

Generalizing from the Instances within “Contextualist” Accounts of Vagueness (first draft)

Delia Graff Fara

Princeton university

Draft 1.0 August 2008

Introduction

The sorites paradox arises for a given vague predicate when its applicability doesn't tolerate large differences of a certain kind, but we believe that the applicability of the predicate does tolerate small differences of that kind. The problem is that small differences can add up to big differences. For example, we believe (before initiation anyway) the following Universal Generalization, which reflects our belief that “short” tolerates small differences in height.

The Universal Generalization: Every man just one millimeter taller than some short man is also a short man.

In addition to this *sorites premise*, we have a couple of true beliefs about men's heights, which together reflect our belief that “short” does not tolerate large differences in height:

Short-Height Claim: Any man whose height is one-thousand millimeters is a short man. (This is about three-foot-three.)

Tall-Height Claim: Any man whose height is *two*-thousand millimeters is *not* a short man.

Our beliefs concerning men's heights are plainly correct. Any man whose height is three-foot-three is a short man and any man whose height is six-foot-six is a tall man; it would be ridiculous to think otherwise. So the Universal Generalization is not true, since it is plainly incompatible with the true beliefs.¹ We should therefore not believe it. The paradox is that it's hard to accept the denial of the Universal Generalization (even after initiation), since the denial doesn't seem like it could be true either, at least not if it amounts to the claim that a difference of one millimeter in height can make for the difference in being short or not. It's very difficult

¹Some implicit premises are involved: for any number of millimeters between one-thousand and two-thousand, we implicitly assume that there is a man whose height is that number of millimeters. Even if they are not true, these implicit premises might as well be accepted for three reasons: the paradox can easily be reformulated in ways that don't require these implicit premises; the premises could easily have been true, even if they're not in fact true; and for all we know, they might in fact be true.

to accept that a short person could grow one millimeter over night then wake up not short in the morning.

The paradox raises a number of questions that should be answered if we're to remain secure in the thoughts that we know what the words in our own language mean and that we're able to use those words to communicate effectively with one another. Among those questions is a psychological one requiring an explanation of our misunderstanding of a core feature of the meaning of vague predicates:²

The Psychological Question: If the Universal Generalization is not true, why are we so inclined to accept it in the first place? In other words, what is it about vague predicates that makes them seem tolerant to us?

Rosanna Keefe (2007) has forcefully presented a number of important criticisms of what she calls *contextualist* solutions to the sorites paradox. My preferred solution is among these. I place it within a class of “contextualist” solutions that I’ll call *instance based*. These instance-based solutions explain our mistaken belief in the Universal Generalization by appealing to our correct beliefs in its instances. They say that each belief in an instance is correct *when that instance is considered*. For example, switching predicates, suppose the “tall”–“not-tall” boundary divides Paul and John, who are known to differ in height by just one millimeter. If you then consider their heights together, that boundary shifts somewhere else. If n millimeters is not a tall height and $n + 1$ is, then when we consider whether this is so, the tall-height boundary ceases to occur there.³ Our mistaken acceptance of the Universal Generalization is therefore understandable; we generalize from some instances that we know to be true.⁴

Keefe argues against these theories. She argues that our mistaken generalization has not thereby been made understandable: instance-based theories provide no good answer to the Psychological Question—not as stated above, anyway. She correctly requires (278) that we distinguish three different versions of the Psychological Question: (i) one about the Universal Generalization, (ii) one about its instances, and (iii) one about a claim of “no sharp boundaries”—by which I take it she means a negative-existential equivalent of the Universal Generalization, in symbols, $\neg\exists n(Fn \wedge \neg Fn + 1)$. She argues that the instance basers have at best answered Psy-

²See Fara (2000, 51).

³As with any brief presentation of a plausible theory, there are important details that are left out and salient exceptions that aren't discussed.

⁴There are surely two further good questions about how the solution might work in the case of reasoning to a universal generalization from an *arbitrary* instance or from a sample of representative instances. But I blur the distinctions here so as to avoid addressing what might be ancillary issues.

chological Question (ii), the one about the instances, and have no basis for taking that answer to provide, *straightaway*, an answer to Psychological Question (i), the one about the Universal Generalization. We can call this the *generalizing-from-the-instances problem*. To give a feel for the problem, as well as one sort of argument she presses in order to entrench it, recall the paradox of the preface. The author of a non-fiction book might well write in his preface that there are sure to be mistakes in the main text. Yet when considering any given claim in the main text, he persists in affirming that claim. He assents to each instance of a universal ascription of truth to the sentences in his book, but denies the universal itself. How then, the thought goes, could beliefs in instances suffice to explain a generalization over those instances?

