Vagueness, Uncertainty and Degrees of Clarity

Paul Égré   Denis Bonnay

Abstract

The focus of the paper is on the logic of clarity and the problem of higher-order vagueness. We first examine the consequences of the notion of intransitivity of indiscriminability for higher-order vagueness, and compare different theories of vagueness understood as inexact or imprecise knowledge, namely Williamson's margin for error semantics, Halpern's two-dimensional semantics, and the system we call centered semantics. We then propose a semantics of degrees of clarity, inspired from the signal detection theory model, and outline a view of higher-order vagueness in which the notions of subjective clarity and unclarity are handled asymmetrically at higher orders, namely such that the clarity of clarity is compatible with the unclarity of unclarity.

1 Intransitivity and introspection

One central and debated aspect of the notion of inexact knowledge concerns the non-transitivity of the relation of indiscriminability and how it should be represented. On the epistemic account of vagueness put forward by Williamson, the intransitivity of the relation of indiscriminability is presented as the main source for vagueness ([16]: 237). In [15] and in the Appendix to [16], Williamson formulates a *fixed margin for error* semantics for propositional modal logic in which the relation of epistemic uncertainty, based on a metric between worlds, is thus reflexive and symmetric, but non-transitive and non-euclidian.[1] An important consequence of Williamson's semantics is that it invalidates the principles of positive introspection (if I know p, then I know that I know p) as well as negative introspection (if I don't know p, then I know that I don't know p).

In an earlier paper [1], we argued against Williamson that models of inexact knowledge that preserve the introspection principles may sometimes be desirable, and we presented an alternative epistemic semantics for the notion of inexact

---

[1]A relation R is transitive if xRy and yRz imply xRz for every x, y and z. A relation is euclidian if xRy and xRz imply yRz for every x, y, z.

knowledge, in which non-transitive and non-euclidian Kripke models can nevertheless validate positive as well as negative introspection. In [5], Halpern also argued against Williamson that an adequate model of vague knowledge need not invalidate the introspection principles, but following a different route. Instead of taking intransitivity as a primitive, and proving that the introspection principles can be preserved for a logic with one epistemic operator, as we did in [1], Halpern proposes a bimodal account of inexact knowledge that preserves the introspection principles, and he shows that there is a way to derive intransitivity. For Halpern, the intransitivity of vague knowledge is more characteristic of our reports on what we perceive than about our actual perception.

Despite these differences, one can establish a precise correspondence between Halpern's semantics and the semantics presented in [1]. One aim of this paper is to spell out the details of this correspondence, and more generally, to compare three related accounts of the notion of inexact knowledge in epistemic logic. Williamson's original account was designed to establish a logic of clarity, namely a logic of the operator "it is clear that" and of its iteration properties, in particular in the face of the problem of higher-order vagueness. By contrast, the account we presented in [1] was primarily intended to state a logic of the operator "I know that" and of its iteration properties. Halpern's own account, to a large extent, holds a position intermediate between Williamson's and ours, and is compatible with both. In particular, Halpern's framework encapsulates both an "internalist" or subjective logic of knowledge that maintains the introspection principles we find desirable, and an "externalist" or objective logic of clarity for which such principles are not necessarily valid.

In the first half of this paper (sections 2 and 3), we take the opposite perspective, namely we examine to what extent the family of semantics introduced in [1] is susceptible to offer a compromise between Williamson's stance and Halpern's approach on the link between vagueness, introspection, and non-transitivity. Like Halpern, but contra Williamson, we think it does make sense to preserve the introspection principles within a logic of inexact knowledge; unlike Halpern, but in agreement with Williamson, we are ready to see non-transitivity as a property of perceptual knowledge proper. The divergence with Halpern's view is more conceptual than technical, however, since we will see that we understand "perceptual" in a broader sense than Halpern, in a way that encompasses what he calls "reports about perception". More fundamentally, however, our approach and Halpern's both rest on the idea that the non-transitivity of phenomenal indiscriminability can be made transitive by reference to a particular context, but in our approach this contextual parameter remains implicit, thereby affording a more intuitive characterization of the uncertainty associated to inexact knowledge.

A central issue in the paper concerns the problem of higher-order vagueness. Williamson's logic of clarity is essentially a logic of infinite higher-order vague-

ness. The first semantics for inexact knowledge that we introduce, on the other hand, makes room only for first-order vagueness when the operator is interpreted as "it is clear that". In section 2, we show how to obtain a non-trivial logic of $n-$order vagueness generalizing our logic of first-order vagueness. While theoretically interesting, these results remain fairly abstract, however. One reason for that is that throughout sections 2 and 3, the models under discussion are fixed margin for error models, in which vagueness is associated to a notion of *constant imprecision* or constant uncertainty. In particular, the frameworks do not make room for degrees of clarity and unclarity, and they only remotely relate to actual psychophysical tasks of discrimination.

In the second half of this paper (section 4), we therefore propose a comparison between margin models and the model of error and uncertainty underlying signal detection theory (SDT). In that section we introduce a semantics for degrees of clarity, based on what we call likelihood models, namely structures representing the probability that a stimulus be misrepresented. These structures can be seen as generalizing margin for error models, and as formalizing a notion of variable margin of error. A further specificity is that they link the notion of vagueness as epistemic directly to the notion of probabilistic error, rather than to the notion of intransitivity of indiscriminability. The main effort of that section is in the relation established between signal detection theory and the notion of degree of clarity. We outline, in particular, a treatment of higher-order vagueness in which the notions of subjective clarity and unclarity are handled asymmetrically at higher orders, namely such that clarity can be clear, and unclarity unclear.

2   Centered Semantics

2.1   Margin models

A good grip on the relation between vagueness and epistemic uncertainty is offered by the consideration of an idealized task of visual discrimination. We shall first repeat the description of a task we discussed in earlier work (see [1], [2]), as it will serve as a recurring theme for the different variations we wish to discuss. Consider a discrete series of pens linearly ordered by size, such that all and only pens that are less than 4 cm fit in a certain box. A subject sees the pens and the box at a certain distance and is asked which pens will fit in the box. We make the supposition that from where she is, the subject cannot perceptually discriminate between pairwise adjacent pens, namely between pens whose size differs by less than 1 cm. However, the subject is able to discriminate between non-adjacent

pens.[2] For instance, when looking at the pen of size 2, the subject cannot discriminate it from the pen of size 1, nor from the pen of size 3, but she can discriminate it from a pen of size 4. Furthermore, we make the idealized assumption that the inability of the subject to detect differences is constant throughout the series.

The scenario can be represented by means of the following linear Kripke model, in which p represents the objective property of fitting in the box, with worlds indexed by sizes. The important fact about the model, reflected in the accessibility relation between worlds, is that the model is reflexive and symmetric, but non-transitive (and non-euclidian). Letting R stand for the relation of perceptual indiscriminability, this represents the fact that the subject does not discriminate any object from itself, nor any two adjacent items in the series, but can indeed discriminate between any two non-adjacent items.
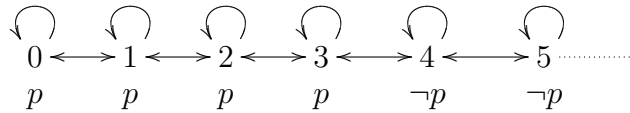

Figure 1: A discrete margin model

We use the language of propositional modal logic, where $\Box$ is interpreted as a knowledge operator: $\Box\phi$ stands for: "the subject knows that ϕ". We consider the usual semantics for propositional modal logic, in which given a Kripke model $\langle W, R, V \rangle$, and a world $w \in W$, $w \models \Box\phi$ iff for every w' such that wRw', $w' \models \phi$. Relative to this model, $n \models \Box p$ means that when looking at a pen of size n, the subject knows that it fits in the box. In the present case, this means that when looking at a given object, the subject knows that it has some property if and only if every object that is indiscriminable from it has that property. Alternatively, the above model can be seen as a particular case of what Williamson calls a fixed margin model. A *fixed-margin* model is a model $\langle W, d, \alpha, V \rangle$, where d is a metric over W, and α a real valued margin for error parameter, such that $w \models \Box\phi$ iff for every w' such that $d(w, w') \leq \alpha$, $w' \models \phi$. That is, at a given world, the subject knows that some proposition ϕ holds if ϕ holds at every world included within the margin α from w. The above model is a discrete margin model, such that $W = \mathbb{N}$ and $\alpha = 1$.

In the above model, for instance, $2 \models \Box p$, and $3 \models \neg\Box p \wedge \neg\Box\neg p$: thus, the subject knows that an object of size 2 will fit in the box, and does not know whether an object of size 3 fits in the box. Crucially, however, $2 \models \neg\Box\Box p$, that is the subject doesn't know that he knows that the pen fits in the box, since 4 is a

---

[2]The size unit is not relevant for the discussion, and the reader may replace centimeters by millimeters if it helps make the scenario more plausible. Likewise, sizes may be shifted for more plausibility (shifting 0 to 18, 1 to 19 and so on).

$\neg p$-world accessible in two steps from 2. For Williamson, this result is a welcome prediction of the model and semantics, since iterations of knowledge operators are seen by Williamson as a "process of gradual erosion" in the case of vague knowledge ([16]: 228). Indeed, each iteration of knowledge is seen as a step by which a margin of error is added. According to Williamson, the subject knows that he knows that p only if his knowledge is "safely safe", namely if the epistemic context of the subject is at least two steps away from the boundary between p and $\neg p$.[3]

However, one may argue that, looking at a pen of size 2, my knowing that I know that it will fit in the box supervenes only on my knowing whether it fits in the box, and not on epistemic alternatives that are further away. One important motivation to suppose so concerns higher-order iterations of knowledge: for instance, the standard semantics makes the prediction that at 0 in the above model, it holds that $\Box\Box\Box p$, and yet that it is not the case that $\Box\Box\Box\Box p$. However, it is hard to make sense of such fine-grained distinctions between levels of knowledge: indeed, if the subject knows that she knows that she knows that the pen fits in the box, how could she fail to know that she knows that she knows that she knows?

2.2 Centered semantics

In [1], we formulated an alternative semantics (CS, for Centered Semantics), in which every fact concerning the knowledge of the agent is decided solely on the basis of worlds that are not distinguishable from the actual world. At a conceptual level, the semantics aims at implementing the principle of supervenience of higher-order knowledge on first-order knowledge we just mentioned: there should not be a difference in what the agent knows that she knows about the world (or knows that she knows that she knows, etc.) without a difference in what the agent knows about the world. At a technical level, the gist of it consists in ensuring that the epistemic alternatives relevant for iterated modalities remain the worlds accessible in one transition from the world of evaluation. Thus, every fact concerning the knowledge of the agent should be decided solely on the basis of worlds that are not distinguishable from that world, without having to move further along the accessibility relation.

Given a model $M = \langle W, R, V \rangle$, we first define the notion of truth for couples of worlds:

Definition 1. *CS-satisfaction for couples of worlds:*

_____

[3]See [18], p. 123 on the topological understanding of the notion of safety.

(i) $M, (w, w') \vDash_{\text{CS}} p$ iff $w' \in V(p)$.

(ii) $M, (w, w') \vDash_{\text{CS}} \neg\phi$ iff $M, (w, w') \nvDash_{\text{CS}} \phi$.

(iii) $M, (w, w') \vDash_{\text{CS}} (\phi \wedge \psi)$ iff $M, (w, w') \vDash_{\text{CS}} \phi$ and $M, (w, w') \vDash_{\text{CS}} \psi$.

(iv) $M, (w, w') \vDash_{\text{CS}} \Box\phi$ iff for all w'' such that wRw'', $M, (w, w'') \vDash_{\text{CS}} \phi$.

