Belief Expressivism

In Part I (*Chapter 1*), I argued that there must be some common explanation for the striking number of grammatical and logical ("grammatological") properties shared by moral judgments (e.g. 'stealing is wrong'), ordinary descriptive judgments (e.g. 'the sky is blue'), and taste judgments (e.g. 'sushi is tasty').[1] To offer *different* explanations for the properties they share fails to explain *why* they are shared. It follows that *if* taste judgments are not primarily used to describe or report facts, then moral, ordinary descriptive, and taste judgments cannot share the grammatological properties that they do in virtue of being descriptive. In Part II (*Chapters 2-4*), I defended the antecedent of this conditional – that taste judgments are not primarily used to describe or report facts. In *Chapter 2*, I examined why taste realism (the view that taste judgments are primarily used to describe mind-independent facts) is not widely defended in the taste literature. In *Chapter 3*, I presented a series of arguments against taste contextualism (the view that taste judgments are primarily used to describe the preferences of contextually salient individuals and groups). These arguments lead away from contextualism and towards its two serious alternatives: expressivism and genuine relativism. Finally, in *Chapter 4*, I established that expressivism and genuine relativism are two sides of the same coin; if they are not strictly identical, then they are strikingly similar views – indeed, each benefits from the theoretical resources of the other.

My arguments in Parts I and II jointly suggest that moral, ordinary descriptive, and taste judgments do not share the grammatological properties that they do in virtue of being descriptive. To reiterate, *some* common explanation of these properties is necessary. But if the explanation cannot be that all three classes of judgment are descriptive (and it obviously cannot

[1] I reserve the term 'judgment' to refer to declarative sentences in order to remain (terminologically) neutral in conflicts over whether certain declarative sentences (e.g. 'rollercoasters are fun') are primarily used to report facts (descriptivism), or not (nondescriptivism).

Declarative sentences are often defined as sentences which report facts – as opposed to interrogative sentences, which ask questions, and imperative sentences, which issue commands. While these definitions are obviously helpful (and fine for ordinary use), they evidently render nondescriptivism about any class of declarative sentences incoherent. Since the debate between descriptivists and nondescriptivists are often substantive, it is necessary to use a neutral term like 'judgment.'

be that all three are nondescriptive), then the explanation must be *indifferent* to what is at stake between descriptivism and nondescriptivism. In other words, any adequate explanation must (i) explain the grammatological properties in question *without* invoking uniquely descriptivist (or uniquely nondescriptivist) resources, and (ii) be compatible both with nondescriptivism about taste judgments, and with descriptivism about ordinary descriptive judgments.

In this chapter, I will defend a solution along the following lines: that all judgments share the grammatological properties that they do because they all express *beliefs* (i.e. commitments with propositional, truth-evaluable contents); the grammatological properties of all judgments inhere in the fact that for a speaker to assert a judgment *J* is for them to present themselves as having a specific kind of mental state – a *belief* – with the same propositional, truth-evaluable content as *J*. I dub any solution along these lines a form of *belief expressivism*. Belief expressivism is compatible with (a) taste nondescriptivism, so long as the propositional contents of taste beliefs do *not* reflect (are not typically supposed to represent) objective ways the world might be. It is likewise compatible with (b) ordinary descriptivism, if the propositional contents of ordinary descriptive beliefs *do* reflect (are typically supposed to represent) objective ways the world might be.

For my solution to be complete, I will of course need to show (i) how the grammatological properties of all judgments derive from the nature and justificatory structure of belief. Moreover, I will need to say much more about (ii) how it is that descriptive and nondescriptive beliefs both qualify *as beliefs*, in spite of their differences, (iii) what nondescriptive beliefs represent, if not objective properties, (iv) how it is possible to believe a content without intending to represent a way the world might be, and more.

As a preliminary step, I will demonstrate that the account I offer is grounded in five independently motivated theoretical constraints:

> *Expressive Assertion* (**EA**): To assert the judgment '*p*' is, at a minimum, to present oneself as having the mental state required to agree with the content *p* – typically, belief that *p*.
>
> *Propositional Taste Nondescriptivism* (**PTN**): Mature taste discourse is nondescriptive; ordinary language users do not intend for their assertions of taste judgments to describe (or report) objective facts. Be that as it may, taste

judgments have propositional content that is truth-evaluable (albeit only relatively), and taste disagreements are propositional disagreements.

Phenomenal Indistinguishability (**PI**): The license afforded to taste beliefs by taste appearances is phenomenally indistinguishable from the license afforded to most beliefs by appearances.

A posteriori Non-objectivity (**APNO**): It is only possible to learn that a class of judgments is non-objective by *a posteriori* means. It is discovered over time – through reflection on our practices and experiences – that certain phenomenal appearances (i.e. "apparent qualities") do not correspond to objective properties.

Finite Geometric Representationality (**FGR**): Primary (or basic) representational mental states – "views" – are finite in number, and rarely map cleanly onto propositions expressible using the predicates of a natural language. Views and propositions can be modeled as regions in geometric conceptual spaces (attributed to objects). Beliefs are secondary (or derivative) states: a person believes every proposition whose associated region fully contains one of their views. The exact relation between views and beliefs is given as:

> S believes that *x* is P iff the region (in the contextually salient domain **D**) that models S's view about *x* is a subset of the region (in **D**) that models P.

I will begin, in Section 5.1, by independently motivating (or reviewing the motivation for) the five theoretical constraints listed above. (**EA**), (**PTN**), and (**PI**) jointly make conceptual space for the possibility of nondescriptive beliefs. (**APNO**) then explains how it is possible for a discourse to be nondescriptive (as (**PTN**) maintains), given (**PI**). And if there are nondescriptive beliefs with propositional content, a solution to the grammatological problem comes into view: the grammatological properties of all judgments – be they descriptive or nondescriptive – inhere in the fact that all judgments express propositional, truth-evaluable content. From there, any standard account of propositional logic (e.g. possible worlds semantics) suffices to explain why all judgments share the same core logical features.

The fifth and final constraint – (**FGR**) – is my account of propositional content and belief. Although (**FGR**) is not required to solve the grammatological problem itself, it provides the broader theoretical framework in which my belief expressivist solution is situated. (**FGR**) treats belief as derivative of more fundamental representational states ("views"), and offers a new way to model propositional content apart from – but compatible with – standard possible worlds semantics. Though (**FGR**) is optional in one sense, any *belief* expressivist had better say what beliefs are – and (**FGR**) is my account. It provides a natural way to model graded belief,

make sense of the *de dicto* / *de re* distinction, and explain how belief can be closed under entailment – and thus potentially infinite in number – despite finite cognitive resources.

Taken together, these five constraints motivate belief expressivism: the view that the grammatological properties of judgment inhere in the fact that all judgments express beliefs – where belief is understood through the lens of (**FGR**).

In Section 5.2, I explain how the key grammatological properties of judgments derive from (**FGR**). I focus in particular on two features that have historically posed problems for nondescriptivism: negation and deductive validity. I show how belief expressivism can accommodate and explain the meaning and use of predicate negation by identifying negated predicates (e.g. 'not red') with all the conceptual space (in a contextually salient domain) *not* picked out by the predicate under negation. I likewise demonstrate how the validity of *modus ponens* arguments is preserved within (**FGR**). These applications are illustrative of how the grammatological properties of judgments follow naturally from belief expressivism, quite independent of descriptivist assumptions.

Finally, in Section 5.3, I examine the differences between moral and taste judgments. While I defend forms of belief expressivism about both – I argue that both are primarily used to express beliefs – the beliefs they express differ in content, licensing conditions, and function. Taste beliefs are typically licensed by immediate sensory experiences, whereas moral beliefs are generally licensed by *a priori* rational intuitions, or by reflecting on the consequences of actions. Additionally, moral beliefs are plausibly action-guiding in ways which taste beliefs are not. I argue that these differences are compatible with a unified account of judgment, so long as we distinguish the belief-expressive function of judgments from the nature of the beliefs they express. Ultimately, I hope to clear the ground for a more pluralistic view of assertoric discourse – one that recognizes the diverse kinds of beliefs different judgments express (on a predicate-by-predicate basis), rather than forcing them into rigid, *sui generis* categories.

—

5.1 The five constraints that point to belief expressivism

In this section, I will review five commitments, each of which constrains the space of possible solutions to the grammatological problem (the challenge of explaining why moral, ordinary descriptive, and taste judgments share so many grammatical and logical properties). Taken together, these five constraints entail the solution that I propose: belief expressivism. I will first independently motivate (or briefly review the motivation for) each commitment, and conclude by demonstrating how they jointly entail a belief expressivistic solution to the grammatological problem.

5.1.1 *Expressive Assertion* (EA)

> To assert the judgment '*p*' is, at a minimum, to present oneself as having the mental state required to agree with the content *p* – typically, belief that *p*.

Expressive Assertion – (**EA**), for short – is neither new[2] nor particularly controversial.[3] As I noted at the end of *Chapter 4*, (**EA**) is expressivism's uncontroversial positive thesis. What matters for present purposes is that (**EA**) follows from basic norms of sincerity, and other hard-to-deny platitudes about assertion.