In addition to pressing the generalizing-from-the-instances problem Keefe also marshals a battery of arguments against the success even of the instance-based “contextualists’ ” answer to Psychological Question (ii), the one about the instances. But I will address those only when doing so supports my response to the generalizing-from-the-instances problem, which is my main concern.

Note that each of Keefe’s psychological questions corresponds to its own version of the sorites paradox: (i) a generalized version, (ii) a particularized version, and (iii) a no-sharp-boundaries version.⁵ So in effect, Keefe argues that the instance baser has at best solved the particularized version of the paradox, not the other versions. For the predicate “short” these different versions could be presented as follows.

Generalized Version

One-thousand millimeters is a short height;
Every number of millimeters that’s just one greater than a short height is also a short height;

Therefore, two-thousand millimeters is a short height.

Particularized Version

One-thousand millimeters is a short height;
If one-thousand millimeters is a short height, then so is one-thousand-and-one;
If one-thousand-and-one millimeters is a short height, then so is one-thousand-and-two;
.
.
.
If one-thousand-nine-hundred-and-ninety-eight millimeters is a short height, then so is one-thousand-nine-hundred-and-ninety-nine;

⁵The “generalized”-“particularized” terminology is from Scott Soames (1999). The “no-sharp-boundaries” terminology is from Crispin Wright (1987, 1992).

If one-thousand-nine-hundred-and-ninety-nine millimeters is a short height, then so is two-thousand;

Therefore, two-thousand millimeters is a short height.

No-Sharp-Boundaries Version

One-thousand millimeters is a short height;
No number of millimeters that's not a short height is just one greater than a number of millimeters that is a short height;

Therefore, two-thousand millimeters is a short height.

The thesis to be defended here, then, is that an answer to the particularized psychological question does directly provide an answer to the generalized psychological question, although with more ado than I originally thought necessary.

Terminology Change: Contextualism v. Boundary Shifting

I prefer my own terminology to Keefe's and think that Keefe's blanket use of the term "contextualism" glosses over some important distinctions that are relevant to her criticisms.⁶ What Keefe calls a "contextualist" solution to the paradox is what I call a *boundary-shifting* one.⁷

Boundary shifting is one of the hallmarks of a vague predicate. We make classifications in order to further our goals. Applying or denying predicates to things is a standard way to do this. I say to you, "hand me the blue one" because I want you to hand me a certain grayish-blue book. As it happens, the book is sitting in a pile of grayish-red books, but if it had been in a pile of ultramarine ones I might have said "hand me the gray one" instead. The boundary between what counts as blue and what counts as gray can shift with a change in what best satisfies our purposes.

There will be different views of what this *counting as gray* or *counting as blue* amounts to. I take it that a thing's *counting as blue* on an occasion amounts to its falling under the predicate "blue" on that occasion; to its being, in other words, in the predicate's *extension* on that occasion; that is, to the predicate's being true of it on that occasion.

⁶Throughout, I assume the framework of Kaplan (1989), and will take his inter-related notions of *character*, *context*, *extension*, *world*, *content*, *property* and *proposition* for granted. Unlike Kaplan, I do not take a context to be a sequences of indices, but rather a pairing of occurrences of context-dependent words or variables, overt or "hidden", with semantic values.

⁷These are what I referred to in "Shifting Sands" as *Bare-Bones Solutions*, because they could be fleshed out in a number of different ways, not all of them being what I would call *contextualist*, since I use that term narrowly, in Kaplan's sense.

Boundary shifters are united in thinking that these changes in extension are more common than one would initially have thought and are precisely what lead to the sorites paradox. Instance-based boundary-shifters think that there are changes in extension that render each instance of the universal generalization true at the time that instance being considered. They think, moreover, that the truth of these instances—when considered—to some extent justifies our belief in them and that these somewhat justified beliefs explain our unjustified belief in the bad universal generalization.

Contextualist and Invariantist Boundary Shifting

There is one main fault line dividing boundary shifters: do they think that these changes in extension signal a change in which property is expressed by the predicate or not? Those that do are *contextualist* boundary-shifters; those that don't are *invariantist* boundary-shifters. Contextualist boundary-shifters think that the boundary-shifting changes in extension for a vague predicate (i.e., a linguistic item such as a *word* or a *phrase*) result not from a change in the extension of any property, but rather from a change in which property the predicate expresses. Invariantist boundary-shifters think that the relevant changes in the vague predicate's extension result from a change in which things satisfy the single property invariantly expressed by the predicate in question.