The use of double-indexing allows us to represent both the perspective of the agent (through the first index, which we may call the *perspective* point), and also the information relevant relative to the agent's perspective (through the second index, which bears the atomic information, and which we may call the *reference* point).[4] The perspective point w is nothing but the evaluation world: the world the agent is in. The reference points w' are the worlds that are considered as epistemic alternatives by the agent, *given what her perspective is*. To evaluate a subformula which is not boxed, the information is provided by the reference point, which says what is true according to the epistemic alternative w' under consideration. Thus, truth clause (i) for atoms evaluates them by looking at what is the case in w', and truth clauses (ii)-(iii) for boolean connectives abide by the usual recursive definition of propositional connectives. All this is standard semantics. The crucial clause is clause (iv); it accounts for the "centered" feature of the semantics. In order to evaluate a boxed formula, one looks at the epistemic alternatives from the perspective of the agent, that is one looks at the worlds which are accessible from the evaluation world w. This is *not* standard semantics for epistemic logic. In Kripkean semantics, a box would prompt a move to worlds w'' that are accessible from the reference point w'. By contrast, in Centered Semantics, everything goes as if the successor set of an arbitrary reference point w' was the successor set of the perspective point w.

Now we shall extract the definition of truth for single worlds:

Definition 2. $M, w \vDash_{\text{CS}} \phi$ *iff* $M, (w, w) \vDash_{\text{CS}} \phi$

Truth with respect to single worlds is defined by diagonalization: at the beginning, the perspective point and the reference point coincide, they diverge when moves are made along the accessibility relation to treat formulae with modalities. Hence, for unboxed formulas, Centered semantics and Kripkean semantics will be equivalent. Of course, this is not so for the boxed ones. In Kripkean semantics, stacks of $\Box$ make more and more distant worlds relevant to the truth or falsity of a modal formula, which is of course the reason why $2 \vDash \neg\Box\Box p$ in the previous model. In Centered Semantics, instead of looking at worlds that are two steps away to check whether $\Box\Box\phi$ is satisfied, one backtracks to the actual world to see whether $\Box\phi$

---

[4]See [2] for details on the relation of CS to other double-indexing frameworks, in particular Rabinowicz & Segerberg's in [11]. The terminology of *perspective* vs *reference* points is from Rabinowicz & Segerberg.

already holds there. Thus, in the previous model, one can check that $2 \vDash_{\mathrm{CS}} \Box\Box p$. To see this, note that $(2, i) \vDash_{\mathrm{CS}} p$ for all $i \in \{1, 2, 3\}$, that is for all the worlds directly accessible from 2. This ensures that $(2, i) \vDash_{\mathrm{CS}} \Box p$ for all $i \in \{1, 2, 3\}$, hence $(2, 2) \vDash_{\mathrm{CS}} \Box\Box p$, which is $2 \vDash_{\mathrm{CS}} \Box\Box p$.

In [1], we proved that the normal logic K45 is sound and complete with respect to CS, namely the logic of positive and negative introspection.[5] An important aspect of the relativization of formulae to two indices in centered semantics furthermore concerns the iteration of what Williamson calls *margin for error principles* for knowledge (see [16], [4]). Margin for error principles play a central role in Williamson's epistemic theory of vagueness, since they are used to explain the status of borderline cases as cases for which an agent is uncertain about whether the predicate applies or not. For instance, let p_i be the (precise) proposition "the pen is of size exactly i", true at world i and false everywhere else. A margin for error principle is a principle such as $\Box\neg p_i \rightarrow \neg p_{i+1}$, which says that in order to know that the pen is of size i, the pen cannot be of size $i + 1$ either. In this case, the principle entails that if the pen is of size $i + 1$, the agent cannot exclude that it is of size i. It is easy to see that every instance of this particular principle is true at every world in the model of Figure 1, both under classic Kripke semantics and under centered semantics. However, the iterated or reflexive version of this principle, $\Box(\Box\neg p_i \rightarrow \neg p_{i+1})$, which also comes out true over the entire model with the standard semantics, is no longer true everywhere using centered semantics (it fails at 4 for $i = 2$). This divergence between the two semantics is discussed and motivated in much detail in [1] and [2], and we shall not dwell on it here. What is important for our purpose is that we take on board the motivation for basic margin principles, even if we handle their iteration differently.

2.3 Clarity and higher-order vagueness

When the $\Box$ operator is interpreted as "the subject knows that" or "I know that", our view is that Centered Semantics gives a more plausible view of higher-order knowledge than the standard semantics. When $\Box$ is interpreted as "it is clear that", however, as Williamson originally considered in his Logic of Clarity, the situation may be different. In particular, the main motivation to conceive of iterations as a process of gradual erosion is to have an account of higher-order vagueness: the idea is that clear cases of some property, like unclear instances, may in turn have borderline cases. For instance, using the standard semantics and interpreting "$\Box$" as "clearly", $2 \models \Box p$, which means that 2 clearly fits in the box, but $2 \nvDash \Box\Box p$,

---

[5]Axiom K is $\Box(p \rightarrow q) \rightarrow (\Box p \rightarrow \Box q)$, 4 is $\Box p \rightarrow \Box\Box p$ (positive introspection) and 5 is $\neg\Box p \rightarrow \Box\neg\Box p$ (negative introspection). S5 is the extension of K45 obtained by adding T, namely $\Box p \rightarrow p$. In [1], we also formulate a version of Williamson's fixed margin semantics, called CMS, for which the system S5 is sound and complete.

which means that it isn't clear that 2 clearly fits in the box. If we use Centered semantics instead, then $2 \vDash_{\text{CS}} \Box p$, but this time $2 \nvDash_{\text{CS}} \Box\Box p$. Relative to this interpretation of the $\Box$, it can therefore be objected that CS makes room only for first-order vagueness, and not for higher-order vagueness, since "it is clear that p" systematically entails "it is clear that it is clear that p" in CS (positive introspection), and furthermore "it is not clear that p" systematically entails "it is clear that it is not clear that p" (negative introspection).

Two main clarifications can be made on this point. Firstly, the operators "it is clear that" and "I know that" should not be taken to be synonymous, even when the kind of knowledge described is inexact.[6] Consequently, taking CS as a logic of inexact knowledge does not commit us to denying higher-order vagueness, so long as the $\Box$ operator is interpreted as "I know that", and not as "it is clear that". The relation between the two operators is not obvious. One issue lies in the understanding of the adverb "clearly": on one interpretation, "clearly" typically means "clearly from the point of view of a given subject". On another, more externalist perspective, "clearly" may be taken to mean more than that, as do related adverbs such as "objectively" or "determinately". To say that p is clearly the case might then mean that the occurrence of p should elicit the same subjective impression of clarity in different subjects for instance, or that it should elicit a stable or reliable response from the same subject upon various trials. A distinction of that kind is central to Halpern's account, in particular, as we will see in the next section.

In our perspective, when $\Box$ is read as "I know that", as we assume, $\neg\Box p \rightarrow \Box\neg\Box p$ should mean that I am aware of my uncertainty at the moment it first arises, namely that I know I am uncertain when I report uncertainty. For instance, consider a situation of "forced march" (see [6]: 173), in which I am forced to answer by "yes" or "no" (or possibly more values) for the application of one and the same predicate of each item in a soritical series. Given the structure of a forced march, a "jump" must occur at some point in my judgments, namely there must be a first item for which I switch my answer from "yes" to "no" or to another value.[7] With regard to such a situation, to say that I know that I don't know whether p holds means that I should be aware of the jump in my judgments as soon as the jump first occurs. In the case under discussion, we could imagine that the choice is between "yes", "no" and "indeterminate". In that case, the subject will say "yes" to "does pen 0 fit in the box?", and likewise for pens 1 and 2. When looking at pen 3, however, the subject starts to hesitate: the subject then shifts her answer to "I am not sure" (namely "indeterminate"). But at the moment the subject makes this

---

[6]See [2] for further developments on this point.

[7]A jump occurs as soon as the subject declares a difference in semantic status for two subsequent items in the series. See [6], 173 sqq.

jump from "yes" to "indeterminate", the subject presumably is aware of making the jump: it is perfectly consistent with the scenario to imagine that the subject knows that she doesn't know whether 3 fits in the box or not.

Note that this is not the same thing as saying that if p is unclearly instantiated, it is clear that it is unclearly instantiated. Sure enough, I declare "indeterminate" because it is not clear to me whether the pen fits in the box; yet I may still hesitate as to whether I might have issued a better judgment in that situation. Another way to put it is to say that I can report "unclear" when asked whether the pen fits in the box, without necessarily being inclined to report "clearly unclear". There is at least one understanding of "clearly clear" and "clearly unclear" on which they mean "very clear" and "very unclear". Thus I can report clarity and unclarity without necessarily reporting a high intensity of clarity or a high intensity of unclarity. This points to another difference between knowledge and clarity, which we shall investigate in greater detail in section 4, namely that clarity is a matter of degree.

A second element of response to the problem of higher-order vagueness concerns the fact that in [1] we show that CS is a particular case of a family of resource-sensitive semantics called TS(n) (for "token semantics with n tokens"), for which iterations of operators is not automatic at the first order, but can happen at some higher level n of iterated modalities, depending on the number n of tokens available. For this family of semantics, the interpretation of $\Box$ as an operator of clarity (rather than knowledge) is compatible with higher-order vagueness. When the number of tokens is greater than 1, in particular, a proposition of the form $\neg\Box p \wedge \neg\Box\neg p$ no longer entails $\Box(\neg\Box p \wedge \neg\Box\neg p)$: interpreted in terms of clarity, this means that if p is such that neither it nor its negation is clear, then this unclarity is not automatically clear, in agreement with second-order vagueness. If the number of tokens is bounded, moreover, the resulting semantics makes room for a limited form of higher-order vagueness beyond second-order vagueness, as we shall see.

Informally presented, the intuition behind Token Semantics is that moving along the accessibility relation has a cost, which is mirrored by the fact that a token is spent for each (non-reflexive) move in a model, and the initial number of tokens available to the agent is finite and non-zero.[8] When all tokens have been spent, just as in CS, the agent backtracks to the position reached before the last move, gets a token back, and can spend it for a new move. Formally, satisfaction is defined with respect to a sequence of worlds and a number of tokens: q is short for an arbitrary sequence of worlds, qw for an arbitrary sequence augmented with

_____

[8]We refer the reader to [1] for a detailed and step by step presentation of the semantics. Non-reflexive moves come at no cost, in particular, to ensure that the T axiom of veridicality of knowledge is preserved over reflexive structures.

w, and n is an arbitrary number of tokens.

Definition 3. *Token satisfaction:*

    (i) $\mathcal{M}, qw \vDash_{\text{TS}} p \; [n]$ iff $w \in V(p)$.

    (ii) $\mathcal{M}, qw \vDash_{\text{TS}} \neg\phi \; [n]$ iff $\mathcal{M}, qw \nvDash_{\text{TS}} \phi \; [n]$.

    (iii) $\mathcal{M}, qw \vDash_{\text{TS}} (\phi \wedge \psi) \; [n]$ iff $\mathcal{M}, qw \vDash_{\text{TS}} \phi \; [n]$ and $\mathcal{M}, qw \vDash_{\text{TS}} \psi \; [n]$.

    (iv) $\mathcal{M}, qw \vDash_{\text{TS}} \Box\psi \; [n]$ iff

          • $n \neq 0$ and for all w' such that wRw', $\mathcal{M}, qww' \vDash_{\text{TS}} \psi \; [n-k]$, where $k = 1$ if $w \neq w'$ and $k = 0$ if $w = w'$

          • or $n = 0$ and $\mathcal{M}, q \vDash_{\text{TS}} \Box\psi \; [1]$.

Since the initial number of tokens required for evaluation is assumed to be strictly positive, this guarantees that q is never empty when $n = 0$, so the second item of clause (iv) is well-defined. To take a concrete example, on the previous model the rules of satisfaction in Token Semantics predict that: $1 \vDash_{\text{TS}} \Box\Box p \; [2]$, as in the standard semantics, since all worlds reachable in two steps from 1 satisfy p. Likewise, as in the standard case, $2 \nvDash_{\text{TS}} \Box\Box p \; [2]$. However, we now have that $1 \vDash_{\text{TS}} \Box\Box\Box p \; [2]$, since with only 2 tokens, the agent cannot visit worlds beyond 3 when her initial context is 1. It is easy to see that CS corresponds to Token Semantics with only 1 token allowed, and standard Kripke Semantics to Token Semantics with an infinite number of tokens available. More generally, $\text{TS}(n)$ and standard Kripke semantics coincide for formulas with less than n embedded modalities (see [1]).