It is widely accepted that to assert a judgment '*p*' is to present the content *p* as true. Moreover, it is a basic norm of assertion that speakers should not assert '*p*' unless they agree with *p*. This norm – the sincerity norm of assertion – reflects the broader principle that speakers

---

[2] Black (1952) and Davidson (1984) make essentially the same point. Black (1952) writes that "In order to use the English language correctly, one has to learn that to pronounce [a sentence] in a certain tone of voice [signifying assertion] is to *represent oneself* as knowing, or believing, or at least not disbelieving what is being said." (p.31) Davidson (1984) constrains the point to belief: "To assert is, among other things, to represent oneself as believing what one asserts." (pp.7-8)

[3] The only challenge I know of to (**EA**), if it can really be called that, is the challenge of clarifying what it is to *present oneself* as agreeing with *p* in a way that distinguishes it from *explicitly saying* that one agrees with *p*. That it is necessary to do so is evident from Moore's Paradox: if all it was for a speaker to *present themselves* as believing that *p* was for them to *say* that they believe that *p*, then sentences of the form '*p*, but I don't believe that *p*' would express inconsistent contents – but most think not. Most have followed Moore (1912, 1944) in thinking that the tension in Moorean sentences should be analyzed as something like the tension between *presenting oneself* as believing that *p* (by virtue of asserting '*p*') and *explicitly saying* that one does not believe that *p* (by virtue of asserting 'I don't believe that *p*'). Moore (1912, 1944) himself formulated the point by saying that for *S* to assert '*p*' is for *S* to "*imply*" that *S* believes that *p*.

should not present *p* as true unless they agree with *p*.[4] And since both speakers and listeners accept these norms – and rely on others to do the same – it follows that to assert '*p*' is not merely to present *p* as true, but also to present oneself as agreeing with *p*.[5] Now, to agree with *p* is just a matter of having a certain mental state. Typically, this mental state is some species of *belief* – namely, belief that *p*.[6] Hence, (**EA**): to assert '*p*' is (at a minimum) to present oneself as having the mental state required to agree with '*p*' – typically, belief that *p*.

To put it plainly, the basic motivation for (**EA**) is simply that we generally assume that speakers agree with what they assert. Indeed, we generally suppose that speakers assert what they do (at least in part) *because* they agree with it. Since agreeing with a content is a matter of having the right kind of mental state, we thereby take speakers to have whichever mental states put them in agreement with what they assert. Most of the time, this mental state is simply a belief.[7] In a phrase, we generally assume that speakers believe what they say, and that they say what they do (at least in part) because they believe it. Or more precisely: we assume that assertions are *sincere*, and therefore that speakers have the mental states that make their assertions sincere.[8]

---

[4] If (1) speakers should not present *p* as true (e.g. by presupposing that *p*, implicating that *p*, etc.) unless they agree with *p*, and (2) to assert '*p*' is to present *p* as true, the sincerity norm of assertion follows: speakers should not assert '*p*' unless they agree with *p*.

[5] It does not iteratively follow that to assert '*p*' is also to present oneself *as presenting oneself* as agreeing with *p*. Listeners expect speakers to be sincere, and speakers appreciate this expectation – but listeners do not interpret speakers as making explicit second-order claims about their own sincerity. That a speaker presents themselves as agreeing with *p* (merely by asserting '*p*') is a pragmatic upshot of the norm of sincerity – not the result of listeners treating every act of *φ-ing* as a presentation of oneself as *φ-ing*.

[6] Sometimes, agreeing with a judgment (or an assertion of one) requires something besides belief. Consider: 'Management announced more *damn* layoffs' and 'James *actually* painted the walls orange.' To agree with the first is, quite plausibly, not merely to believe that management announced more layoffs, but also to have a strong attitude of disapproval or frustration toward management's conduct. Likewise, to agree with the second is not just to believe that James painted the walls orange, but also to be in a state that registers James's choice of color as striking, surprising, or hard to accept.

[7] See footnote 135.

[8] Of course, not all assertions are sincere. But even insincere speakers rely on the norm of sincerity. Deception only works because listeners generally treat assertions as sincere. So whether or not an assertion *is* sincere, to assert '*p*' is (at the very least) to *present oneself* as agreeing with *p*.

Following standard expressivist practice, I will call the relation between a judgment and the mental state required to agree with it *expression.* Thus, I will say that the judgment '*p*' *expresses* the belief that *p* just in case (i) to *sincerely* assert '*p*' is (among other things) to believe that *p*, and for that reason, (ii) to assert '*p*' is to present oneself as believing that *p*. In brief, a judgment *expresses* the mental state that makes assertion of that judgment sincere.

In review, *Expressive Assertion* (**EA**) captures a basic but fundamental feature of assertion: that to assert any judgment is, at a minimum, to present oneself as having a particular mental state – namely, the mental state that puts one in agreement with what they assert. For conciseness, I will say that to assert a judgment is to present oneself as having the mental state *expressed* by the judgment one asserts. (**EA**) thus distills the basic link between assertion, agreement, and sincerity into a single principle, and it will serve as a basic assumption in what follows. With (**EA**) in place, I turn now to (**PTN**).

5.1.2 *Propositional Taste Nondescriptivism* (PTN)

> Mature taste discourse is nondescriptive; ordinary language users do not intend for their assertions of taste judgments to describe (or report) objective facts. Be that as it may, taste judgments have propositional content that is truth-evaluable (albeit only relatively), and taste disagreements are propositional disagreements.

Propositional Taste Nondescriptivism (**PTN**) is the fairly specific view that (i) ordinary language users do not intend for their assertions of taste judgments (e.g. 'sushi is tasty') to report objective facts, but for all that, (ii) taste disagreements (e.g. disagreements over whether sushi is tasty) are cases in which one party accepts a (truth-evaluable but non-objective) propositional content that the other party rejects. More precisely, a taste disagreement obtains between any two parties (according to (**PTN**)) iff the first believes some taste content *p*, the second believes some taste content *q*, and *p* and *q* are inconsistent propositions.

Whereas the first component of (**PTN**) puts the view in opposition to any form of descriptivism about taste judgments – and hence, both realism and contextualism – the second component of (**PTN**) differentiates it from an alternative (and more traditional) form of taste nondescriptivism, according to which (a) taste judgments express non-doxastic mental states

("taste attitudes"), and (b) taste disagreements obtain between two parties just in case they have clashing taste attitudes.[9]

I argued in *Chapter 4* that genuine relativism and minimalist expressivism are best understood as two sides of the same coin: nondescriptivism. (**PTN**) should therefore be understood as a specific kind of genuine relativism and minimalist expressivism – the *propositional* kind. Crucially, not all forms of genuine relativism and minimalist expressivism are consistent with (**PTN**). Some expressivists, including Gibbard (1990, 2003) and Blackburn (1984, 1993), deny that evaluative judgments express genuine beliefs, and instead explain evaluative disagreements in terms of clashing attitudes. Some genuine relativists – like Beršky tė & Stevens (2022) – agree with (**PTN**) that taste judgments express non-objective propositions, but then deny that taste disagreements are propositional in nature.[10] They, too, end up appealing to clashing attitudes to explain intuitions of disagreement. By contrast, many genuine relativists – including Kölbel (2002) and Lasersohn (2005) – and some minimalist expressivists, such as Horgan & Timmons (2006), endorse views equivalent (or analogous) to (**PTN**). On each of their accounts, the propositional character of evaluative judgments is genuine and semantically basic, and evaluative disagreements hinge on the inconsistency of non-objective contents.

The chief motivation for (**PTN**) is that it sidesteps the multitude of problems for its three alternatives – realism, contextualism, and non-propositional (i.e. "clashing attitudes") nondescriptivism. I reviewed these problems extensively in Part II (*Chapters 2-4*), and so I take it that I have already sufficiently motivated (**PTN**). My goal in this subsection is merely to review the advantages (**PTN**) enjoys over these three other views.

Taste realism is the view that there are objective, *mind-independent* taste facts, and that ordinary language users intend for their assertions of taste judgments to describe such facts.[11]

---

[9] The second component of (**PTN**) – that taste disagreements hinge on inconsistent truth-evaluable propositions – also puts (**PTN**) in opposition to various forms of contextualism, according to which paradigm faultless taste disagreements do not involve the expression of inconsistent propositions.

[10] Beršky tė & Stevens (2022) argue that non-objective propositions cannot be genuinely inconsistent with one another. Since propositional disagreements require that two individuals believe inconsistent contents, Beršky tė & Stevens (2022) proceed to deny that taste disagreements are propositional in nature.

[11] A note about '*mind-independence*': as I use the term, a *mind-independent* fact is one that concerns objects apart from their relation to anyone's mental states (e.g. size, shape, color, weight, etc). A

Taste realists suppose, in other words, that just as there are mind-independent facts which settle the exact mass of any given rock, and whether ice is colder than steam – there are analogous facts which settle exactly how tasty sushi is, whether roses are prettier than crocuses, and so on. In addition to defending the existence of mind-independent taste facts, taste realism is the view that ordinary people intend to describe such facts when asserting taste judgments.

The basic problem for taste realism is that almost nobody believes in mind-independent taste facts. More to the point, ordinary language users can hardly *intend* to describe something they don't believe in. A second, deeper problem for taste realism is that there do not seem to be any mind-independent taste facts. Or rather, there seem to be good reasons to doubt that there are any. For one thing, if there were any mind-independent taste facts, we would expect there to be at least one method to decisively settle taste disagreements, independent of what anyone thinks. Moreover, we would be hard-pressed to explain (a) why taste experiences quasi-infallibly license taste beliefs, (b) why there are no natural corrective repercussions for "objectively false" taste beliefs (when beliefs are aligned with genuine preferences)[12], (c) why most taste disagreements are faultless, and (d) why personal preferences are reasonable grounds for taste beliefs (rather than systematic sources of bias).

