

The Sorites is nonsense disguised by a fallacy

The Sorites paradox, usually treated as a problem in logic or formal semantics, also lends itself to experimental investigation. Consider a case in which a human agent confronts a series of 1000 colour cards. The first card is painted with primary red paint, one litre of it in a two-litre pot. For all n , the paint used to paint the n^{th} card is obtained by stirring just one small blob (approx. 1 ml.) of yellow into the pot of paint with which its predecessor, the $(n-1)^{\text{th}}$ card, was painted. The 1000th card is clearly orange; it would be called orange by anyone with normal colour vision and a proper command of the English language.

The agent can be exposed to this series in a number of ways. He or she can be presented with its member cards one by one or in pairs, *seriatim*, or in random order. At each presentation, the agent has to give the verdict 'Red' (when the exhibit appears to be more red than orange) or 'Orange' (when the exhibit appears to be more orange than red) and is allowed to refrain from giving either of these verdicts when an exhibit appears either neither red nor orange, or both.¹ When the subject is taken on a 'forced march' from the first card in the series to the last, there will, for any trial, be a first point (which tends to be different from one trial to another) where the subject switches from a 'Red' verdict on one card to a non-'Red' verdict on the next; this is obvious from the fact that the agent does not end up calling the 1000th card 'Red'.

It is, of course, difficult in practice to control all the variables in a Sorites experiment, but we can do so by arranging short journeys backwards in time. In this science fiction version of the Sorites colour card experiment, a subject is presented, at time t_1 , with the first exhibit and, after giving his or her verdict on it, is transported back to t_1 , when the subject's brain reconfigures to the exact state, b_1 , it was in at that time. Then the next

exhibit is presented to the subject who gives a verdict and is transported back to t_1 and reset to b_1 , and so it goes on until the subject has given his or her verdict on all exhibits. In this version of the paradox, there is literally no time for any change in environment; time stands still, the rest of the world is held constant; only the exhibits change.

In a normal Sorites experiment, a subject may become tired, may have memories of earlier phases of the experiment; hysteresis and other psychological changes may occur (Raffman 2005), his or her purposes and interests or those of the experimenter may fluctuate, external conditions such as lighting and distractions may vary etc.. In the sci-fi 'time travel' version of the experiment, all of these factors are eliminated. Yet, as Tim Williamson rightly observes '[v]agueness remains even when the context is fixed' (Williamson 1994: 215)², for Sorites-susceptibility is preserved — even in the colour card time-travel version, we seem driven to a false conclusion, namely, that if the subject judges the first card in the series to be red, then, if he or she were, *at the same time*, to judge the 1000th (orange) card, the verdict on that would be 'Red' too. Or, to change examples, if I now judge a tadpole to be a tadpole, I would now judge a frog to be a tadpole!

In the sci-fi experiment, at that point in the series where a subject judges one card red and an adjacent card something other than red, the *only* factor that can precipitate this switch is the difference between the two exhibits.³ Yet we know that the difference in colour between neighbouring cards is too small to be perceptible to the unaided eye. The correct conclusion now seems obvious: the subject is affected by a difference that he/she does not consciously perceive. Access denied (Block 2007; 2008). The celebrity example is blindsight (Horsten 2010: 213) but it is a commonplace that processes of which subjects are not conscious frequently play a causal role in their perception and decision-making.

A neuron (like a condenser in an electrical circuit) gradually builds up charge until, with perhaps the tiniest of additions (the straw — or molecule — that breaks the camel's back) it discharges. Very crudely, the sudden discharge of a relevant array of neurons precipitates that switch of verdict (from 'Yes, it's red' on one member of the series to a different verdict on an adjacent exhibit), a switch that must occur somewhere between the first and the thousandth colour patch. It is interesting, but completely nonparadoxical, that neural processes, unavailable to consciousness, are responsible for an individual's switching judgment between one member of the series and a neighbouring member.⁴

Graham Priest writes 'The meanings of vague predicates are not determined by some omniscient being in some logically perfect way. Vague predicates are part of *our* language. As a result, their meanings must answer in the last instance to the use that *we* make of them. It is therefore difficult to see how there could be a semantic cut-off point that is *in principle* cognitively inaccessible to us' (2003: 12). We have already observed that, for an individual subject, in a given trial run of a Sorites experiment, there will, nonparadoxically, be a first cut-off point when a judgment-switch occurs. But subjects differ in where they cut off, and an individual subject will cut off at different points across different trials. When Priest speaks of a *semantic* cut-off point he is talking not of the behaviour of agents but of, for example, a point along the colour continuum where red ends and non-red begins. He is thus talking about the Sorites in its standard (agent-independent form), and it is to this that we now turn attention.

