**Ergo** AN OPEN ACCESS JOURNAL OF PHILOSOPHY

# Defining Definiteness

ALEKSANDER DOMOSŁAWSKI
*University of Leeds & Adam Mickiewicz University*

Epistemicism associates vagueness with ignorance produced by semantic plasticity: the shiftiness of intensions in our language resulting from small changes in usage. The recent literature (Caie 2012; Magidor 2018; Yli-Vakkuri 2016) points to a missing piece in the epistemicist theory of vagueness, namely a clear account of the semantics of the definiteness operator $\Delta$. The fundamentals of the epistemicist theory are well understood. However, the technical work of defining the definiteness operator has proven difficult. There are several desiderata that we would like $\Delta$ to satisfy. For instance, we would like the epistemicist notion of 'definiteness' to interact well with modal operators and validate intuitive principles like 'necessarily, if $\varphi$ is definitely true, then $\varphi$ is true'. Providing an account that would meet all such desiderata has eluded the epistemicists so far. In this paper, I present a novel version of a multidimensional model inspired by the work of Robert Stalnaker and David Kaplan. Using this model, I provide an account of epistemicist definiteness that meets our desiderata.

**Keywords:** two-dimensional semantics; vagueness; epistemicism; definiteness

## 1. Introduction

The aim of this paper is twofold. Firstly, to present a refined version of a Stalnakerian multidimensional model that would allow us to answer broader metasemantic questions. Secondly, to solve a technical problem in defining the definiteness operator for the epistemicist theory of vagueness.

Robert Stalnaker (1978) uses a multidimensional (2D) model to present his metasemantic theory and shows how the expressions that we use might have different content had the metasemantic context been different. He employs a simple metasemantic context and a simple circumstance of evaluation; his analysis is concerned with how the changes in the metasemantic context change

**Contact:** Aleksander Domosławski <aleksander.domoslawski@gmail.com>

https://doi.org/10.3998/ergo.4650 

the content of expressions. On the other hand, David Kaplan (1977/1989) uses a multidimensional model to account for the semantics of demonstratives and indexicals. I propose a multidimensional model, which squares the approaches of Kaplan and Stalnaker: a model that would allow us to handle indexicality (as Kaplan wants) and also answer metasemantic questions.

Furthermore, I make use of the model to solve an important problem for the epistemicist theory of vagueness (Williamson 1994). The recent literature (Caie 2012; Magidor 2018; Yli-Vakkuri 2016) presents difficulties for the epistemicist theory associated with defining the definiteness operator $\Delta$. Attempts explored in the literature often employ a multidimensional model to provide the semantics of $\Delta$; however, all these attempts face big obstacles. The multidimensional model I outline allows us to handle these problems and provide a clear semantics for the definiteness operator.

## 2. Multidimensional Models

Since I will be using multidimensional models to present an account of the definiteness operator, it's worth taking the time to explain how my model will fit with the multidimensional accounts developed in the literature. Using multidimensional models for this purpose is not a new idea: this is also the approach taken for instance by Caie (2012), Litland and Yli-Vakkuri (2016), and Yli-Vakkuri (2016). However, to fully understand what our model achieves and how it fits with the accounts developed in the literature, I will need to start by saying a bit more about some of the uses of multidimensional models present in the literature, in particular their use by David Kaplan and Robert Stalnaker.

The simple idea behind multidimensional models is that formulas in the language are not to be evaluated at a single world, but rather at more complex points of evaluation (e.g., a pair of worlds). This complication in the model allows us to account for some complexities in our language that we wouldn't otherwise be able to. The majority of literature on multidimensional models uses two-dimensional frameworks, which use double-indexed points of evaluation consisting of a context and circumstance of evaluation. Philosophers use multidimensional models for a variety of purposes. Two of the most important uses of multidimensional models come from the work of Kaplan and Stalnaker.

Kaplan (1977/1989) uses two-dimensional models to account for the semantics of demonstratives and indexicals. On his account, formulas are assigned characters—functions from worlds to contents (intensions). The first index, the context of utterance, determines the content of the expression given by the character function. For instance, the character of 'I' is a function that for any context assigns the speaker in that context as the referent of 'I'. Contents are then evaluated relative to the second index (the circumstance of evaluation). Kaplan's account uses the two-dimensional model to explain the semantics of a certain class of expressions: a character of an expression is a certain kind of meaning.

On the other hand, Stalnaker (1978) uses the 2D model to show us how the expressions that we use might have meant something different than they actually do had the external world been slightly different. On this understanding of the 2D model, the function from the first index to intensions is not a kind of meaning like it is for Kaplan—intensions are the only kind of content there is[1]—it merely shows what the content of our utterances might have been.[2] For instance, presumably the metasemantic rule for the term 'water' in our language is that it picks out a substance with chemical composition that is identical to the composition of the stuff that is found in rivers and lakes. In the actual world the chemical composition of the stuff found in rivers and lakes is $H_2O$. However, in some different possible world $w^*$, the stuff in rivers in lakes is some compound XYZ—our metasemantic rule for 'water' would dictate that 'water' refers to XYZ in that world.

One may wonder what I mean by a 'metasemantic rule'. Roughly, I mean a descriptive element of metasemantics, perhaps inscrutable to us, that describes how the facts about the world determine meaning. For instance, in the standard Twin Earth (Putnam 1975) setup, 'water' on Earth and on Twin Earth have different reference: on Earth it refers to $H_2O$ and on Twin Earth to XYZ. However, when thinking about the case, we apply the same metasemantic rule to the term 'water' on Earth and on 'Twin Earth'. Whatever substance is found in the rivers and lakes on each planet (and hence causally regulates the behaviour within linguistic communities at these planets) is the reference of the term 'water' on each planet. We don't have to treat this metasemantic rule as a part of the meaning of the term 'water'. Rather we can treat it as a part of metasemantics: a function that outlines how the external facts determine meaning.

Despite the fact that Stalnaker and Kaplan give different interpretations to the 2D model we don't have to treat their approaches as necessarily at odds: they simply employ the 2D model for different purposes. Stalnaker's interest is in his metasemantic project: he uses the 2D model to show how words in our language would have a different meaning if the external world had been different. On Stalnaker's model there is a simple metasemantic context, which determines the meanings of expressions in the language, and a simple circumstance of evaluation. Kaplan is interested in the analysis of indexicals. This means that

1. Stalnaker refines his approach to handle indexicals in his later work. However, I am concerned here only with his basic insight about the interpretation of the 2D model.

2. Another prominent interpretation of the 2D model is given by David Chalmers (1996; 2004) and Frank Jackson (1998; 2004), i.e., the epistemic interpretation of the 2D model. I will not be concerned with this kind of interpretation in this paper.

he keeps the meanings (characters) of expressions in the language fixed. From a Stalnakerian perspective, Kaplan does not look at the effects of change in the metasemantic context (he keeps it fixed), but instead introduces a complex multidimensional circumstance of evaluation: expressions in the language are first evaluated relative to context, which tells us what intension should be assigned to expressions in our language, and then relative to a circumstance of evaluation (which gives us the semantic values).