All forms of taste irrealism – including (**PTN**) – are well placed to explain these peculiar features of taste discourse and experience. Taste disagreements cannot be decisively settled because, on any irrealist view, there are no mind-independent facts to determine who is right. Without mind-independent taste facts, there are no underlying mind-independent properties for phenomenal taste qualities to misrepresent. Similarly, there cannot be any natural repercussions for "objectively false" taste beliefs if taste beliefs *cannot be* objectively false. And if objective

---

mind-dependent fact, by contrast, is a fact about how an object figures in someone's mental life. For example, whether Alice likes a particular rollercoaster is a mind-*dependent* question, distinct in kind from the mind-*independent* question as to whether that rollercoaster is wooden or steel. Both are objective, factual matters: mind-dependence (as I use the term) is entirely consistent with objectivity.

[12] As a refresher, a "natural corrective repercussion" is a natural consequence of an objectively false belief that can serve as a signal to an epistemic agent that their belief is objectively mistaken. If I dip my hand into hydrochloric acid on the presupposition that it is water – nature will demonstrate the error of my objectively false belief without delay. Although this is a particularly extreme example, the point remains. When taste beliefs are aligned with genuine preferences – as between two groups: one that genuinely prefers apples to oranges, and one that genuinely prefers the reverse – nature is silent. She neither imposes a cost on the former group nor the latter group. In matters of taste, there are no unambiguously discernible costs to being "objectively mistaken" – indeed, it is hard to even conceive of what they might be.

taste mistakes are not possible (and subjective mistakes only arise on occasion), that explains why taste disagreements are often faultless. Finally, if there are no mind-independent taste facts, preferences can hardly distort taste appearances, for they alone infuse taste appearances with their distinctive character.

Short of abandoning taste descriptivism, the only alternative to realism (i.e. mind-*independent* descriptivism) is contextualism (i.e. mind-*dependent* descriptivism). Taste contextualism is the view that ordinary language users intend for their assertions of taste judgments to describe (or report) the preferences of (contextually salient) individuals and groups – most often, the preferences of the speaker.[13] On any given occasion, the context of use determines the person or group whose tastes are described. For example, when Diane is contextually salient, 'sushi is tasty' expresses the proposition that Diane likes sushi. When Americans are contextually salient, the same judgment expresses the proposition that Americans like sushi. And so on.

As a form of taste irrealism, contextualism can explain all of the features problematic for realism, but faces its own challenges. I raised four unique problems for taste contextualism in *Chapter 3*. First, contextualism claims that taste assertions are attitude reports (i.e. express mind-*dependent* propositions). But since attitude reports generally have objective truth-conditions, contextualism implies that taste assertions typically have objective truth-conditions, too. (For example, it is either objectively true that Diane likes sushi, or else objectively false.) The first problem for contextualism is that intuitively, deflationary remarks (e.g. 'that's subjective!') are felicitous as replies to taste assertions – but such remarks are inappropriate as replies to assertions with objective truth-values, like most attitude reports.

---

[13] I am concerned here with *mental state* contextualism, the only form of taste contextualism defended in the literature. At the end of *Chapter 3*, I considered a potential alternative form of contextualism: *abstract* contextualism. According to abstract contextualism, the context of use determines an abstract standard which modifies the content of the taste predicate used. For instance, when the salty standard is salient, 'tasty' means "tasty according to the salty standard." But since objects are "tasty according to the salty standard" iff they contain salt, this analysis of 'tasty' reduces it to a natural property. In general, any abstract contextualist analysis along these lines reduces taste predicates to mind-independent, natural properties. Abstract contextualism thus collapses into a context-sensitive form of taste realism, quite unlike the irrealism of mental state contextualism. As such, abstract contextualism inherits most of the problems for taste realism, in addition to many of the problems for mental state contextualism.

Second, taste contextualism maintains that ordinary language users do not differentiate between taste assertions and explicit descriptions of people's preferences (since both express the same kinds of propositions). As a consequence, taste contextualism falsely claims that some intuitively consistent statements (e.g. 'like everyone else, I enjoy their french fries, but they are terrible') express contradictions, and that some intuitively open questions (e.g. 'are french fries that everyone enjoys good?') are closed.

The third problem for taste contextualism is that it dubiously suggests that the content of any given taste assertion (i.e. the person or group supposedly described) should be indeterminate when the context of use is unknown. Imagine, for instance, that you read a taste judgment (e.g. 'ice cream is not tasty') asserted on a torn piece of an unidentifiable newspaper. In such cases, taste contextualism predicts that it would be inappropriate for you to object to the assertion without first receiving the missing context (on the grounds that the assertion's content depends on its context of use). Moreover, contextualism predicts that it would be appropriate for you to ask who the assertion is about. Neither prediction is true.

The fourth and final problem for taste contextualism is that although it can accommodate the possibility of (i) propositional taste disagreements, and (ii) faultless, but non-propositional taste disagreements, it cannot accommodate (iii) *faultless propositional* taste disagreements. This is problematic because those seem to be the default kind of taste disagreements. Consider, for instance, the following paradigm examples of taste disagreements:

(95) Olive: Matisse is better than Picasso.
Felix: No, Picasso is better than Matisse.

(96) Mary: Sushi is tasty.
John: No, sushi is disgusting.

(97) Ted: Rollercoasters are fun.
Judy: No, they're not – they're too scary to be fun.

All three paradigm examples are intuitively faultless, and resist analysis as forms of non-propositional disagreement. Moreover, the empirical literature suggests that disagreements like these are faultless and very likely propositional.[14]

[14] Cf. Cova & Pain (2012), Foushee & Srinivasan (2017), and Solt (2018).

Propositional forms of taste nondescriptivism (i.e. (**PTN**)) can neatly explain all four features problematic for taste contextualism. Deflationary remarks (e.g. 'that's subjective!') are felicitous replies to taste assertions because, according to any form of taste nondescriptivism, taste assertions do not have objective truth-values. People can distinguish between taste judgments (e.g. 'sushi is tasty') and explicit descriptions of people's preferences (e.g. 'I like sushi') because they *are* different – as (**PTN**) has it, the former express propositions about objects in the external world, whereas the latter express propositions about people. (**PTN**) can easily explain why it is appropriate to object to taste assertions when their context of use is unknown, because (**PTN**) suggests that taste judgments express the same propositions in all contexts of use. Hence, anyone who believes that ice cream is tasty has sufficient grounds to object to anyone who asserts 'ice cream is not tasty,' regardless of the context of use. Finally, (**PTN**) can accommodate the evidence that most taste disagreements are *faultless* and *propositional* because, according to (**PTN**), taste disagreements involve the assertion of inconsistent propositions without objective truth-conditions.

So far, we have seen why *Propositional Taste Nondescriptivism* (**PTN**) is preferable to both realism and contextualism. The falsity of realism (i.e. mind-*independent* descriptivism) and contextualism (i.e. mind-*dependent* descriptivism) entails some form of nondescriptivism about taste. What remains to be seen is why *propositional* taste nondescriptivism (**PTN**) is preferable to *non-propositional* nondescriptivism.

On the non-propositional view – henceforth (**NTN**) – taste judgments express non-doxastic "taste attitudes" like enjoyment and disgust, rather than beliefs. These attitudes are directed at objects and actions rather than propositions, and thus lack propositional content in the standard sense; as such, taste attitudes are not truth-evaluable. According to (**NTN**), taste disagreements involve "clashing" taste attitudes rather than beliefs with inconsistent contents. For two attitudes to "clash" is (supposedly) for them to be incompatible attitudes toward the same object or action. For example, the clash between two attitudes – love and hate – plausibly underwrites the following disagreement:

(76) Gary: I love chili.
Lucy: I disagree. I hate it.

In *Chapter 3*, I called disagreements like (76) *attitude disagreements*.

There are at least four reasons to prefer the propositional account of (**PTN**) to the non-propositional account of (**NTN**). First, disagreement markers like 'no' and 'nuh uh' cannot felicitously appear in mere attitude disagreements. For instance, substituting 'I disagree' with 'no' in (76) renders Lucy's response infelicitous:

(77) Gary: I love chili.
Lucy: #No, I hate it.

Since disagreement markers like 'no' very often appear in taste disagreements (as in (95) - (97)), it is unlikely that taste disagreements merely involve clashing attitudes.

Second, ordinary language users routinely embed taste judgments under propositional attitude verbs – including 'believes that,' but also 'hopes that,' 'wonders whether,' and 'fears that.' Not only do we understand what it is to believe (say) that the cookies are tasty – we also understand what it is to hope that they are, to wonder whether they are, to fear that they are, and so on. The best explanation of this flexibility is that taste judgments have propositional content.

The third reason to prefer (**PTN**) over (**NTN**) is that only (**PTN**) can distinguish between "taste beliefs" and underlying taste preferences. (**NTN**) generally glosses the former *as* the latter, but it is fairly clear that the two are worth differentiating. For example, suppose John is disposed to enjoy the taste of sushi but, never having tried it, asserts 'sushi is disgusting' in (96). This need not make his assertion insincere; we can easily imagine John being in the mental state required to sincerely assert that sushi is disgusting, despite his underlying disposition to enjoy it. So, there are two things here: (i) John's belief that sushi is disgusting (viz. the mental state that makes John's assertion sincere), and (ii) John's underlying taste attitude, or preference (viz. John's enjoyment of sushi). (**NTN**) glosses (i) as a non-doxastic taste attitude – but how can it be, if it does not reflect John's genuine preferences? John's *actual* taste attitude would seem to be his enjoyment of sushi – a state that is neither introspectively available to him nor expressed by his sincere assertion.

