1 Introduction

Little empirical evidence has been discussed so far in the philosophical literature in order to shed light on the phenomenon of vagueness, despite the existence of a large body of psychological literature on categorization and discrimination.¹ Most often, vagueness is discussed from the standpoint of a single thought experiment, namely the sorites paradox. Indeed, philosophical treatments of vagueness can be classified depending on which stance is taken on the sorites paradox and on the status of the main premise of the sorites in particular. My aim in this paper is not to object to the methodology which consists in finding the best logical treatment to the sorites, for ultimately, I consider that such a treatment is needed, and that logical matters cannot be evaded. Nor would I recommend discarding normative intuitions about vagueness, for here as elsewhere they necessarily guide philosophical thinking. However, I am of the opinion that new and potentially fruitful hypotheses for our understanding of sorites series can be gathered from the experimental literature, and more generally that thought experiments themselves can be emulated by the consideration of actual experiments.

In this preliminary essay I thus wish to discuss an intriguing set of stimuli originally designed by psychologist G. Fisher (see Fisher 1967), bearing a striking analogy to a soritical series. The specificity of Fisher’s figures is that they combine two phenomena: vagueness and ambiguity. In the literature on vagueness, the two phenomena of vagueness and ambiguity have been

¹Numerous exceptions ought to be mentioned, of course. What I mean is that few empirical data have been discussed by philosophers in proportion to the vast amount of theoretical work accumulated on vagueness.
mostly opposed, and rightly so in my opinion, but mainly in relation to the phenomenon of lexical ambiguity (see Fine 1975, Keefe 2000, Bromberger 2008). Despite this, Raffman has made the suggestion that within soritical series, borderline cases typically pattern as ambiguous stimuli (see Raffman 1994). For instance, Raffman considers a series of color shades making a smooth transition from red to orange and notes that “borderline cases are cases for which looking red and looking orange are very much alike”. One observation she makes – presumably on the basis of her introspection in this particular case – is that while the color’s quality appears to change as the judgment flips, there is also a more basic sense in which the stimulus remains constant. Her conclusion is that “such an effect may amount to a kind of Gestalt switch: there is a similar respect in which (for example) the duck-rabbit ‘looks the same’ while yet ‘looking different’ as it species fluctuates” (Raffman 1994: 53).

Though highly suggestive, Raffman’s comparison between the kind of instability experienced in sorites series and the sort of multistability experienced in ambiguous stimuli has not received close attention. In this paper, I want to suggest that further evidence might be adduced in favor of her hypothesis from the consideration of Fisher series. Prima facie, what Fisher series may reveal is only that ambiguity is a gradable notion, at least when it comes to perceptual ambiguity, and therefore that ambiguity itself is a vague concept. However, I propose to examine the symmetric and more interesting hypothesis, according to which standard sorites series may share structural features with Fisher series. To the extent that this analogy can be sustained, I shall argue that it is at odds with one understanding of the epistemic conception of vagueness, on which a vague predicate must have a sharp boundary along a sorites series. At the same time, the analogy will allow us to give a precise articulation of the idea that borderline cases are cases for which contradictory judgments are permissible (see Wright 1994).

2 Sorites series and hysteresis effects

Sorites series are series of stimuli that are gradually altered in such a way that the first stimulus in the series clearly and unambiguously instantiates a given category $A$, while the end stimulus clearly and unambiguously instantiates a distinct and exclusive category $B$. Adjacent stimuli in the series are assumed to be hard to discriminate when taken pairwise. The puzzle raised by sorites
series is that although the first individual and the end individual in the series are very distinct when taken pairwise, there seems to be no way to draw a non-arbitrary boundary between the categories A and B along the series.

When exposed to a sorites series and forced to categorize one way or the other (a situation T. Horgan called forced march), however, subjects do draw a boundary, as can be expected on logical grounds. What is interesting is how subjects tend to do this. In recent work (see Lindsey, Brown and Raffman 2005 and Raffman 2009), Raffman and colleagues have undertaken a series of experiments on color perception which show that soritical transitions between two color categories give rise to hysteresis effects, namely to the longer persistence of one judgment over the other, depending on which color category one is coming from. Thus, within subjects and across subjects, the point at which subjects switch their judgment from “blue” to “green” is significantly displaced relative to the point at which they switch their judgment from “green” to “blue”. In other words, the direction of the sorites matters to the way people draw the boundary between the two categories. Hysteresis effects are interesting from a theoretical point of view because they suggest that there is a range of cases for which subjects can equally respond in two ways.

Several interpretations of this phenomenon are conceivable, however. Standardly, vagueness is seen as a phenomenon of semantic indefiniteness, such that central shades in a soritical series instantiate neither of the categories that are instantiated at the end of the series. Alternatively, vagueness has been described as a phenomenon of ignorance, namely as a form of uncertainty regarding the location of a precise boundary. On both of these interpretations, one may expect subjects who are forced to categorize one way or the other to adopt a conservative strategy for the range of cases that are indefinite relative to the categorization task, or to maintain their previous classification as far as possible for the range of cases for which they are

\[2\] See Raffman (2009) for details. In their experiment, subjects were given the choice between three answers, “blue”, “green” and a “?” answer for where they would feel dissatisfied with the “blue” and “green” answers. Lindsey, Brown and Raffman’s data show clear instances of a range of stimuli on which the “blue” and “green” judgment overlap, or so that the switch to “?” is shifted depending on the order of presentation. In her interpretation of the data, Raffman suggests interpreting “?” as meaning “borderline”. But if borderline cases are seen, as C. Wright suggests, as cases where subjects can judge either way, one may be tempted to characterize the borderline area more broadly as the whole overlap area on which hysteresis occurs.
This, however, may fall short of explaining on what grounds subjects would choose to be conservative. Another interpretation for the phenomenon is that along a sorites series, the central shades might bear cues that support both interpretations rather than neither. Not only would subjects be conservative for those cases, but arguably they would exploit cues that favor the category from which they come until cues to the contrary become more prevalent. This view of the hysteresis phenomenon, arguably, may be more readily compatible with a general conception of vagueness either as the outcome of an overlap between categories (as in glut theories of vagueness, or degree theories) or as a form of boundarylessness (see Sainsbury 1990).