Therefore, we could devise a model that attempts to square the approaches of Stalnaker and Kaplan: a model that allows us to answer metasemantic questions and is also able to handle indexicality. One way of doing this would be to introduce a three-dimensional model, on which we have a simple metasemantic context like Stalnaker, but have a double-indexed point of evaluation. For each expression, the metasemantic context gives us its character, which allows us to evaluate the expressions relative to a context and circumstance of evaluation like on Kaplan's model. Therefore, on a three-dimensional model the function from the metasemantic context of utterance to characters is not a kind of meaning (analogously to the way that the 2D matrices don't represent a kind of meaning for Stalnaker). Such function merely tells us what character an expression has in different metasemantic contexts.

However, such a three-dimensional model will not be sufficient for my purposes. Just like substituting the unidimensional circumstance of evaluation for a double-indexed point is needed to handle some complex features of semantics like indexicality, we also need to substitute a simple metasemantic context for a double-indexed point to handle complexities in the metasemantics. Stalnaker's model is used to show how different metasemantic contexts would produce different meanings in our language. Stalnaker is concerned with the way that differences in the external world may affect the content of one's utterance. For instance, the metasemantic rules for assigning a semantic value for the term 'Hesperus' are such that 'Hesperus' is a rigid designator, referring to whichever celestial body is the brightest one in the evening. Actually, Venus is the brightest celestial body in the evening, so 'Hesperus' rigidly designates Venus. However, if there is a possible world $w^*$ where the brightest celestial body in the evening is Mars then our metasemantic rules for assigning the content to 'Hesperus' would assign Mars as the reference of 'Hesperus'. So in sum, Stalnaker is concerned with exploring the way in which, by applying our metasemantic rules, the expressions in our language would get assigned different contents had the external world turned out differently.

My aim here is different. Whereas Stalnaker wants to keep the metasemantic rules (such as the rule that 'Hesperus' refers to the brightest celestial body in the evening) fixed, I am also interested in investigating cases where the metasemantic rules are different. For instance, we could be interested in investigating how changing the metasemantic rules while keeping the external world fixed would affect the contents of expressions in our language. On this approach, the question regarding 'Hesperus' would be: had our metasemantic rules been different, what would be the content assigned to 'Hesperus'? For instance, if in a world $w^*$ the metasemantic rules regarding 'Hesperus' are such that it is a rigid designator that refers to the Moon, then by the metasemantic rules for 'Hesperus' refers to the Moon in the actual world.

For this reason, we need a double-indexed metasemantic context. The first index gives us the metasemantic rules that govern the expressions in the language (such as the rule that 'Hesperus' refers to the brightest celestial body in the evening): it gives us the function from worlds to characters. The second index gives us what we may dub the *metasemantic circumstance of evaluation*: it provides an input to the function determined by the first index. For instance, if we apply the actual metasemantic rule for 'Hesperus' to a world $w^*$ where Mars is the brightest celestial body, we get a constant character picking out Mars.

My aim in what follows will be to outline the epistemicist theory of vagueness and present a multidimensional model that would allow us to define a definiteness operator for epistemicism. In Section 3, I outline the epistemicist theory of vagueness and I present my understanding of epistemicist definiteness. In Section 4, I present the desiderata for defining definiteness. In Section 5, I present a multidimensional model where I define definiteness and show that this account can deal with some problematic cases.

## 3. Safety and Definiteness

The epistemicist project uses safety-based epistemology to explain the ignorance that arises in borderline cases. On the safety conception, knowledge is incompatible with a possibility of an easy error. The epistemicist identifies a feature of vague expressions in our language, namely *semantic plasticity*, and argues that semantic plasticity gives rise to a possibility of easy errors in borderline cases. An expression is semantically plastic if and only if small shifts in the use of language across close possible worlds unbeknownst to us slightly shift the intension of the expression. It's easy to see how semantic plasticity could give rise to ignorance in borderline cases. If Michael is a borderline case of 'tallness', then even if Michael is *tall*, he is not *tall** (where *tallness** is a slightly different property from *tallness* such that we could easily refer to it had we used the language slightly differently). This means that if I express my belief that Michael is tall using the sentence 'Michael is tall', there is a possibility for me to make an easy error: if we used the term 'tall' so that it refers to *tallness** (and it could easily be the case that we used the language that way), 'Michael is tall' would say something false.

What is definiteness on the epistemicist account? If ignorance in borderline cases is associated with the easy possibility of error due to semantic plasticity, then definite cases are ones where semantic plasticity does not pose such a threat to safety. This does not mean that definite cases are ones that are *overall* safe, it only means that they are safe relative to semantic plasticity (as an ignorance inducing factor). This is to be expected; after all, 'definitely p' does not entail that we are in a position to know p, just that vagueness alone cannot prevent us from being in a position to know p. Therefore, on the epistemic treatment of vagueness, it's definitely φ when the variation in semantic facts (facts having to do with what intensions are assigned to expressions in our language) is insufficient to produce an error across the space of nearby worlds.

There is a question about how important it is for the epistemicist to give an interpretation of the definiteness operator. On a certain reading of epistemicism, this could be seen as a superfluous addition. One of the important advantages of epistemicism is its parsimony. The theory employs (a) the independently established safety-based account of knowledge and (b) the postulate that expressions in our language are semantically plastic, which is more controversial but also independently motivated (Dorr & Hawthorne 2014). Using these two pieces of theoretical machinery, the epistemicist wants to explain, fundamentally, why there is ignorance in borderline cases. If the epistemicist is able to explain ignorance in borderline cases, why is there a need for additional category of definiteness? If a proposition is definitely true, it just means that it's safe relative to semantic plasticity. However, semantic plasticity is one of many ignorance inducing factors. I could be ignorant as to whether Michael is tall because 'Michael is tall' is semantically plastic, but I could also be ignorant because I cannot see Michael very well from a distance. It's not clear that the epistemicist should treat *safety relative to semantic plasticity* as a particularly important category as compared to safety relative to imperfect eyesight.

However, going forward I will assume that defining the definiteness operator is a worthwhile task for the epistemicist. There are certainly dialectical advantages to having such an interpretation, not the least of which is being able to paraphrase the discourse employed by other theories of vagueness (such as the supervaluationist or metaphysical accounts) for whom definiteness does constitute a special category. If we are able to provide a clear interpretation of the definiteness operator, there is no harm in doing so.

## 4. Desiderata for an Account of Definiteness

There are several desiderata that the epistemicist definiteness operator should meet. The accounts presented in the literature so far struggle to meet one or more of these criteria. My aim is to outline these desiderata and show that my account meets them. Then, in Section 6, I will show that alternative formulations of the epistemicist notion of definiteness fail to meet the desiderata.

### 4.1. Isolation of Semantic Plasticity

The epistemicist associates borderlineness with ignorance produced by semantic plasticity. Consider the following puzzle.

> *Definiteness Puzzle* (Magidor 2018). Suppose that Michael is 190 cm in height. The actual cut-off point for 'tall' is 185 cm in height and the margin for error is 3 cm. Since Michael's height is not within the margin for error, Michael counts as being definitely tall. Consider a world $w^*$ where Michael is 3 cm shorter and where the cut-off point for 'tall' is 188 cm. Such a world should count as close: Michael is only slightly shorter than in the actual world and the cut-off point is within the margin of error. However, Michael is not in the extension of 'tall' at $w^*$ as it is used at $w^*$. So Michael cannot be definitely tall in the actual world, because there is a close possible world where 'Michael is tall' is false at a close possible world; thus, a belief that Michael is tall is not safe and does not constitute knowledge.