The difference can be demonstrated by a routine example. Consider how a sentence may change its truth value over time.

- (1) Nikola has never been to Ithaca:
 - (a) (said on 14 July 1999, true);
 - (b) (said on 14 July 2004, false).

An utterance of (1) may be true or false depending on when it takes place.

A contextualist account of this difference would say that the two utterances of (1) have different *content*: they attribute different properties to Nikola. One falsely attributes to Nikola the property of *having never been to Ithaca as of 14 July 1999*; the other correctly attributes to Nikola the different property of *having never been to Ithaca as of 14 July 2004*. The properties attributed here are not ones that Nikola could lack and then acquire; he lacks the first forever and always; he possesses the second forever and always.

An invariantist account of the difference would maintain, in contrast, that the two utterances have the same content—they attribute the same property to Nikola, the property of *having*

never been to Ithaca. The invariantist says that something can lack this property at one time and acquire this very property by some later time—this is what happened to Nikola—and that, correspondingly, a proposition attributing such a property may change *its* truth value over time.

Contextualists about this issue are called *eternalists* because they think that a proposition only ever has its truth value eternally. Invariantists about this issue are called *temporalists* because they think that a proposition can have its truth value temporarily. I prefer temporalism, though that won't be very important in what follows. It will be nonetheless important to retain the two views in mind.

There are further fault lines dividing boundary shifters on both sides of the main divide: to what do they attribute the relevant changes in extension for vague predicates? *Interest relativists* think that the changes in extension result from changes in what best satisfies (or furthers, or piques, etc.) our interests. Interest-relative *contextualists* think that these changes in what best satisfies our interests may affect which property is expressed by a vague predicate on a given occasion.⁸ Interest-relative *invariantists* think that changes in what best satisfies our interests may affect which things have the invariantly-expressed property in question. Both think that the respective contributions of interests are the ones most responsible for the relevant extension changes.

It is open to an interest relativist—be she contextualist or invariantist—to account for these changes in two different ways: as a change in what our interests are; alternatively, as a change in facts that are external to what our interests are. You could change from being someone I wanted to dance with to being someone I don't want to dance with either because I used to be interested in dancing only with short people, but now I'm interested in dancing with tall people as well; or because I've wanted to dance only with short people all along, but you have grown from short to tall.

By the way, “invariantist” is not a completely apt moniker, since even the “invariantist” boundary shifter would be hard pressed to deny that vague predicates are context sensitive in at least some ways. Which property a vague predicate expresses can, for example, be sensitive to a comparison class that's left implicit. Either of the properties *tall for a jockey* or *tall for a basketball player* could be expressed by a bare use of “tall”—when the context is right. The invariantist boundary shifters merely think that the changes in extension which they appeal to in answering the Psychological Question(s) need not mark changes in which property is expressed by the predicate; it's not change in context *per se* that causes *these* changes in extension. By

⁸Not every position adumbrated here has been advocated in print.

analogy, the temporalist need not deny that claims that lack explicit reference to times, like (1), are completely insensitive to context; she will surely allow that (1) may be used to ascribe different properties to Nikola depending on whether we are talking about the Ithaca in upstate New York or the Ithaca in the Ionian Sea. The temporalist merely thinks that the changes over time in the extension of predicates like “has never been to Ithaca” need not mark a contextual change in the property expressed by the predicate; it’s not change in context *per se* that causes *these* changes in extension.

About vagueness, I prefer interest-relative invariantism, and I prefer to think that the relevant changes in what best satisfy our interests are due to changes in external facts rather than to changes in our interests. The external facts in question are facts about which things it’s least efficient for us to have differential attitudes towards with respect to our purposes at hand.⁹ When we are considering whether to classify n millimeters as short and $n + 1$ millimeters as not short, it is very inefficient to treat them as different for present purposes since they are right before us, and it is more inefficient to treat them as different than it is to treat $n + 50$ and $n + 51$ as different given that the latter pair is not under consideration. Because we have a standing interest in efficiency, our interest is then better satisfied if we treat n and $n + 1$ as the same for present purposes than it is if we treat $n + 50$ and $n + 51$ as the same for present purposes. And the comparative degree of inefficiency reverses when our comparative attention to the two pairs reverses.

Furthermore, we can’t in general hold the extension of a vague predicate fixed just by holding our own interests fixed because we can’t control which things are presented to us for our consideration or as live options.¹⁰ This phenomenon, that of being unable to hold extensions fixed at will, is familiar to contextualists and invariantists alike. We usually cannot hold the referent of “I” fixed because we usually cannot prevent other people from speaking. And we cannot hold the extension of the verb phrase “is shorter than one-hundred-and-fifty millimeters” fixed because we cannot prevent a man from growing or time from passing.