Interestingly, hysteresis effects have been reported independently in the psychological literature on the perception of bistable stimuli (see Hock et al. 1993, 2004, Gregson 2004), which help to clarify this issue. Hock and colleagues discuss in particular the case of motion quartets. Motion quartets consist of imaginary rectangles whose opposite corners along the diagonals are materialized by dots that twinkle alternately (imagine a rectangle whose corners are Top Left, Top Right, Bottom Left, Bottom Right: TL-BR get illuminated together, then BL-TR together, and so forth; see Hock et al. 1993, 2004 for figures). Motion quartets are bistable stimuli, since one can perceive the opposite points either as moving vertically along the left and right sides of the rectangle, or as moving horizontally along the top and bottom sides. Which of these two percepts is seen first depends in large part on objective cues, namely on the ratio of the vertical path length to the horizontal path length of the rectangle (also called aspect ratio). When the aspect ratio is near 1 to 1.25, namely when the rectangle is a square, both percepts appear equally likely to be first perceived (see Hock et al. 1993: 66). When the ratio is less than 1, the vertical path motion is more likely, and conversely for the horizontal path motion. Thus the motion perceived appears to depend on a “shortest path” principle. In series of trials in which the aspect ratio gradually increases from 0.5 to 2.0, or gradually decreases from 2.0 to 0.5, hysteresis effects are observed. This indicates, once again, that confronted with an ambiguous stimulus, subjects categorize not directly as a function of the aspect ratio statically presented, but also as a function of their previous choices and of the dynamics of the presentation of the stimuli.

\[3\text{See Raffman (2009) for a detailed discussion of the hypothesis that hysteresis might be evidence for the epistemic theory of vagueness.}\]
It would be well beyond the scope of this paper to investigate the nature and origin of hysteresis effects in general (see Kelso 1995 and Raffman 2009 for extensive discussions). However, one relevant question posed by Hock and colleagues for our own inquiry is whether hysteresis effects can be considered properly perceptual, or whether they are rather “judgmental”, namely to be imputed to a higher level of representation such as a judgment bias in the presence of uncertainty. The conclusion they reach in the case of motion quartets is that, because subjects are never uncertain of what they see in motion quartets (they clearly see either horizontal motion or vertical motion, exclusively), hysteresis can be said to be properly perceptual, rather than grounded in a judgmental bias (see Hock et al. 1993, p. 70).

The situation may be different in the presence of borderline cases of an arbitrary sorites series, however. When confronted with a series of colors, for example, each individual color presents a particular quale. Borderline cases are characterized by the fact that they give rise to a form of judgmental uncertainty, so that classifying some of these qualia is more difficult than classifying others. Hysteresis effects may therefore reflect only a judgmental rather than perceptual decision bias in those cases. If so, this may contradict the second hypothesis we put forward above, which is that in the case of standard sorites series, hysteresis would not simply be the persistence of the recent history of a category for uncertain cases across some unknown boundary, but rather the expression of a competition or rivalry between overlapping categories. In the rest of this paper, however, I shall attempt to make a case for the latter hypothesis, namely for the view that for soritical series in general, hysteresis effects arise from more than simply ignorance of a sharp boundary. To do this, I turn to a closer examination of series of ambiguous stimuli of the kind originally designed by G. Fisher.

3 Fisher series

The stimuli to be discussed in this section were originally designed by G. Fisher in order to measure perceptual ambiguity. Fisher wanted to question the assumption that in a two-way ambiguous figure, such as the Necker cube, or Jastrow’s duck-rabbit, the two percepts are always equally likely to be perceived. In order to challenge this view, Fisher designed several sets of ambiguous figures in which individual features are altered so as to gradually favor one percept over the other. He then asked individual subjects to report
which percept they first saw when presented with a given figure.

3.1 Fisher’s “Gypsy and Girl”

For instance, Fisher’s “Gypsy and Girl” set of cards comprises 15 cards, each of which represents an alteration of the same ambiguous figure, which can be seen either as representing a Man’s profile (percept A) or as representing a Girl holding a mirror (percept B) (see Fisher 1967, Figure 1, here reproduced in the Appendix). The cards are designed in such a way that card 1 strongly favors the perception of a man’s face, while card 15 strongly favors the perception of a woman’s figure. As in a sorites series, adjacent cards in the series are hard to discriminate when considered pairwise, and the difference between adjacent pairs is made to be “approximately the same” (see Fisher 1967: 542); but each card \( n + 1 \) is designed to make percept B slightly more salient than percept A relative to card \( n \).

Indeed, what Fisher found is that the central figures in his series (Figure 7 in the case of the “Gypsy-Girl” set) are those for which the subjects’s responses come closest to the mean value of the corresponding binomial distribution for two equally probable events. For such central figures, not only is the split between subjects statistically the greatest in comparison to other figures, but it therefore comes closest to the theoretical maximum of ambiguity.\(^4\) For the end figures in the series, on the other hand, subjects’ judgments converge in a proportion of more than 80%, and the corresponding percept is reversed between the first and the last figure. Indeed, as expected, the proportion of answers reporting percept A (the Man’s face) as first percept decreases monotonically from its maximum in Figures 1-2 until it reaches a minimum in Figure 15 (see Fisher’s Table I, reproduced in the Appendix). As a consequence, each card in Fisher’s set of cards appears to determine a different objective probability for any of the two percepts to be seen first, as reflected by the changing distribution of answers along the series.

\(^4\)See Fisher 1967, p. 545, who writes: “It may be considered justifiable to accept a figure as being ambiguous, in the sense that the appearance of each of its two alternative aspects is equally probable, if the number of responses indicating one of them to become apparent falls within the range plus or minus two standard deviations about the mean of the sampling distribution.”
3.2 Connection with the sorites

An aspect we must emphasize is that Fisher presented his stimuli in random order, as is customary in categorization and discrimination experiments, precisely to minimize order effects. A second aspect that deserves emphasis is that Fisher was not concerned with the investigation of vagueness or with sorites phenomena. Fisher’s focus, once again, was the phenomenon of perceptual ambiguity. Nevertheless, Fisher’s study is undeniably of interest when thinking about vagueness and sorites phenomena.

