by cases, $r \vee \neg r \vDash_R p \leftrightarrow q$ from [c] and [f], so $\vDash_R p \leftrightarrow q$. But that is not so, for the semantics makes p and q mutually independent.

The other meta-rules fail similarly. For instance, conditional proof and contraposition fail, since $r \vDash_R p$ but neither $\vDash_R r \rightarrow p$ nor $\neg p \vDash_R \neg r$. Similarly, reductio ad absurdum fails because $\neg p, r \vDash_R \bot$ but not $\neg p \vDash_R \neg r$. Under regional validity, the contested meta-rules enable one to exploit the ambiguity of r to treat its two readings as equivalent to each other.

The failure of the meta-rules does not depend on the simplicity of the example. The proof can be adapted to much more complex cases of ambiguity, involving other grammatical categories, more than two readings, pre-existing ambiguity and vagueness in the object-language, and so on. Moreover, as already emphasized, it does not depend on the presence of any special operators in the object-language for dealing with vagueness or ambiguity.

3.2. Quantifier rules

Although the failure of meta-rules for regional validity is usually discussed in the setting of propositional logic, it also affects some meta-rules for quantifiers. An example is a standard elimination rule for the existential quantifier: if from $\Gamma \cup \{\alpha(c)\}$ one can derive β, and the constant c occurs nowhere in Γ, $\alpha(x)$, or β, then from $\Gamma \cup \{\exists x \, \alpha(x)\}$ one can also derive β (with the usual restriction to avoid clash of variables in $\exists x \, \alpha(x)$). In effect, this rule permits one to use an 'arbitrary name' in working out the consequences of an existential premise. For regional validity, the rule says that if $\Gamma \cup \{\alpha(c)\} \vDash_R$ β then $\Gamma \cup \{\exists x \, \alpha(x)\} \vDash_R$ β, subject to the same restrictions.

To assess the rule, we must extend the model theory to handle quantifiers and individual variables. The main choice-point in doing so is whether we allow variation in the quantifier domain from one point of evaluation to another, corresponding to vagueness in how far the quantifiers range. To keep things simple, we assume that there is no such vagueness. In particular, we may specify that the quantifiers range precisely over the natural numbers. If the result is a non-logical consequence relation, that is consonant with Graff Fara's argument. However, for the time being, we keep the operator D in the language. Let F be a monadic atomic predicate, and j an arbitrary name. Thus, for the same reason as in §1, Fj and $\neg DFj$ are never both supertrue at a point in a model (on an assignment of values to variables). Since $\neg \exists n \, DFn$ is true only at those points where $\neg DFj$ is true, $\neg \exists n \, DFn$ is supertrue only at those points where $\neg DFj$ is supertrue. Thus Fj and $\neg \exists n \, DFn$ are never both supertrue at a given point, so vacuously $\{\neg \exists n \, DFn, Fj\} \vDash_R \bot$. Since j occurs nowhere in $\neg \exists n \, DFn$ and $\bot$, the existential elimination rule at issue for regional validity tells us that $\{\neg \exists n \, DFn, \exists n \, Fn\} \vDash_R \bot$. But that is not so. For consider a simple model where $W = \{j, ..., j + k\}$ $(1 \le k)$, R is the universal relation over W, and the extension of F at i is $\{i\}$ $(j \le i \le j + k)$. Then $\exists n \, Fn$ is true at every point in the model. But, on any assignment of natural numbers to variables, DFn is false at every point because Fn is false at some point, so $\neg \exists n \, DFn$ is true at every point. Thus the argument from $\neg \exists n \, DFn$ and $\exists n \, Fn$ to $\bot$ has supertrue premises and a superfalse conclusion at every point in the model.

The model has a natural supervaluationist interpretation. For we can read Fn as 'n is small and $n + 1$ is not small', where 'small' is of course vague. Thus each point in the model corresponds to a potential cut-off for 'small'. On the standard supervaluationist view, it is definite both that there is a cut-off for smallness and that no number is definitely the cut-off. Thus the counter-model to existential elimination respects the spirit as well as the letter of supervaluationism. Admittedly, the model rules out higher-order vagueness, but one can easily tweak it to admit that phenomenon while still providing a counter-model to existential elimination.

That example depends on the D operator. But we can adapt Graff Fara's own example to build a counterexample to existential elimination with no such operator. For consider these two sentences:

> CUTOFF   Mary is tall and John is not tall and Mary is not more than a millimetre taller than John.
>
> ∃CUTOFF   For some X, for some Y, X is tall and Y is not tall and X is not more than a millimetre taller than Y.

