Vagueness, Logic and Use: Some Experimental Results

Phil Serchuk, Ian Hargreaves and Richard Zach

Although the use of experimental data is a somewhat new methodology for philosophers, research on vagueness is poised to make excellent gains from the approach. This is because any theory that purports to explain the vagueness of natural language can benefit from data about how ordinary speakers use vague words. Nonetheless, we must carefully consider the kind of evidence that we're interested in. It wouldn't be helpful to solicit naïve opinions about the plausibility of different philosophical theories: we aren't interested in what non-philosophers think about epistemicism or degree theory. Rather, the goal is to collect data that will tell us which theories of vagueness comport with linguistic use. Of course, theories of vagueness cannot be evaluated solely on the basis of empirical evidence. But if data shows that speakers do not use vague predicates as a given theory predicts, then proponents of that theory will need to explain why their theory accurately describes the semantics of vague predicates but not our use of them. This is a new kind of objection that must be overcome. Traditionally, appeals to common practice were made from the armchair, and even then philosophers rarely gave a detailed account of the relationship between their theory and our linguistic use of vague predicates. This is a shortcoming that empirical methods promise to help us overcome. In this paper we present data from three experiments. Each tests a specific claim in the literature which, in our view, data on linguistic use has an especially helpful role in evaluating. We do not suggest that our experiments definitively establish or refute any of these claims; rather, we hold that they provide valuable data that can be used when assessing intuitions about the use of vague language.

The first experiment tests a claim made by Brian Weatherson concerning the intuitive plausibility of the treatment of borderline cases by fuzzy logic. At least some of the arguments for fuzzy logic claim that notions like degrees of truth and fuzzy boundaries are rooted in common sense.[1] To that extent fuzzy logic has more to prove in the empirical arena than other theories which rely less on appealing to (purportedly) intuitive concepts. According to Weatherson, a “good theory of vagueness should tell us ... what is wrong with Sorites arguments ... [and] why the premises looked plausible to start with”.[2] The fuzzy logician answers that the Sorites premises “look plausible because we confuse near truth for truth”.[3] As Edgington puts it: “The difference between clear truth and almost clear truth – between 1 and 0.99 – is an insignificant difference upon which, normally, nothing hangs”.[4] But, Weatherson claims, this account is plausible only in Sorites arguments that use conditionals. The Sorites' inductive premise usually takes the following form: (SI): P(n) → P(n-1). But in classical logic, (SI) is equivalent to (SA): ~(P(n) ^ ~P(n-1)) and (SO): ~P(n) v P(n-1). Fuzzy logic treats (SI) having a greater or equal truth-value than (SA) and (SO), which are assigned identical truth-values. Weatherson gives a thought experiment purporting to show that (SA) is the most plausible version of the inductive step, followed by (SI) and (SO). So, his argument goes, fuzzy logic fails to capture the relative plausibility of the inductive step's different logical forms. This is a strong argument against the degree theorist, for her argument depends on there being an at least rough correlation between degrees of truth and plausibility.

We tested Weatherson's claim. We asked 243 subjects, via paper survey, to rank the relative plausibility of the different logical forms of the Sorites' inductive step. Subjects were divided into two groups, one of which was asked about 'heap' and the other about 'rich'. Each group was presented with natural language translations of the three logical forms and asked to rank them in order of persuasiveness. We made two predictions. First, we predicted, contra Weatherson, that respondents would rank (SI) as being more persuasive than (SA). The second was that respondents would, as Weatherson predicts, rank (SA) as being more persuasive than (SO). Both predictions were supported by the data, which showed a statistically significant and identical ranking in both groups. In sum: 63% of respondents ranked (SI) as being more persuasive than (SA), and 67% ranked (SA) as being more persuasive than (SO). So fuzzy logic's treatment of (SI) as being 'truer' than (SA) parallels speakers' intuitions about their relative persuasiveness; its treatment of (SA) as being equally true to (SO) does not.

Our second experiment tests speakers' use of 'definitely' in natural language. Many theories of

1 This is especially true among the so-called fuzzy logic community in computer science; philosophical accounts are more tempered. For example, Machina claims that his “inclinations [about degrees of truth] are at least verbally in agreement with the common sense view,” though he recognizes “that agreement cannot be taken at face value as an indication that the common man thinks of degrees of truth in the same way” (Machina, Kenton F. “Truth, belief and vagueness,” *Journal of Philosophical Logic*, Vol. 5 (1976), p. 54).