The first respect in which it turns out to be relevant is the idea that perceptual ambiguity can be graded, in such a way that a figure is more or less likely to be perceived as ambiguous. In the case of lexical ambiguity, for instance, we usually consider that an ambiguous word has the same potential of conveying two distinct meanings in principle. Of course, it remains possible that the word “bank” has a most salient meaning, namely that out of the blue, it will tend to convey one of its meanings first (for instance the meaning “money bank” might come to mind more readily than “river bank”). It is reasonable to think that the relative salience of each concept relative to an ambiguous word is furthermore different from one ambiguous word to the other, as a function of uses and contexts of use. Still, a difference between ambiguous words and ambiguous figures is that it is hard to imagine how one would modify the acoustic or written form of the word “bank” so as to modulate the perceived ambiguity between its two conventional meanings. Pinkal (1995: 76), for instance, opposes lexical vagueness and lexical ambiguity on considering that a vague word usually has a continuous range of potentially distinct extensions, while an ambiguous word only has a finite and discrete range of possible meanings. Fisher’s experimental design, on the other hand, suggests that a tighter connection between vagueness and ambiguity can be conceived when considering perceptual, rather than lexical ambiguity, precisely because perceptual ambiguity seems compatible with a fine-grained and even continuous modulation of the stimulus itself, in a way that simply does not happen with words.

The second main element of connection one can see between Fisher’s framework and the vagueness phenomenon concerns the analogy one may establish between Fisher’s stimuli, when considered in ascending or descending order, and an ordinary soritical series. At this point the ingredients of the analogy should be obvious, but they are worth repeating. As in a sorites series, a Fisher series makes a smooth transition of a given figure into a dif-
ferent figure, by means of small alterations. As in a sorites series, adjacent pairs in the series are hard to discriminate. More crucially, these changes are such that when focussing one’s attention only on pairs of adjacent figures, the same percept seems to come to mind, whichever it is. Finally, the first figure and the last figure in the series appear sufficiently distinct when considered pairwise, in such a way that the first figure and the last figure elicit quite distinct percepts.

3.3 Limits of the analogy

One should carefully qualify the scope of the analogy we are drawing. One element of disanalogy is that in a paradigmatic soritical series, such as a series of homogeneous color shades gradually altered, for instance, shades are not ambiguous when seen in isolation. At any rate, they do not seem to be ambiguous in exactly the way Fisher’s figures are. Individually and in normal lightning conditions, a homogeneous color shade gives rise to only one percept, whatever its quality. If ever there is ambiguity in a series of such color shades gradually altered, this ambiguity seems to be contextual rather than internal to the stimulus, and to be constrained primarily by the similarity relation of a given shade to the surrounding shades.

This does not necessarily mean that the color’s ambiguity is not perceptual in this case, and that it should be purely “judgmental”. For instance, the same red-orange shade can sometimes appear more red than orange when seen next to only redder shades, and more orange than red when seen next to shades that are all more orange.\(^5\) By talking of contextual ambiguity, however, what I have in mind is that a color shade seen in isolation does not appear to pattern as a bistable stimulus. Even if the color would be described as “orange-reddish” when seen in isolation, in particular, there is reason to doubt that what that means is that the same color would be perceived alternatively as red and then as orange while the subject keeps her

---

\(^5\)In many cases, context will enhance contrast rather than similarity between colors. The case of assimilation I have in mind is one in which a set of 13 shades gradually going from clear red (1) to clear orange (13) is presented in two sub-series ending in the same intermediate shade (7). In one occasion, only shades 1-7 are simultaneously presented from left to right; in the next, shades 13-7 are simultaneously presented in place of 1-7. While both series end in the same rightmost shade #7, the quality of that shade appears to change between the two environments: in my experience the shade tends to appear rather red in the first series, and rather orange in the second.
attention fixed at the stimulus. Nevertheless, depending on the context, the same orange-red shade can be perceived as slightly more red than orange, or as slightly more orange than red (we return to this point below). But prima facie an orange-reddish color is more accurately characterized as a fusion of red and orange, or as intermediate between a clearly red shade and a clearly orange shade, rather than as something that flickers alternatively between a clear orange and a clear red.

A second element of disanalogy between Fisher’s series and a standard sorites series is that, as we said earlier, the ends of a sorites series are supposed to clearly and unambiguously instantiate categories that are distinct and exclusive. In the case of Fisher’s series, on the other hand, each card in the series is such that both percepts remain available in principle, even for the end cards. Thus, in Fisher’s experiment, 7 subjects out of 200 still report a Man’s face as first percept seen on card 15, and 29 out 200 first see a Woman on card 1. The same phenomenon would most likely be evidenced if Fisher had made repeated trials within subjects. For instance, individual inspection of Fisher’s stimuli is enough to see that both percepts remain available in principle. When I look at Fisher’s Figure 1, which more readily presents a Man’s face to me, I can nevertheless strain my attention to see an imperfectly drawn Girl holding a mirror. When looking at Figure 15, I can likewise still see the contours of an imperfectly drawn face. Some of the cues, in particular, remained constant from Figure 1 to Figure 15 (such as the shadow under the Girl’s arm on Figure 15, or the left contour of the Man’s nose in Figure 1). The invariance of these cues, combined with the memory of the most salient percept, appears to be sufficient to make the percept available in principle.

3.4 Discussion

The two elements of disanalogy we emphasized between Fisher’s series and an ordered series of color shades do not necessarily compromise the project of building a tighter analogy between Fisher’s series and sorites series. Both elements of disanalogy concern the notion of ambiguity. Thus we saw (i) that in a Fisher series, individual stimuli can be perceived ambiguously to various degrees, and (ii) that the end stimuli in the series too have this property.

Regarding (ii), however, we may easily imagine to extend Fisher’s set of stimuli on both ends in a way that would make the probability of perceiving a Man’s face higher than what it is for the first card and even close to 1, and similarly for the probability of perceiving a Girl’s figure on the last card.
In other words, in principle we can imagine to extend Fisher’s series so as to reach stimuli that are unambiguously perceived, or very unlikely to be perceived ambiguously. Seen in this light, feature (ii) of Fisher’s series may therefore strengthen rather than weaken the analogy we are seeking. Thus, if the penumbral area of a sorites series can be characterized as the area in which a subject is likely to categorize one and the same stimulus in two opposite ways, though possibly to different degrees, then what we are saying is that an analogy can be established between a series of ambiguous figures like Fisher’s and the penumbral area of a more extended sorites series.