It is clear that the falsity at $w^*$ of 'Michael is tall' as used at $w^*$ is due to both the variation in non-semantic facts (that have to do with Michael's height) and the semantic facts (that have to do with the boundary for 'tall') between the actual world and $w^*$. This is not the kind of variation that we are interested in. We are interested in knowing whether in a given case it is semantic plasticity alone could produce an easy error. The epistemicist account of definiteness should be formulated as to isolate the effect of semantic plasticity.

Since semantic plasticity is variation in semantic facts, what we would ideally like to do is to keep the non-semantic facts fixed and only allow for variation in the semantic facts. If it were possible, we would like to evaluate sentences at nearby worlds which differ from the actual world at most with regards to the semantic facts (but not with regards to the non-semantic facts). Unfortunately, this is not possible: the semantic facts supervene on the non-semantic facts.[3] Keeping the non-semantic facts fixed would also have the unintended

3. Kearns and Magidor (2012) argue that semantic facts are sovereign: they don't supervene on the non-semantic facts. Some additional avenues of defining definiteness could be available to an epistemicist who embraces semantic sovereignty. However, I will proceed via a more traditional route that accepts the supervenience of the semantic on the non-semantic.

consequence of keeping the semantic facts fixed. Therefore, we need another approach.

The Definiteness Puzzle allows us to make the following observation: a definition of Δ should incorporate the insight that '$\Delta\varphi$' ought to be false only if variation in the intension of φ alone could easily make φ false.

### 4.2. *Logic of Δ*

The definiteness operator should satisfy the typical constraints.

(Factivity) $\Delta\varphi\rightarrow\varphi$
(Distribution) $\Delta(\varphi\rightarrow\psi)\rightarrow(\Delta\varphi\rightarrow\Delta\psi)$
(Intransitivity) $\neg\forall\varphi(\Delta\varphi\rightarrow\Delta\Delta\varphi)$
(Taut.) For all logical tautologies φ, $\Delta\varphi$ is true.
(Necessitation) For all φ such that $\vdash\varphi$, $\Delta\varphi$ is true.

### 4.3. *Interaction with Modal Operators*

When giving semantics for Δ we should be careful to define an operator that interacts well with the metaphysical modality operators. There are cases where our definiteness talk interacts with modal talk. For instance, even though I am actually definitely not bald, it's possible for me to be definitely bald. There are certain conditions we would like the definiteness operator to satisfy. For instance, on the correct definition of Δ the following should come out as true:

(i) $\Box(\Delta\varphi \rightarrow \varphi)$

On the other hand, the following principles should not be universally true:

(ii) $\Box(\Delta\varphi \vee \Delta\neg\varphi)$
(iii) $\Delta\varphi \rightarrow \Box\Delta\varphi$
(iv) $\Box\varphi \rightarrow \Delta\varphi$
(v) $\Delta\varphi \rightarrow \Box\varphi$

### 4.4 *Knowability Constraint*

The defined operator should have an appropriate relation to knowledge—if we are in a position to know φ then $\Delta\varphi$. In particular, the following disquotational schemas should come out as definitely true:

(T-Schema) '$\varphi$' is true $\leftrightarrow \varphi$
(R-Schema) '$\tau$' refers to $\tau$

The task for the epistemicist is to find an account of definiteness that would satisfy all these desiderata.[4]

## 5. Epistemicist Definiteness

### *5.1. How to Build a Model of Epistemicist Definiteness?*

My approach to defining the epistemicist definiteness operator mainly draws inspiration from three sources: Stalnaker's metasemantic interpretation of multidimensional semantics, John Hawthorne's (2006) theoretical approach to defining epistemicist definiteness and Juhani Yli-Vakkuri's (2016)[5] technical work on defining definiteness. The plan is to use Hawthorne's insights regarding the understanding of definiteness in an epistemicist setting and use a multidimensional model to define $\Delta$.

John Hawthorne (2006) provides a good illustration of what we should be after in defining the epistemicist notion of definiteness. The epistemicist claims that in order for $\varphi$ to be definitely true, we need to evaluate $\varphi$ relative to how it is used in all nearby worlds. Hawthorne (2006: 293) illustrates this by imagining that a speaker of a non-actual close variant of the language (a Twinglish speaker) is with us in the actual world and we are evaluating his utterances made in the actual world. This does not mean that there are actually speakers of non-actual variants of English in the actual world. However, Hawthorne's picture gives us a good intuitive understanding of what the epistemicist is after. By pretending that there is a community of Twinglish speakers in the actual world, we are able to isolate semantic plasticity as an ignorance inducing factor. We keep the worldly (non-semantic) facts fixed and let the language vary, which allows us to determine whether a shift in the linguistic facts alone is able to produce an easy error. Consider Michael's case from the Definiteness Puzzle. Michael is 190 cm in height and in English the boundary for 'tall' (185 cm) is below that height, so he

4. The epistemicist is not committed to absolutely every instance of the T-Schema coming out as definitely true. After all, we might want to restrict the T-Schema because of paradoxes like the Liar. Nevertheless, if we restrict ourselves to only 'well-behaved' instances of the T-Schema, such as '"Michael is tall" is true if and only if Michael is tall', we would want them to come out as definitely true.

5. See also Litland and Yli-Vakkuri (2016).

counts as tall. In Twinglish the boundary for 'tall' is 188 cm. Had the Twinglish speaker been with us in the actual world, his utterance 'Michael is tall' would be true as Michael's actual height is over the boundary.

Using the terminology I employed in Section 2, we may say that the epistemicist is interested in discovering how changing the metasemantic rules of our language would impact its semantics (while keeping the metasemantic circumstance of evaluation fixed). One might naturally wonder why the epistemicist needs a double-indexed rather than a simple metasemantic context. This need is demonstrated by the following case. Suppose that that there is a lottery with 1000 tickets and only one winner. Let the name 'Chancy' denote the winner of the lottery. Even if I don't know which ticket won the lottery, it seems that I am still in a position to know that Chancy is the winner of the lottery, because I have introduced the name 'Chancy' so that it refers to the winner. Even if another ticket would have won, I would be in a position to know that Chancy is the winner, because then 'Chancy' would refer to that (other) winner. If I know that Chancy is the winner, then it's definitely the case that Chancy is the winner. However, if we go with a simple metasemantic context that determines the intension of 'Chancy' we will get the wrong result. That is because in some close world $w^*$, some ticket other than Chancy, let's call it 'Schmancy', is the winner. Therefore, 'Chancy' as uttered in the metasemantic context $w^*$ will refer to Schmancy. But Schmancy is not the winner in the actual world (Chancy is). Therefore, we need a double-indexed metasemantic context: we change the first index, which gives us the metasemantic rules for 'Chancy', but we keep the metasemantic circumstance fixed.