Using $\text{TS}(n)$ semantics, one can in principle account for n-order vagueness when the $\Box$ operator is interpreted as "it is clear that". But again, $\text{TS}(n)$ cannot be a logic of $n + 1$-order vagueness, since in $\text{TS}(n)$ the schema $\Box^n p \to \Box^{n+1} p$ comes out as a validity (see [1] and the Appendix for technical details): thus "it is clear ($n - 1$ times)... that p" does not necessarily imply "it is clear (n times) that p" (that is $\nvDash_{\text{TS}(n)} \Box^{n-1} p \to \Box^n p$), but "it is clear ($n$ times) that p" implies "it is clear that it is clear (n times) that p" (namely $\vDash_{\text{TS}(n)} \Box^n p \to \Box\Box^n p$). The question, here, is whether one can plausibly conceive of higher-order vagueness without being committed to higher-order vagueness at all orders.

Williamson, in ([17]:136), shows that in KTB,[9] namely the basic Logic of Clarity, a proposition p either is *precise* (such that $\Box p \vee \Box\neg p$ holds), or has *first-order vagueness* but not higher-order vagueness (namely is such that $\Box p \to \Box\Box p$ and $\neg\Box p \to \Box\neg\Box p$ necessarily hold), or has *vagueness of all orders* (for instance, $\Box\Box p$ is compatible with $\neg\Box\Box\Box p$, and similarly for any level of embedding). However, Williamson remarks that if "B is abandoned, p can have vagueness of all orders below n and precision thereafter for any $n \geq 1$" ([17]:138).

---

[9]B is the axiom $p \to \Box\neg\Box\neg p$. KTB axiomatizes Williamson's fixed margin semantics. Williamson also presents a variable margin semantics, axiomatized by the logic KT.

Conversely, Williamson notes that in a logic stronger than KTB, like S5, there can be first-order vagueness but no second-order vagueness.

The relevant point on this issue is that for every $n \geq 2$, TS(n) is axiomatized by a logic weaker than K45 that fails to yield B, even when axiom T is systematically included. Conversely, when T and B are assumed, the resulting logic, which includes the schema $\Box^n p \rightarrow \Box^{n+1} p$ (see Appendix) is then stronger than KTB, but still weaker than S5 for $n \geq 2$. Either way, the resulting logic for TS(n) remains a plausible candidate for n-order vagueness when $\Box$ is interpreted as an operator of clarity. Williamson ([17]: 134) notes that "intuitively, there is a strong connection between the non-transitivity of indiscriminability and higher-order vagueness". What matters in this respect is that while transitivity and euclideanness are invalid in TS(n) for $n > 1$, weaker forms of both properties are preserved for the logic (namely n'-transitivity and n'-euclideanness, as described in the Appendix).

More fundamentally, we do consider that higher-order vagueness running out at some finite order is plausible enough, much for the same reasons which concern the iterations of knowledge operators. In our view, instances of a property may become clearly clear instances at some point, in much the same way in which a man with 0 hair on his head is a clear case of baldness, and a clearly clear case thereof, and so on at all levels (see [4]). As we shall see later in section 4, this view calls for several refinements, in particular because clarity and unclarity need not pattern in the same way, but the view of limited higher-order vagueness we consider here will be an important component in the view we want to argue for.

3   Halpern's semantics

Halpern takes a different approach to the problem of inexact knowledge, since his logic makes room for distinct syntactic and semantic representations of the operators "I know that" and "it is clear that". His logic (in the one-agent case) uses two primitive operators, namely R and D, where $R\phi$ is taken to mean that the agent "reports ϕ", and $D\phi$ to mean that "according to the agent, ϕ is definitely the case". As we shall see in greater detail, the operator R is the equivalent of the operator "I know that" and implements a notion of internal and subjective uncertainty. The definiteness operator D, on the other hand, gives a description of the same uncertainty, but from the perspective of the objective stimulus, as it were, rather than from the viewpoint of the agent. Williamson's "it is clear that" operator, finally, is not expressed directly by D in Halpern's system, but simulated by a combination of the two operators D and R.

3.1 Subjective and objective uncertainty

A model, relative to Halpern's language, is a structure $\langle W, P, \sim_s, \sim_o, V \rangle$, with $W \subseteq O \times S$, where S is intended to denote a set of subjective indices and O a set of objective indices. A subjective index s is taken to represent the subjective estimate or measurement that an agent makes of a target value o. For the same value o, the agent's measurement or estimate can vary. Dually, for the same subjective estimate, the target value is not uniquely determined, but lies in a certain interval of error. Either way, the relation of one parameter to the other is therefore many-to-many. Relative to a state (w, v) of W, it is natural to define two uncertainty relations. The relation $\sim_s$ relates states with the same subjective index: it describes the agent's subjective uncertainty about the target value. Accordingly, an agent "reports that ϕ" relative to his subjective estimate. Conversely, the relation $\sim_o$ relates states with the same objective index, and describes the fluctuation of the agent's estimate around the target value. Accordingly, something is "definitely the case, according to the agent" relative to the target value (which is known, for instance, to the experimenter).

The relations $\sim_o$ and $\sim_s$ both are supposed to be equivalence relations over W, and V is a valuation over W. P, finally, is a subset of W, intended to denote the states that the agent considers plausible. For simplicity, we shall assume that $P = W$ here, and therefore we shall omit reference to P in the definition of satisfaction. With that simplification, the satisfaction clauses for the modal operators are the expected ones, namely $M, (w, v) \models R\phi$ iff for every (w', v') such that $(w, v) \sim_s (w', v')$, $M, (w', v') \models \phi$, and similarly for $D\phi$ with respect to $\sim_o$. As a consequence, each operator is axiomatized by the logic S5.[10] The point of Halpern's approach, however, is that although each operator separately obeys transitivity (and euclidianness), their combination DR need not (if two binary relations A and B are equivalence relations, it does not follow that their composition $A \circ B$ is transitive or euclidian).[11] Intuitively, $DR\phi$ holds at a world when the agent's subjective estimation is sufficiently reliable, namely when the agent would report ϕ for all subjective evaluations compatible with the objective stimulus. From $DR\phi$, it follows that $R\phi$, but the converse does not hold. Thus, we can understand $DR\phi$ to mean that the agent "reports ϕ reliably". More generally, the complex operator (DR) plays exactly the role of Williamson's "clearly" operator in margin for error semantics: like Williamson's operator, the operator (DR) need not be iterative.

To make the link concrete between the two approaches, let us consider a

---

[10]In Halpern's full version of the semantics for the multi-agent case, each modality is actually a KD45 operator (where D is $\Box\neg p \rightarrow \neg\Box p$), and for each agent the corresponding D operator satisfies a weakened version of axiom **T**.

[11]For two binary relations A and B over W, $A \circ B =_{df} \{(v, w); \exists u : (v, u) \in A \wedge (u, w) \in B\}$.
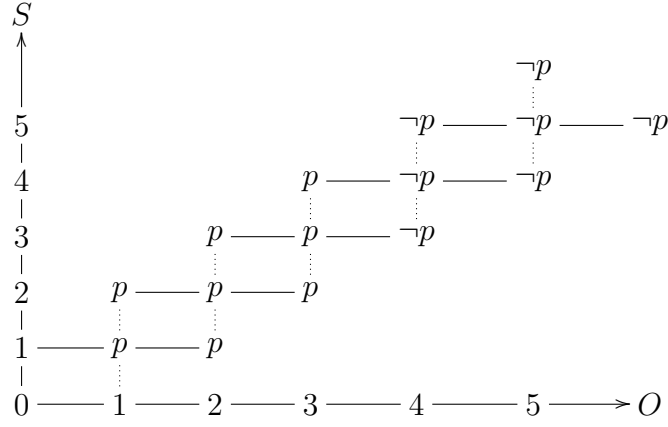
$$S$$

Figure 2: A layered margin model

model, depicted in Figure 2, in which W is the subset of $\mathbb{N} \times \mathbb{N}$ consisting of couples (n, m) such that $|n - m| \leq 1$. Let us suppose that n is the objective size of some object, or the objective value of some parameter, and m its subjective estimate. The constraint on n and m represents the fact that the subject's estimate cannot deviate from more than 1 on the objective value, namely that the subject's margin of error is 1. Let us suppose moreover that $(n, m) \sim_o (n', m')$ iff $n = n'$, namely if they agree on their objective indices, and likewise $(n, m) \sim_s (n', m')$ iff $m = m'$, namely if they agree on their subjective indices. It is easy to verify that both relations are equivalence relations over W.

In the above figure, each cell of the partition determined by $\sim_o$ corresponds to the points connected by a vertical dotted line, and each cell of the partition determined by $\sim_s$ corresponds to the points connected by a horizontal straight line. Let us suppose moreover that whether a point w is a member of $V(p)$ depends only on the objective part of w. For instance suppose that $(n, m) \in V(p)$ if and only if $n < 4$ (as in our previous example, p may stand for "fitting in the box"). It is easily checked that $(1, 2) \models DRp$, that is for all worlds $(1, x)$ subjectively compatible with $(1, 2)$, the agent reports p. But $(1, 2) \nvDash DRDRp$, since $(4, 3)$, for instance, is a $\neg p$-state accessible by a D-R-D-R transition. Concretely, this means that if the agent measures a 2, while the objective state is 1, the agent makes a measure that is subjectively compatible with the objective state being 3: but if the objective state had been 3, the agent might have measured 4, and then would not have reported p anymore. Thus, if the size of the object is 1 and the measurement made by the agent is 2, with a threshold for $\neg p$ that is between 3 and 4, then the agent "reliably reports" that p, but this reliable report is not itself "reliably reliable".

By contrast to DR, R is an S5 modality, satisfying negative and positive intro-

spection at any point in the model. The way we understand this is the following: R is taken to characterize the agent's subjective certainty about his own response. If the agent reports that ϕ, the agent knows that she reports ϕ. This is not necessarily so with DR: at $(1, 2)$, the agent reliably reports that p, namely DRp is true, but $RDRp$ does not hold: what is predicted is that the agent need not know that her report is reliable, even when it is reliable.

3.2 Correspondence with centered semantics

If we consider only the relation of *subjective* equivalence for R, a model like the model of Figure 2 may be called a *layered margin model*, since each horizontal equivalence class (namely the classes for $\sim_s$) contains the possible objective values that are compatible with the agent's subjective parameter, and the horizontal projection of these classes onto the O-axis of the model would yield a linear structure of inexact knowledge of exactly the kind with which we started in Figure 1.

This notion of layering can be made precise. Thus, given a Kripke model $M = \langle W, R, V \rangle$, let us call $L(M) = \langle W', R', V' \rangle$ a *layering* of M if it satisfies: $W' = \{(w, w') \in W \times W; w'Rw \lor w' = w\}$; $(w, w')R'(u, u')$ iff $w' = u'$ and $w'Ru$; and finally, $(w, w') \in V'(p)$ iff $w \in V(p)$ (note that it is the first index here, namely the objective index, which specifies the atomic information, whereas in Centered Semantics the first index, or perspective point, comes first: this explains the inversion of indices in Lemma 1 below). It can be checked that R' in $L(M)$ is necessarily transitive and euclidian for every R, and is an equivalence relation if R is reflexive. For instance, consider a non-transitive and non-euclidian structure such that $W = \{0, 1, 2\}$ and $R = \{(0, 1), (1, 2)\}$. Then $W' = \{(0, 0), (1, 1), (2, 2), (1, 0), (2, 1)\}$, and $R' = \{((0, 0), (1, 0)), ((1, 1), (2, 1)), ((1, 0), (1, 0)), ((2, 1), (2, 1))\}$. R' is trivially transitive and euclidian in this case. Likewise, over the model of Figure 2, the equivalence relation $\sim_s$ is exactly equal to the layering R' of the reflexive and symmetric relation R of Figure 1.

It is easy to establish that, relative to the basic modal language in which $\Box$ is the single modality:

Proposition 1. $M, w \vDash_{\text{CS}} \phi \iff L(M), (w, w) \models \phi$

by proving that for all (w', w) in $L(M)$:

Lemma 1. $M, (w, w') \vDash_{\text{CS}} \phi \iff L(M), (w', w) \models \phi$.