Objection (i), on the other hand, is directly relevant to assess the validity of Raffman’s suggestion that within soritical series, borderline cases might pattern as bistable stimuli. Let us consider Fisher series again. There, each figure in the series can be seen either as a Man’s face or as a Girl holding a mirror, alternatively. Raffman’s 1994 claim is that in the case of a series of color shades, the shades that are borderline can be seen as red or as orange in a similar fashion. The content of objection (i) is that when looking at a particular shade in isolation, we do not have the impression that two percepts overlap. There is a single percept, even though this percept can be referred to two distinct categories, “red” or “orange”.

I think this objection is correct, but I surmise that it does not undercut Raffman’s point, it only forces us to make it more precise. For one thing it could be that although the percept is relatively stable upon a single occasion, the underlying stimulus might elicit distinct percepts and therefore distinct judgments upon sufficiently distant occasions. Besides, we acknowledge that a distinction must be made between a stimulus being ambiguous when seen in isolation, and a stimulus being ambiguous due to contextual effects. Once that distinction is made, however, it remains that contextual ambiguity can be properly perceptual. Again the same red-orange color shade, seen next to only redder ones, will look congruent and sufficiently so to be seen as red. When seen next to only more orange shades, it will look congruent too and sufficiently so to be seen as orange. The hypothesis I am making is that this “seeing as” is indeed a matter of perception, and not merely of the

---

6 Raffman subsequently rejected this view of borderline cases, and now considers borderline cases as instantiating neither of the other categories, rather than both (Raffman, p.c.).

7 As suggested to me by J-L. Schwartz, one would need to check, then, whether the categorization dynamics would or not pattern like the perceptual dynamics of bistability evidenced for more rapidly alternating stimuli (see Pressnitzer and Hupé 2006).
decision to categorize one way rather than the other. Hence, although the rivalry between categories may not be directly present within the stimulus in this case (because “orange-reddish” here designates a shade of its own), the assumption I am making is that in a series of shades that go from clear red to clear orange, each shade has a different potential of being classified as red or as orange because ultimately it has a different potential of being seen as congruent to either of the two end colors.

Incidentally, one may wish to reflect more thoroughly about the very status of semantic categories such as “orange-reddish”, “blue-greenish”, and others of this kind. If indeed we can make sense of such categories, this also suggests that some such colors bear cues that are simultaneously congruent with more “pure” instances of categories like “red” or “orange”. The consideration of Fisher’s stimuli is interesting in this particular respect. When looking at Figure 1 in Fisher’s series, I almost inevitably see a Man’s profile. When looking at Fisher’s Figure 15, I likewise see first and foremost a Girl holding an object. When looking at Figure 7, I feel that I am switching between either of these two interpretations. Now, if I look at 1 and then at 7, I will indeed tend to see clearly a Man’s profile in 7. If I go the other way, I will tend to see clearly a Girl’s face in 7. Still, I am aware that Figure 7 is different from both Figure 1 and Figure 15. While entirely self-administered, this small experiment seems to exemplify both the kind of hysteresis phenomenon we discussed in section 1, but also the potential for Gestalt switch that Raffman talks about.

Now, my own impression is that the same is true to a significant extent of color shades. If I happen to consider a fine-grained series of 13 color shades that go from a specific quality of red to a specific quality of orange by similarly small shifts, which seem to make no difference for pairwise adjacent pairs: when I look at shade 1 first, and then at shade 7, I perceive a form of “color echo” in 7, and then a different color echo when I then look at shade 13 and then at shade 7. In that particular case, shade 7 may be described as “orange-reddish”, and indeed, despite the fact that it is homogeneous, it bears a resemblance to the other two shades. Of course, when I look at 7, I can discriminate it from either of shade 1 or 13. But in fact the same is true of Figure 7 in Fisher’s series, which I can discriminate from Figure 1 and 15 respectively, and which is such that I could learn to memorize it distinctively.

The upshot of these considerations, therefore, is that neither of the dis-analogies we pointed out seems to run against the project of using Fisher’s series as a template for the analysis of soritical series more generally. Of
course, there are as many sorites series as there are varieties of stimuli to
start from, and as we saw, colors and structured figures are different kinds
of objects with different psychophysical properties. The purpose of the next
section, however, will be to draw general consequences about the nature of
sorites series, based on the idea that the analogy with Fisher series is suffi-
ciently safe.

4 Cut-offs and tolerance in soritical series

The aim of this section is to cast light on the nature of sorites series, reflecting
on features that we find in play in Fisher’s particular series. On the one
hand, we propose to raise and discuss an objection to what appears as a
consequence of the epistemic solution to the sorites paradox. On the other,
we shall argue that the paradox-inducing principle at the heart of the sorites
cannot be literally correct. Both objections, ultimately, are grounded in the
consideration of the phenomenology of Fisher’s series.

4.1 Epistemicism

As is well-known, the sorites paradox results from the assumption of a par-
ticular premise, commonly referred to as the tolerance principle (see Wright
1976), which states that if an arbitrary individual \( n \) in the series instanti-
ates property \( A \), and if \( n+1 \) is indiscriminable from \( n \), then \( n \) ought to
instantiate property \( A \) as well. I here put emphasis on “ought to” because
the normative aspect conveyed by this expression is usually not reflected
in the way the premise is formalized, no more than the intended meaning
to be given to “ought” (should it mean “must”, “is likely”? ). Indeed, usu-
ally this principle is formalized as a standard induction principle, of the
form \( \forall n (A(n) \rightarrow A(n+1)) \). A more abstract version of the principle is
\( \forall xy (A(x) \land R(x, y) \rightarrow A(y)) \), where \( R \) is a relational predicate intended to
express a specific similarity relation (such as indiscriminability), but here too
the normative component is left implicit.