Objection I: If equivocating here then equivocating everywhere!

According to Keefe (284), instance-based boundary-shifters charge us with “something akin to a fallacy of equivocation” when we generalize from the instances to the universal claim in the

⁹For elaboration and development of the view see Fara (2000) and (*forthcoming*).

¹⁰Keefe suggests that a “contextualist” view of vagueness should render it possible for us to hold “contexts” fixed whenever we intend to do so (2007, 280).

Generalized Version of the sorites paradox. *How akin* is the important question.

We normally take an argument to involve equivocation when a single word is used with more than one meaning over the course of the argument. If you say that you want an iron for your birthday and I say that I got you an iron for your birthday, I cannot on the basis of our two comments then legitimately maintain that I've gotten you something you wanted if you meant a clothes iron (and I know this) and I meant a golf club. To do so would be to equivocate. The root of the problem here is that the change in meaning allows for an extension change for "iron". On one meaning it applies to some of the things you want; on the other meaning it applies to none of the things you want.

There is also what we might think of as a weak, extended sense of *equivocation* that deems equivocal any argument that allows for just the sort of extension change that renders meaning-change equivocations fallacious. Many more arguments are weakly equivocal than are traditionally equivocal since extension change can result not only from meaning change but also from either of the following two sources: (i) the context might shift in the course of making the argument so that the premises might be true, but not all within a single context (ii) there might be a shift in what things are like, so that the premises might be true, but not all at the same time.

Instance basers claim that our mistake in coming to believe the Universal Generalization stems from one of the two sources of weak equivocation just mentioned. (When I speak of *instance basers* I always mean *instance-based boundary-shifters*.) The contextualist thinks that it's an equivocation of type (i); the invariantist thinks that it's an equivocation of type (ii).¹¹ The contextualist might, for example, think that in each utterance "short for a man" expresses the property of bearing relation *R* to height *h*, where *R* and *h* are assigned contextually-given values. The values might vary from utterance to utterance if they are rigidly fixed by non-rigid descriptive conditions. For example, *R* might at one time be the relation of *being two inches shorter than* and at another time be the relation of *being three inches shorter than* where these are fixed by the interest-relative description "is significantly shorter than, given current purposes". And *h* might be the height *five-feet-ten* at one time and then *five-feet-nine* at another time, where these are fixed by the description "the (currently) average height". The invariantist, on the other hand, might think that "short for a man" invariantly expresses the non-rigid interest-relative property of being *significantly shorter, given current purposes, than*

¹¹The issues here would be further complicated if we threw the temporalism–eternalism divide into the mix. Rather than outline all four combinations of view here, I'll just assume that the contextualist is an eternalist and that the invariantist is a temporalist.

the currently average height. This is a property with an easily changed extension. Which differences are significant differences can change, and which height is currently the average can change.

Keefe objects (285–87) that if the instance baser were correct about our generalizing from the instance’s involving equivocation in the sorites case, then all sorts of other cases of what’s normally deemed good reasoning would involve equivocation and hence be fallacious. Let’s call this the *too-much-fallacious-reasoning objection*. The idea is this: (a) an argument is fallacious *whenever* there is extension shift due to context change or to the passage of time; (b) if the boundary shifter were right, then equivocation-producing extension shift would be rampant; but (c) fallacious reasoning is not in fact rampant; so (d) the boundary shifter is wrong.

Keefe offers illustration of the extent to which the objection is supposed to be damaging by presenting a couple of examples. I’ll specifically address only one of these examples, in order to tease out a separable issue that underlies a number of Keefe’s other criticisms. Then I’ll proceed to give a general response to the objection.

Example: Inductive reasoning about individuals

I’ve come to England for the first time and notice after two weeks that every post box that I’ve seen has been red. I come to believe, correctly, that all English post boxes are red. But, complains Keefe:

... the context changes with each observation and, since they are marginally different shades of red and different objects are salient along side them, the standards required to count as “red” will also change. When I go on to conclude that all post-boxes are red, I do not mean the same thing by “red” in the conclusion as I mean in all the instances, and so, again, I am guilty of equivocation.¹² (285)

There is an easy reply to this argument. The boundary shifter does not say that every perception of a red in changed surroundings or of a red of a slightly different shade will induce a shift in boundary.¹³ Formally, only two potential boundary sites are required for the instance baser’s

¹²Keefe notes in a footnote (nt.11, 285) that the argument against the invariantist must differ from that against the contextualist by making reference to change in standards over time rather than across contexts.