2 Weatherson, Brian. “True, Truer, Truest,” *Philosophical Studies,* Vol. 123 (2005), p. 61.

3 Weatherson, p. 61.

4 Edgington, Dorothy. “The Philosophical Problem of Vagueness,” *Legal Theory*, Vol. 7 (2001), p. 375.

vagueness, particularly in their account of higher-order vagueness (e.g. Fine[5]), appeal to a 'definitely' operator. Sometimes this operator is taken in a technical sense, and in these cases correspondence with natural language may not be important. But theorists also appeal to the operator when explaining our use of vague language. Keefe argues that speakers confuse disjunctions like 'Fa v ~Fa' with 'definitely Fa v definitely ~Fa': “it is so common for our judgments of both sentences to be dictated by our judgments of [the latter]”.[6] Williams generalizes Keefe's argument in his explanation of the confusion hypothesis. “The confusion hypothesis maintains that we confuse an utterance of 'There is something that is *F*' with the claim that *there is something that is definitely F:* our intuitions about the former track the truth-values of the latter.”[7] We tested this hypothesis in two ways. For the first, we stipulated that some item *a* was a borderline case for a vague predicate 'F'. As before, there were two groups of subjects: one was asked about 'heavy', the other about 'rich'. We then asked respondents to assess the truth of a sentence of the form 'F*a*', and then later for one of the form 'definitely F*a*'. Participants had to choose one of the following options: 'true', 'false', 'not true, but also not false', 'partially true and partially false', 'both true and false', and 'true or false, but I don't know which'. We predicted, contra the hypothesis, that speakers would distinguish between the sentences by giving them different assessments: in particular, that sentences of the form 'F*a*' would be described as something other than 'true' or 'false' and that sentences of the form 'definitely F*a*' would be described as 'false'. Our prediction was supported by the data: approximately 75% of respondents assessed 'F*a*' as something other than 'true' or 'false', and approximately 75% answered 'definitely F*a*' as 'false'. Although we did not ask subjects to compare 'Fa v ~Fa' with 'definitely Fa v definitely ~Fa', such a method could be used to test Keefe's version of the confusion hypothesis.

The second method used a modified version of the questions asked by Bonini et al. (1999) in their experimental work on vagueness. Bonini et al. asked participants to give the boundaries to particular applications of vague predicates; for example, to give the smallest height that would count a man as 'tall'.[8] Building on this approach, we posed the same question to one group of respondents and an identical question, substituting 'definitely tall' for 'tall' (*mutatis mutandis* for the other predicates), to a separate group. Our hypothesis was that respondents in the 'definitely' group would attribute greater numbers to the cut-off points than those in the other group. The data proved inconclusive. Although the responses were consistently larger, the difference was statistically significant only in two of the five cases.

The third experiment tests speakers' use of negation in borderline cases. However negation works, we can assume that it must satisfy the following conditions: the negation of 'F*a*' is clearly true when '*a*' is a clear non-case and is clearly false when '*a*' is a clear case. But several different options have been proposed for how to treat the negation of 'F*a*' when 'F*a*' is neither clearly true nor clearly false, including strong, weak and intuitionistic/Gödel negation.[9] Certainly there are real technical differences between these proposed variants. But it is an open question whether these differences are reflected in speakers' use of negation. To test this, we had subjects assess the truth-values of six different sentences. As before, respondents were divided into two groups: one group was asked about 'rich', the other about 'heavy', and '*a*' was stipulated to be a borderline case. In our analysis, sentences were divided into pairs: (1) 'not F*a*' and 'it is not the case that F*a*', (2) 'Either F*a* or not F*a*' and 'Either F*a* or it is not the case that F*a*', and (3) 'F*a* and not F*a*' and 'F*a* and it is not the case that F*a*'.

For each pair we made two hypotheses: one which tracked the truth-value ascribed by the logician, and another which tested for differences in responses between negation types. For example, 'not F*a*' is supposed to be the linguistic form of strong negation and 'it is not the case that F*a*' the form of weak negation. So our first hypothesis was that 'not F*a*', the strong negation of a borderline case, would be described as something other than 'true' or 'false', and 'it is not the case that F*a*', the weak negation of a borderline case, would be described as 'true'. We generated similar hypotheses for the other four sentences. The second hypothesis was that, for pairs (1) and (2), respondents would describe the strong negation differently than the weak one; the two forms in pair (3) were predicted to receive similar descriptions. The data supported the first hypothesis and was partially

---

5 Fine, Kit. “Vagueness, Truth and Logic,” *Synthese*, 30 (1975), pp. 265-300.

6 Keefe, Rosanna. *Theories of Vagueness*, Cambridge: Cambridge University Press (2000), p. 164.

7 Williams, J. Robert G. “An Argument for the Many,” *Proceedings of the Aristotelian Society,* Vol. 103:3, (2006), p. 412.

8 Nicolao Bonini, Daniel Osherson, Riccardo Viale and Timothy Williamson. “On the Psychology of Vague Predicates,” *Mind & Language*, Vol. 14:4 (1999), pp. 377-93.

9 The weak negation of P, $\neg$P, is clearly true iff it's not the case that P is clearly true, and false otherwise. P's strong negation, ~P, is clearly true iff P is clearly false and clearly false iff P is clearly true. And P's intuitionistic (Gödel) negation, $\sim_g$P, is clearly true iff P is clearly false, and false otherwise.

consistent with the second. Although many of the predictions were observed, there were some puzzling results. For instance, a plurality of respondents (39%) classed 'P or not P' as false; this contrasts with almost every theory's predicted result, though it is consistent with Keefe's confusion hypothesis. Although it was not the focus of our research, we use this data (along with some from related questions) to make observations about the presence of bivalence and the law of non-contradiction in speakers' use of vague language.