In the version of (96) that I have described, suppose that Mary believes that sushi is tasty *and* genuinely enjoys it. In that case, Mary and John intuitively seem to disagree, and yet their

disagreement cannot be in virtue of having "clashing" preferences (since Mary and John both enjoy sushi). Their disagreement must be located between their taste beliefs – or rather, between whichever mental states make both of their assertions sincere. To preserve this disagreement, (**NTN**) would have to analyze Mary and John's "taste beliefs" as clashing, non-doxastic taste attitudes – attitudes which are distinct in kind from actual preferences. But once it does, the notion of a "taste attitude" becomes fragmented and mysterious. Taste attitudes were supposed to be the mental states that make speakers' taste assertions sincere *and be* constitutive of underlying preferences. But if they cannot be both, (**NTN**) is now committed to a theoretically unstable class of mental states, apparently united only by their exclusion from doxastic explanation. At that point, the very idea of a "taste attitude" no longer does any real explanatory work. A cleaner and more natural alternative would be to differentiate (as (**PTN**) does) between (i) taste beliefs, which are propositional and expressed in assertion, and (ii) underlying dispositional preferences, which may or may not align with those beliefs. In that light, Mary and John have inconsistent taste beliefs – and the same preferences.

Fourth and finally, (**NTN**) cannot adequately address the grammatological problem. Even if an account of the grammatological properties of taste judgments (and discourse) could be developed on the assumption that they express non-doxastic attitudes without propositional content, it would still fail to explain why nearly all of these properties are shared with ordinary descriptive judgments (and discourse). As I argued in *Chapter 1*, to offer a fundamentally different kind of explanation for the same structural features is to offer no real explanation for the striking overlap. (**PTN**), by contrast, can at least hope to explain this overlap by tracing it to the underlying structure of belief. While further work is needed to show how a common foundation in belief can account for the grammatological properties shared by descriptive and nondescriptive beliefs (without collapsing the distinction between the two), (**PTN**) at least offers a viable path forward.

In sum, *Propositional Taste Nondescriptivism* (**PTN**) is well-motivated by its ability to accommodate the numerous features of taste experience and discourse which undermine realism, contextualism, and non-propositional nondescriptivism. Unlike realism, (**PTN**) respects the irrealist data – and unlike contextualism, (**PTN**) can do so without crudely recasting taste judgments (e.g. 'sushi is disgusting') as context-sensitive attitude reports with objective

truth-conditions. If both realism and contextualism are mistaken, some form of nondescriptivism must be correct. Here, (**PTN**) distinguishes itself from non-propositional (i.e. "clashing attitudes") nondescriptivism by better respecting the observed grammatological properties of taste discourse, and their deep resemblance to those of ordinary descriptive discourse. Taken together, these comparative advantages strongly motivate (**PTN**). In conjunction with (**EA**), (**PTN**) forms the semantic core of belief expressivism. The remaining three constraints aim to clarify how it is possible for beliefs to be nondescriptive, and yet still count as genuine beliefs.

5.1.3 *Phenomenal Indistinguishability* (PI)

> The license afforded to taste beliefs by taste appearances is phenomenally indistinguishable from the license afforded to most beliefs by appearances.

On a fairly basic conception of phenomenal epistemology, most beliefs are ultimately grounded in the way things *phenomenally appear* (or "seem"[15]) to us.[16] For example, the belief that the sun is bright is licensed by its phenomenal appearance *as bright*; the belief that sandpaper is rough is licensed by its phenomenal appearance *as rough*; and so on. Of course, appearances can mislead – a pencil partially submerged in water may appear *as bent*, though closer inspection (e.g. by retrieving or grasping it) reveals otherwise. In general, the license afforded to beliefs by phenomenal appearances (i.e. by "apparent qualities") is defeasible – it can be overturned, outweighed, or otherwise defeated – but always by other, more trustworthy appearances. Still, the basic point remains: if an object *x* appears *as P* to some subject *S*, then *S* is, *ceteris paribus*, licensed to believe that *x* is P.

Against this background, *Phenomenal Indistinguishability* (**PI**) is the view that the justificatory force of taste appearances is phenomenally indistinguishable from the justificatory force of most other appearances. (**PI**) claims, in other words, that there is nothing *phenomenally* unusual or deficient about the license which taste appearances (or apparent qualities) afford to taste beliefs. According to (**PI**), there is nothing within phenomenal experience itself that marks

---

[15] As I intend to use the term, a phenomenal appearance need not be *visual* (as in 'the sky *seems* blue,') or even *sensory* (as in 'this cup *seems* hot'). Any phenomenal "seeming" will do, up to and including *intellectual* appearances (as in '2 + 2 *seems* to equal 4,' and '*x seems* to logically follow from ~~*x*'). In this sense of what a "phenomenal appearance" is, rational intuitions also count as phenomenal appearances. For a more sustained defense of this position, see Huemer (2005, 2007).

[16] Cf. Huemer (2001, 2005, 2007), Chisholm (1989), Bealer (1999), Tucker (2010), and Skene (2013).

apparent taste qualities (e.g. tastiness, fun, prettiness, etc.) as illusory, non-objective, or otherwise unreal.

To be sure, I maintain that apparent taste qualities *do not* correspond to objective properties, and in that sense, are illusory – but this does not contravene (**PI**). For (**PI**) merely states that there is nothing directly in phenomenal taste experiences to indicate that apparent taste qualities are illusory. (**PI**) is therefore fully compatible with nondescriptivism about taste – and thus with (**PTN**) – so long as ordinary language users determine by some other means that objects cannot be objectively tasty, fun, pretty, etc.

The motivation for (**PI**) comes primarily from two sources: introspection, and developmental psychology. The introspective case is relatively simple. In phenomenal experience, objects generally appear as having multiple different qualities. For instance, the sun appears not only *as bright*, but also *as round*, *white, warm*, and so on. The introspective motivation for (**PI**) is simply that, when we take stock of the apparent qualities of objects, we include taste qualities. For those (like me) who enjoy mint ice cream, it appears *as cold*, *sweet*, *minty*, *tasty*, *creamy*, *melty*, and so on. Its tastiness is simply one more way it appears. My beliefs about mint ice cream – that it is cold, sweet, minty, tasty, creamy, melty – are grounded in these appearances. When prompted to describe mint ice cream, these are the predicates I reach for. If asked why I like it, I might well cite the same list. What introspection reveals is that taste qualities are no exception to the general rule: if an object appears to me *as P*, then I am, *ceteris paribus*, licensed to believe that the object is *P* – regardless of whether '*P*' is a taste predicate. And importantly, this appearance-based license operates without indicating that taste qualities are metaphysically or epistemically suspect.

This is hardly a peculiarity of taste. Phenomenal experience in general does not disclose the underlying nature of apparent qualities. When something appears a certain way, that appearance may or may not correspond to a metaphysically simple or natural property – and phenomenal experience itself offers no guidance either way. Consider the apparent quality of purple, for instance. Purple appears as just another color, like red and green. But unlike red and green – which correspond to discrete segments of the visible light spectrum – purple is a composite color, corresponding to a roughly equal mix of red and blue light (wavelengths from

opposite ends of the visible spectrum).[17] This structural fact about purple is not revealed by how it appears: purple appears as unified and basic, inasmuch as red and green do. And if phenomenal experience does not discriminate between apparent qualities that correspond to simple properties, and those which correspond to composite properties, it is natural to wonder whether it likewise fails to distinguish between apparent qualities that correspond to objective properties and those that do not. (**PI**) suggests that even if taste qualities turn out not to correspond to objective properties (as (**PTN**) suggests), we should not expect that fact to be evident from the phenomenology of taste qualities alone.

A second source of motivation for (**PI**) comes from developmental psychology. (**PI**) predicts that young children – who (i) are capable of recognizing regularities in the apparent qualities of objects, but (ii) are not yet conceptually equipped to grasp the possibility of *merely* subjective appearances – should initially treat all apparent qualities as corresponding to objective features of the world. That is, they should behave like realists about taste (i.e. about tastiness and fun), just as they do about color, texture, and shape. This prediction is borne out by Foushee & Srinivasan's (2017) third and fourth experiments, which investigated children's (aged 4 - 9.5) willingness to attribute faultlessness to taste disagreements.

Participants watched two puppets disagree over whether a novel object was, for instance, tall, spotted, or pretty. The gradable adjective used in each disagreement varied across trials: some were "*absolute*" (e.g. 'spotted'), some "*relative*" (e.g. 'tall'), and others "*subjective*" (e.g. 'pretty'). After each disagreement, participants judged whether each speaker was "wrong" or "could be right." Responses in which participants reported that both speakers "could be right" were coded as indicating faultlessness.

---

[17] To really appreciate how unique the apparent quality of purple is, contrast the fact that the visible light spectrum is *linear* with the fact that human color perception is topologically *round*. Newton, who first introduced the color wheel in 1665, mistakenly thought that red and violet (the two ends of the visible spectrum) were perceptually connected in the manner musical notes an octave apart are. In fact, the rounded geometry of human color perception is the result of the visual system computing chromaticity by comparing two opponent channels: one axis reflects the red-green (L-M) opponent channel, and the other axis reflects the blue-yellow (S-(L+M)) opponent channel. This yields a two-dimensional chromatic space (such as that modeled in CIELAB) which plots the visible spectrum on a continuous horseshoe-shaped arc from red (+, - : lower right) to yellow (0, - : lower center), to green (-, - : lower left), to blue (0, + : upper center). The human mind then "closes the loop" by mapping phenomenal purples onto the missing segment of the space: (+, +/0 : upper right to center right), corresponding to simultaneous L-cone (red) and S-cone (blue) activation without M-cone (green) activation.