Proof. $M, (w, w') \vDash_{\text{CS}} p$ iff $w' \in V(p)$ iff $(w', w) \in V'(p)$ iff $L(M), (w', w) \models p$. The boolean cases are immediate. $M, (w, w') \vDash_{\text{CS}} \Box\phi$ iff for every w'' such that wRw'', $M, (w, w'') \vDash_{\text{CS}} \phi$ iff, by induction hypothesis, for every w'' such that

14

wRw'', $L(M), (w'', w) \models \phi$, iff, by definition of $L(M)$, for every (u, u') such that $(w', w)R'(u, u')$, we have $L(M), (u, u') \models \phi$, iff $L(M), (w', w) \models \Box\phi$.

$\Box$

Proposition 1 applies also to Williamson's fixed margin models $M = \langle W, d, \alpha, V \rangle$ (see [4]) in which two worlds w, w' are accessible iff $d(w, w') \leq \alpha$, where d is a metric over W. A layered margin model necessarily is a model in which the induced epistemic accessibility is an equivalence relation, since margin models are reflexive by definition of a metric between worlds. With respect to margin models, Proposition 1 shows that Halpern's operator R therefore plays exactly the role of the knowledge operator $\Box$ in the framework of centered semantics.

On the one hand, these results makes clear that centered semantics really is a standard two-dimensional semantics in disguise. On the other hand, the operation of layering shows how it is possible to recover transitivity (and euclideanness) from a non-transitive (and non-euclidian) relation. In Halpern's approach, the intransitivity characteristic of qualitative comparison is simulated by means of two operators. In our semantics, by contrast, only one relation of uncertainty is given, which is supposed to match the intransitivity of judgments of discrimination directly. Despite this, the non-transitivity of indiscriminability does not impair the principle of positive introspection.

3.3 Intransitivity

According to Halpern, if one conceives of epistemic accessibility relations as indistinguishability relations, "there is a strong intuition that the indistinguishability relation should be transitive", "as should the relation of equivalence on preferences" in the case of rational choice theory ([5]: 2). Halpern's argument in favor of this intuition is grounded in part on a wish to preserve partitional models of information as a general framework (see [5]:2). At the same time, however, we may note that Halpern's view converges with arguments given independently by Fara in [3] and Raffman in [12], in favor of the idea that phenomenal indiscriminability is indeed transitive, despite appearances to the contrary.[12] For Fara, in particular, the apparent non-transitivity of "looking the same as" is due to a surreptitious shift of the context with regard to which judgments of resemblance are made. Raffman reaches the same conclusion as Fara, by pointing to the fact that in a soritical series, sameness is context-dependent, and that one and the same color

---

[12]See Fara for the characterization of *phenomenal* sorites. Phenomenal indiscriminability (the relation of "looking the same as") is supposed to apply to *observational* predicates, whose "applicability to an object (given a fixed context of evaluation) depends only on the way that object appears" ([3]: 907).

patch #4, for instance, may upon reflection look different in two different acts of comparison, even though it looks the same as #3 in one context, and the same as #5 in a different context.

The correspondence between Halpern's approach and ours suggests that we can perfectly agree with Halpern, Fara or Raffman on the idea that phenomenal indiscriminability is transitive, *once it is explicitly contextualized*. Indeed, in a layered margin model such as the model of Figure 2, the relation of subjective indistinguishability is between ordered pairs with the same subjective index. In the non-transitive margin model of Figure 1, by contrast, the relation of indistinguishability is not relativized in this way. However, when epistemic sentences are evaluated with respect to Centered Semantics over this model, what obtains is in fact an *implicit relativization* of exactly the same kind. Thus, to say that 1 looks the same as 0, and that 1 looks the same as 2, when understood relative to our framework as well as to Halpern's framework, can in fact mean different things. Typically, it will mean that 1 and 0 look the same *from the perspective of 1*, and that 1 and 2 look the same *from the perspective of 1*. If indeed "looking the same *from a given perspective*" is transitive, then 0 and 2 look the same *from the perspective of 1*. However, 0 and 2 need not look the same from the perspective of 0 or from the perspective of 2. Note that this distinction suggests some empirical consequences. Consider a discrimination task in which a subject must say whether two color patches are identical or distinct. If what we say is right, it should be easier for the subject to say that shade 0 and shade 2 are distinct when presented next to each other than when presented with intermediate shade 1 in between.

Despite this, one could still ask which of the two indistinguishability relations should be considered as more primitive in order to describe knowledge: should we take as primitive the transitive relation of indistinguishability of the layered model? or rather, should it be the non-transitive relation of the original margin model? Here Proposition 1 can be interpreted in two opposite ways. Like Halpern, we may consider that the relation of perceptual indiscriminability fundamentally is transitive, and that the non-transitive linear model of Figure 1 in fact results from the transitive product model of Figure 2 (by an operation inverse to the operation of layering). On that view, the "true" relation of indistinguishability behind the operator $\square$ is the relation R' of the layered model, not the indistinguishability relation R of the linear margin model. But one could interpret the situation the other way around, and consider that the non-transitive relation R of the model of Figure 1 is the primitive relation.

The choice between these two options depends in part on what one takes to be the best representation of the notion of epistemic context, and then on what one takes "perceptual" to mean when talking of "perceptual indistinguishability" between contexts. The layered model of Figure 2 gives a way of having an infinite set of distinct accessibility relations, one for each individual context, which

is equivalent in turn to defining a single accessibility relation between richer contexts defined as ordered pairs, as in Halpern's semantics. In comparison, the margin model of Figure 1 gives a simpler representation of the notion of context: on our view, this representation gives a more intuitive rendering of the perceptual *experience* that subjects might have of a soritical series (like a series of pens ordered by height, or a series of hues cleverly shaded, and so on).

The difference of granularity between the two notions of context reflects itself in a difference between two ways of understanding the notion of "perception". When Halpern opposes "perception about sweetness" and "reports about sweetness", for instance, Halpern understands "perceptual" in a narrow sense, namely by reference to the subpersonal processes by which an agent organizes his or her basic sensory information. The example given by Halpern is that of a robot with sensors, whose perception of sweetness goes by unambiguous thresholds. By contrast, what Halpern describes as "reports about sweetness" is supposed to correspond to the qualitative experience the robot has of sweetness, which Halpern sees as potentially biased and ambiguous, leaving room for non-transitivity in discrimination. This notion of "report" is in fact the qualitative notion of "perception" that we intend when we talk of perceptual indistinguishability.

There should be no misunderstanding, consequently, when we assert that in our view "perceptual indistinguishability" can be conceived as non-transitive: by this, we have in mind exactly the kind of indistinguishability which Halpern would associate to reports about perception. But the difference is that we can take such a non-transitive relation as primitive, without having to match it to a complex operator in order to describe knowledge. We do not think, therefore, that any epistemic uncertainty relation should necessarily be transitive. Concerning introspection, however, what Proposition 1 shows is that Halpern's approach in terms of product models and our approach in terms of a non-standard semantics follow essentially the same inspiration, by making higher-order knowledge depend only on the states that are in the immediate ken of the agent, and by avoiding spurious dependencies to alternatives that lie beyond those on which knowledge of the first-order supervenes.

4 Vagueness and degrees of clarity

So far we have compared three accounts of the notion of vague or inexact knowledge in epistemic logic. In so doing, we have placed our account under the following two assumptions, put forward by Williamson: firstly that "vagueness issues from our limited powers of conceptual discrimination", and secondly that "it is often associated with the expression in logic of such limits: the non-transitivity of indiscriminability" ([16], 237). In the present section, we propose to separate

those two assumptions more clearly from each other, and to discuss the epistemic account of vagueness in the light of another model of inexact knowledge, the signal detection model widely used as a general framework in psychophysics (see [10], [9]).

An essential reason for undertaking this comparison is that the notions of epistemic certainty and clarity that we have discussed so far do not admit of degrees. Intuitively, however, clarity comes in degrees. We say that something is *more clearly* red than something else, or that something is *very clearly* the case. Another reason is that signal detection theory offers an account of imprecision that does not rest fundamentally on the notion of intransitivity in indiscriminability, but on a more abstract notion of probability of error (represented by the notion of signal to noise ratio). We start this section with a brief overview of signal detection theory. We relate it to the notion of degree of clarity, and conclude with a fresh discussion of higher-order vagueness.

4.1 Signal detection, sensitivity and discrimination

One motivation to make the comparison between SDT and the epistemic account of vagueness is that in section 2 we started from the description of an imaginary psychophysical experiment, a yes-no task in which a subject is asked whether objects of various sizes will fit in a certain box. The central assumption behind this scenario is that if indeed we are dealing with a situation of vagueness, the vagueness in question is neither in the world (pens and box have precise sizes), nor in language ("fitting in the box" has a perfectly precise extension relative to the setting), but it is wholly imputable to the subject's inability to perfectly discriminate between various lengths.

An important caveat at this juncture is that we do not mean to endorse the thesis that vagueness is fundamentally an epistemic phenomenon in general. Our perspective in this paper is much more limited, namely to account for situations in which it is at least *consistent* to reduce vagueness to a form of epistemic uncertainty, or to identify vagueness with a form of inexact or imprecise knowledge. Within this limited perspective, what we aim at here is to see to what extent the different models of inexact knowledge we compared fit with signal detection theory as the prevalent model underlying actual psychophysical experiments.

Psychophysical tasks are typically designed to measure the subject's sensitivity in the detection of a particular signal, namely the subject's ability to correctly discriminate between stimuli that vary gradually or continuously. In a standard yes-no task, in particular, the experimenter must be able to classify the subjects' answers into *hits* and *correct rejections* on the one hand, vs. *false alarms* and *misses* on the other. This implies that the experimenter can fix a clear boundary between signal and noise, or more generally between the categories to be discrim-

inated from each other, exactly as in the imaginary scenario discussed in section 2. To take a concrete example, in a version of an actual task of visual density discrimination run by D. Smith & al. (on humans and on monkeys), the subjects have to select between a "Dense response" and a "Sparse response" relative to a number of pixels illuminated on a screen's rectangle box, so that "the Dense response (choosing the box) was correct if the box contained exactly 2950 pixels. The Sparse response (choosing the S) was correct if the box had any fewer pixels" ([14], 321, capitals ours). As a consequence, the boundary between the two target categories "Sparse" and "Dense" is fixed and completely precise. It is stipulated by the experimenter in that case, but it needn't be stipulated most of the time, for instance it will correspond to objective identity and objective difference in a Same vs. Different discrimination task (for instance one in which the subject must say if two boxes have the same number of illuminated pixels or not, as in experiments by Shields et al., reported in [14]: 325). If indeed there is vagueness in such scenarios, it is only to be seen as uncertainty on the side of the subject.

An assumption common to signal detection theory and to the epistemic theory of vagueness, therefore, lies in the idea that the distribution of answers is primarily a factor of the subject's limited capacity to perform discriminations. On top of that, however, the distribution of answers in the signal detection model is also supposed to depend on the subject's response strategy or *decision criterion*, an aspect that is not at all reflected in the kind of epistemic model we presented so far and to which we shall return. However, a typical way of measuring the subject's sensitivity is to hold the strength of signal and noise constant, and to encourage the subject to change her decision criterion (see [10], chap. 2), by manipulating the rewards and costs associated to correct and wrong answers. In what follows, we make the implicit assumption that costs and rewards are identical, and that the subject we are talking about is an ideal observer, capable of minimizing incorrect answers and maximizing correct ones. Concretely, this means that the main ingredient in signal detection for our purposes is the notion of subject's sensitivity. What we shall focus on, in other words, is not the decision part of signal detection theory, but rather the underlying theory of errors and imprecision.