The epistemicists are interested in the effect that change in metasemantic rules would have on the reference of 'Chancy' while keeping the external facts (such as which ticket is the winner) constant. What we need instead is an interpretation that would mimic Hawthorne's setup. If the Twinglish speaker were with us in the actual world, their utterance 'Chancy is the winner' would be true, because 'Chancy' in Twinglish would refer to Chancy. The reason for this is that we have a clear rule of determining the intension of 'Chancy': it picks out the actual winner of the lottery. Since we have a clear rule of determining the intension of 'Chancy', we can expect that this rule applies to every close variant of the language. Furthermore, if the Twinglish speaker were with us in the actual world, then 'Chancy' in Twinglish would refer to Chancy, because there is only one winner of the lottery in the actual world and it is Chancy. Consequently, instead of letting a simple metasemantic context determine the intension (or character) of 'Chancy', we need something slightly more complicated.

In contrast to 'Chancy', a term like 'bald' does not have a clear metasemantic rule that would be the same in every nearby world. Suppose that a metasemantic rule for 'bald' is a function from the external facts about the number of hairs on everyone's head to the intension (or character) of 'bald'. Furthermore, suppose that the actual metasemantic rule for determining the boundary for 'bald' is that all and only those in the bottom 20% (when it comes to the number of hairs) count as 'bald'. However, there is no uniform rule across nearby worlds that determines where the boundary for 'bald' is. Even if in the actual world, the metasemantic rule is such that it places the cut-off at 20%, our usage could easily be slightly different, so that the rule would place the cut-off at 19%. Thus, even if we kept all the facts about the number of hairs on everyone's head fixed, we would still get variation in the intension (or character) of 'bald', as there is no clear uniform rule that guides the metasemantics of 'bald'.[6] Therefore, the intension of 'bald' varies across the nearby worlds (even if we keep fixed the facts about the number of hairs on everyone's heads).

The setup proposed by Hawthorne requires a complex metasemantic context. Firstly, we need a metasemantic function that would tell us how to assign characters to expressions given the facts in the external world. Secondly, we need a world relative to which the characters are to be assigned to expressions. We need these two elements of metasemantic context to explain why it's definitely the case that Chancy is the winner of the lottery. In case of 'Chancy', the first element of metasemantic context gives us the rule that tells us that the winner of the lottery is the referent of 'Chancy' and the second element gives us the world from which the winner is drawn. It's definitely the case that Chancy is the winner, because (a) all nearby variants of the language give us the same metasemantic rule that assigns the winner of the lottery as the intended referent of 'Chancy' and (b) we keep the second element of the metasemantic context fixed on the actual world. This is essentially what Hawthorne does when he imagines that the Twinglish speaker is with us in the actual world: the first element of metasemantic context changes (the element that gives us the function from worlds to characters), but the second element of metasemantic context is kept fixed (so the external facts which determine meanings/characters in our language are kept fixed).

Switching from a Stalnakerian 2D model to model with a complex circumstance of evaluation (a pair of worlds) allows us to handle indexicality. Furthermore, what is needed for handling Hawthorne's approach to definiteness within the multi-dimensional model is a switch from a simple metasemantic context (a single world) to a complex one (a pair of worlds). The first element in the pair gives use the 'descriptive' metasemantic rule that tells us how the facts in

6. One could argue that there is a uniform metasemantic rule for a vague term like 'bald' across the nearby: the rule is that the intension of 'bald' is such that it applies only to those who have little hair. However, this is not a clear rule: it's vague what counts as little hair. In contrast, a rule 'the winner of the lottery' that guides the metasemantics of 'Chancy' is not relevantly vague in this way.

the external world determine the characters of expressions in the variant of the language, that is, it gives us a function from worlds to characters. The second element is the external world that is an argument in the function. For instance, in the example of 'Chancy', the first element would give us the rule that the constant character of 'Chancy' picks out the winner of the lottery; the second element gives us the winner of the lottery (and determines which ticket in particular 'Chancy' will refer to).

Therefore, what we need is a multi-dimensional model with quadruple-indexed points of evaluation $\langle w_1, w_2, w_3, w_4 \rangle$ where:

- the first index, the *metasemantic context*, gives us the function from worlds to characters for the expressions in the language
- the second index, the *metasemantic circumstance*, gives us the character for the expressions in the language
- the third index, the *semantic context* (or context simpliciter), gives us the intensions
- the fourth index, the *circumstance of evaluation*, gives us the semantic values

In the next section, I briefly outline the 4D model.

## *5.2. The 4D model*

In the 4D model, formulas are evaluated relative to quadruple-indexed points of evaluation $\langle w, v, u, z \rangle$. Intuitively we may think about this as a refinement of a two-dimensional model where the simple metasemantic context and circumstance of evaluation is replaced by a pair of worlds (to handle metasemantic and semantic complications outlined above). The first index is the metasemantic context, which determines the variant of the language relative to which we are to evaluate formulas. The metasemantic context is something that the definiteness operator operates on, which allows us to see how small shifts in the interpretation (metasemantic rules) of an expression shift the semantics in the language.

A 4D-model is a triple $M = \langle W, R, [\![.]\!] \rangle$, where W is a set of worlds, R a reflexive accessibility relation[7] and $[\![.]\!]$ is a function from atomic sentences to sets of four-dimensional points (subsets of $W \times W \times W \times W$). The bivalent valuation of

7. The accessibility relation R should be interpreted as the epistemic 'closeness' relation: $w$ is R-related to $v$ if $w$ is close to (epistemically indiscriminable from) $v$. For simplicity, I assume that the accessibility relation relevant to the operation of standard metaphysical modality operators is universal, so I will use quantification over worlds (universal for necessity and existential for possibility) to give the interpretation to modal operators.

sentences at 4D-points of evaluation $\langle w, v, u, z\rangle$ is defined as follows (P stands for an atomic sentence and φ and ψ are schematic variables):

$$
\begin{array}{lll}
|P|_{w, v, u, z} = 1 & \text{iff} & \langle w, v, u, z\rangle \in [\![P]\!] \\
|\neg\varphi|_{w, v, u, z} = 1 & \text{iff} & |\varphi|_{w, v, u, z} = 0 \\
|\varphi\wedge\psi|_{w, v, u, z} = 1 & \text{iff} & |\varphi|_{w, v, u, z} = 1 \text{ and } |\psi|_{w, v, u, z} = 1 \\
|\Box\varphi|_{w, v, u, z} = 1 & \text{iff} & \text{for all } z^* \in W,\ |\varphi|_{w, v, u, z^*} = 1 \\
|A\varphi|_{w, v, u, z} = 1 & \text{iff} & |\varphi|_{w, v, u, u} = 1 \\
|\Delta\varphi|_{w, v, u, z} = 1 & \text{iff} & \text{for all } w^*Rw,\ |\varphi|_{w^*, v, u, z} = 1
\end{array}
$$

We can say that φ is true *simpliciter* in $w$ if $|\varphi|_{w, w, w, w} = 1$. Logical consequence is defined as truth preservation in all proper points of evaluation (points of the form $\langle w, w, w, w\rangle$) in all models.