Let us temporarily set aside the normative intent of the tolerance principle
and focus on its standard form, namely \( \forall n (A(n) \rightarrow A(n+1)) \). As is well-
known, a contradiction results from the tolerance principle if it is assumed
that \( A(m) \) and \( \neg A(n) \) hold for some \( n \) and \( m \) with \( n \) greater than \( m \). One
way to envisage the strength of the paradox is to present it in the form of
a dilemma. One horn of the dilemma is that the tolerance principle seems entirely plausible: two objects that differ imperceptibly are such that if one is categorized one way, the other should be too. Yet we see that accepting the principle leads to contradiction. The other horn of the dilemma is that if we reject the tolerance principle, and keep our reasoning classical, then we should endorse its negation, namely $\exists n (A(n) \land \neg A(n + 1))$. In that case, we must endorse the view that there is a sharp cut-off in the series, a demarcation between the last object $n$ to which $A$ applies and the first object $m$ to which $A$ does not. But the problem is that, in general, we do not see such an $n$ and find it very hard to imagine that there is such a demarcation.

Faced with the dilemma, epistemicists like Sorensen and Williamson go for the second horn (see Williamson 1994, Sorensen 2001). On the epistemic view of sorites series, there is such a sharp cut-off point, but we are simply ignorant as to its precise location. The epistemic solution suggests that if classical logic is to be preserved, then it is more sensible to revise and explain away our intuition that the tolerance principle is plausible and true, rather than to end in contradiction. Of course, several other ways out of the paradox are conceivable if one is ready to forsake classical logic. For now, however, let us restrict our attention to the epistemicist view and what it commits us to.

4.2 Variations in the boundary

The epistemic conception of vagueness has been criticized on the ground that it severs the ontology of vagueness and the psychology of vagueness in a manner that appears too drastic (see Wright 1994, Schiffer 1999, Shapiro 2006). I sympathize with this line of criticism, but my opinion is that to make it compelling, it is necessary to articulate the tolerance principle and its negation in a way that makes it normative dimension explicit. So let us consider a sorites series of colors ranging from a clear red to a clear orange, where $A$ is a unary predicate for “red” (we use $\text{Red}$ where it is clearer). Let us assume it holds that $\exists n (A(n) \land \neg A(n + 1))$. If so, this conclusion seems to entail that as a matter of fact, there is a shade $n$ such that $n$ ought to be judged red, and such that the consecutive shade $n + 1$ ought to be judged not red. A weaker normative requirement might be that there is a shade $n$ such that $n$ ought to be judged red while shade $n + 1$ ought not to be judged red. In what follows I shall focus on the former and stronger requirement, however, mostly for the sake of simplification. For the main
part, however, our argument could directly be adapted to a discussion of the weaker requirement by appropriate adjustments (see fn. 12).

The sense of *ought* I have in mind here means that subjects *rationally* ought to issue particular judgments. The formulation I used remains syntactically ambiguous, however. To sort it out, let us introduce an epistemic operator $\Box_s$ to mean “$s$ judges that”, and a deontic operator $\mathcal{O}$ to mean “it ought to be the case that”. One possible normative understanding of the negation of the tolerance principle is the following:

\[
(1) \quad \forall s \mathcal{O}\exists n \Box_s (A(n) \land \neg A(n + 1))
\]

This principle says that for every subject, there ought to be a shade $n$ such that $s$ judges that $A(n)$ and $\neg A(n + 1)$. This principle is not implausible, for if we think of situations of forced march where subjects have to judge shades consecutively, the principle can be taken to mean that every subject that would judge $A(0)$ and $\neg A(k)$ logically ought to switch category at some point between 0 and $k$. This principle is weaker, in particular, than the following:\footnote{See Sweeney 2008, who draws a similar distinction between a *contextual* and an *acontextual* version of the affirmation of sharp boundaries in soritical series, depending on the scope of quantifiers. On the contextual version, for every judgment context there is a cut-off, whose location can vary with the context; on the acontextual version, there is a cut-off, the same for every context.}

\[
(2) \quad \forall s \exists n \mathcal{O}\Box_s (A(n) \land \neg A(n + 1))
\]

This says that for every subject, there is a shade $n$ such that $s$ ought to judge that it is red and its successor is not. Yet a stronger principle, finally, is the following:

\[
(3) \quad \exists n \forall s \mathcal{O}\Box_s (A(n) \land \neg A(n + 1))
\]

This says that there a shade $n$ such that every subject ought to judge that it is red and its successor isn’t.

When we judge the conclusion reached by the epistemicist exorbitant, I think that what we intuitively feel reluctant to accepting is primarily a principle such as (3). For (3) entails that subjects ought to draw boundaries between the categories of “red” and “non-red” at exactly the same point. But this does not sound right. Why cannot this be right?

One argument I can offer here rests on the consideration of Fisher’s series.
One characteristic of Fisher’s series is that every figure is such that it can be perceived as a Man’s face or as a Girl’s face. Even if the likelihood or objective probability of that happening varies from one stimulus to the next, we saw that even the end figures in the series make room for this variation. Indeed, ambiguous figures are such that they can be conceptualized in opposite and mutually exclusive ways. As a result, all cards in Fisher’s series are legitimate candidates for category shift. But at the same time, no card mandates a particular shift.

This fact suggests that even within a given subject, a weaker principle such as (2) remains too strong in the case of Fisher’s series. Indeed, if we consider the 15 figures of Fisher’s series, because all figures are such that they sustain either of the two percepts to some degree, no single figure seems to be such that a given subject ought to judge that it is A rather than not A, but also no figure is such that one ought not to judge it A or to judge it not A.

Interestingly, the particular configuration of Fisher’s stimuli backs up a more general characterization of borderline cases that Wright puts forward in his defense of the permissibility of opposite verdicts for borderline cases (see Wright 1994: 139, Wright forthcoming). For Wright, borderline cases are cases for which competent speakers can judge in opposite ways. One reason to think so, on Wright’s proposal, rests on the view that borderline cases are cases for which it has not yet been determined whether they are A or not A. I think Wright’s characterization of borderline cases is exactly right. On the present account, however, the reason verdicts can go either way for borderline cases is grounded in the consideration that the stimulus itself sustains both interpretations even within a single subject, and already at the perceptual level. This, in fact, is a far-reaching property of ambiguous figures and bistable stimuli: by definition, the categorization of such stimuli is response-dependent. There is no fact intrinsic to the stimulus that dictates that one interpretation should be absolutely preferred to the other.