¹³This can be so even for perception-based boundary shifters. See Diana Raffman (1994). On Raffman’s (1994) view an extension change occurs when the concept of red is stretched to its limit—when what’s counted as red becomes just too orange (gray, pink, *et cetera*) for any slightly more orange shade to still count.

solution to work. Whatever the current extension boundary for “red” is, it need shift only when I examine it. A second shift from the new site then need occur only when I examine *it*. The boundary might well at that point shift right back to its initial site, and then back and forth between just those two sites as one or the other is examined.

This is not to say, of course, that an instance baser would in fact commit to no more than two potential boundary sites for red/not-red. It is safe to assume that a plausible theory of boundary shifting would require that any point between two potential boundary sites also be an potential boundary site. In any case, with or without this formal requirement, viewings of the various post boxes will not by themselves induce an equivocation-producing boundary shift.

The General Case

There is, nevertheless, an equivocation “problem” with the post-box example, but it’s not the one Keefe elaborates. Let’s revise the example so as to make it seem deductively rather than inductively valid, just in order to shed any potential distractions.

I’ve seen only red post boxes on my street;
I’ve seen every post box on my street;

Therefore, every post box on my street is red. ¹⁴

I’ll happily agree with Keefe that this is indeed a seemingly unimpeachable argument involving workaday, if insipid, reasoning. I’ll also concede that the argument does, on the boundary shifter’s theory, involve equivocation *in the weak sense*. In fact, *all* seemingly unimpeachable arguments involving workaday reasoning are infected, on the boundary shifter’s theory, with type-(i) or type-(ii) equivocation—at least when offered over time, as they are in everyday life. (For an example of Keefe’s to work, note, it would have to be a case of reasoning taking place over time, since only then could there be either a change in context or a change in facts sufficing to yield an extension change.)

So to what extent does this ruin the boundary shifter’s theory? I say, to none at all. The reason is that the problem here is completely general and it’s completely independent of the boundary shifter’s view about when extension changes take place. To see this, consider the following pretty good argument.

¹⁴I take it that one sees a red post box when one sees a post box that is red, not when one sees a post box that *looks* red.

Paul is taller than John;
John is taller than George;
George is taller than Ringo;

Therefore, Paul is taller than Ringo.

Were we to make an argument such as this one, we would make it over time. We would ascertain the truth of the premises sequentially, over longer or shorter stretches of time. We would utter the sentences sequentially. Would this spoil the argument? Unlikely. But return to the distinction between eternalism and temporalism about the propositions expressed by these utterances. The eternalist says that four different relations are ascribed in the argument, one for each of the different times at which an utterance in it is made: *taller than at t_1* , *taller than at t_2* , *taller than at t_3* , *taller than at t_4* . The temporalist says that the argument involves a single relation, *taller than*, but that which things are related by it may change over time. Either way, there is extension change in the un-adorned phrase “taller than”. Which things are taller than which is constantly in flux as grandparents shrink; babies grow; and seedlings, then saplings, grow even faster.

Equivocation! —Type-(i) equivocation (due to change in context) if the eternalist is right, type-(2) equivocation (due to changes in which things have which properties) if the temporalist is right. Why then would we let the argument stand when made in real time?

We let it stand as long as there are safeguards against our argument’s being damaged by the changes in extension. There need to be assurances that the truth of our premises guarantees the truth of our conclusion despite the changes in extension over time. The validity of the argument statically presented, as above, ensures the acceptability of the argument as made in real time, as it would be in everyday life, whenever the comparative heights of the particular people involved are not changing more quickly than the real-time argument is being made. When the comparative heights of the people involved do not change from time t_1 to time t_4 , then the extension of the relation expression “taller than” does not change *as restricted to the people in our argument*. This is all the safeguard we need. In the case at hand, we can be confident of safety when we are confident that there is a propitious balance of (a) the relative rates of growth and shrinkage among the people involved in the argument, (b) the extent of the pairwise differences between their heights at the times that the truth of the premises is ascertained, and (c) the duration of the event of the argument’s being made.

More generally, if there are no extension changes of the predicates in a real-time argument *as restricted to the objects under discussion in the argument* then the real-time argument is as good as its static version. In the general case, we have safety to the extent that the objects under discussion aren't likely to change so quickly that there would be change in these restricted extensions over the course of the argument's being made.

It is not equivocation *per se* that's problematic. Since one of temporalism and eternalism is correct, whoever makes an argument in real time equivocates in the weak, extended, change-of-extension sense at issue. When we're confident of safety we don't object or fret. When we know that there's safety, we know that our argument is a good one.