The youngest children surveyed (aged 4 - 5.5) attributed faultlessness to taste disagreements (e.g. disagreements over whether an object is pretty) less than 10% of the time. The proportion of children willing to attribute faultlessness to taste disagreements rose gradually with age, but even the oldest children surveyed (aged 8 - 9.5) attributed faultlessness to taste disagreements far less often than adults (~30% vs ~100%). The fact that young children almost never attribute faultlessness to taste disagreements – and remain hesitant to do so long after they are considered fluent and to have an intact theory of mind (ages 4 - 6) – suggests not only that most children are taste realists, but also that they discover the non-objective character of taste discourse by *a posteriori* means (rather than by learning *a priori* discursive norms). I will discuss this point in more detail in the next subsection.

For now, take note that Foushee & Srinivasan's (2017) empirical results dovetail with (**PI**): if the license afforded by taste appearances (to taste beliefs) is phenomenologically indistinguishable from the license afforded by most other appearances, then it makes sense that children treat apparent taste qualities no differently than other apparent qualities. And since most apparent qualities correspond to objective properties, it is unsurprising that most children respond as taste realists. Consider a case where (i) ice cream appears *as cold* and *tasty*, (ii) the license afforded by these two appearances are phenomenally indistinguishable (i.e. there are equal grounds to believe that ice cream is cold as there are to believe that ice cream is tasty), and (iii) like most apparent qualities, coldness corresponds to an objective property. From a child's point of view – why not tastiness, too? Adults implicitly know the answer: they implicitly understand that taste appearances are irreducibly subjective, and that not all appearance-licensed beliefs track objective properties. But such distinctions take conceptual maturity to appreciate – and children are still in the process of learning the different ways that appearances and reality can part ways.

In conclusion, *Phenomenal Indistinguishability* (**PI**) draws support from both introspection and developmental psychology. Introspectively, apparent taste qualities like tastiness and fun present themselves just as straightforwardly as other apparent qualities like color or texture. They fall under the same general epistemic rule: if an object *x* appears *as P*, then *ceteris paribus*, one is licensed to believe that *x* is P. Introspection reveals that the manner in which taste beliefs are licensed by appearances is utterly mundane – if apparent taste qualities

turn out not to correspond to objective properties (as (**PTN**) suggests), it will *not* be because they are self-evidently unreal.

Developmental psychology complements this picture. Almost all young children respond like taste realists by indicating that taste disagreements are faultful. A supermajority persists in doing so even after acquiring theory of mind. This suggests that, from the point of view of children, there is nothing in their experience that sets taste qualities (e.g. tastiness, fun, prettiness, etc.) apart as non-objective or unreal. *A fortiori*, there is nothing in phenomenal experience that immediately sets taste qualities apart as non-objective or unreal. If (**PI**) were false, and taste appearances *did* signal their illusory or subjective nature, this pattern of responses would be difficult to explain.

The upshot of (**PI**) for belief expressivism is that it helps explain why taste beliefs should count as genuine beliefs, even if – as (**PTN**) would have it – the apparent qualities that license them do not, in the final analysis, correspond to objective properties. According to (**PI**), the license afforded to taste beliefs is *phenomenally indistinguishable* from the license afforded to most beliefs. That explains why, as children, we initially form taste beliefs – and naively assume they describe objective features of the world. But (**PI**) also helps explain why we continue to hold onto such beliefs after withdrawing from taste realism. Whether or not we are taste realists, objects persist in appearing to us *as tasty*, *fun*, *pretty*, and so on – vividly, and with striking regularity. These appearances command a kind of epistemic respect, and we continue to show that respect by forming taste beliefs. Even if we no longer take them to represent objective properties, our taste beliefs remain licensed by appearances in the ordinary way – and so remain, quite naturally, beliefs.

5.1.4 *A posteriori Non-objectivity* (APNO)

> It is only possible to learn that a class of judgments is non-objective by *a posteriori* means. It is discovered over time – through reflection on our practices and experiences – that certain phenomenal appearances (i.e. "apparent qualities") do not correspond to objective properties.

A Posteriori Non-Objectivity (**APNO**) is the view that the non-objectivity of a discourse is something we can only grasp through *a posteriori* reflection on the anomalous features of non-objective judgments. Non-objectivity is not a matter to be known *a priori* merely by

consulting the linguistic and social conventions of use from which we derive our concepts. Rather, we come to know that a discourse is non-objective when the familiar marks of objectivity – independent verifiability, theoretical convergence, and natural consequences for error – fail to materialize. Norms about disagreement may influence which discourses people *regard* as non-objective, but, as (**APNO**) maintains, the non-objectivity of any given discourse is ultimately a substantive matter; it is possible for norms about disagreement to mislead one into thinking that an objective discourse is non-objective, and vice versa. In this respect, (**APNO**) contrasts with Kölbel's (2002) view, according to which objectivity and non-objectivity are determined by *a priori* discursive conventions that "specify the topics disagreement on which indicates a mistake and warrants discussion, and other topics disagreement on which does not indicate a mistake[.]" (p.105)

Specifically, Kölbel (2002) defends the view that it is part of "the function of our concepts" – about *red* and *tasty*, say – that disagreements about the former indicate a mistake, and that disagreements about the latter do not. Since competent speakers inherit their concepts (and *a fortiori* "the function" of their concepts) from discursive conventions of use which mark the topics on which disagreements indicate a mistake and which not – it follows (on Kölbel's (2002) view) that whether a discourse is objective is a matter to be known *a priori*, by consulting the functions of our concepts and the discursive norms of use from which they derive. As Kölbel (2002) writes,

> Now, how do we know that a proposition is objective? We seem to know this a priori, that is, we need to know only what a disagreement is about, and that is enough for us to know whether error must be involved. But what is the source of this knowledge? … Why do we know a priori that if two representational devices, two cameras say, yield conflicting representations, then one of them must be malfunctioning? It's because we know [that it] is part of the concept of a camera that well-functioning cameras cannot yield diverging outputs. … That's why we can know a priori what it is for a camera to function properly. The same goes for our a priori knowledge that disagreements on certain topics must be the result of some mistake. For this is part of the function of certain beliefs, and of their linguistic expression. It is part of the function of beliefs on whether the lights are on that if two such beliefs are contradictory, a mistake has occurred. It is not part of the function of judgments of taste that divergences are the symptom of some mistake. …
> On this theoretical background, I now want to put forward a new criterion of objectivity[.] … For any *p*: it is an objective matter whether *p*, just if:
>
> (CO) For all thinkers *A* and *B*: it is a priori that if *A* believes that *p* and *B* believes that not-*p* then either *A* has made a mistake or *B* has made a mistake. (pp.30-31)

Crucially, Kölbel (2002) intends (CO) to be the *criterion* – or litmus test – by which we determine that a discourse is objective. On his view, then, we come to know whether a topic is objective by *a priori* reflection – and ultimately, by attending to the linguistic conventions of use which inform our concepts, and which "specify the topics disagreement on which indicates a mistake and warrants discussion[.]" (p.105)

The flaw with Kölbel's (2002) criterion is simply that linguistic conventions can *mislead* us about which discourses are objective (and which not) – but so far as Kölbel's (2002) criterion is concerned, the objectivity of a discourse *just is* a matter of linguistic conventions. Kölbel's (2002) criterion is thus either inaccurate, or implies (absurdly) that whether a discourse is *objective* is itself a *relative* matter – as Kölbel's (2002) criterion would have it, for a discourse to be objective is little more than for it to be *regarded as* objective by linguistic conventions.

Imagine, for instance, a society whose linguistic conventions mark disagreements over the tastiness of borscht as *faultful* (i.e. indicating a mistake) and disagreements about the mass of the moon as *faultless* (i.e. *not* indicating a mistake). Kölbel's (2002) criterion predicts that a member of this society can come to know that the tastiness of borscht is objective – and that the mass of the moon is non-objective – *a priori*, simply by reflecting on the functions of the concepts *tasty* and *mass*, and on the linguistic conventions of use from which they derive.

Suffice to say, Kölbel's (2002) criterion is untenable. The kernel of truth at its core is that the source of much of our knowledge is authority – specifically, the various claims which a society disseminates (explicitly or otherwise) *as* facts, especially to children. By extension, a not insignificant source of our knowledge about which discourses are objective (and which not) is likely testimonial – for instance, we are consistently *told* that tastiness is subjective and that mass is objective. More to the point, our linguistic conventions specify that disagreements over mass indicate a mistake, and that disagreements over tastiness do not – and we take these conventions, on authority, as decent reasons to believe (or at the very least, to entertain the possibility) that tastiness is subjective and that mass is objective. But the *ultimate* source of knowledge is not authority; we understand well enough that objectivity is ultimately an objective matter. We understand that if the mass of the moon is objective, then it is the same for everyone, regardless of who you ask, and regardless of whether some set of discursive norms – however popular –

regards mass as non-objective (viz. by specifying that disagreements over mass do not indicate a mistake).

The fundamental problem with Kölbel's (2002) criterion can perhaps be put this way: whether a linguistic community *should* mark disagreement over a topic as indicating a mistake is itself an open question. The answer to that question, for any given topic *T*, depends on whether *T* is objective. No linguistic community can determine the objectivity of *T* by *a priori* means – they certainly cannot do so by circularly reflecting on their own discursive norms. And so the ultimate grounds for our knowledge about which topics are objective is not *a priori*. According to (**APNO**), our knowledge of which topics are objective (and which not) ultimately rests on our *a posteriori* (and often implicit) understanding of the ways in which certain discourses are anomalous.