In SDT, the notion of sensitivity is represented by the distance or separation between two probability distributions that represent the prior probability of observing a given stimulus value along a dimension of subjective impressions, depending on whether there is noise or signal respectively. For instance, in the case of the imaginary task described in section 2, for each size x that the subject perceives, we have to imagine that there is a prior probability $P(x|s)$ of observing that size assuming that the pen fits in the box (so assuming it is less than 4, the signal condition), and a prior probability $P(x|n)$ of observing that size assuming that it does not fit in the box (so assuming it is greater than 4, the noise condition). These two probability distributions in turn permit to determine the posterior prob-
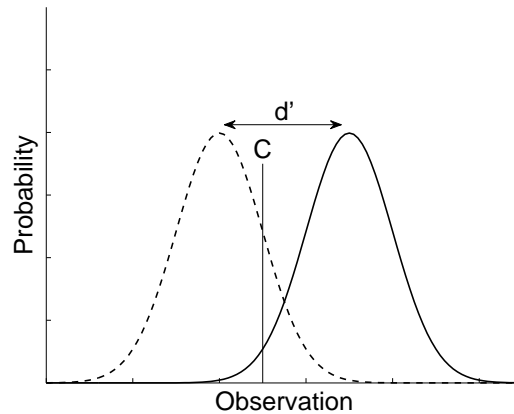
Figure 3: A typical SDT plot

abilities $P(s|x)$ and $P(n|x)$ that the subject perceives signal or noise respectively, given the sensory value x.

A typical plot for the two prior probability distributions of signal and noise is represented in Figure 3, where the dashed curve represents noise and the continuous curve signal, and on which we have moreover represented some hypothetical position of the decision criterion (indicated by C on the schema). Most of the time, the distributions of signal and noise are assumed to be standard normal distributions. The range of values for which the two probability distributions of signal and noise overlap corresponds to the range of values for which the subject is likely to make mistakes and in which she will experience uncertainty. Quite generally, the sensitivity of the subject is represented by the distance d' between the mean value of the two distributions. In particular, if those two distributions had no overlap, then it would entail that the subject would have perfect discrimination in principle. If the distance d' was null, conversely, and the distributions were to completely overlap, this would mean that the ratio $P(x|s)/P(x|n)$ is always equal to 1, and that the odds of having noise or signal are identical. In that case, the subject would make no difference between signal and noise, and this would correspond to a state of maximal uncertainty, a state in which the subject's answers are typically at chance level.

4.2 Degrees of indiscriminability

If we briefly compare the signal detection model to the margin for error model, we can see that the latter is more coarse-grained. In particular it does not represent the degree of confidence the subject might have that a pen of such and such size will fit in the box. For instance, when the actual size is 3 or 4 and the margin of error is 1 (see Fig. 1), margin semantics as well as centered semantics would predict that the subject does not know whether p or not p, namely is uncertain as to whether the pen fits in the box. Conversely, when the size of the pen is 1 or 2, the subject is predicted to be certain that the pen fits in the box. But no distinction is made concerning the degree of certainty or uncertainty of the subject relative to these different states.

On a more realistic interpretation, however, the subjective probability $P(p|x)$ that the pen fits in the box, given that the subject perceives the pen as of size x, will most likely vary more gradually with the perceived size x of the pen. We may expect, for instance, $P(p|2)$ to be less than $P(p|1)$. That is, we may expect that the subject will be less likely to say that the pen fits as the perceived size of the pen increases. In practice it might also be the case that $P(\neg p|2)$ is positive rather than 0.

If we think about this in relation to the notion of margin of error, this would mean that even when the subject perceives a pen as of size 2, the subject may not be entirely confident that the pen will fit in the box. In particular, there could be a slight possibility (though not necessarily conscious) that the pen be misrepresented as of 5 for instance. This probability might be lower than that of mistaking a pen of size 2 with a pen of size 3, but this would stand in contrast to the assumption, implicit behind the margin model of Fig. 1, that all worlds within a margin of 1 cm from 2 are equally indiscriminable from 2, and that all worlds more remote are equally discriminable from it.

To represent this possibility using Kripke models, we can imagine to label indiscriminability relations between worlds with various weights, or more generally, to have for each world a distribution of probability over worlds, indicating the degree with which it is likely to be represented as different from what it is. Doing this will allow us to implement something like a *variable* margin of error, or more accurately, to implement a notion of *degree* of indiscriminability (intermediate between 0 and 1).[13]

Let us imagine, for instance, that when the subject sees a pen as of size 3, there is a probability 1/3 that it be correctly perceived, a probability 1/5 of misrepresenting it as of size 4, a slightly lower probability of misrepresenting it as of

---

[13]On variable margins, see also Williamson's recent work in [19]. Variable margins are not to be confused with Williamson's variable *margin semantics* in [16], which is defined over fixed margin models.

size 2, and various lower probabilities to see it as of size 1, 5 or 6 (as specified on Figure 4). Given a probability distribution, let us stipulate that a world w' is accessible from w if and only if the probability of mistaking w' for w is greater or equal to 0 (this will make every world accessible to every other world).

Relative to each world in the structure, one can define the probability of a given proposition as the sum of the probabilities of accessible worlds that satisfy the proposition. In that way, it is possible to introduce a degree-theoretic clarity operator $\Box_\alpha$ in the language, such that $w \models \Box_\alpha \phi$ iff $\sum_{wRw', w' \models \phi} P_w(w') = \alpha$. Intuitively, the operator of clarity $\Box_\alpha$ we are here defining can be seen as describing the relative confidence the subject might have in reporting that a condition is instantiated.[14]

In the Kripke structure below, for instance, we represented a possible probability distribution relative to world 3. One can check that $3 \models \Box_{3/5} p$, meaning that when the subject perceives a pen as of size 3, it is clear to degree 3/5 to the subject that it will fit in the box. More generally, it can be checked that the standard operator of clarity $\Box$ can be defined as $\Box_1$ in this context.
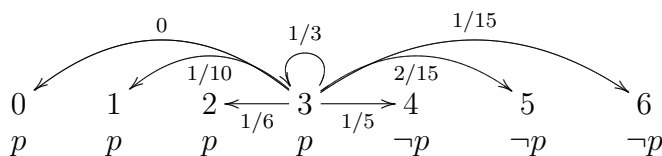

Figure 4: A likelihood model

Let us call likelihood model such a structure.[15] We may impose that each world should have some positive probability relative to itself. An important point, on the other hand, is that $P_w(w')$ need not be equal to $P_{w'}(w)$: for instance, when seeing a pen as of size 4, there may be a slight probability of mistaking it for a

---

[14]One way to think about the operator of relative clarity is as expressing what Schiffer calls *vagueness-related partial belief*, or a particular case thereof (see [13], and [8] for a recent discussion). On our perspective, in particular, degrees of clarity are as unproblematic as degrees of belief, unlike degrees of truth. A difference of our approach with Schiffer's account is that for Schiffer, vagueness-related partial beliefs are "not a measure of uncertainty or ignorance", at least not so in general ([13]: 202). In our framework, degrees of clarity are by definition a measure of epistemic uncertainty, since we derive them from structures of perceptual error.

[15]More formally, a likelihood model relative to a propositional language may be defined as a structure $\langle W, R, V, (P_w)_{w \in W} \rangle$, where W and V are defined as usual for a Kripke model; for each world w, P_w is a probability distribution over W; wRw' iff $P_w(w') \geq 0$ (hence R is automatically universal; we could also let wRw' iff $P_w(w') > 0$, but this would introduce unnecessary complications in what follows). In [19], Williamson also discusses Kripke structures with a probability distribution on the set of worlds. However, relations of accessibility there are defined independently of the distribution, namely to characterize an operator of knowledge, and with no reference to degrees of clarity.

pen of size 1, but when seeing the pen as of size 1, the probability of mistaking it for a pen of size 4 may well fall to 0. Thus, we may conceive that the closer we get to the boundary between pens that fit in the box and pens that do not, the more we are at a risk of error. Various other structural constraints are conceivable on such probability assignments. In the figure above we only assumed that the greater the distance between two worlds (as defined by the difference between the numbers labelling them), the shorter the probability of mistaking them for one another, even though equidistant worlds can have different weights.

A further difference with fixed margin models here is that the interval of error (namely the set of worlds that receive a positive probability of error around a given world) may vary both in amplitude and in probabilistic structure from world to world. The notion of relative indiscriminability here introduced may appear dubious if one considers that pairwise adjacent pens should indeed be treated as equally indiscriminable throughout, or that worlds within a given radius should receive identical weights.[16] What we are after, however, is in fact a generalization of the notion of fixed margin of error. To make it more plausible, we need to look more closely at the signal detection model.

4.3 An example of signal detection

In order to illustrate more explicitly the link between SDT and the family of epistemic models we presented earlier, let us consider an abstract task of signal detection. The task is based on a simple card-sorting problem presented by McNicol in ([10]: 15), intended to make clear the link between signal detection theory and decision theory more generally. The problem is simple enough in that it involves discrete probability distributions of signal and noise.

Let us imagine that a subject must sort 450 cards which originally had 1 to 5 spots painted on them. Before the experience, however, the experimenter adds one spot on half of the cards. The subject is then shown the cards and asked, for each card, to tell whether it received an extra spot or not. The original distribution of spots on the cards, which is known to the subject, is as follows:

---

[16]In particular, using the margin model of Figure 1, we could impose that for all worlds $w \geq 1$, worlds within the unit margin from w are equiprobable, with probability 1/3. This would make the prediction that $2 \models \Box_0 \neg p$, $3 \models \Box_{1/3} \neg p$, and $4 \models \Box_{2/3} \neg p$, $5 \models \Box_1 \neg p$. This prediction would already give a more realistic description of the situation. However, the prediction may still appear unrealistic, considering that in general, the certainty of the subject does not vary linearly with the intensity of the stimulus.

Number of spots on card	Number of cards in pack
1	50
2	100
3	150
4	100
5	50

The distribution of extra spots is also known to the subject, namely the experimenter uniformly added one extra spot to half of the cards in each group (so 25 cards that had only one spot now have two spots, 50 cards that had two spots now have three, and so on).

In this scenario, we may call *signal* the condition s in which a card received one extra spot, and *noise* the condition n in which it did not (we could as well make the opposite choice). For each number x of spots that the subject observes on a given card, one can compute the prior conditional probabilities $P(x|s)$ and $P(x|n)$ to observe x, given that the card received an extra spot or not. The two distributions are given on the left hand side of the table below, and are plotted in the figure below it. Based on that, one can compute the posterior probability $P(s|x)$ of signal (resp. noise) of having signal (or noise) given an observed stimulus value, as represented on the right hand side of the table.

x	$P(x|n)$	$P(x|s)$	$P(s|x)$	$P(n|x) = 1 - P(s|x)$
1	25/225	0	0	1
2	50/225	25/225	1/3	2/3
3	75/225	50/225	2/5	3/5
4	50/225	75/225	3/5	2/5
5	25/225	50/225	2/3	1/3
6	0	25/225	1	0

As can be seen from the table and the plot, only when there is 1 spot on a card can the subject declare with certainty that no spot was added. Conversely, only when there are 6 spots on a card can the subject declare with certainty that a spot was added. Between these two values of the stimulus, however, namely for those values for which the two distributions of noise and signal overlap, the subject necessarily experiences uncertainty. In those states, the subject simply cannot tell, just from the observation that a card has 2, 3, 4 or 5 spots, whether a spot was added to the card or not. Nevertheless, if the subject were challenged to sort out cards *as best as she can*, given her uncertainty, then her best strategy (or optimal *decision criterion*) would be to declare that there is signal whenever the number of spots is at 4 or more, and to declare that there is noise whenever the number of spots is less than 4. Indeed, as can be checked from the right hand
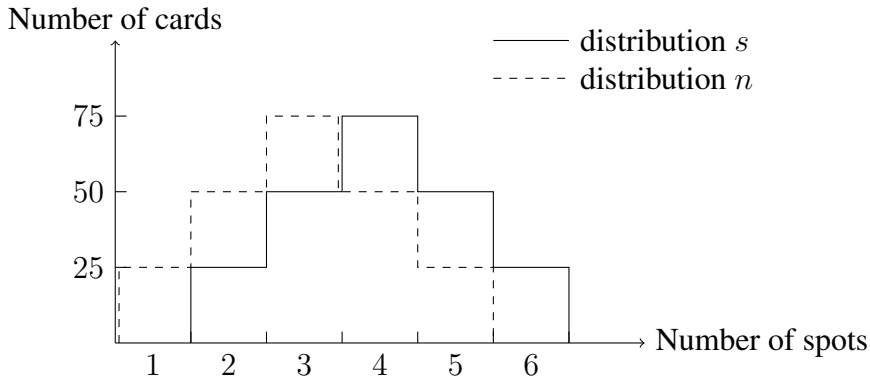
Number of cards

Figure 5: Signal and noise distributions

side of the table above, the conditional probability of signal $P(s|x)$ exceeds a half when $x \geq 4$, and conversely, that of noise $P(n|x)$ exceeds a half when $x < 4$. It is easy to check that by choosing such a strategy, the subject would then be able to correctly sort out two thirds of the cards, thereby doing better than chance, and better than by sorting out only the cards for which she is absolutely certain (namely those with 1 or 6 spots).