So this is where we are. There is indeed extension-change equivocation everywhere—none the more or less so if the boundary shifter's view is correct. But there is not fallacious reasoning everywhere. So the too-much-fallacious-reasoning objection does not stand. This is how the objection went:

Too-Much-Fallacious-Reasoning Objection:

- (a) an argument is fallacious whenever there is equivocation;
- (b) if the boundary shifter were right, equivocation would be rampant; but
- (c) fallacious reasoning is not in fact rampant; so
- (d) the boundary shifter is wrong.

Let's agree that (c) is true: fallacious reasoning is not in fact rampant. If by *equivocation* we just mean equivocation in the weak, extension-change sense, then equivocation is rampant, whether or not the boundary shifter is right; it occurs nearly everywhere arguments are made in real time. In this case, (a) is not true. But if by *equivocation* we mean equivocation in the stronger, more traditional sense of *meaning* change over the course of an argument, then (a) is true but (b) is not, since the boundary shifter appeals not to meaning change but only to extension change—due to a change in context if contextualist, due to a change in which things have which properties if invariantist.

This is not to say that the boundary shifter deems generalizing from the instances in sorites cases a good thing to do. She doesn't. But nor does she think that it involves equivocation, traditionally conceived. She thinks the generalizing involves a mistake. She attributes a kind of goodness (truth at the time of consideration) to each instance so as to explain why we would be tempted to make the mistake of generalizing. The mistake here is not the mistake

of equivocation; rather, it is the mistake of being confident that there is safety in our real-time inference when in fact there is none. We think that there's safety when there isn't because we're not aware of the subtle causes of extension change that have been teased out by philosophers investigating the matter.

Let's spell this out a bit more. We have an argument involving restricted universal generalization such as the following.

If one-thousand millimeters is a short height, then so is one-thousand-and-one;

.
. .
.

If one-thousand-nine-hundred-and-ninety-nine millimeters is a short height, then so is two-thousand;

Therefore, for any height h less than two-thousand millimeters, if h is a short height then so is h plus one millimeter.¹⁵

The argument is valid. (Indeed, it is logically valid given some basic facts about numbers, which we take for granted.) Furthermore, we believe the premises of the argument (perhaps subconsciously) as well as its conclusion. To answer the universalized psychological question is to explain why we believe the conclusion of this argument. I claim that our belief in the premises explains our belief in the conclusion precisely because the premises entail the conclusion. I claim too that our belief in the premises is explained by the fact that each is true when we assess it. The validity of the argument, however, does not ensure that the inference will lead from true premises only to a true conclusion when it is made in real time. This is because the validity of the argument does not ensure the *safety* of the inference. A real-time inference is safe when it is based on a valid argument *and* the extensions of the predicates in it do not change for the duration of the inference *when restricted to the objects under discussion*. Our generalization in this case leads to error because the inference is not safe. It is not safe because which things have the property of being *short* changes as the extension boundary of the property changes, which happens as different pairs of heights are brought to our attention. We make the generalization, however, because we mistakenly think that the inference is safe.

¹⁵Please don't quibble. When I say speak of the heights being quantified over here are of course heights that are integer multiples of one millimeter.

Objection II: If generalizing here why not generalizing everywhere?

We now turn to another sort of generalizing-from-the-instances problem that often gets raised. It goes like this. The instance baser's explanation of why we generalize from the instances doesn't hold any water because there are a number of cases where we believe all of the instances of a universal claim and yet don't believe the universal itself. That there are such cases is supposed to undermine the basic idea about beliefs in instances explaining belief in the corresponding universal generalization. Let's look at examples.

Example: Unconscious Beliefs

When we consider any one of our beliefs, that belief then becomes conscious. Considering whether b is a conscious belief immediately makes us conscious of b . For each belief b , the claim that b is a conscious belief is true at the time that we consider that claim. There is no time at which we can identify a belief b and truly say, at that time, that b is not among our conscious beliefs. In that sense, we cannot identify any of our conscious beliefs.

The situation is clearly parallel in these ways to the sorites situation—as the instance baser conceives it. Considering whether a very similar pair of objects falls on the same side of the extension boundary of a vague predicate F immediately makes them fall on the same side of it. For each such pair x and x' , the claim that if x is F then so is x' is true at the time we consider that claim.¹⁶ There is thus no time at which we can identify a pair x and x' and truly say, at that time, that x and x' straddle the extension boundary for F . In that sense, we cannot identify the extension boundary for vague predicates.