In the context Graff Fara envisages, CUTOFF is guaranteed not to be supertrue. Thus the argument from CUTOFF to a contradiction is materially regionally valid. Hence, if the ex-

istential elimination rule for English preserves material regional validity, then the argument from ∃CUTOFF to a contradiction is also materially regionally valid, by two applications of the rule. But it is not. Given a sufficiently dense distribution of heights in the relevant population between the definitely tall and the definitely not tall, ∃CUTOFF is supertrue. Thus the existential elimination rule fails for material regional validity even in the absence of operators that trade on vagueness.

3.3. Mathematical notation

Contrary to stereotype, some standard mathematical notation is deliberately left vague and context-dependent. For instance, the symbol ≈ is often used for *approximate* equality, typically of real numbers. For mathematical convenience, the required degree of approximation is left unspecified. Such uses yield further examples of Graff Fara's phenomenon.

We can see this by building a toy supervaluationist model of how ≈ works in some simple contexts. Suppose that on each assignment there is a constant δ such that for all real numbers x and y, $x \approx y$ if and only if the absolute difference $|x - y| < \delta$.[13] The value of δ is left vague: it varies from assignment to assignment, subject to the constraint that $0 < a \le \delta \le b$, where the model fixes a and b assignment-independently. The assignments correspond one-one with the members of the closed interval $[a, b]$. All assignments are mutually accessible; for simplicity, we exclude higher-order vagueness, but including it would only slightly complicate the argument. The width of the interval $[a, b]$ measures the degree of vagueness of 'δ'. We suppose that it is reasonably wide, in the sense that $2a \le b$.

Now consider the set of formulas $\{x \approx y, y \approx z, \neg x \approx z\}$, which describes a counter-example to the transitivity of approximate equality. Suppose that all three formulas are supertrue together at some point in the model. Thus for all δ in $[a, b]$, $|x - y| < \delta$, $|y - z| < \delta$, and $|x - z| \ge \delta$. Therefore, in particular, $|x - y| < a$, $|y - z| < a$, and $|x - z| \ge b$. But then, by the triangle inequality for absolute difference:

$$b \le |x - z| \le |x - y| + |y - z| < 2a \le b$$

That is impossible, so the three formulas cannot be supertrue together. Thus the argument from $\{x \approx y, y \approx z, \neg x \approx z\}$ to a contradiction is regionally valid, in the material sense restricted to the context just described. Consequently, if this version of regional validity obeyed reductio ad absurdum, then the argument from $\{x \approx y, y \approx z\}$ to $x \approx z$ would also be regionally valid. But it is not, for when x, y, and z take the values 0, $2a/3$, and $4a/3$ respectively, the argument's premises are supertrue while its conclusion is not. Indeed, multiple instances of the argument from $\{x \approx y, y \approx z\}$ to $x \approx z$ generate a sorites paradox, just as multiple instances of Graff Fara's argument IIA did. From the same example one can easily construct counterexamples to the meta-rules of conditional proof, contraposition, argument by cases, and existential elimination too.

Another mathematical symbol deliberately left vague and context-dependent is <<: '$x << y$' is read 'x is much less than y'. It generates more examples of Graff Fara's phe-

---

[13] Pedantically speaking, the three object-language variables x, y, and z will be used in the meta-language for their values on the given assignment.

nomenon. That is unsurprising, for '$x << y$' is similar in meaning to '$x < y$ and not $x \approx y$'. We can treat '$x << y$' as equivalent to '$x + \delta < y$' in models like those above, though perhaps for different values of a and b.

Now consider the set of formulas $\{\neg x << y, \neg y << z, x << z\}$. Suppose that all three formulas are supertrue together at some point in such a model. Thus for all δ in $[a, b]$, $x + \delta \geq y$, $y + \delta \geq z$, and $x + \delta < z$; hence $x + 2\delta \geq y + \delta \geq z$. Therefore, in particular, $x + b < z$ and $x + 2a \geq z$, so $b < 2a$, contrary to hypothesis. Hence the three formulas cannot be supertrue together. Thus the argument from $\{\neg x << y, \neg y << z, x << z\}$ to a contradiction is regionally valid, in the material sense restricted to this context. Consequently, if this version of regional validity obeyed reductio ad absurdum, the argument from $\{\neg x << y, \neg y << z\}$ to $\neg x << z$ would also be regionally valid. But it is not, as the previous assignment of values to the variables shows. Indeed, multiple instances of the argument from $\{\neg x << y, \neg y << z\}$ to $\neg x << z$ generate a sorites paradox, in the same way as before. Once again, from the example one can easily construct counterexamples to the meta-rules of conditional proof, contraposition, argument by cases, and existential elimination too.

4. Morals

The evidence from §3 confirms Graff Fara's claim that the potential for violations of standard meta-rules by material forms of validity is extremely widespread. It extends to ordinary cases of ambiguity, to quantifier rules, and to many mathematical contexts. What morals should we draw from this?