On the epistemic conception of vagueness, on the other hand, if our discriminative capacities were more fine-grained than they actually are, then we would be able to locate an objective boundary between A and not A within a soritical series. In the specific case of a series of stimuli such as Fisher’s,

---

9 On a related view, see Schiffer 2009, who writes: “borderline applications aren’t merely ones that fall between two stools; they are ones that bounce around between them because of the resemblance-based attraction each stool exerts”.
such a view cannot be correct. For we should expect that even a subject with perfect discriminative capacities is likely to experience a phenomenon of Gestalt switch when confronted with any of the figures in the series. If indeed Fisher’s series can be seen as revealing of the structure of sorites series more generally, this would suggest that the vagueness of the boundary in a soritical series is not wholly reducible to a phenomenon of imperfect discrimination, but it is a phenomenon of ambivalence and shiftiness in categorization proper.

4.3 Discussion

As we pointed out earlier, however, Fisher’s series may be highly specific in comparison to other soritical series. As a result several objections can be raised against our point. To articulate them, it may be worth repeating our argument in the following form:

1. In a Fisher series, category-shifting is permissible at every point in the series, and no figure mandates a particular shift.
2. One can establish a structural analogy between a Fisher series and the penumbral area of an arbitrary soritical series.
3. Therefore, in an arbitrary soritical series, category-shifting is permissible at every point in the penumbral area of the series, and no figure there mandates a particular shift.

In response to this, the epistemicist may object to the soundness of premise 2. Or he may accept premise 2, but argue from the denial of 3 to the denial of 1, making use of the very same analogy. He may also grant the conclusion 3, finally, but argue that it leaves epistemicism unscathed.

In what follows I shall focus only on the first objection. The second objection is basically the idea that our modus ponens can be reverted into the epistemicist’s modus tollens.\(^{10}\) This is fair enough, but as we will see, the objection to premise 2 already has the seeds for simultaneously weakening premise 1 and conclusion 3 of our argument. Concerning the third objection, namely the idea that epistemicist can accept the argument but relativize its impact, one way to articulate it is to say that even if boundaries between

\(^{10}\)R. Cook and M. Werning both formulated this objection during a presentation of this paper at Kirchberg.
perceptual categories are indeed variable within and across subjects, boundaries between linguistic categories are determined by a complex process of aggregation of individual judgments, for which stronger normative principles such as (2) or (3) remain valid anyway. A discussion of this point would take us too far afield, however, and here too we will see that already from our discussion of premise 2 a point in favor of some mild form of epistemicism is conceivable.

So let us consider premise 2. First one may object that the analogy we are suggesting is vacuous. Indeed, if really Fisher’s stimuli are ambiguous all along, then one can argue that for such stimuli, the induction premise $\forall n (\text{Man}(n) \rightarrow \text{Man}(n + 1))$ simply holds without exceptions, and likewise the induction premise $\forall n (\text{Girl}(n) \rightarrow \text{Girl}(n + 1))$. More generally, the series supports $\forall n (\text{Man}(n) \land \text{Girl}(n))$, namely every figure represents both a Man and a Girl (irrespective of the degree to which they do so). Granted, because of the exclusiveness of the percepts, the series also supports $\forall n (\square s \text{Man}(n) \rightarrow \neg \square s \text{Girl}(n))$, and conversely, namely if I categorize shade $n$ as representing a Man, then I do not categorize it as representing a Girl. But this shows nothing, for in the case of a standard soritical series in which one starts from a non-ambiguous figure, for instance a clear red shade, until one reaches a clear non-red shade, then it is simply false that $\forall n (\text{Red}(n) \rightarrow \text{Red}(n + 1))$, and it is also false that $\forall n (\text{Red}(n) \land \neg \text{Red}(n))$.

Another way to put the objection is that if indeed we could extend Fisher’s series to the left and to the right, so that it starts out with a figure that has objective probability 1 of being perceived as a Girl’s face, and finishes with a figure that has objective probability 1 of being perceived as a Man’s face (and probability 0 of being perceived as a Girl), then the series would no longer support $\forall n (\text{Man}(n) \land \text{Girl}(n))$, and it would not support, in particular $\forall n (\text{Girl}(n) \rightarrow \text{Girl}(n + 1))$. At any rate there ought to be a last non-ambiguous shade $k$ such that $\text{Girl}(k) \land \neg \text{Man}(k)$, and a first ambiguous shade $k + 1$ such that $\text{Girl}(k + 1) \land \text{Man}(k + 1)$.

The force of this objection is undeniable. What the objection urges us to acknowledge is that to the extent that one can establish a structural analogy between a Fisher series and the penumbral area of a soritical series, what this analogy shows is only, at best, that the penumbral area of a standard soritical series can be seen as one where contrary categories overlap and pattern ambiguously. But the analogy leaves open the possibility that there remains a sharp and objective cut-off between the last non-ambiguous stimulus and the first ambiguous one. For instance, there may be a range of shades in
a color series from clear red to clear orange that are ambiguously red and orange. But on this account there must be a last non-ambiguously red shade.

Let us assume that there is such an objective and yet imperceptible demarcation. I think even so, the analogy we are drawing casts light on the nature of soritical series. Indeed, even if there is a sharp cut-off between the last non-ambiguous stimulus and the first ambiguous one, and we do not know where it is, it remains that the first ambiguous stimulus is one for which it will remain permissible to judge that it is $A$ and permissible to judge that it is not $A$.

To justify this view, let us picture a soritical series of 8 color shades ranging from a clear red to a clear orange. Drawing inspiration from Fisher’s series, we may suppose that each shade in the series comes with a different potential of being perceived as red or as orange. This potential should be seen as the extent to which the stimulus constrains perception and categorization (as reflected for instance in Fisher’s survey, see Appendix, table I). Mathematically, potentials can be described by means of prior probabilities $p(A(x))$ and $p(O(x))$ of perceiving shade $x$ as red or as orange. In keeping with the idea of bistability, we make the assumption that each percept inhibits the other, so that $p(O(x)) = p(\neg A(x)) = 1 - p(A(x))$. On the following figure, we are assuming that the first two shades have the same potential and can only be seen as red, while the last two have the same potential and can only be seen as orange. Each intermediate shade $n$ is assumed to be such that its potential $\alpha_n$ of being seen as red lies strictly between 0 and 1, and is less than for the previous shade.\footnote{See Hampton 2007, MacFarlane 2008 and Lassiter 2009 for further discussions of probabilistic aspects of vagueness and categorization. See Lassiter 2009, in particular, for the idea that all vague predicates should be treated probabilistically.}

\begin{center}
\begin{tabular}{c|cccccccc}
$x$ & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\
p($A(x)$) & 1 & 1 & $\alpha_3$ & $\alpha_4$ & $\alpha_5$ & $\alpha_6$ & 0 & 0 \\
p($O(x)$) & 0 & 0 & $1 - \alpha_3$ & $1 - \alpha_4$ & $1 - \alpha_5$ & $1 - \alpha_6$ & 1 & 1 \\
\end{tabular}
\end{center}