Although there is ample independent motivation for (**APNO**), it's worth briefly highlighting that it follows as a matter of logical consistency from the conjunction of (**PTN**) and (**PI**). If taste discourse is nondescriptive, then mature speakers recognize that apparent taste qualities fail to correspond to objective properties. But if the license afforded to taste beliefs by appearances is phenomenally indistinguishable from the license afforded to ordinary descriptive beliefs by appearances – as (**PI**) claims – then recognition of the non-objectivity of taste discourse cannot be immediate, or concomitant with competently grasping the meanings of taste judgments. The only remaining option is for the non-objectivity of taste discourse to be determined later – perhaps by reflection on the anomalous features of taste judgments and discourse.

Perhaps unsurprisingly, the empirical evidence that supports (**PI**) also supports (**APNO**). Foushee & Srinivasan's (2017) third and fourth experiments show that even among older children (aged 8 - 9.5), only ~30% attribute faultlessness to taste disagreements. This is striking, given that by age 6, children are generally considered both to have developed theory of mind and to be fluent speakers. These results suggest, in other words, that most children respond as taste realists despite grasping that others have different taste experiences and beliefs, and despite competently understanding how to use concepts like 'tasty' and 'fun' – well enough, anyways, to understand which taste beliefs are licensed by which experiences. In sharp contrast, Foushee &

Srinivasan's (2017) first two experiments demonstrate that nearly all young adults attribute faultlessness to taste disagreements. Taken together, these findings suggest that the non-objectivity of taste discourse is discovered *a posteriori* – indeed, somewhere between the ages of 10 and 20.

It is likely during this period that we take notice of the anomalous features of taste discourse – features which tell against taste realism (some of which I reviewed in Section 5.1.2). For example, we might reflect on the fact that there are no established methods for decisively settling taste disagreements. Or, perhaps, on the fact that there are no natural corrective repercussions for "objectively mistaken" taste beliefs: if one group believes that apples are tastier than oranges, another claims the reverse, and the beliefs of both groups cohere with their preferences – we can scarcely even imagine how either group could be made to suffer the sorts of consequences nature reserves for objectively false beliefs. Moreover, we may observe that objects may share the same apparent qualities (e.g. tastiness, or fun) without sharing any stable set of objective properties. For example, ripe tomatoes and peanuts may both appear *as tasty* (to the same person) despite differing radically: ripe tomatoes may appear *as tasty* because they are juicy, slightly acidic, and savory, while peanuts may appear *as tasty* because they are nutty, earthy, and slightly sweet. In each case, the apparent quality of tastiness supervenes a *different* set of objective properties – the two foods are tasty for *different* reasons. We may notice that tastiness does not, without exceptions, track any particular objective property or combination of properties – such as salt, sugar, or fat content. That some apparent qualities supervene on objective features without corresponding to them is a distinct possibility, and one which underwrites (**APNO**): for it is only by gradually recognizing peculiarities like these that we come to appreciate that some discourses are non-objective.

5.1.5 A new theory of belief

What are beliefs? First, the background truisms: beliefs are representational mental states that aim at truth. They are propositional attitudes: to believe that *p* is to take a particular stance *towards* the truth-evaluable proposition that *p* – in particular, the stance that *p* is true. Moreover (as I reviewed in Section 5.1.1), the belief that *p* is (typically) the mental state that a person is required to have to agree with an assertion of the judgment '*p*' – it is the mental state that a

person is required to have in order for an assertion of '*p*' to be sincere. Listeners assume that assertions are sincere, and speakers understand this fact, and this yields (**EA**): that to assert '*p*' is (typically) to present oneself as believing that *p*.

In the lead up to (**PI**) (in section 5.1.3), I discussed another feature of belief, from phenomenal epistemology: most beliefs are licensed (i.e. justified) by phenomenal appearances. The general rule is that if an object *x* appears *as P* to some subject *S*, then *S* is, *ceteris paribus*, licensed to believe that *x* is P. In other words, appearances have a certain phenomenal character – objects *seem* to be a certain way. The way that objects *seem* (i.e. their apparent qualities) command our epistemic respect, and we show that respect by forming beliefs about objects – by attributing predicates to them that capture how they seem.

A natural picture of belief – call it *Atomic Representationalism* (**AR**) – treats beliefs as primary (or basic) representational states that stand in a one-to-one correspondence with lexicalized propositions. On this view, to believe that *x* is P is simply to stand in a *sui generis* mental relation to the proposition that *x* is P, much as desiring or hoping might be seen as standing in distinct attitudinal relations to that same proposition. Beliefs, on this model, are not grounded in or constructed from more basic representational elements. They are *atomic*: each belief is its own self-standing mental state, tied to a specific truth-evaluable content, typically presumed to be expressible using the predicates of a natural language.

While few theorists explicitly endorse (**AR**) in its pure form, aspects of it are widespread. In symbolic AI and cognitive science, belief systems are often treated as databases of discrete propositions, each stored and manipulated atomically. Formal semantic frameworks – such as Hintikka's (1962) treatment of belief as an accessibility relation over possible worlds, and Stalnaker's (1984) modeling of belief via context sets – abstract away from questions about the psychological structure of belief, but nevertheless encourage a picture where beliefs are treated as basic commitments to specific propositions. These theorists model agents in ways that suggest that beliefs are discrete commitments to a set of propositions, available for inference in a way that makes belief seem representationally basic – even though they acknowledge that these are idealizations, not psychological theories. The theoretical appeal of (**AR**) is that it simplifies

representation, supports formal modeling of belief systems, and pairs naturally with logical and probabilistic models of inference (see Fagin et al. (1995)).

Even theorists like Fodor (1975), who argue that internal thought must have a compositional structure, still assume that beliefs correspond to discrete mental representations – something like a sentence in an internal "language of thought" (viz. "Mentalese"). Fodor highlights the *productivity of belief*: our ability to generate an infinite number of beliefs from a finite base. This suggests that our internal representations must have a compositional structure. In Fodor's view, to believe that *x* is P is to have a specific Mentalese sentence in the mind that represents the proposition that *x* is P. Although Mentalese sentences are syntactically structured entities, each belief corresponds to a commitment to a particular Mentalese sentence. This preserves a key feature of (**AR**): that beliefs are discrete commitment states – albeit, commitments to internally structured Mentalese sentences, rather than to lexicalized propositions.

For these reasons, (**AR**) should be seen less as a fully endorsed theory than as a widespread and useful idealization – one that captures important aspects of our intuitive and theoretical thinking about belief, even if it ultimately oversimplifies the architecture of cognition. In what follows, I will highlight two major difficulties for (**AR**) – difficulties that suggest it cannot serve as an adequate model of belief's underlying structure.

The first cost is theoretical incoherence. (**AR**) forms an inconsistent triad with two plausible claims about belief:

1. *Atomic Representationalism* (**AR**): Beliefs are basic, *sui generis* representational states.
2. *Indefinitely Many Beliefs*: We possess a potentially infinite number of beliefs. This follows from any plausible version of epistemic closure: if I believe that *x* weighs ten pounds, then intuitively, I also believe that *x* weighs at least five pounds, that *x* does not weigh twenty pounds, that *x* has weight, and so on, across indefinitely many valid entailments.
3. *Finite Architecture*: Our cognitive architecture is finite. We do not have the capacity to maintain an infinite number of *sui generis* representational mental states.

At least one of these claims must go. One way to resolve this tension is by denying (2), either by rejecting logical closure (Dretske (1970); Cherniak (1986); Harman (1986)), or by drawing

distinctions between beliefs one "has" and those one is merely "disposed to infer" in appropriate contexts (Fodor (1975), Dennett (1978)).[18] But such distinctions are *ad hoc* and run against our ordinary understanding of belief: if I say that my dog is four years old (and I am sincere) – it would be bizarre for you to ask whether I also think she is less than ten years old. And it would be equally strange for me to respond, "Well, I hadn't thought about it." We ordinarily assume that people have all manner of beliefs that are necessitated by what they think. Rejecting (3) is out of the question – cognitive finitude is a psychological and biological fact. That leaves us with (1): the idea that beliefs are basic must be rejected. Beliefs must be derived from something more fundamental – some underlying representational structure from which indefinitely many propositional commitments can be generated. That underlying structure, I will later suggest, is a system of "views."

The second problem for (**AR**) is phenomenological and linguistic. (**AR**) cannot accommodate the fact that our internal representations are often more fine-grained than the predicates and sentences we use to express them. In our phenomenal experiences, the world regularly appears in ways that outstrip our vocabulary. Suppose I look at a wall painted in a complex teal that lies somewhere between turquoise and cyan – a color for which I have no precise word. Still, I form a stable and specific representation of how it looks, and consequently, I am disposed to affirm certain propositions (e.g. "the wall is blue-green") and reject others (e.g. "the wall is navy") based on how well they match that internal representation. This suggests that I have a determinate "view" of the wall's color – a specific representational state – even if I lack the resources to express it precisely. The same is true for taste, texture, musical tone, emotional expression, and more. In such cases, we reach for predicates that approximate how things seem – and ultimately, what we actually think – knowing full well that predicates fail to capture either exactly.

This everyday phenomenon undermines (**AR**)'s assumption that beliefs are lexicalized propositional attitudes, and points again toward a more flexible picture of how thought and

[18] Some version of the distinction between beliefs one "has" and those one is merely disposed to infer is worth preserving. But construing it in the standard way risks putting ordinary talk of what a person believes unnecessarily in the crossfire: it is often natural to attribute a belief to a person even when they are merely disposed to infer its content. (**FGR**) preserves the relevant distinction between stored propositions (views) and propositions one is merely disposed to infer (beliefs). This formulation of the distinction better conforms with ordinary linguistic practice.

language interact. On this picture, (i) our actual cognitive states – "views" – are more fine-grained than the sentences we use to express them, and are commitments to the truth of propositional contents that are generally unlexicalized; (ii) propositional contents themselves are structured sets of possibilities, situated in domains that model degrees of similarity and difference between contents, rather than treating propositions as mere truth-sets; and (iii) we count as believing all the propositions which are supersets of any view we have – even if we never recognize or articulate those propositions. Such a framework makes it possible to reconcile our finite cognitive architecture with the possibility of having indefinitely many beliefs. It also explains why assertion so often feels like an exercise in approximation: we reach for the most apt combination of words available to triangulate our actual view. Rather than picturing the mind as a sprawling database of atomic beliefs, this suggests a layered system of stored views that license indefinite propositional commitments. These considerations point toward the need for a representational framework that can accommodate fine-grained mental states and cognitive finitude.