As we mentioned previously, the type of probabilistic decision taken in the uncertainty interval by the subject is not reflected in the frameworks we presented in the previous section.[17] Despite this, the uncertainty of the subject can readily be pictured using a Kripke structure of the kind we are familiar with. In particular, we can easily represent the subject's epistemic situation using Halpern's two-dimensional framework. Thus, in Figure 6, the vertical axis represents the stimulus value, namely the number of spots the subject sees on a given card. The horizontal axis, on the other hand, represents the number there was on the card before a spot was added. From the perspective of the subject, the O-axis encodes the potential original number of spots that there was on the card. For instance, (2,3) is a state in which the subject sees a card with three spots and in which the card received an extra spot. We label n the noise states, namely the states in which the observed number of cards does not differ from the original number, and s the signal states, those for which the observed number and the original number differ.

As reflected in the figure, states of subjective uncertainty are those whose stimulus value or subjective index is between 2 and 5. On the other hand, states with subjective index 1 and 6 are states in which signal and noise no longer overlap, and in which the subject can report the original number of spots with certainty and accuracy. In particular, it follows from Halpern's semantics that $(1,1) \models Rn$, and

---

[17]As explained in section 3, Halpern (2004) incorporates a notion of plausible worlds in his model. However, he does not incorporate probabilities explicitly.
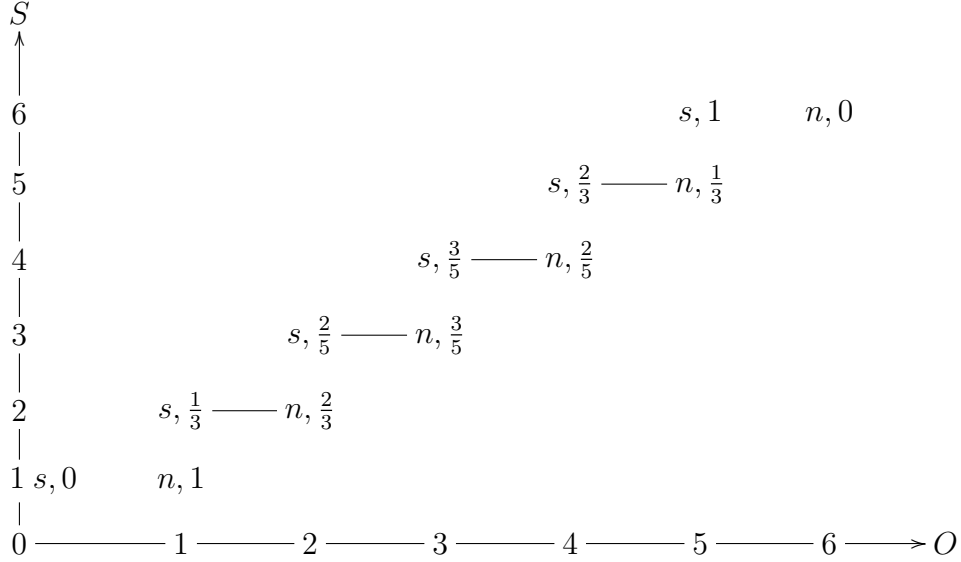
Figure 6: Relative uncertainties of signal and noise

$(5, 6) \models Rs$.

Figure 6 differs from Figure 2 in several respects. First, we have indicated next to the states in each subjective equivalence class the posterior probability of that state, as it results from the table. The probability of $(2, 3)$, for instance, is the probability that there is a 2 given that a 3 was observed (namely $P(s|3)$), and similarly the probability of $(3, 3)$ is the probability that there is a 3 given that a 3 was observed (namely $P(n|3)$). As we are about to see, this representation of relative uncertainty can be generalized, based on the likelihood models introduced earlier.

Secondly, we can see that the model here cannot be seen as the layering of a simple linear model of the type presented in Figure 1. The reason for that is that states with the same objective index do not necessarily convey the same atomic information (witness (1,1) which is n and (1,2) which is s). A related difference is that we followed Halpern's convention of making objectively similar only states with the same objective index that carry the same atomic information. As a consequence, each objective equivalence class is reduced to a singleton, that is no dotted line relates two distinct states. In this case, it follows from Halpern's semantics that $(1, 1) \models DRn, DRDRn$, and so on. Likewise, $(6, 6) \models DRs, DRDRs$ and so forth. By contrast, for $1 < x < 6$, $(x, x) \models DR(\neg DRn \wedge \neg DRs)$. In the terms of Halpern's DR operator for "clearly", and quite plausibly with regard to the scenario, there is either clearly clearly noise, clearly clearly signal, or clearly neither clearly noise nor clearly signal. No higher-order uncertainty arises in that situation.

26

4.4   Degrees of clarity and higher-order vagueness

The interest of presenting an abstract signal detection task like the one put forward by McNicol is that it encapsulates the essential features of more realistic psychophysical tasks. At the same time, it displays some significant analogies with other discrimination tasks like the pen-and-box scenario we started with. Thus, the card sorting task is one in which we go from *clear noise* to *clear signal*. Upon observing 1 spot on a card, the subject can discriminate the card clearly as having received no extra spot. Likewise, upon observing 6 spots, the subject can discriminate the card clearly as having received an extra spot. In between lies a penumbral area, in which the probability of signal gradually increases, with the probability of noise gradually decreasing, but in which neither noise nor signal comes out clearly, as represented explicitly on Figure 6.

Similarly, the pen-and-box task is one in which the subject is supposed to see that a pen of size 1 clearly fits in the box, and that a pen of size 6 does not, while being uncertain about intermediate sizes like 3 and 4. As emphasized previously, however, the transition from "the pen clearly fits in the box" to "the pen clearly does not fit in the box" is more gradual than actually displayed in fixed margin models. To make the parallel completely clear, we need to tighten up the link between the SDT model and margin models, putting together the different elements we introduced.

Let us consider again the pen-and-box scenario, and let us call *signal* the property $\neg p$ of not fitting in the box, and *noise* the property p of fitting in the box. To give the parallel substance, we may imagine that for each value x of the evidence variable (namely the observed size of the pen), the posterior probabilities of signal and noise are distributed exactly as in McNicol's scenario. A question that immediately arises is: where could such probability assignments come from? In the McNicol's card sorting problem, the answer is simple: the subject is supposed to found her prior subjective probabilities of signal and noise over the objective proportions of signal and noise in each card group, from which posterior probabilities of signal and noise are derivable. In actual detection tasks, however, distributions of signal and noise cannot be directly observed, and are obviously not transparent to the subject.

To derive posterior conditional probabilities of signal and noise in the present setting, however, we may appeal to the likelihood structures previously introduced. For each value of the stimulus along the dimension of observed sizes (axis S on Figure 7), we can suppose that there is a probability distribution representing error, namely the propensity for the observed stimulus to encode a potentially distinct actual value (represented on axis O of the figure). In Figure 7, we thus present another Halpern style structure in which each subjective equivalence class (namely each set of points connected by a straight line) receives a certain proba-
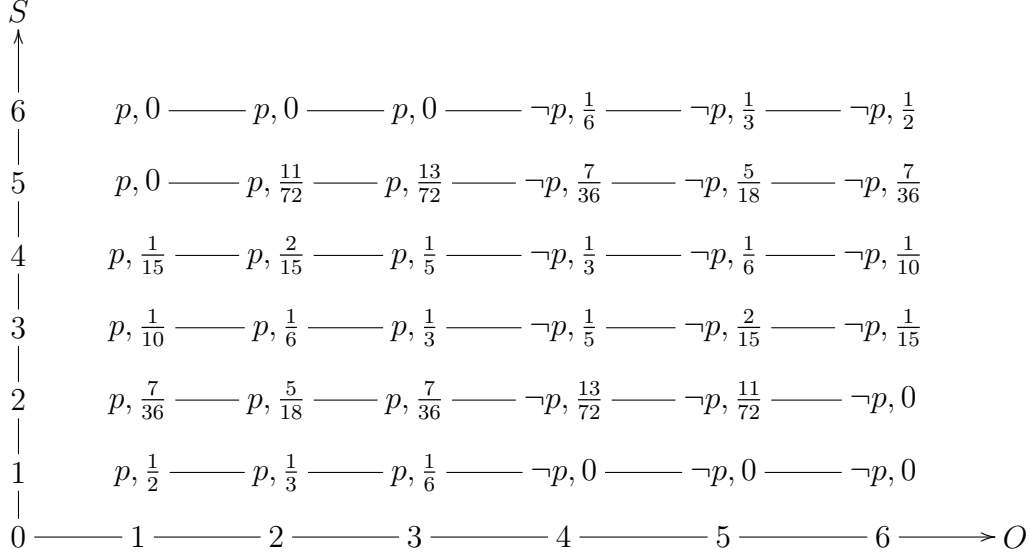
S

6     $p,0$ —— $p,0$ —— $p,0$ —— $\neg p,\frac{1}{6}$ —— $\neg p,\frac{1}{3}$ —— $\neg p,\frac{1}{2}$

5     $p,0$ —— $p,\frac{11}{72}$ —— $p,\frac{13}{72}$ —— $\neg p,\frac{7}{36}$ —— $\neg p,\frac{5}{18}$ —— $\neg p,\frac{7}{36}$

4     $p,\frac{1}{15}$ —— $p,\frac{2}{15}$ —— $p,\frac{1}{5}$ —— $\neg p,\frac{1}{3}$ —— $\neg p,\frac{1}{6}$ —— $\neg p,\frac{1}{10}$

3     $p,\frac{1}{10}$ —— $p,\frac{1}{6}$ —— $p,\frac{1}{3}$ —— $\neg p,\frac{1}{5}$ —— $\neg p,\frac{2}{15}$ —— $\neg p,\frac{1}{15}$

2     $p,\frac{7}{36}$ —— $p,\frac{5}{18}$ —— $p,\frac{7}{36}$ —— $\neg p,\frac{13}{72}$ —— $\neg p,\frac{11}{72}$ —— $\neg p,0$

1     $p,\frac{1}{2}$ —— $p,\frac{1}{3}$ —— $p,\frac{1}{6}$ —— $\neg p,0$ —— $\neg p,0$ —— $\neg p,0$

0 —— 1 —— 2 —— 3 —— 4 —— 5 —— 6 —→ O

Figure 7: Layered Likelihood model

bility distribution. For instance, on line 3 of the array, the value 1/6 for the state $(2,3)$ represents the probability $P(2|3)$ that when the observed stimulus value is 3, it is perceived as if it were a 2. As the reader will recognize, moreover, the subjective equivalence class for line 3 can be seen as the layering of the likelihood model presented above in Figure 4.[18] Because the layering of the initial reflexive model is necessarily reflexive, symmetric, and transitive, one can label states instead of arrows with their respective weight (letting $P((y,x)) = P_x(y)$).

Letting $P(p|x) = \sum_{(y,x)\sim_s(x,x),(y,x)\models p} P((y,x))$, one can check from the diagram that the conditional probability of noise and signal relative to each subjective value comes out exactly as specified in McNicol's scenario. Note that we assigned weights to states in an ad hoc manner, so as to fit the conditional probability of signal (resp. noise) resulting, in McNicol's example, from the noise and signal distributions. Of course, other assignments would equally work. Something we lose here is the representation of the noise and signal prior distributions $P(x|p)$ and $P(x|\neg p)$. However, it can be seen that we made the assignment exactly symmetric between noise and signal as we go from 1 to 6. Moreover, the probability $P((x,x))$ of correctly identifying the stimulus can be seen to decrease as the value of x gets nearer to the boundary between signal and noise.