Epistemicists (whom Priest is here attacking) hold that there must be a cut-off, or borderline, because, in the standard Sorites, we make a transition from 'X₁ is F' to 'X_{last} is not F' so there must, so they think, be a transition point somewhere between the first and the last member of the series. Now, while it is perfectly true that, for any given run of a Sorites

experiment, there will be a point at which the subject switches from an 'F'- to a non-'F' verdict, it is a fallacy to infer that because, for each Sorites experiment, there is such a borderline between 'F'-judgments and non-'F'-judgments, there must be a borderline between F and non-F *simpliciter*. A particular metal, such as copper, lead or mild steel has a tensile strength, but it is obviously incorrect to infer that there is such a thing as the tensile strength of metal. Unrelenting noise at 80 decibels is annoying and, if exposed to it, at some point even the most patient of us will snap (some much earlier than others). But to ask 'What is the snapping point, *simpliciter*, for unrelenting noise at 80 decibels?' would be nonsense. And similarly, it is nonsense — a category mistake — to ask where the red/non-red borderline is, not for a particular agent but objectively. The same can, of course, be said about the heap/non-heap borderline, the infant/non-infant borderline, the wealthy/non-wealthy borderline etc.: no such objective borderlines exist.

This conclusion should occasion little surprise. Many discussions of the Sorites start from noting that vague predicates are 'tolerant', and this implies no sharp borderlines. Consider also true blue. A subject, presented with various shades of blue, is asked to say which one seems true blue — pure blue untainted by any other colour. Subjects who, on independent tests, are established as having equally excellent colour vision, tend to differ about which shade is true blue. This might be thought to be faultless disagreement, except that the appearance of disagreement is illusory. It's simply that a particular shade of colour looks true blue to one competent observer and not true blue to another. True blue is a response-dependent property. There is no objective true blue about which these observers are disagreeing (Triplet 2007). For exactly the same reasons, there is no objective red/non-red borderline. As Haim Gaifman (2010: 7) remarks, '[b]orderline vagueness is not mere absence of semantic determination, but recognition, shared by competent speakers, that a

certain divergence in usage is *legitimate and to be expected*'. Where, for vague predicate 'F', some competent observers judge it to be true of a particular object, while other equally competent observers judge it to be false of the same object, then there can be nothing to warrant the claim that that object is F (or that it is not). It appears F to some, not F to others; it is neither objectively F nor objectively not-F; this is determined by the meaning — by the use — of 'F'.

We have argued that when competent observers differ in their assessment of a shade of colour (assuming good lighting conditions etc.) then there is no observer-independent correct judgment of the colour of the shade that would render one of the two observers correct, the other incorrect. We may wonder what the colour of the object *really* is in itself, but, as Michael Dummett points out, 'the question is nonsensical. It is nonsense because its appearance to us is an interaction between it and us: it can look a certain way only *to* a viewer. The object itself is not an interaction, but something that interacts with us and with others: it is nonsense to speak of its looking this way or that save to one who looks at it'.⁵

We know a great deal about the visual discriminatory powers of various animals, but do not know how the colour of a ripe tomato looks to a bee, any more than we know what it is like for a bat to be a bat. When we ascribe a particular colour to something, we therefore are saying how that thing looks to us competent human perceivers; to be red is to be perceived as red by us; the objectivity of the ascription is just a matter of our overwhelming intersubjective agreement. Hence, where there is significant *disagreement* among competent observers about whether a shade is red or is non-red, there is no objective fact of the matter in dispute and it is nonsensical to say of any such shade that it is really, objectively red (or that it is really non-red). With judgments of colour, as with all vague

predications where there are clear cases of F-ness and also contested cases, there are both quasi-objective judgments (the appearance of objectivity being due to massive agreement of judgment) and other judgments, akin to expressions of personal taste, where a rational subject will acknowledge that there is no question of securing agreement.⁶

The standard (forced march) version of our colour card Sorites is usually formulated as a series of modus ponens arguments, each containing a conditional of the form

If Exhibit_n is red then Exhibit_{n+1} is red.