While shades come with different potentials of being categorized as red or orange, we are making the assumption that $R$ and $O$ are true of a shade whenever this potential is non-zero, and false whenever this potential is 0. Under this assumption, the series can be pictured by means of the following classical model, on which the $O$ and $R$ categories overlap, except for end
shades. The model now satisfies \( \exists n (A(n) \land \neg A(n + 1)) \), since by assumption there is a first non-ambiguously orange shade:

<table>
<thead>
<tr>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>A</th>
<th>\neg A</th>
<th>\neg A</th>
</tr>
</thead>
<tbody>
<tr>
<td>\neg O</td>
<td>\neg O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
<td>O</td>
</tr>
</tbody>
</table>

Our main assumption, finally, is that a subject *ought to judge* \( A(n) \) if and only if \( p(A(n)) = 1 \), and ought to judge \( \neg A(n) \) if and only if \( p(A(n)) = 0 \). In other words, only shades that are non-ambiguous in their potential dictate a particular judgment.\(^{12}\) Under those assumptions, an ideal and rational subject ought to start out judging the first two shades are red. But when reaching the third shade, the subject can go either way. The subject can judge the shade red or judge it orange, and therefore not red. Now assume several trials are performed. In principle an ideal subject whose judgments would obey exactly the shades’ potentials should remain confident about the first two shades, but from the third shade onward, the subject is likely to switch her judgment, in accordance with the various probabilities that determine the saliency of \( O \) and \( A \).

Obviously, the assignment of relative probabilities to \( A \) and \( O \) is reminiscent of degree-theoretic treatments of vagueness. But on our account, as soon as the potential of a predicate is non-zero, the predicate applies and is true *simpliciter*, and it is false *simpliciter* exactly when the potential is zero.\(^{13}\) What is interesting regarding the distinctions we made earlier about normative judgments, however, is that such a structure would support both the following normative judgments:

\[
\begin{align*}
(4) & \quad \exists n \diamond s, A(n) \\
(5) & \quad \exists n \diamond s, \neg A(n)
\end{align*}
\]

That is, there are shades that \( s \) ought to judge red, and other shades that he ought to judge not red, namely those for which the corresponding potential is 1 or 0 respectively. However, the crucial fact is that \( \neg \exists n \diamond s, (A(n) \land \neg A(n + 1)) \), that is, no shade is such it ought to be judged red, while its successor ought to be judged not red, because no two consecutive shades in the series

\(^{12}\) Alternatively, for the case of wide scope negation, we could stipulate that a subject ought not to judge \( A(n) \) exactly if \( p(A(n)) = 0 \).

\(^{13}\) I do not see potentials as degrees of truth, but rather as the degree to which a property is expressed or salient in some stimulus. Nevertheless, the present discussion may easily be transposed to the degree-theoretic framework.
are such that \( p(A(x)) \) goes directly from 1 to 0.\(^{14}\) As a consequence, even when a series such as Fisher’s is extended on both ends in a way that restores \( \exists n(A(n) \land \neg A(n + 1)) \), normative principles like (2) and (3) need not follow. This, of course, should be seen as good news for the epistemicist. For what we have established is that from \( \exists n(A(n) \land \neg A(n + 1)) \) it does not follow analytically that \( \exists n\Box_s(A(n) \land \neg A(n + 1)) \) (contra principle (2) above).\(^{15}\)

By way of consequence, this discussion also gives us a hint as to what an explicitly normative formulation of the tolerance principle might be. Here my conclusion will be exactly consonant with that of Raffman (forthcoming), who distinguishes the false principle whereby small increments in a sorites series are ever insufficient to make a difference in the way we categorize consecutive individuals in the series, and the correct principle whereby small increments in a sorites series are “sufficiently small as to make any differential application of \( \Phi \) as between them (either incorrect) or arbitrary.”

As we suggested earlier, the tolerance principle has both a descriptive and a normative dimension. The descriptive dimension, on our account, corresponds to the idea that if the probability for a given stimulus \( n \) to be seen as \( A \) is \( \alpha \), then the probability for a sufficiently similar stimulus \( n + 1 \) to be seen as \( A \) should be sufficiently close to \( \alpha \).\(^{16}\) On this view, if two objects are very similar, then they are very likely to be categorized alike, although the need not be. The normative dimension of the principle, on the other hand, can be articulated from the negation of principle (2), which we take to be incorrect. Principle (2) is equivalent to \( \exists n(\Box_s A(n) \land \Box_s \neg A(n + 1)) \). Its negation is equivalent to:

\[
(6) \quad \forall n(\Box_s A(n) \rightarrow \neg \Box_s \neg A(n + 1))
\]

This says that if an individual ought to be judged \( A \), then it is not the case

\(^{14}\)See Lassiter 2009, who makes exactly this point about probabilistic versions of the sorites more generally. Lassiter’s observation is that a natural probabilistic understanding of the tolerance principle, for a vague predicate such as tall, is that there is usually no \( n \) such that the probability of judging tall\((x_n)\) is 0 while that of judging tall\((x_n)\) is 1.

\(^{15}\)Of course, one may insist that where the potential switches from 0 to more than 0, one ought to make that distinction (a form of the higher-order vagueness problem). But on my view the corresponding sense for ought would be much too fine-grained.