As in a margin model, therefore, the present model is one in which the

---

[18]For the sake of readability, we intentionally dropped the connecting edge and labelling for states of the form $(0,x)$, which would otherwise be $p,0$.

boundary between signal and noise is precise, but not known to the subject (here the cut-off between p and $\neg p$ is between 3 and 4). The difference is that we can now assign degrees of clarity to propositions. Let us introduce again a probabilistic operator of relative clarity $\square_\alpha$, such that $(y,x) \models \square_\alpha \phi$ iff $\sum_{(y',x')\sim_s(y,x),(y',x')\models\phi} P((y',x')) = \alpha$. In the model of Figure 7, $(3,3) \models \square_{2/3}p$. Moreover, $(3,3) \models \square_1\square_{2/3}p$, since all states $(x,3)$ in the same equivalence class satisfy $\square_{2/3}p$. That is, when the observed size of the pen is 3, it is clear to degree 2/3 to the subject that the pen fits in the box, and moreover it is fully clear that it clear to degree 2/3 that the pen fits in the box. Note that this prediction differs from the one we would get with the one-dimensional semantics introduced in section 4.2 over the structure in Figure 4. In that case, $3 \models \square_{3/5}p$, but since 3 is the only accessible world from 3 satisfying $\square_{3/5}p$, with a weight of 1/3, $3 \models \square_{1/3}\square_{3/5}p$, that is, it is only a third clear that it is clear to degree 3/5 that the pen fits in the box.

Not surprisingly, therefore, the switch from a one-dimensional semantics to a two-dimensional semantics occasions the same distinctions for embeddings of probabilistic operators that we had obtained for categorical operators. And similarly, therefore, we encounter the same difference between a Williamsonian view of higher-order clarity gradually fading off, and the view we exposed earlier with centered semantics, in which clarity automatically occurs at the second order.

Note that both predictions may appear overly precise. Indeed, it may sound implausible altogether to assign precise numerical degrees of clarity to express a notion of subjective clarity. For as explained above, the various probabilities of error for any given stimulus need not be conscious to the subject. One way to go around this problem is to use large orderings in the semantics of our operator of clarity, and to assume that degrees can be mapped to a more qualitative description of the degree of clarity. Thus, let us replace the operator $\square_\alpha$ by the operator $\square_{\geq\alpha}$, to mean that "it is clear to degree at least α", assigning identical truth-conditions. Relative to a one-dimensional likelihood model, let $w \models \square_{\geq\alpha}\phi$ iff $\sum_{wRw',w'\models\phi} P_w(w') \geq \alpha$. For a layered likelihood model, let $(y,x) \models \square_{\geq\alpha}\phi$ iff $\sum_{(y',x')\sim_s(y,x),(y',x')\models\phi} P((y',x')) \geq \alpha$. We shall write $\square_{<\alpha}p$ to abbreviate $\neg\square_{\geq\alpha}p$, to mean "it is clear to degree less than α that p".

Under those assumptions, if it is clear to degree at least α that ϕ, then it will also be clear to degree at least α' that ϕ, for $\alpha' \leq \alpha$. This is a welcome fact, since it entails that if something is clear to degree at least 2/3, for instance, it is also clear to at least any lesser degree. Correlatively, we obtain different predictions for embedded operators with regard to the likelihood model of Figure 4. Let us assume, in particular, that the model of Figure 4 is completely decorated, in agreement with the layered model of Figure 7. In that case, since the probability of p relative to 1 and 2 is at least 3/5, as in 3, it will follow that $3 \models \square_{\geq3/5}\square_{\geq3/5}p$.

In a more qualitative language, this predicts that if at 3 it is is *fairly clear* that p, then it is also fairly clear that it is fairly clear that p. For that particular case, clarity at the second order has the same degree as that of the first order. On the other hand, it will still be the case that $1 \models \Box_{\geq 1/2}\Box_{\geq 1}p$, so that even under the large interpretation of degrees, clarity is likely to fade off at higher-orders.

In the two-dimensional case, on the other hand, the prediction remains that $(3,3) \models \Box_{\geq 1}\Box_{\geq 3/5}p$, and more generally, that $(x,x) \models \Box_{\geq 1}\Box_{\geq \alpha}p$ whenever $(x,x) \models \Box_{\geq \alpha}p$. Likewise, when $(x,x) \models \Box_{<\alpha}p$, $(x,x) \models \Box_{\geq 1}\Box_{<\alpha}p$. As a consequence, it seems we are forced to deny higher-order vagueness. But this seems too strong a prediction. Indeed, let us consider again the state $(3,3)$, in which a pen of size 3 is observed. In that state it is predicted that $\Box_{\geq 3/5}p$. From the diagram, however, there is a probability of 1/5 to treat 3 as if it was a 4. But if a 4 was observed, then it should be the case that $\Box_{<3/5}p$. In other words, upon seeing a 3, the subject should not be entirely confident that it is fairly clear that p, contrary to what we argued above.

4.5 Clear clarity, unclear unclarity

Which semantics, therefore, should we select as ultimately most plausible? In our view, the correct view about higher-order vagueness may well lie half-way between these two models. On the one hand, we find it plausible that if a proposition is fully clear, then it is also fully clear that it is fully clear. On the other hand, we consider indeed that when a proposition is less than fully clear, it will typically be the case that it is not fully clear that this proposition is not fully clear.

To put it in a slogan, the view of higher-order vagueness we have in mind is one in which clarity is clear, and unclarity unclear. There are several ways of articulating this view, however. One option would be to come up with a natural semantics for clarity that satisfies positive introspection, but not negative introspection (typically, a logic of the type S4, but not S5 for "clearly"). On that option, some propositions that are not clear might still be clearly not clear in specific cases, but this connection would no longer be automatic. None of the semantics we discussed so far shows such an asymmetry, however, and we see the definition of such a semantics as an open problem. A different approach to the asymmetry between clarity and unclarity might then be the following. Typically, we say that a proposition p is *unclear* not only if it is not clear that p (which may be compatible with its being clear that $\neg p$), but moreover if it is not clear that $\neg p$. Another way of articulating the view that clarity and unclarity pattern non-symmetrically, therefore, would be to require that propositions that have this status of unclarity (clarity neither way) are propositions that never become clearly unclear at higher orders.

Prima facie, such a view may seem too strong. For certainly, we can judge

some cases to be "pretty clearly unclear".[19] Nevertheless, "pretty clearly unclear" does not mean the same thing as "fully clearly unclear". The intuition we want to promote is that if indeed there is an asymmetry between clarity and unclarity, it concerns the notion of full clarity and of less than full clarity. Let us imagine a series of colors ranging from a paradigm case of blue to a paradigm case of green. Certainly, some cases along the transition are neither clearly blue, nor clearly green. But if really the idea of higher-order vagueness makes sense, none of these cases is such that it is *fully clearly* less than fully clearly blue or green. The intuition we have is that while we can know where full clarity begins, we do not necessarily know where it ends. In particular, there may be a first stimulus value for which things become less than fully clear. But this first stimulus value, if indeed it can be postulated, is such that it cannot be fully clearly unclear. The sort of asymmetry we are looking for, therefore, will not express itself directly in terms of the axiomatic properties of a logic, but rather, should be linked to the kinds of models of discrimination we have exposed so far.

Before going to the formal details, we may add that one motivation to entertain this asymmetrical view between clarity and unclarity is based on the consideration of actual experimental results in psychophysics, independently of the issue of sorites series. A number of psychophysical experiments show that for some values of the stimulus that are sufficiently remote from the objective boundary between target categories, answers will be 100% correct on average. For instance, in the target version of the Sparse/Dense task of Smith & al., subjects can also use a third response, materialized by selecting a star on the screen, to decline the choice between "Sparse" and "Dense". This third response is called the "Uncertain" response by Smith & al., as its use matches the increase in uncertainty felt by human subjects (see [14]: 321). What can be observed is that when the number of pixels illuminated on the screen is less than 1650 (the actual threshold for "Dense" being at 2950 pixels), the number of "Sparse" responses to the task by subjects (humans and monkeys alike) is invariably of 100% (see [14], p. 322, Figure 3). For those values, moreover, subjects never choose the response "Uncertain". This means than upon repeated trial, subjects show the same confidence in their correct categorization for such low values. There is no sense, in our view, in which the answer may not be clearly clear to the subject for that range of values.

However, as the number of pixels is gradually increased, there is typically a critical value of the stimulus at which responses start to falter, and for which the subject begins to feel uncertainty (e.g. around 2000 pixels in Smith & al. plots). Interestingly, for some isolated values that are only slightly further toward the boundary for "Dense", the subject will occasionally categorize again with no error (e.g. at 2100 pixels in [14]). But from a further critical point onward, uncertainty

---

[19]We are indebted to A. Everett for this remark.

significantly increases, until reaching a plateau (at around 2650 pixels), before the "Dense" response constantly increases in turn. These experimental data suggest the existence of a fluctuation zone between the range of values for which subjects would use only the "Sparse" response, and those where the "Uncertain" response becomes prevalent. One way to conceive of this fluctuation would be to say that the first moment where subjects experience unclarity and respond "Uncertain" cannot be "clearly unclear" for that matter. In other words, the first occurrence of less than full clarity is not itself fully clear.

By parity, however, the last stimulus for which subjects correctly respond "Sparse" on all trials may likewise be considered "not clearly clear", contrary to the intuition of an asymmetry between clarity and unclarity that we are seeking. In keeping with our intuition, however, we may take a more dynamic perspective and consider as "fully clear" all and only those moments that are consecutively reported as "Sparse" with no error. While still speculative, such a perspective allows us to restore the intuition of an asymmetry between full clarity and less than full clarity. In the experiments by Smith & al., indeed, a typical feature of the uncertainty response is that unlike the "Sparse" response, the uncertainty response never reaches 100%, but culminates at about 70% (see monkey Abel's curve in [14]: 322, Figure 3.C).[20] Thus, everything happens as if the subject was fully certain to be certain when she is certain. But when the subject fails to be certain, on the other hand, the subject never seems to be fully uncertain (in particular because she may well have used a better answer). In other words, everything happens as if unclarity about the correct way to categorize the stimulus was in turn unclear, unlike clarity.

The foregoing considerations remain speculative, of course, but to some extent they sustain our intuition that clarity and unclarity might pattern asymmetrically at higher orders. How then are we going to implement this asymmetry between clarity and unclarity in our framework? Halpern's solution to the higher-order vagueness predicament, we remember, was to define "clearly p" in terms of the "definitely" operator D, namely as $D\square p$ (transposed to the present language). The same may be done here for the operator $\square_{\geq \alpha}$, using $D\square_{\geq \alpha}$ to mean "it is fully clear with degree at least α". In particular, we shall define the notion of full subjective clarity by means of the operator $D\square_{\geq 1}$ rather than $\square_{\geq 1}$, since as explained above, $\square_{\geq \alpha} p$ entails $\square_{\geq 1} \square_{\geq \alpha}$ in the two-dimensional framework underlying Figure 7, thereby blocking the path to higher-order vagueness.

Let us call p *fully clear* if $D\square_{\geq 1} p$. Now consider again the model of Figure 7. On that model, $(1,1) \models D\square_{\geq 1} p$, and likewise, $(6,6) \models D\square_{\geq 1} \neg p$. That is, 1 and

---

[20]The contrast is less pronounced with the "Dense" response, but presumably because Smith & al. match "Dense" with an exact number of pixels (namely 2950), whereas "Sparse" covers all pixel configurations sparser than this number.

6 are stimulus values for which p and $\neg p$ are fully clear. For values x between 2 and 5, on the other hand, $(x,x) \models \neg D\square_{\geq 1} p$, and $(x,x) \models \neg D\square_{\geq 1} \neg p$, namely for these values both p and $\neg p$ are less than fully clear.