As we have seen, for some exhibits, there is irresolvable difference of judgment among competent observers about whether they are red or are non-red and so, as we have tried to show, to claim of any such exhibits that they are (objectively) red is not false but nonsensical. Hence there will be some instances of the above conditional with *nonsensical* antecedent and/or consequent, rendering the conditional as a whole nonsensical. The Sorites is not solved by showing a premise or premises to be false but is dissolved by showing some to be nonsensical. There is not a precise answer to the question of which is the first nonsensical conditional in the series, for how much dissent is needed from an 'F'-verdict on an object for it to be unreasonable to call that object (quasi-) objectively F is not an objective matter. The standard Sorites paradox is, then, nothing more than a nonsense resting on a simple fallacy.⁷

REFERENCES

Åkerman, J. and Greenough, P. 2010. Hold the context fixed — vagueness still remains. In Dietz, R. and Moruzzi, S. (eds), 275-88.

Alxatib, S. and Pelletier, F.J. (forthcoming). The psychology of vagueness: borderline cases and contradictions. *Mind and Language*.

Beall, J.C. (ed.) 2003. *Liars and Heaps: New Essays on Paradox*. Oxford: Clarendon Press.

Block, N. 2007. Consciousness, accessibility and the mesh between psychology and neuroscience. *Behavioral and Brain Sciences* 30: 481-548.

Block, N. 2008. Consciousness and cognitive access. *Proceedings of the Aristotelian Society* 108: 289-317.

Cohen, J. 2009. *The Red and the Real: An Essay on Color Ontology*. Oxford: Oxford University Press.

Dietz, R. and Moruzzi, S. (eds). 2010. *Cuts and Clouds: Vagueness, its Nature and its Logic*. Oxford: Oxford University Press.

Dummett, M. 2010. Sense-data and senses. In Muntz, V., Puhl, K. And Wang, J. (eds), 97-102.

Gaifman, H. 2010. Vagueness, tolerance and contextual logic. *Synthese* 174: 5-46.

García-Carpintero, M. and Kölbel, M. (eds) 2008. *Relative Truth*. Oxford: Oxford University Press.

Heck, R. 2003. Semantic accounts of vagueness. In Beall, J.C. (ed.), 106-27.

Horsten, N. 2010. Perceptual indiscriminability and the concept of a color shade. In Dietz, R. and Moruzzi, S. (eds): 209-27.

Iacona, A. 2008. Faultless or disagreement. In M. García-Carpintero and M. Kölbel (eds): 287-95.

Kaneoke, Y., Urakawa, T., Hirai, M., Kakigi, R., Murakami, I. 2009. Neural basis of stable perception of an ambiguous apparent motion stimulus. *Neuroscience* 159: 150-60.

Munz, V., Puhl, K. and Wang, J. (eds). 2010. *Language and World. Part Two: Signs, Minds and Actions*. Frankfurt: ontos.

Priest, G. 2003. A site for Sorites. In Beall, JC. (ed.), 9-23.

Raffman, D. 2005. How to understand contextualism about vagueness: reply to Stanley. *Analysis* 65: 244-8.

Tripllett, T. 2007. Tye's missing shade of blue. *Analysis* 67: 166-70.

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

NOTES

¹ For experimental evidence and critical discussion, see Alxatib and Pelletier (forthcoming).

² See also Heck (2003: 118-20), Åkerman and Greenough (2010).

³ I say 'precipitate' rather than 'cause' because, as John Bigelow pointed out, the difference between two cards may simply increase the probability of my shifting from a 'Red' to a non-'Red' verdict in a non-deterministic way.

⁴ There is a considerable body of experimental work on cortical activity associated with perceptual switching. See, for example, Kaneoke et al. (2009).

⁵ Dummett (2010: 99-100). For an extended defence of colour relationism, see Cohen (2009).

⁶ Andrea Iacona (2008: 291) calls this 'an underlying ambivalence in our way of talking'.

⁷ I am grateful to Michael Morreau for inspiring this paper with the suggestion of a 'counterfactual Sorites' that stymied an earlier attempt of mine to solve the paradox, and to Lynda Burns, Paul Egré, David Ripley and Zach Weber for lengthy discussions. And I should like to thank audiences at the universities of Hong Kong, Kent and Monash, and at Birkbeck and the London School of Economics, for valuable criticism.