A central drawback of possible worlds semantics, in its canonical form, is that it builds objectivity into the very notion of truth: a sentence *S* is true in some world *w* just in case *S* describes *w* accurately – that is, its content obtains in *w*. This conflation of truth with obtaining in a world makes the framework ill-suited for contents that are plausibly truth-evaluable yet non-objective, like taste, aesthetic, and moral judgments. These judgments do not describe possible worlds in the straightforward way ordinary descriptive judgments do. As discussed in Section 4.3, this tension led Gibbard (1990) and Lasersohn (2005) to supplement possible worlds semantics with additional coordinates – systems of norms, perspectives, or judges – in order to model truth in moral and taste domains, respectively. However, these extensions are notoriously *ad hoc* and difficult to interpret: what does it mean, for example, for a moral content to obtain relative to a factual-normative world pair <***w***, ***n***>, or for a taste content to obtain relative to a world-judge pair <***w***, ***j***>? That possible worlds semantics must be bent out of shape in this way underscores the need for an alternative representational framework – one that treats non-objective contents as truth-evaluable without first shoehorning them into a metaphysics of obtaining-in-a-world. (**FGR**) is designed precisely for this role.

5.1.6 *Finite Geometric Representationality* (FGR)

> Primary (or basic) representational mental states – "views" – are finite in number, and rarely map cleanly onto propositions expressible using the predicates of a natural language. Views and propositions can be modeled as regions in geometric conceptual spaces (attributed to objects). Beliefs are secondary (or derivative) states: a person believes every proposition whose associated region fully contains one of their views. The exact relation between views and beliefs is given as:
>
> > S believes that *x* is P iff the region (in the contextually salient domain **D**) that models S's view about *x* is a subset of the region (in **D**) that models P.

Finite Geometric Representationality (**FGR**) is a model of mental representation and the structure of belief that builds directly on the conceptual spaces framework developed by Gärdenfors (2000, 2014, 2025). Like Gärdenfors, (**FGR**) models predicates as regions in geometric spaces (e.g. color, height, time, evaluative taste) structured by cognitively salient quality dimensions (e.g. hue, length, tastiness). I follow Gärdenfors in calling these geometric spaces "domains."

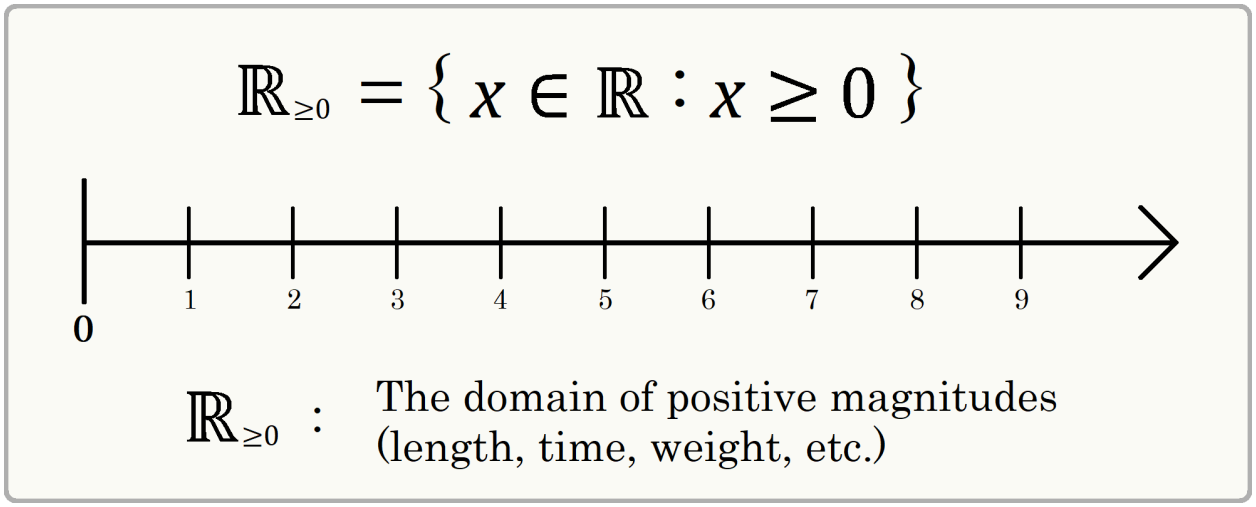

Figure 5: $\mathbb{R}_{\geq 0}$: The set of positive real numbers. This mathematical structure underpins several one-dimensional quality domains – such as length, time, weight, speed, loudness, force, pressure, energy, etc.

Where (**FGR**) goes beyond Gärdenfors' framework is in modeling propositional content[19] and belief. (**FGR**) treats propositional content as a pairing between regions and objects; that is, a proposition corresponds to the attribution of a region (a predicate) to a particular object. Actual representational mental states – what I call "views" – are cognitive commitments to the truth of a finite set of propositional contents: in effect, a view is an attribution of a region to an object cognitively stored as true. But instead of identifying beliefs with these stored views, (**FGR**) models belief as a *derivative* state: to believe a proposition *p* is to have a view whose associated region is fully enclosed by the region associated with *p*.

This conception of belief as a geometric relation between views and propositions promises to explain numerous puzzles about belief – including why it is natural to attribute indefinitely many beliefs to people with finite representational states, why assertions are often imprecise, how it is possible to almost (and barely) believe propositions, and how to formally distinguish *de dicto* from *de re* belief. To see how these explanatory advantages hang together, I will now reintroduce each core component of (**FGR**), beginning with domains.

I follow Gärdenfors (2000, 2014, 2025) in modeling a domain as the set of all points in an *n*-dimensional geometric space, generally defined as the set of all points yielded by a Cartesian or cross product (×) of cognitively salient quality dimensions, such as length, time, temperature, weight, and so on. I say "generally," because (**FGR**) also permits (i) any proper subset of an *n*-dimensional quality space (a subdomain), and (ii) any augmentation of an *n*-dimensional quality space with additional values not already included in the space (an enriched domain) to serve the semantic function of domains.[20] While all quality dimensions constitute

[19] Gärdenfors (2014, 2025) eschews the Fregean view that sentences express propositional contents, and claims instead that sentences express *events*: a structured quadruple that consists in (i) an agent (viz. the object performing the action), (ii) a patient (viz. the object acted upon), (iii) a force vector (viz. the action performed on the patient), and (iv) a result vector (viz. the change induced on the patient). To borrow Gärdenfors' (2025) example: 'Oscar pulls a sled to the top of the hill' refers to the event wherein Oscar (agent) pulls (force vector) a sled (patient) up the hill (result vector). Unary copular sentences – like 'the ball is red' – are said to express *states*: a special form of event, in which the agent, force vector, and result vector are void, leaving only a static snapshot of the patient and its qualities.

[20] By an "enriched domain," I mean any set that contains a structured region (e.g. color space) and one or more additional elements that fall outside that structure (e.g. the singleton {colorless}). Enrichment in this sense refers to a disjoint union – not a Cartesian product or dimension extension.

their own 1-dimensional domain (e.g. time, weight, length), familiar examples of multi-dimensional domains include color space (L* × a* × b*), physical space (length × width × height), temperature sensation (warmth × humidity × wind speed), and non-evaluative taste (sweetness × sourness × bitterness × saltiness × umami × texture × spiciness).

A region of a domain is defined as any subset of that domain. In (**FGR**), predicates – including those expressed by basic adjectives – are modeled as regions of contextually salient domains.[21] For example, the basic adjective 'red' corresponds to a specific region of color space, and 'burgundy' picks out an even smaller region of that space (itself a proper subset of 'red'). 'Hot' and 'cold' pick out non-overlapping regions at opposite ends of the temperature domain. 'Saccharine' corresponds to a region of the non-evaluative taste domain which models relatively high sweetness, while 'tasteless' indicates relatively low values across all dimensions in that space.