Let us now define an appropriate similarity relation $\sim_o$ for the operator D, suitable to do justice to the asymmetry we have outlined between full clarity and less than full clarity. Let every couple $(x,1)$ be $\sim_o$-related only to itself. Likewise, let every couple of the form $(x,6)$ be $\sim_o$-related only to itself. Under those assumptions, it will follow that all couples of the form $(x,1)$ satisfy $D\square_{\geq 1} p$, $D\square_{\geq 1} D\square_{\geq 1} p$, and so forth. Similarly, all couples of the form $(x,6)$ satisfy $D\square_{\geq 1} \neg p$, $D\square_{\geq 1} D\square_{\geq 1} \neg p$, and so forth. In words, it means that 1 and 6 are evidential values of the stimulus at which full clarity iterates. For such values, full clarity comes with no first- or higher-order vagueness.

Now, let us suppose that for $2 \leq y \leq 5$, every state of the form (x,y) is $\sim_o$-related to every other state with the same index x. This implies, in particular, that the relation $\sim_o$ that we defined is no longer an equivalence relation. For instance, $(1,2) \sim_o (1,6)$, but as we defined it $(1,6) \not\sim_o (1,2)$. Under those assumptions, it holds that every state of the form (x,y) for $2 \leq y \leq 5$ satisfies $\neg D\square_{\geq 1} p$, and likewise $\neg D\square_{\geq 1} \neg p$. The same holds at higher levels, namely $\neg D\square_{\geq 1} \neg D\square_{\geq 1} p$ and $\neg D\square_{\geq 1} \neg D\square_{\geq 1} \neg p$, and so on. This is due, basically, to the fact that states of the form $(x,6)$ and $(x,1)$ are $\sim_o$-accessible from every world with the same objective coordinate, but only have access to themselves. As a consequence, for worlds with subjective coordinate between 2 and 5, it is not fully clear that p, not fully clear that it not fully clear that p, and so on and so forth, and similarly for $\neg p$. Hence unclarity about p and $\neg p$ propagates to higher levels.

With enough assumptions on the model of Figure 7, and using Halpern's machinery, we see that it is possible to formalize the idea that full clarity can be fully clear, and less than full clarity less than fully clear. This view of higher-order vagueness differs from the views we examined earlier, since in those clarity and unclarity either both become clear at some level (as in our family of token semantics), or both are never clearly clear in principle (as in Williamson's account of higher-order vagueness).

We must nevertheless emphasize that we have not provided a proper logic in which to derive these asymmetries in terms of logical validities. Our effort has been only to build a model consistent with these desiderata, and plausible enough from a psychological point of view.

5 Conclusion

In this paper we examined several accounts of vagueness understood as a form of epistemic uncertainty, that is, essentially, as deriving from our limited powers

of discrimination. We have proposed different ways of making sense of the idea of limited higher-order vagueness, both in our account of the notion of categorical clarity in sections 2 and 3, and in the account of relative clarity outlined in the previous section. Finally, we have seen how margin for error models can be generalized into probabilistic models of error, bearing a tighter connection to the signal detection model used in psychophysics.

The main idea we have tried to defend here is that the phenomenon of higher-order vagueness should not necessarily threaten the full clarity of some properties relative to some contexts. On the view we advocate, propositions that are the object of inexact knowledge can still be clear and clearly clear all the way down without impugning on the idea of higher-order vagueness, so long as unclarity itself falls short of full clarity at higher orders. At the same time, we have tried to show that the introduction of degrees of clarity can afford a more realistic picture of the transition from clarity to unclarity, and back again.

A few more words should be added regarding the probabilistic treatment of vagueness and clarity outlined here. By putting emphasis on signal detection theory, we have worked under the assumption that unclarity and degrees thereof are wholly dependent on the subject's limited powers of discrimination. Correlatively, we have used as working hypothesis the assumption that concepts have sharp boundaries in our scenarios, just as signal and noise in psychophysical tasks. Nevertheless, we have not committed ourselves on the problem of the origin of vagueness more generally. In particular, the present contribution remains entirely neutral regarding the question of whether or not all vagueness, including linguistic vagueness, is adequately described as epistemic.

Independently of the issue we addressed here, however, we may imagine to draw inspiration from signal detection theory in support of a different, non-epistemic theory of vagueness. Let us imagine for instance that "red" is a concept with no absolutely precise boundaries, such that when a subject judges whether something is red or not, she does it according to prior subjective distributions analogous to those of signal and noise, based on various objective cues, and on some implicit decision criterion used to categorize. A picture of vague predication along those lines might be able to explain why, where the two prior distributions overlap, the subject can in principle go one way or the other, without there necessarily being a fact of the matter as to whether the object seen is intrinsically red or not. The examination of the implications of such a view would obviously be the topic for a different paper. Pending further work in this direction, at least we hope to have established a fruitful connection between logical models and psychological models of imperfect discrimination.[21]

References

[1] BONNAY D. & P. ÉGRÉ (2009), Inexact Knowledge with Introspection, the *Journal of Philosophical Logic*, 38, 179-227.

[2] BONNAY D. & P. ÉGRÉ (2008), Margins for Error in Context, in M. Garcia-Carpintero & M. Kölbel (eds.), *Relative Truth*, Oxford University Press, pp. 103-127.

[3] FARA GRAFF D. (2001), Phenomenal Continua and the Sorites, *Mind*, vol. 110(440) pp. 905-935. Originally published under the name "Delia Graff".

[4] FARA GRAFF D. (2002), An Anti-Epistemicist Consequence of Margin for Error Semantics for Knowledge, *Philosophy and Phenomenological Research*, 64, 127-142. Originally published under the name "Delia Graff".

[5] HALPERN J. (2004), Intransitivity and Vagueness, *Proceedings of the Ninth International Conference on Principles of Knowledge Representation and Reasoning (KR 2004)*, pp. 121-129.

[6] HORGAN T. (1994), Robust Vagueness and the Forced-March Sorites Paradox, *Philosophical Perspectives*, 8: *Logic and Language*, 159-88.

[7] KAMP H. (1971), Formal Properties of "Now", *Theoria* 37, 227-273.

[8] MACFARLANE J. (2008), Fuzzy Epistemicism, manuscript (as of 09/10/2008), in R. Dietz & S. Moruzzi (eds), forthcoming Oxford volume.

[9] MACMILLAN N.A. & C.D. CREELMAN (1991), *Detection Theory: a User's Guide*, Cambridge University Press.

[10] MCNICOL D. (2005), *A Primer of Signal Detection Theory*, Lawrence Erlbaum Associates, Publishers, London, originally published 1972.

[11] RABINOWICZ W. & K. SEGERBERG (1994), Actual Truth, Possible Knowledge, *Topoi* 13, 101-15.

[12] RAFFMAN D. (2000), Is Perceptual Indiscriminability Nontransitive?, *Philosophical Topics*, 28, 1, *Vagueness*.

[13] SCHIFFER S. (2003), *The Things We Mean*, Oxford University Press: Oxford.

[14] SMITH D., W.E. SHIELDS & D.A. WASHBURN (2003), The comparative psychology of uncertainty monitoring and metacognition, *Behavioral and Brain Sciences*, 26, 317-373.

[15] WILLIAMSON T. (1992), Inexact Knowledge, *Mind*, 101, pp. 217-42.

[16] WILLIAMSON T. (1994), *Vagueness*, Routledge.

[17] WILLIAMSON T. (1999), On the Structure of Higher-Order Vagueness, *Mind*, vol. 108, 127-43.

[18] WILLIAMSON T. (2000), *Knowledge and its Limits*, Oxford University Press.

[19] WILLIAMSON T. (2007), Improbable Knowing, manuscript (online notes, as of 28/06/2007).

Appendix: Token Semantics

Token Semantics is a generalization of Centered Semantics by which only worlds that are n steps away from the actual world can be visited for a given n. In Token Semantics, the evaluation of a formula is parameterized by a number n of tokens: each non-reflexive move along the accessibility relation costs one token, and once all tokens have been spent, no worlds that are essentially further away can be accessed.

Formally, satisfaction is defined with respect to a sequence of worlds and a number of tokens: q is short for an arbitrary sequence of worlds, qw for an arbitrary sequence augmented with w, and n is an arbitrary number of tokens. The definition is here repeated as Definition 4:

Definition 4. *Token satisfaction:*

(i) $\mathcal{M}, qw \vDash_{\text{TS}} p\ [n]$ iff $w \in V(p)$.

(ii) $\mathcal{M}, qw \vDash_{\text{TS}} \neg\phi\ [n]$ iff $\mathcal{M}, qw \nvDash_{\text{TS}} \phi\ [n]$.

(iii) $\mathcal{M}, qw \vDash_{\text{TS}} (\phi \wedge \psi)\ [n]$ iff $\mathcal{M}, qw \vDash_{\text{TS}} \phi\ [n]$ and $\mathcal{M}, qw \vDash_{\text{TS}} \psi\ [n]$.

(iv) $\mathcal{M}, qw \vDash_{\text{TS}} \Box\psi\ [n]$ iff
- $n \neq 0$ and for all w' such that wRw', $\mathcal{M}, qww' \vDash_{\text{TS}} \psi\ [n-k]$, where $k = 1$ if $w \neq w'$, and $k = 0$ if $w = w'$.
- Or $n = 0$ and $\mathcal{M}, q \vDash_{\text{TS}} \Box\psi\ [1]$.

Definition 5. *Let n be such that $1 \leq n \leq \omega$ (we assume $\omega - 1 = \omega$). TS(n) is the modal semantics defined by the following satisfaction relation: $\mathcal{M}, w \vDash_{\text{TS}(n)} \phi$ iff $\mathcal{M}, w \vDash_{\text{TS}} \phi\ [n]$.*

By definition, a formula is TS(n)-valid if it is TS(n)-true at every world of every model. CS is equivalent to TS(1), and standard Kriple semantics is equivalent to TS(ω).

We have seen that CS validates (4) and (5). Similarly, TS(n) validates the following schemata, which generalize (4) and (5) respectively:

(4.n') $(\neg p_1 \wedge \Diamond(p_1 \wedge \neg p_2 \wedge \Diamond(p_2 \wedge \ldots \wedge \Diamond(p_{n-1} \wedge \Diamond\Diamond r)\ldots))) \rightarrow \Diamond(p_1 \wedge \Diamond(p_2 \wedge \ldots \wedge \Diamond(p_{n-1} \wedge \Diamond r)\ldots))$

(5.n') $(\neg p_1 \wedge \Diamond(p_1 \wedge \neg p_2 \wedge \Diamond(p_2 \wedge \ldots \wedge \Diamond(p_{n-1} \wedge \Diamond r)\ldots))) \rightarrow \Diamond(p_1 \wedge \Diamond(p_2 \wedge \ldots \wedge \Diamond(p_{n-1} \wedge \Box\Diamond r)\ldots))$

where $\Diamond^n$ is short for $\Diamond\ldots\Diamond$, n times. In particular, the following schema (4.n) follows from (4.n'):

(4.n) $\Box^n p \rightarrow \Box^n \Box p$

In [1] we provide a complete axiomatization for the semantics TS(n) with respect to the logic $K(4.n')(5.n')$. With respect to standard Kripke semantics, the two schemata axiomatize the class of n'-transitive and n'-euclidian frames respectively, where an n'-transitive frame is a frame satisfying:

$\forall w, w', w''\ (\exists x_1, ..., x_n[x_1Rx_2 \wedge \ldots \wedge x_{n-1}Rx_n \wedge x_1 \neq x_2 \wedge \ldots \wedge x_{n-1} \neq x_n \wedge x_n = w] \rightarrow (wRw' \wedge w'Rw'' \rightarrow wRw''))$

and an n'-euclidian frame is a frame satisfying:

$\forall w, w', w''\ (\exists x_1, ..., x_n[x_1Rx_2 \wedge \ldots \wedge x_{n-1}Rx_n \wedge x_1 \neq x_2 \wedge \ldots \wedge x_{n-1} \neq x_n \wedge x_n = w] \rightarrow (wRw' \wedge wRw'' \rightarrow w'Rw''))$

PAUL ÉGRÉ
Institut Jean-Nicod (EHESS/ENS/CNRS)
Département d'Etudes Cognitives - 29, rue d'Ulm
75005 Paris - France
paulegre@gmail.com

DENIS BONNAY
Université Paris Ouest-Nanterre-La Défense
Ireph
& IHPST, Département d'Etudes Cognitives, ENS
denis.bonnay@ens.fr