Graff Fara is clearly right that in practice we rarely care whether our deductions are valid by a purely logical standard or a looser, context-dependent one, reliant on background assumptions. We are fine with arguments like 'She lives in Paris, so she lives in France', despite their invalidity by a strictly logical standard. Even working mathematicians rarely care whether their proofs are *logically* valid, so long as they are *mathematically* valid. For example, they do not normally differentiate in status between standard set-theoretic principles, such as the power-set axiom, and purely logical principles, such as the law of excluded middle. They treat them all as part of the assumed background. A non-logical, material standard of consequence is operative.

Suppose that, in a vague and sometimes ambiguous language, our deductive reasoning tends to accord with material regional validity: we tend to accept materially regionally valid deductions and reject materially regionally invalid ones. We must therefore be implicitly avoiding numerous nearby instances of the meta-rules that would take us from materially regionally valid arguments to materially regionally invalid ones. Yet, on the face of it, we freely apply conditional proof, reductio ad absurdum, contraposition, argument by cases, and existential elimination, without specially checking whether the instance in question is problematic for supervaluationist reasons. Did you monitor the purely technical parts of this paper for such fallacies? (Supervaluationists would in fact accept the informal proofs.)

Obviously, in practice we rarely insert 'definitely' into the input or output of the meta-rules to ensure (material) regional validity. Do we instead apply qualified versions of the meta-rules with extra conditions on the status of the input argument, such as McGee and McLaughlin and others impose? Such proposals typically require the input argument to meet a more demanding standard of validity than regional validity. Thus, if ordinary rea-

soners pay attention to regional validity at all, they must keep track of *two* standards of validity, and of which arguments meet which standard, if they are to be in a position to apply the qualified versions of conditional proof, reductio ad absurdum, contraposition, argument by cases, and existential elimination. In particular, they must be sensitive to the status of some of their arguments as merely regionally valid: good enough for the truth of their conclusions given the truth of their premises, but not good enough to serve as inputs to the qualified meta-rules. That we habitually do any such thing is far from obvious. If mathematicians assign such a differential status to some of their proofs, why do they never speak of it?

One hypothesis is that we accept regionally valid arguments, and make free but illicit use of the unqualified meta-rules, because we are unaware of the fallacies. On that hypothesis, our reasoning should get us into serious trouble. Does it? A first thought is that we do get into serious trouble: slippery slopes and sorites paradoxes. However, it is not clear that our problems there depend on illicit use of the meta-rules. For if they do, why are the problems so limited, when their cause is almost ubiquitous? Why do mathematicians not get into trouble with their use of symbols like ≈ and <<? After all, mathematicians are the people who give our patterns of deductive reasoning their most systematic and severe stress-testing.

Consider ambiguity as a toy model of vagueness, as supervaluationists such as Fine have done. Of course, ambiguity *does* sometimes get us into trouble: there are fallacies of equivocation. Sometimes we argue from α to β(*e*), using the expression *e* in one sense, and then from β(*e*) to γ, using *e* in another sense, and put the two arguments together to make an argument from α to γ that is not truth-preserving on any sense of *e* (which is absent from both α and γ). But our problem there does not depend on any of the meta-rules that fail for material regional validity. Once we recognize the ambiguity in an expression, we are not tempted to reason with it according to material regional validity. This argument is an example (with no shift of reference in 'that'):

> *Argument BANK*
> That is a bank.
> Therefore, that is both a financial bank and a river bank.

BANK is untempting, to say the least. However one disambiguates the premise, the conclusion does not follow, even materially. Yet BANK is materially regionally valid, on a supervaluationist treatment of ambiguity: if the premise is true on both disambiguations, the conclusion is true (compare §3.2, [a] and [b]). Of course, committed supervaluationists may be able to talk themselves into accepting the validity of BANK. But it does not come naturally to reason according to BANK.

The same goes for vagueness proper. For instance, when I think as a mathematician, I feel no temptation to regard $\{x \approx y, y \approx z, \neg x \approx z\}$ as an inconsistent set. After all, the symbol ≈ means *approximate* equality, and who expects approximate equality to be transitive? Similarly, Graff Fara's argument IIB does not come naturally, despite its material regional validity in a suitable context. The obvious hypothesis as to why our free, unqualified use of the meta-rules does not get us into trouble is that we do not accept the materially regionally valid input arguments that make the trouble. Supervaluationism is far more revisionary of our current practice than supervaluationists like to assume.

Ought we to tailor our deductive practices to material regional validity? Its failure to validate conditional proof, reductio ad absurdum, contraposition, argument by cases, and existential elimination is a strong reason not to. As already noted, using the qualified versions instead of them carries a heavy cognitive burden, for no clear gain.[14]

I will not labour the case. Graff Fara's point goes too deep for any technical fix. Supervaluationists need to undertake a far more searching investigation of what our deductive practice would really be like if properly tailored to context-dependent material regional validity. In this, as in so much else, she has left us a rewarding challenge for the future.[15]

REFERENCES

Bobzien, Susanne. 2011. If it's clear, then it's clear that it's clear, or is it? Higher-order vagueness and the S4 axiom. In Ben Morison and Katerina Ierodiakonou, eds., *Episteme, etc.: Essays in honour of Jonathan Barnes*, pp. 189-212. Oxford: Oxford University Press.