## 5.3. *Satisfaction of Desiderata*

We are mostly interested in the definition of Δ.[8] We have given the following truth conditions for Δ:

(Epistemicist Definiteness) $|\Delta\varphi|_{w, v, u, z} = 1$ iff for all $w^*Rw$, $|\varphi|_{w^*, v, u, z} = 1$

The rationale for the semantics is as follows. We keep the metasemantic circumstance, the context, and the circumstance of evaluation fixed and we vary only the metasemantic context, because we are interested in the change that would result from using different metasemantic rules to assign characters to our formulas. To borrow a term from Kaplan (1977/1989), the definiteness operator is a metasemantic monster[9] on my view. We allow the metasemantic context (the

8. The account presented above treats Δ as an operator. This is also an approach taken, for instance, by Williamson (2003), Litland and Yli-Vakkuri (2016), and Yli-Vakkuri (2016). However, there is a question whether we should treat Δ as an operator or a predicate (Bacon 2018). I follow the standard approach in treating Δ as an operator. One reason for this approach is that operators, in contrast to predicates, are easily iterable. The second reason is that they are easily combinable with other operators, e.g., the modal operators □ or A. I think this gives us good reasons to prefer the operator approach to a predicate approach. However, it is worth noting that the operator approach may come at a cost once we use a language that includes quantifiers. We may need to reject some principles of quantification theory (such as Extensional Beta Equivalence: $(\lambda v.P)(a)\leftrightarrow P(a/v)$) if we define 'definiteness' as an operator. My approach is that such principles are dubious when reasoning about vagueness due to the famous result by Evans (1978): they are not worth giving up iterability and the elegance of combination with modal operators for.

9. According to Kaplan (1977/1989) monsters are operators that operate on contexts of utterances. Kaplan claims that English does not actually have monstrous operators, but that they nevertheless could exist in some other (perhaps merely possible) languages.

first index) to shift across the space of close possible worlds. This allows us to see how the shift in usage across close linguistic communities influences the semantic value of expressions in our language.

The account has a few initial advantages going for it. Firstly, it does not fall prey to the Definiteness Puzzle: the account keeps the 'worldly' facts fixed as Epistemicist Definiteness (ED) does not shift the circumstance of evaluation. This allows us to isolate semantic plasticity as an ignorance-inducing factor: if the truth value of φ changes as we change the metasemantic context, then we can say that shifts in language can produce an easy error (and consequently ignorance). This means that ED meets the first desideratum for an epistemicist account of definiteness: isolation of semantic plasticity as an ignorance inducing factor.

Secondly, ED allows for a very natural interaction between the definiteness and metaphysical modality operators and it validates the formulas using the necessity operator that we would intuitively want validated. For instance, our semantics validates the intuitive formula '$\Box(\Delta\varphi \rightarrow \varphi)$'. Furthermore, it correctly does not validate $\Box(\Delta\varphi \vee \Delta\neg\varphi)$ nor '$\Box\varphi \rightarrow \Delta\varphi$' nor '$\Delta\varphi \rightarrow \Box\varphi$' nor '$\Delta\varphi \rightarrow \Box\Delta\varphi$' nor '$\Box\varphi \rightarrow \Delta\Box\varphi$'.

Thirdly, the account allows us to handle important problematic cases that posed problems for some accounts presented in the literature. For instance, consider the sentence 'definitely, Chancy is the winner of the lottery'. This is validated by the account. It's definitely the case that Chancy is the winner, because the metasemantic rule that says that the ticket that wins the lottery is the same in all nearby worlds. Therefore, for all nearby worlds $w$, 'Chancy is the winner of the lottery' is true as evaluated at $\langle w, @, @, @\rangle$.

Furthermore, let's consider the case of disquotational sentences like:

(1) 'Tallness' refers to tallness.

Generally we want sentences like (1) to come out as definitely true. Suppose that (1) is uttered at $\langle @, @, @, @\rangle$ and we want to check whether it is definitely true. We know that "Δ 'tallness' refers to tallness" is true at $\langle @, @, @, @\rangle$ only if for all close worlds $w$, "'tallness' refers to tallness" is true at $\langle w, @, @, @\rangle$. Whether (1) is definitely true will depend on our interpretation of metalinguistic vocabulary such as 'refers'. For instance, (1) will not be definitely true, if we interpret the utterance of (1) at a point $\langle w, @, @, @\rangle$ as if it were the case that a speaker from the metasemantic context $w$ is talking about the actual variant of the language (because @ is the circumstance of evaluation, so (1) seems to be about how things are at @). As uttered in $w$, 'tallness' refers to *tallness** (some property distinct from *tallness*), but in the actual variant of the language 'tallness' refers to *tallness*. So if the speaker from $w$ were talking about the actual variant of the language, they would be saying something false (as 'tallness' doesn't refer to *tallness** in the actual variant, it refers to *tallness*).

However, I don't think that interpreting (1) in this way makes much sense. (1) seems to be definitely true because no matter which close variant of the language we use to utter (1) it should come out as true; no matter what exact property is referred to by 'tallness', the fact that we use the word 'tallness' to talk about the property that 'tallness' refers to, guarantees that we are telling the truth. However, there is no intuition that says that no matter what variant of the language we use to evaluate the actual variant of the language '"tallness" refers to tallness' should come out as true. Suppose that we apply Hawthorne's setup to analysing definiteness to the case. If there were a speaker of the language coming from $w^*$ (Twinglish) in the actual world, their Twinglish utterance '"tallness" refers to tallness' made in the actual would be about Twinglish not English. By uttering (1) we are not using any expressions identifying English specifically as the language about which the sentence is uttered. Rather, we are using the fact that (1) is naturally interpreted as being about the language variant in which the sentence is made. In other words, when we ask whether "'tallness' refers to tallness" is definitely true, we are not looking to find out what speakers of other variants of the language would say about us. Rather, what we want to know is this: if we had used a different language variant to talk about ourselves (in that variant), would what we say be true?.

Therefore, what we want is for (1) as evaluated relative to $\langle w^*, @, @, @\rangle$ to be true. We can make this happen by assigning the appropriate truth conditions to metalinguistic vocabulary such as 'true' and 'refers'. This is a solution suggested by Juhani Yli-Vakkuri (2016: 822–23) to an analogous problem for his model of definiteness. We can stipulate that sentences of the form '"$\tau$" refers to $\tau$' and '"$\varphi$" is true if and only if $\varphi$' are true whenever the metasemantic circumstance, context, and circumstance of evaluation are identical. Suppose that we call the function from worlds to characters a *metasemantic character*. Furthermore, call a metasemantic character f *diagonally true* if and only if for all $w \in W$, $f(w)(w)(w) = 1$. The proposal is that in all the nearby worlds, instances of disquotational schemas like '"$\tau$" refers to $\tau$' and '"$\varphi$" is true if and only if $\varphi$' have diagonally true metasemantic characters.[10] This would make them definitely true.[11]

---

10. One issue that remains is what exact semantics should be given to metalinguistic vocabulary such as 'true' and 'refers'. On the one hand we want to validate sentences like '"$\tau$" refers to $\tau$', so we want diagonally true metasemantic characters. Consequently, if $w^*$ is a nearby world we want '"$\tau$" refers to $\tau$' to be evaluated as true at $\langle w^*, @, @, @\rangle$. However, we also want our metalinguistic vocabulary to be sufficiently expressive to be able to talk about other language variants. This is not a problem. We can introduce additional vocabulary into the language that would allow us to talk about other language variants by referring to them specifically. For instance, we can talk about the language in some world $w$ and say things like '"$\tau$" refers to $x$ at $w$'. The details of the treatment of metalinguisitic vocabulary like 'true' and 'refers' requires a separate project.