But that's where the parallel ends. We don't think that all of our beliefs are conscious ones, but we do believe that no pair of very similar objects straddles the extension boundary for a vague predicate. In other words, we don't generalize from the instances in the belief case, while we do generalize from the instances in the sorites case. According to the objector, this difference wrecks the instance baser's theory.¹⁸ If we don't generalize in the case of unconscious beliefs, why do we generalize in the case of the sorites? The objector is presumably saying that if our beliefs in the instances don't lead to a belief in the generalization in the former case,

¹⁶My own version of the view is more specific: when we consider any adjacent pair in a sorites series, it becomes strenuously inefficient, because of their great similarity, to treat that pair differentially for the purposes at hand. Then given our standing interest in efficiency, one member of the pair will depart significantly from the norm, in the relevant respect, just in case the other will. But, on my analysis, this is what it is for them to be on the same side of the extension boundary for the predicate in question.¹⁷

¹⁸See Keefe (2007, 282–283). I have also heard the objection in discussion from Timothy Williamson.

then it can't be that our beliefs in the instances are leading to a belief in the generalization in the sorites case, and that our beliefs in the instances therefore does explain our belief in the generalization in the sorites case.

Let's evaluate this objection by making an analogy to event causation. In our case there is an event-type *I*, that of coming to believe all of the instances of a (perhaps restricted) universal generalization, and an event-type *G*, that of coming to believe a universal generalization. The question analogous to the one we're faced with is how could an *I* event ever cause a corresponding *G* event if there are *I* events that do not cause a corresponding *G* event.¹⁹ The general form of the questions is: how could a *C*-type event ever cause a corresponding *E*-type event if there are cases in which a *C*-type event occurs without the corresponding *G*-type event occurring at all.²⁰

Let's particularize the question by considering a stock example from the literature on causation. We have an event-type *T*, an event of throwing a brick towards a window, and an event-type *S*, an event of a window's shattering. The question analogous to the question we're faced with is this: how could a *T* event ever cause a corresponding *S* event if there are cases in which a *T* event occurs without a corresponding *S* event occurring at all? The relevant situation is of this sort: Suzy throws a brick towards a window and it shatters; Billy throws a brick towards a different window and it doesn't shatter. The question to consider is: how could Suzy's throw cause her window to shatter if Billy's throw doesn't cause his window to shatter? The answer to this question is perhaps all too obvious. Brick-throwings at windows cause shatterings of those windows only *defeasibly*. Billy might have a weak arm or bad aim. Pammy might hurl a shot put towards Billy's brick in order to throw it off course because Billy's brick was heading towards her bedroom window. There are lots of potential defeaters here; none of them casts doubt on the idea that Suzy's throw caused her window to shatter.

The same holds true of explanation. Why did that window shatter? Because Suzy threw a brick at it. Billy's failure to shatter a window with his throw in no way vitiates this explanation.

The instance baser appeals to a similar principle of explanation: we can explain why someone believes a certain claim by citing their beliefs in some claims that entail it.²¹ In particular,

¹⁹We say that an event of type *G* *corresponds* to an event of type *I* when the universal generalization from the *I* event is the same generalization as the one in the *G* event (and the agent is the same).

²⁰For some notion of correspondence appropriate to the type of events involved.

²¹There's a delicate issue here that's being finessed. The claims being spoken of here must not be propositions since the proposition believed is not entailed by any other propositions that are believed. By *claims* we mean something more like sentences. Some premise-claims entail a conclusion-claim when the argument from those premises to that conclusion is valid.

we can explain why someone believes a certain universal generalization by citing their beliefs in all of its instances (along with their belief that those are all of the instances). That there are cases when people do not believe claims that are consequences of claims they do believe in no way vitiates this explanation.

The reason the explanation works as an explanation is that belief in some group of claims does often lead to belief in claims that they entail. This is defeasible, however, as witnessed by the examples discussed in the section.

What then are the defeaters in the case of unconscious beliefs? We can think of the case as involving an inference, made in real time, taking the form: my belief b_1 is not unconscious; my belief b_2 is not unconscious; . . . ; my belief b_n is not unconscious; therefore, none of my beliefs b_1 through b_n is unconscious. The statically presented argument $C(b_1), \dots, C(b_n)$, therefore $\forall b_i, 1 \leq i \leq n, C(b_i)$ is valid, but the real-time inference is not good at all since it is obviously not *safe*: consciousness and unconsciousness for beliefs are fleeting properties. Since we know this, we know that the argument is completely unsafe and therefore are not tempted to make it. In this case we have what we can call a *known-danger defeater*.