Cobreros, Pablo. 2008. Supervaluationism and logical consequence: a third way. *Studia Logica* 90: 291-312.

—. 2010. Varzi on supervaluationism and logical consequence. *Mind* 120: 833-843.

Dorr, Cian. 2015. How vagueness could cut out at any order. *Review of Symbolic Logic* 8: 1-10.

Fara, Michael [Graff] and Timothy Williamson. 2005. Counterparts and actuality. *Mind* 114: 1-30.

Fine, Kit. 1975. Vagueness, truth, and logic. *Synthese* 30: 265-300.

Graff [Fara], Delia. 2000. Shifting sands: an interest-relative theory of vagueness. *Philosophical Topics* 28: 45-81.

—. 2001. Phenomenal continua and the sorites. *Mind* 110: 905-935.

—. 2002. An anti-epistemicist consequence of margin for error semantics for knowledge. *Philosophy and Phenomenological Research* 64: 127-142.

—. 2003. Gap principles, penumbral consequence, and infinitely higher-order vagueness. In J.C. Beall, ed., *Liars and heaps: New essays on paradox*, pp. 195-221. Oxford: Clarendon Press.

—. 2010. Scope confusions and unsatisfiable disjuncts: two problems for supervaluationism. In Richard Dietz and Sebastiano Moruzzi, eds., *Cuts and Clouds: Vagueness, Its Nature, and its Logic*, pp. 373-382. Oxford: Oxford University Press.

—. 2011. 'Truth in a region', in Paul Égré and Nathan Klinedinst, eds., *Vagueness and language use*, pp. 222-248. London: Palgrave Macmillan.

Graff [Fara], Delia and Timothy Williamson. 2002. Introduction. In Delia Graff [Fara] and Timothy Williamson, eds., *Vagueness*, pp. xi-xxviii. Dartmouth: Ashgate.

Hjortland, Ole Thomassen. 2017. Anti-exceptionalism about logic. *Philosophical Studies* 174: 631-658.

Kamp, Hans. 1975. Two theories about adjectives. In E.L. Keenan, ed., *Formal semantics of natural language*, pp. 123-155. Cambridge: Cambridge University Press.

Keefe, Rosanna. 2000. *Theories of vagueness*. Cambridge: Cambridge University Press.

---

[14] Cobreros (2010) describes a procedure for obtaining proof systems for global validity from proof systems for local validity, where the global proof differs from the local proof in only one step. However, Graff Fara's examples and mine suggest that no such mechanical procedure will extend to material global and regional validity.

[15]

Lewis, David. 1972. General semantics. In D. Davidson and G. Harman, eds., *Semantics of natural language*, 169-218.

Mahtani, Anna. 2008. Can vagueness cut out at any order? *Australasian Journal of Philosophy* 86: 499-508.

McGee, Vann and Brian McLaughlin. 1994. Distinctions without a difference. *The Southern Journal of Philosophy* 33 (sup): 203-251.

—. 1997. Review of *Vagueness*. *Linguistics and Philosophy*, 21: 221-235.

—. 2004. Logical commitment and semantic indeterminacy: a reply to Williamson. *Linguistics and Philosophy* 27: 123-136.

Varzi, Achille. 2007. Supervaluationism and its logics. *Mind* 116: 633-676.

Williams, J. Robert G. 2008. Supervaluationism and logical revisionism. *The Journal of Philosophy* 105: 192-212.

Williamson, Timothy. 1994. *Vagueness*. London: Routledge.

—. 1999. On the structure of higher-order vagueness. *Mind*, 108: 127-143.

—. 2002. Epistemicist models: comments on Gómez-Torrente and Graff. *Philosophy and Phenomenological Research* 64: 143-150.

—. 2004. Reply to McGee and McLaughlin. *Linguistics and Philosophy* 27: 113-122.

—. 2013. *Modal logic as metaphysics*. Oxford: Oxford University Press.

—. 2017. Semantic paradoxes and abductive methodology. In Bradley Armour-Garb, ed., *Reflections on the liar*, pp. 325-346. Oxford: Oxford University Press.

TIMOTHY WILLIAMSON is the Wykeham Professor of Logic at the University of Oxford, a Fellow of New College Oxford, and the A. Whitney Griswold Visiting Professor at Yale University.

ADDRESS: Faculty of Philosophy, University of Oxford. New College, Oxford OX1 3BN, U.K.
E-mail: timothy.williamson@philosophy.ox.ac.uk