11. This is not an ad hoc solution. The metasemanitic characters of '"$\tau$" refers to $\tau$' and '"$\varphi$" is true if and only if $\varphi$' are diagonally true because when we evaluate them at a point $\langle w^*, w, w, w\rangle$

In the next section, I will outline alternatives to my account that have been presented in the literature. I will argue that each of these accounts suffers from issues that my account does not.

## 6. Alternatives and Their Issues

### *6.1. Simple Diagonal Account*

The epistemicist wants ultimately to explain the ignorance due to vagueness by the failure of the safety principle. The basic idea is that even if some borderline sentence φ is true in the variant of the language we are actually speaking, it could easily be the case that we would be speaking a very similar variant of the language on which φ is false. We cannot discriminate between very similar variants of the language, so we could easily make an error, for example, by asserting φ. This possibility of error is something that stands in the way of our knowledge. There is a natural proposal of defining definiteness in the light of the above remarks:

> (Simple Diagonal Account) φ is definitely true (at @) if and only if for all close worlds $w$, φ as used at $w$ is true at $w$.

Unfortunately, there is a simple problem with this account: it does not isolate the effect of semantic plasticity as an ignorance inducing factor. Suppose that φ is actually true but at some nearby world $w^*$, φ (as used at $w^*$) is false at $w^*$. However, the account does not tell us why φ is false at $w^*$. For all we know it could be false because of factors other than semantic plasticity.

Furthermore, there are problems with the interaction of definiteness and metaphysical modality on the Simple Diagonal Account. Suppose that we are evaluating a sentence 'I could be definitely bald'. This means that there is a world $w^*$ such that 'definitely I am bald' is true as evaluated at $w^*$. However, in fact I am definitely not bald. If we were to evaluate the sentence 'I am bald' so that for all worlds $w$ (such that $w$ is close to @), 'I am bald' as uttered at $w$ is to be evaluated at $w$, then 'definitely I am bald' would come out as false (since I am actually definitely not bald). However, if we were instead to evaluate the sentence 'I am bald' relative to all the worlds that are close to $w^*$, we also get the wrong result, as 'bald' as uttered at $w^*$ may mean something completely different from what it actually means ($w^*$ could be any world for all we know; it does not have to be a close one).

(where $w^*$ is a nearby world), we interpret the situation alongside Hawthorne's picture: we imagine a speaker of the variant of the language from $w^*$ who is in $w$ uttering the sentence.

Therefore, the Simple Diagonal Account does not meet our desired desiderata: it does not isolate the effect of semantic plasticity as an ignorance inducing factor and does not interact well with metaphysical modality operators.

### 6.2. Complex Diagonal Account

Suppose that we define definiteness like this:

(Diagonal) $|\Delta\varphi|_{w, v, u, z} = 1$ iff for all $w^*Rw$, $|\varphi|_{w^*, w^*, w^*, w^*} = 1$

Unfortunately, there are multiple problems with the definition of Δ. One such problem is that on the above definition, metasemantic circumstance (second index), the context (third index) and world of evaluation (fourth index) are irrelevant to the evaluation of Δφ at any 4D point ⟨*w, v, u, z*⟩, that is, no matter what worlds *v*, *u* and *z* are, the truth value of Δφ will depend only on the truth value of φ in worlds close to *w* (the metasemantic context). This is of course undesirable: it matters a great deal to the truth value of 'it's definitely the case that Napoleon lost the battle of Waterloo' whether the world of evaluation is one where Napoleon lost the battle or one where he won it.

We could of course revise the account. For instance, on a natural revision Δ could be defined as follows:

(Diagonal*) $|\Delta\varphi|_{w, v, u, z} = 1$ iff for all $z^*Rz$, $|\varphi|_{z^*, z^*, z^*, z^*} = 1$

On this definition, the initial world of evaluation does matter to the truth value of Δφ, but other problems emerge. For instance, the initial context of utterance (third index) is still irrelevant to the evaluation of the sentence. This also clearly problematic, as it makes intuitively knowable (and therefore definitely true) propositions borderline. For instance, 'Aφ↔φ' seems to be definite, but turns out to be false on the revised account. Similarly, the above account falsifies the intuitively true sentence '□(Δφ→φ)'. For suppose that φ is the sentence 'Hesperus ≠ Phosphorus'. It is necessarily the case that Hesperus is identical to Phosphorus (both are identical to the planet Venus) so for all *z**, $|\text{Hesperus} \neq \text{Phosphorus}|_{@,@,@,z^*}=0$. However, there are worlds (perhaps far away ones) where 'Hesperus' is used to refer to Mars and Phosphorus is used to refer to Venus. Suppose that *w** is such a world and additionally that the use of 'Phosphorus' and 'Hesperus' at all the worlds close to *w** is the same as in *w**. This means that $|\Delta\varphi|_{@,@,@,w^*}=1$. However, it is still the case that $|\varphi|_{@,@,@,w^*}=0$ because φ relative to the actual variant of the language still expresses the proposition that Venus is distinct from Venus. So the conditional 'Δφ→φ' comes out as false.

The general technical problem with the diagonal strategy is that the definiteness operator is given the power to shift the context of utterance and the world of evaluation. The problem arises when we attempt to use both the definiteness operator and the modal operators in the same sentence, as the competences of the different operators conflict. Having multiple operators that shift the same index is not always problematic of course, for example, in case of multiple temporal operators, when it is clear that these different operators really should operate alongside the same dimension (i.e., time). However, in the case of definiteness, there is no reason why the definiteness operator should have the power to shift the world of evaluation since definiteness is not a metaphysical modality operator.

### *6.3. Ground Fixing Account*

Ofra Magidor (2018) has proposed an interesting account of definiteness that makes use of the notion of grounding. Magidor's proposal is that 'definite' should be given the following truth conditions:

> (Ground Fixing Account) Let $p$ be the proposition that is expressed by φ at @ and let Q be the ultimate grounds for $p$ in @. φ is definitely true at @ if and only if for all nearby worlds $w$ where Q holds, φ as used at $w$ is true at $w$.

This proposal is intuitively plausible. What went wrong with the Simple Diagonal Account, as demonstrated by the Definiteness Puzzle, is that we allowed the non-semantic facts such as ones related to Michael's height to vary across the set of close possible worlds. However, on Magidor's proposal the relevant non-semantic facts should be kept fixed. Facts about Michael's height are what grounds the truth of the sentence 'Michael is tall'; thus, on Magidor's proposal we only consider close worlds where the grounding facts that are relevant to the sentence 'Michael is tall' are kept fixed. This solution allows to isolate semantic plasticity as an ignorance inducing factor and avoid the issues brought to light by the Definiteness Puzzle.

Magidor's proposal has no problem in dealing with cases such as:

(1) 'Tallness' refers to tallness.

which correctly comes out as definitely true. However, the problem for the proposal is that it does not treat as relevant certain worlds that intuitively ought to count as relevant. This is visible in examples such as:

(2) 'Tallness' refers to the property of being over 185 cm in height.

which comes out as definitely true on Magidor's proposal. By supposition, in the actual world 'tallness' refers to the property of being over 185 cm in height. On Magidor's proposal, when evaluating the definite status of (2), we should only look at close possible worlds where the grounding facts relevant to the truth of (2) are the same as in the actual world. However, if all the facts that ground the truth of (2) are kept fixed across the relevant set of worlds, then (2) is true in all these worlds. If we only look at worlds where the grounding facts for the truth of '"tallness" refers to the property of being over 185 cm in height' are the same as in the actual world, then 'tallness' refers at these worlds to the same property as in the actual world, namely the property of being over 185 cm in height. So again (2) is definitely true on Magidor's proposal.