\(^{16}\)That is: if \( d(x, y) \leq \varepsilon \), then \( |p(A(x)) - p(A(y))| \leq \delta \) (for some specified \( \delta \in [0, 1] \)), namely: whenever two stimuli \( x \) and \( y \) are sufficiently similar given a suitable metric \( d \), the probability of judging the second \( A \) must be close enough to the probability of judging the first \( A \). Lassiter’s constraint can be seen as a particular case of this constraint.
that an individual that differs only very slightly ought to be judged not \( A \).\footnote{This version of the tolerance principle should be compared to the version proposed by Shapiro in terms of \textit{competent judgment} (I am indebted to L. Horsten for this remark). For Shapiro, a tolerant predicate \( A \) is such that if one competently judges \( A(n) \), then one cannot competently judge \( \neg A(n + 1) \) whenever \( n \) and \( n + 1 \) “differ only marginally in the relevant respect” (see Shapiro 2006:8). If we equate \textit{competent judging} with \textit{ought to judge}, then (6) stands as a possible paraphrase for Shapiro’s version of tolerance. This is no longer the case, however, if we define a \textit{competent judge} as someone who never judges \( \neg A(n) \) when \( p(A(n)) = 1 \), or \( A(n) \) when \( p(A(n)) = 0 \). When \( p(A(n)) \) and \( p(A(n + 1)) \) lie strictly between 0 and 1, then given our semantics one may competently judge both \( A(n) \) and \( \neg A(n + 1) \). Prima facie, this makes room for more tolerance than Shapiro’s principle. But this is also less tolerant, since when \( p(A(n)) = 0.99 \) and \( p(A(n + 1)) = 1 \), one can competently judge \( \neg A(n) \) by those standards, but one cannot competently judge \( \neg A(n + 1) \).} Under suitable assumptions (see fn. 12), this principle can be strengthened to:

\[
\forall n (O\Box_A n(n) \rightarrow \neg O\neg \Box_A n(n + 1))
\]

which says that if an individual \( n \) ought to be judged \( A \), then it is \textit{permissible} to judge an individual \( n + 1 \) that differs only very slightly as \( A \) as well. Seen in Kantian terms, the resulting idea of tolerance is now expressed as a principle about \textit{judgment}, and no longer directly as a standard induction principle constraining the category membership of \textit{things in themselves}.\footnote{See Shapiro 2006, chap. 1 for more on the centrality of the notion of judgment in the treatment of vagueness, following insights from Wright and Raffman in particular.} Thus, while we agree that the standard soritical premise cannot hold unrestrictedly in classical logic on pain of contradiction, the present discussion suggests that a safer view of the intent of this premise should be placed at the psychological rather than the strictly logical level.

5 Conclusions

Several claims have been made in this paper. The first is the idea that by re-considering the relation between vagueness and ambiguity at the perceptual level, further insights can be gained for our understanding of soritical series. On our account, the penumbral area of a soritical series is primarily an area of competition, ambivalence and rivalry between overlapping categories. We are prone to judging that \( A \) and to judging that not \( A \) for borderline cases because those stimuli come with cues that support distinct representations,
and ultimately opposite judgments. The hypothesis that borderline cases correspond to an area of overlap, where cues toward $A$-judgments coexist in different proportion with cues toward $\neg A$-judgments, needs further elaboration, but it presents several explanatory virtues. First it suggests that hysteresis effects in front of soritical series do not simply reflect pure judgment biases, but originate at the perceptual level, from the potential of a stimulus to elicit contrasting representations. Secondly, it supports a probabilistic theory of judgment, where those probabilities would express these various potentials of being perceived and categorized one way or the other. Finally, it suggests that borderline cases, as argued in particular by Wright, Schiffer, Raffman and Shapiro, are indeed adequately viewed as cases for which it is permissible to draw boundaries and categorize in variable and opposite ways, thereby affording us a clearer view of the tolerance principle. In agreement with epistemicism, some ignorance remains in the picture, in particular about the structure and amplitude of these potentials. But much of the error theory that goes with epistemicism is left behind.

**Acknowledgments**

I am indebted to V. de Gardelle, R. Gregson, D. Pressnitzer, J-L. Schwartz, J-M. Hupe for references and for enlightening discussions on the psychology of bistability and hysteresis phenomena, which sparked this research. Special thanks go to D. Raffman, S. Bromberger, S. Schiffer, and D. Lassiter, each of whose recent work on vagueness has been a source of insights and inspiration, and whom I thank for stimulating conversations. For helpful discussion and comments, I particularly thank D. Raffman, P. Schlenker, M. Cozic and S. Bromberger, as well as D. Bonnay, J. Chandler, R. Cook, W. Davies, J. Dokic, L. Horsten, J. Sackur, B. Spector, M. Werning, and audiences at the Bistability Workshop held at the ENS in Paris in July 2008, at the 2008 Wittgenstein symposium in Kirchberg, at the Philosophy and Psychology of Vagueness Workshop held in Paris in November 2008, and at the Center for Logic and Analytic Philosophy in Leuven. This research was supported by the ANR project “Cognitive Origins of Vagueness” (ANR-07-JCJC-0070-01) and the ANR project “Multistap” (ANR-08-BLAN-0167-01), which are gratefully acknowledged.
References

[1] S. Bromberger (2008), Vagueness, Ambiguity and the “Sound” of Meaning, manuscript, MIT.


[23] P. Sweeney (2008), Contextual Intolerance, manuscript, University of St Andrews.


Appendix

A. Fisher’s “Gypsy and Girl” set of stimuli (from Fisher 1967: 542, smaller reproduction)

![Figure 1: The 'Gypsy and Girl' Set of Ambiguous Figures](image)

B. Fisher’s results

Fisher’s subjects were shown figures separately on a screen and asked to indicate the first aspect they saw on a piece of paper, “by writing an appropriate descriptive word, or phrase” (the descriptions “Gipsy” and “Girl” are from Fisher, who notes that subjects distinguished two percepts but varied in their descriptions, see p. 544, fn. 14). The following data, reproduced from Fisher’s table I, report the number of responses (for 200 subjects) indicating the “Gypsy” as first aspect seen.

<table>
<thead>
<tr>
<th>Figure</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
<th>11</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>Responses</td>
<td>171</td>
<td>178</td>
<td>171</td>
<td>163</td>
<td>144</td>
<td>132</td>
<td>103</td>
<td>79</td>
<td>68</td>
<td>53</td>
<td>43</td>
<td>18</td>
<td>11</td>
<td>10</td>
<td>7</td>
</tr>
</tbody>
</table>