Example: The Lottery Paradox

Five million tickets have been sold for a lottery drawing. You have bought one of them. You believe that yours won't win, and you have other beliefs that depend on this one, for example, that you won't quit your job anytime soon. We stipulate that you believe of each of the other individual tickets that it won't win either. (At least some of these beliefs are subconscious, presumably.) Nevertheless, you don't believe that *none* of the tickets will win; in fact, you believe the contrary. You know that one of the tickets *will* win.

We have a paradox if we think that (i) for each of the beliefs just described, it is rational for you to have that belief—on the grounds that your evidence makes its truth overwhelmingly likely; that (ii) rationality of belief agglomerates in the sense that the rationality of belief in each of p and q requires the rationality of belief in their conjunction; and that (iii) belief in a contradiction is never rational.

The standard response is to give up the agglomeration of rationality on the grounds that a belief *simpliciter* can be rational even when it is not believed with absolute certainty, that is, rationality requires only something less than degree of belief one on a scale of zero to one. I find agglomeration of rationality attractive, however, and therefore incline towards a different response, namely, that (i) should go: it is not the case that each of your beliefs is rational.

In any case, the rationality of your beliefs is not what is at issue here. We have stipulated that you do believe (irrationally, I think) of each of the tickets that it won't win and yet that you don't believe that none of them will win.

Again, we have an example where you believe each of the instances of some universal generalization (and you know that they are all of the instances), but you don't believe the generalization itself. As before, since you don't believe the generalization, there is no belief-in-a-generalization that needs to be explained. Our principle of explanation is that *if* a universal generalization is believed then the belief in it can be explained by citing belief in every instance of it.

Since I do claim that beliefs in instances lead to belief in a universal, albeit defeasibly, and that this is why the explanation principle is a good one, it behooves me to explain what might be the defeater in the case of the lottery paradox. Assuming that you're not absolutely certain with respect to each ticket that it will win but rather believe that only to degree slightly less than one, your beliefs should not agglomerate. As n gets bigger, you have less and less confidence of any given group of n tickets that it will contain only losers. This is what your evidence dictates anyway. In other words, although the five-million premises "ticket #1 won't win" through "ticket #5,000,000 won't win" entail the conclusion "none of tickets # 1–5,000,000 will win", you do not come to believe that conclusion despite your belief in each of those premises because you have very low credence in the conjunction of those premises and hence do not believe that conjunction. This is what your evidence requires. In this case we have what we can call a *probabilistic defeater*.

Example: The Paradox of the Preface

The paradox of the preface has the same general form as the lottery paradox. In fact, we could think of them as special cases of the very same paradox if we stipulated that our author, although he affirms each claim made in the main text of his book, is not absolutely certain of each of these claims.

To make the case interestingly different, then, let's imagine that the author is absolutely certain of the truth of each of the claims made in the main text. In this case he should be absolutely certain of the conjunction of claims that each is true, and therefore should be absolutely certain of anything that the truth of those claims entails (as long as he's actually drawn the inference: he has sufficient logical acumen, and he exercised it). By stipulation, he is not absolutely certain of everything he realizes that the truth of the claims in his book entails, since

he denies the universal generalization that all of the claims in his book are true. We can agree that he is irrational, as well as insecure. Probably it is an unfitting combination of insecurity and brash cockiness that leads to his irrationality. He should either become absolutely certain that he has made no mistake in his book, or he should lower his credence in at least one of those claims. If he can't find a weak link, but remains certain that there must be at least one, then he should lower his credence somewhat in each of the claims. If he doesn't do this then he simply refuses to bring his beliefs into line with one another and there is just no explaining why he believes what he does by citing any rational justification for them.

Conclusion

⋮

References

- Fara, Delia Graff (2000), 'Shifting Sands: An Interest-Relative Theory of Vagueness', *Philosophical Topics* **28**(1): 45–81. Published under the name 'Delia Graff'.
- Fara, Delia Graff (*forthcoming*), 'Profiling Interest Relativity: A Reply to Stanley'. *Forthcoming in Analysis*.
- Kaplan, David (1989), 'Demonstratives', in J. Almog, J. Perry & H. Wettstein, eds., *Themes from Kaplan*, Oxford University Press, Oxford, pp. 481–563.
- Keefe, Rosanna (2007), 'Vagueness without Context Change', *Mind* **116**(462): 275–292.
- Raffman, Diana (1994), 'Vagueness without Paradox', *Philosophical Review* **103**(1): 41–74.
- Soames, Scott (1999), *Understanding Truth*, Oxford University Press, New York.
- Wright, Crispin (1987), 'Further Reflections on the Sorites Paradox', *Philosophical Topics* **15**: 227–290.
- Wright, Crispin (1992), *Truth and Objectivity*, Harvard University Press, Cambridge, MA.