Magidor (2018) is happy to accept this consequence.[12] However, there are good reasons to think that Magidor's proposal doesn't pick out the notion of definiteness the epistemicist is after. What the epistemicist is after is the notion of safety relative to semantic plasticity. Is semantic plasticity an obstacle in coming to know that (2) is true? It is, but it's not the semantic plasticity of (2) that is responsible for our ignorance, it's the semantic plasticity of 'tallness'. Magidor's approach assumes that if semantic plasticity is an obstacle to knowing that φ is true, then it is the semantic plasticity of φ in particular that is responsible. However, it may be the case that the semantic plasticity of some other piece of vocabulary is responsible for our ignorance of the truth value of φ (e.g., an expression merely mentioned in φ). We need a global notion of safety relative to semantic plasticity that would look at the variation in the interpretation of a language as a whole.

To put the point another way, the Ground Fixing Account goes against the spirit of the epistemicist view, because it does not treat as relevant worlds that intuitively should count as such. Expressions in our language are semantically plastic according to the epistemicist, which means that we are not sensitive to slight shifts in reference between close possible variants of the language. That is to say that we cannot discriminate between variants of the language that differ only slightly in the assignment of semantic values to expressions. Naturally this means that worlds which differ slightly, but differ nevertheless, from the actual world with respect to what property the word 'tallness' refers to should count as relevant for our analysis, because these worlds are semantically indiscriminable from the actual world. However, on Magidor's proposal such worlds are not

12. An interesting consequence of Magidor's proposal is that the definiteness operator would fail to satisfy Distribution: $\Delta(\varphi \rightarrow \psi) \rightarrow (\Delta\varphi \rightarrow \Delta\psi)$. For instance, if we substitute (2) for φ and "'tallness' is the property of being over 185 cm in height" for ψ, then $\Delta(\varphi \rightarrow \psi)$ and $\Delta\varphi$ are true but $\Delta\psi$ is false.

taken to be relevant, because the grounding facts for the truth of (2) are different there. Magidor's proposal is a good attempt at solving the problem at hand, but unfortunately it fails.

### *6.4. 3D Model*

Litland and Yli-Vakkuri (2016) develop a three-dimensional model to provide an account of definiteness. On the 3D semantics, formulas are evaluated relative to triple-indexed points of evaluation $\langle w, v, u\rangle$. The first index is the metasemantic context, which determines the variant of the language relative to which we are to evaluate formulas; the second index is the context of utterance and the third index is the circumstance of evaluation. In short, instead of having a complex metasemantic point like my 4D model employs, Litland and Yli-Vakkuri (2016) have a single metasemantic context. The second and third index are used in the usual way: they allow us to handle indexical expressions in a Kaplanian way.

A 3D-model is a triple $M = \langle W, R, [\![.]\!]\rangle$, where as usual W is a set of worlds, R a reflexive (epistemicist) accessibility relation and $[\![.]\!]$ a function from atomic sentences to sets of 3D points of evaluation (each point being a subset of W × W × W). The bivalent valuation of sentences at 3D-points of evaluation $\langle w, v, u\rangle$ is defined as follows (P stands for an atomic sentence and φ and ψ are schematic variables):

| | | |
|---|---|---|
| $\lvert P\rvert_{w,v,u} = 1$ | iff | $\langle w, v, u\rangle \in [\![P]\!]$ |
| $\lvert\neg\varphi\rvert_{w,v,u} = 1$ | iff | $\lvert\varphi\rvert_{w,v,u} = 0$ |
| $\lvert\varphi\wedge\psi\rvert_{w,v,u} = 1$ | iff | $\lvert\varphi\rvert_{w,v,u} = 1$ and $\lvert\psi\rvert_{w,v,u} = 1$ |
| $\lvert\Box\varphi\rvert_{w,v,u} = 1$ | iff | for all $u^* \in W$, $\lvert\varphi\rvert_{w,v,u^*} = 1$ |
| $\lvert A\varphi\rvert_{w,v,u} = 1$ | iff | $\lvert\varphi\rvert_{w,v,v} = 1$ |
| $\lvert\Delta\varphi\rvert_{w,v,u} = 1$ | iff | for all $u^*Rw$, $\lvert\varphi\rvert_{u^*,v,w} = 1$ |

Again, φ is true *simpliciter* in $w$ if $\lvert\varphi\rvert_{w,w,w} = 1$. Logical consequence is defined as truth preservation in all proper points of evaluation (points of the form $\langle w, w, w\rangle$) in all models.

The first problem for the 3D semantics for definiteness comes from cases like "'Chancy' refers to the winner of the lottery" or "'tallness' refers to tallness". The solution that worked for the 4D model (insisting that 'refers' and 'Chancy' have diagonally true metasemantic character) comes at a price if we adopt the 3D model. The reason why Yli-Vakkuri (2016: 823) rejects the solution is that on his 3D model, the above solution implies that 'Chancy' and 'refers' are indexicals. On a 3D model, the diagonally true metasemantic character becomes a diagonally true character (i.e., a function f such that for all $w \in W$, $f(w)(w) = 1$). However, postulating a diagonally true character for expressions like '"τ" refers to τ' and 'Chancy is the winner of the lottery' would mean that 'Chancy' and 'refers' are indexicals. Suppose that we are using Yli-Vakkuri's 3D model: postulating a diagonally true character for 'Chancy is the winner of the lottery' would mean that 'Chancy' refers to whichever ticket is the winner in the context world, for example, 'Chancy' as evaluated at $\langle @, v, v\rangle$ would refer to the ticket winner at $v$. However, this would mean that 'Chancy' is an indexical. This is problematic because terms like 'Chancy' are not normally taken to be indexicals: they are rigid designators introduced by reference fixing descriptions.

These issues arise because of the limitations of the 3D model that are not present in the 4D model. In the 3D model, the second index (the context) is overloaded: it has to play the standard role of determining the intensions of indexical expressions as well as the role of determining the intensions of non-indexical expressions like 'Chancy'. This is the reason why expressions like 'Chancy' look like indexicals on the 3D model. However, on the 4D model these two roles are played by two separate indexes: the intension-determining role for indexicals is played by the context (the third index) and the character-determining role for expressions like 'Chancy' is played by the metasemantic circumstance (the second index). On a 4D model, neither 'true', nor 'refers', nor 'Chancy' are indexicals because their characters are constant: the proposal requires that they have diagonally true *metasemantic characters* (functions from worlds to characters) and not *characters* simpliciter (functions from worlds to contents). By moving to a 4D model, we are able to avoid the issues that a 3D model faces.

Another challenge to the 3D model comes from some examples of use of the actuality operator. The actuality operator is an indexical operator, which allows us to refer to the actual world. Yli-Vakkuri (2016: 824) presents an argument against epistemicist treatment of definiteness on his 3D model using a cleverly constructed non-indexical operator which is used to refer to the actual world.

Suppose that there is a non-indexical way of referring to the actual world. For instance, suppose that instead of using the standard indexical operator A, we can simply name the actual world 'Worldy' and refer to it by simply using that proper name. We can then define a different actuality operator A*; 'A*φ' would roughly translate to 'In Worldy, φ'. Using his 3D model, Yli-Vakkuri (2016: 824) gives the following truth conditions to both the standard and the novel actuality operators:

$$|A\varphi|_{w,v,u} = 1 \quad \text{iff} \quad |\varphi|_{w,v,v} = 1$$
$$|A^*\varphi|_{w,v,u} = 1 \quad \text{iff} \quad |\varphi|_{w,w,w} = 1$$

The problem for the 3D-account of definiteness is that we seem to be in a position to know that in Worldy φ if and only if actually φ, that is, our two ways of referring to the actual world should yield the same result. Therefore, ‘A*φ↔Aφ’ should come out as definitely true. However, on the 3D semantics it does not come out as definitely true:

$$
\begin{array}{lll}
|\Delta(A^*\varphi \leftrightarrow A\varphi)|_{@,@,@} = 1 & \text{iff} & \text{for all } w^*R@,\ |A^*\varphi \leftrightarrow A\varphi|_{w^*,@,@} = 1 \\
 & \text{iff} & \text{for all } w^*Rw,\ |A^*\varphi|_{w^*,@,@} = |A\varphi|_{w^*,@,@} \\
 & \text{iff} & \text{for all } w^*Rw,\ |\varphi|_{w^*,w^*,w^*} = |\varphi|_{w^*,@,@}
\end{array}
$$

As long as there is a close world $w^*$ such that $|\varphi|_{w^*,w^*,w^*} \neq |\varphi|_{w^*,@,@}$ (which is guaranteed if there are to be differences between close possible worlds) the sentence will be false.

However, the issue does not arise for the 4D model. The reason again is that in a 4D model we have a complex metasemantic index consisting of a metasemantic context (first index) and a metasemantic circumstance. Firstly, the metasemantic context determines the metasemantic rule that the name ‘Worldy’ is to refer to whatever world is given by the metasemantic circumstance. Secondly, the metasemantic circumstance fixes the world that ‘Worldy’ is to refer to. Therefore, the truth conditions for the two actuality operators on the 4D model are given as follows:

$$
\begin{array}{lll}
|A\varphi|_{w,v,u,z} = 1 & \text{iff} & |\varphi|_{w,v,u,u} = 1 \\
|A^*\varphi|_{w,v,u,z} = 1 & \text{iff} & |\varphi|_{w,v,v,v} = 1
\end{array}
$$

However, now ‘A*φ↔Aφ’ comes out as definitely true:

$$
\begin{array}{lll}
|\Delta(A^*\varphi \leftrightarrow A\varphi)|_{@,@,@,@} = 1 & \text{iff} & \text{for all } w^*R@,\ |A^*\varphi \leftrightarrow A\varphi|_{w^*,@,@,@} = 1 \\
 & \text{iff} & \text{for all } w^*Rw,\ |A^*\varphi|_{w^*,@,@,@} = |A\varphi|_{w^*,@,@,@} \\
 & \text{iff} & \text{for all } w^*Rw,\ |\varphi|_{w^*,@,@,@} = |\varphi|_{w^*,@,@,@}
\end{array}
$$

As long as the second and third indexes (metasemantic circumstance and metasemantic context) are identical, ‘A’ and ‘A*’ will be interchangeable.

This is the result given by the formal model, but it also makes sense given our intended interpretation. Suppose that the Twinglish speaker is with us in the actual world and is making the utterance ‘In Worldy, φ’. What does ‘Worldy’ in their mouth refer to? Since in our scenario the Twinglish speaker lives in the actual world, ‘Worldy’ in Twinglish refers to the actual world. Like with ‘Chancy’, there is a clear metasemantic rule to determine the reference of ‘Worldy’. Thus, in every close variant of the language ‘Worldy’ refers to the actual world: if we used a different close variant of the language, the reference of ‘Worldy’ would not change as all nearby language variants use the same metalinguistic rule to pick out the reference for 'Worldy'. Therefore, the results given by the formal model are in accordance with what we would expect from it.

## 7. Conclusion

I presented a multidimensional model inspired by the work of Stalnaker and Kaplan. Using that model I outlined an epistemicist proposal for the semantics of the definiteness operator. I showed that the presented account meets the desiderata and I argued that it is superior to the accounts presented in the literature.

## Acknowledgments

My deepest gratitude goes to Daniel Elstein, Ed Elliott, Miriam Bowen, Ofra Magidor, and Robbie Williams for their insightful comments on the paper. Special thanks go also to audiences at the Leeds Postgraduate Seminar and the Warsaw Indexicality Workshop as well as to two anonymous referees at *Ergo*. This work was supported by the Arts & Humanities Research Council (grant number AH/L503848/1) through the White Rose College of the Arts & Humanities, the School of Philosophy, Religion and History of Science at the University of Leeds and the Royal Institute of Philosophy.

## References

Bacon, Andrew (2018). *Vagueness and Thought*. Oxford University Press.

Caie, Michael (2012). Vagueness and Semantic Indiscriminability. *Philosophical Studies, 160*(3), 365–77.

Chalmers, David (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.

Chalmers, David (2004). Epistemic Two-Dimensional Semantics. *Philosophical Studies, 118*, 153–226.

Dorr, Cian and John Hawthorne (2014). Semantic Plasticity and Speech Reports. *Philosophical Review, 123*(3), 281–338.

Evans, Gareth (1978). Can There Be Vague Objects? *Analysis, 38*(4), 208.

Hawthorne, John (2006). Epistemicism and Semantic Plasticity. In Dean Zimmermann (Ed.), *Oxford Studies in Metaphysics* (Vol. 2, 289–322). Oxford University Press.

Jackson, Frank (1998). *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford University Press.

Jackson, Frank (2004). Why We Need A-Intensions. *Philosophical Studies, 118*, 257–77.

Kaplan, David (1989). Demonstratives. In Joseph Almog, John Perry, and Howard Wettstein (Eds.), *Themes from Kaplan* (481–563). Oxford University Press. (Original work published 1977)

Kearns, Stephen and Ofra Magidor (2012). Semantic Sovereignty. *Philosophy and Phenomenological Research*, *85*(2), 322–50.

Litland, John and Juhani Yli-Vakkuri (2016). Vagueness & Modality—An Ecumenical Approach. *Philosophical Perspectives, 30*(1), 229–69.

Magidor, Ofra (2018). Epistemicism, Distribution, and the Argument from Vagueness. *Noûs*, *52*(1), 144–70.

Putnam, Hilary (1975). The Meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science*, *7*, 131–93.

Stalnaker, Robert (1978). Assertion. *Syntax and Semantics, 9*, 315–32.

Williamson, Timothy (1994). *Vagueness*. Routledge.

Williamson, Timothy (2003). Vagueness in Reality. In Michael Loux and Dean Zimmermann (Eds.), *The Oxford Handbook of Metaphysics* (690–716). Oxford University Press.

Yli-Vakkuri, Juhani (2016). Epistemicism and Modality. *Canadian Journal of Philosophy*, *46*(4–5), 803–35.