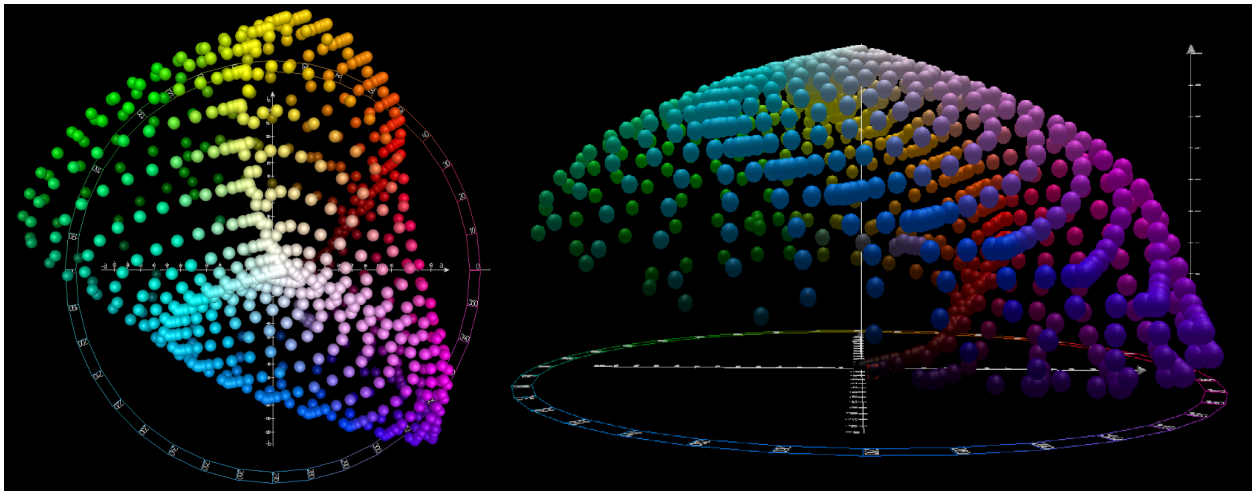

Figure 6: *L*a*b** color space: a three-dimensional space modeling human color perception. *L*a*b** was defined by the International Commission on Illumination (CIE) in 1976. The vertical axis, *L**, represents human perceptual lightness. On the chromaticity plane, *a** (the

[21] Most – but importantly not all – predicates (and the adjectives that express them) will be modeled as *continuous* and *convex* regions: as regions containing every point within a boundary, such that any line segment connecting two points in the region lies entirely within it. Many predicates, however, are not convex. In fact, for any predicate P modeled as a convex region (e.g. 'red') in some domain D, the negation of P (e.g. 'not red') will pick out the complement of P's region in D – and the complement of any convex region will generally be concave. Moreover, some positive predicates are neither convex nor continuous. For example, 'odd' and 'even' pick out non-continuous, concave regions of the domain of countable numbers.

"red-green" axis) and *b** (the "blue-yellow" axis) approximate the two chromatic opponent channels in human vision: L-M (red-green) and S-(L+M) (blue-yellow), respectively.

By modeling predicates as regions of domains, similarity and inclusion relations between predicates emerge naturally from geometric structure. If Tree A is tall and Tree B is short (and these uses of 'tall' and 'short' correspond to regions in the same contextually-determined domain), then we can infer that A is taller than B, merely by virtue of the fact that 'tall' corresponds to a range of values (in the domain of height) greater than those picked out by 'short.' Similarly, we can know that nothing that is red all over is blue because the regions of color space corresponding to 'red' and 'blue' do not intersect. And we know that a hot day is more similar to a warm day than to a cold day (with respect to temperature) because we know that in the temperature domain, the region corresponding to 'hot' is closer to the region corresponding to 'warm' than it is to the region corresponding to 'cold.' And we know that burgundy is a form of (or way of being) red because the region corresponding to 'burgundy' is a proper subset of the region corresponding to 'red.' And so on.

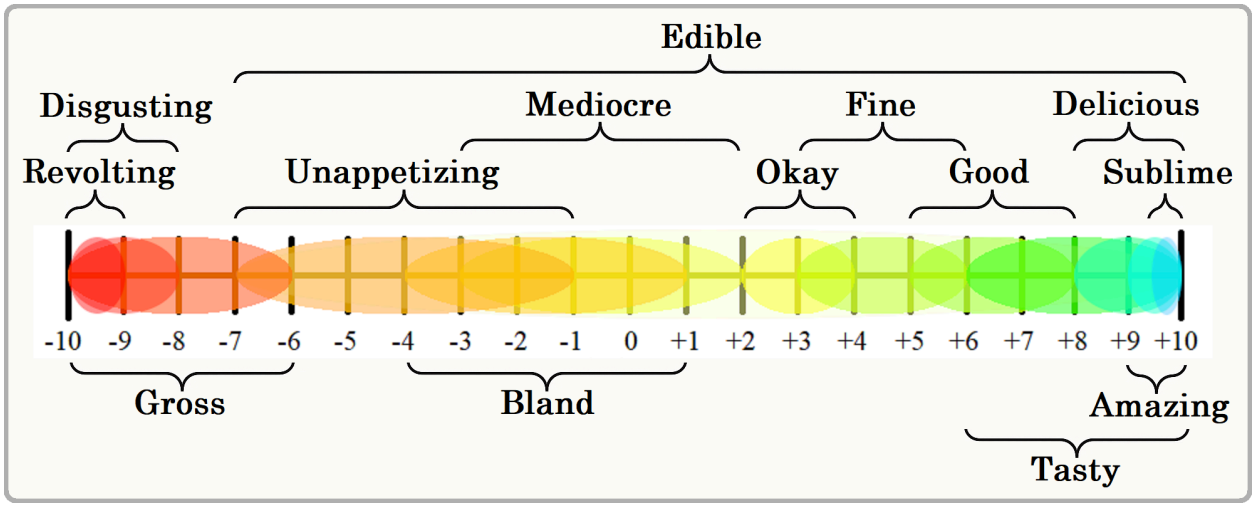

Figure 7: (Mock-up) Evaluative Taste: a one-dimensional space modeling the valence of evaluative taste. Higher values indicate tastier; lower values indicate less tasty. The elliptical shape and color-coding of regions are exclusively for readability. Normalization of the domain range from -10 to +10 is likewise for readability. Placement of predicate regions is inexact, but illustrative of how predicate regions (i) vary in size, (ii) may be nested, (iii) may partially overlap, and (iv) may fail to intersect.

If adjectives are regions in domains, the existence of any meaningful adjective implies a domain in which it is represented as a region. This is clearest with families of inferentially-related, gradable adjectives (e.g. height, color, temperature, and size adjectives). But

the same principle extends to more exotic cases, when the domain is not pre-theoretically obvious. For instance, 'running,' 'jogging,' 'sprinting,' and 'walking' plausibly correspond to different regions in a shared domain for terrestrial locomotion, given that jogging and sprinting are kinds of running, no form of running is walking, and no form of jogging is sprinting.

Let the region picked out by predicate P in domain **D** be denoted as $[\![P]\!]^{\mathbf{D}}$. As an example of how this notation works, recall that according to (**FGR**), to say that burgundy is a form of red is to say that the region of color space which corresponds to 'burgundy' is a proper subset of the region which corresponds to 'red.' This is expressible as $[\![\text{burgundy}]\!]^{\mathbf{C}} \subset [\![\text{red}]\!]^{\mathbf{C}}$, where **C** is the domain of color space. Likewise, 'jogging and sprinting are both forms of running, but no form of jogging is sprinting' is expressible as $(([\![\text{jogging}]\!]^{\mathbf{TL}} \cup [\![\text{sprinting}]\!]^{\mathbf{TL}}) \subset [\![\text{running}]\!]^{\mathbf{TL}}) \wedge ([\![\text{jogging}]\!]^{\mathbf{TL}} \cap [\![\text{sprinting}]\!]^{\mathbf{TL}} = \varnothing)$, where **TL** is the domain of terrestrial locomotion.

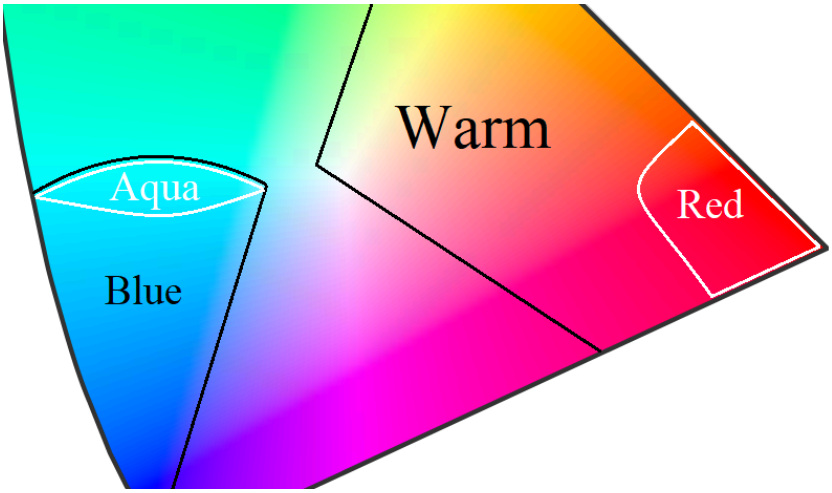

Figure 8: Partial CIE 1931 *xy* chromaticity diagram with semantic annotations. Let **X** denote the CIE *xy* chromaticity space. The regions in **X** picked out by the predicates 'aqua,' 'blue,' 'warm,' and 'red' are denoted $[\![\text{aqua}]\!]^{\mathbf{X}}$, $[\![\text{blue}]\!]^{\mathbf{X}}$, $[\![\text{warm}]\!]^{\mathbf{X}}$, and $[\![\text{red}]\!]^{\mathbf{X}}$, respectively. The judgment 'aqua is a shade of blue' is true iff $[\![\text{aqua}]\!]^{\mathbf{X}} \subset [\![\text{blue}]\!]^{\mathbf{X}}$. Likewise, 'red is a warm color' is true iff $[\![\text{red}]\!]^{\mathbf{X}} \subset [\![\text{warm}]\!]^{\mathbf{X}}$.

Now, if predicates are modeled as regions of domains, propositions can be modeled as the *attribution* of a region (i.e. predicate) to an object. In (**FGR**), the content of 'the ball is red' is modeled as a region in color space – $[\![\text{red}]\!]^{\mathbf{C}}$ – attributed to the ball. To attribute a region $[\![P]\!]^{\mathbf{D}}$ to an object *x* is to claim that *x* has a quality corresponding to some point within $[\![P]\!]^{\mathbf{D}}$. This can be modeled extensionally as the set of all the ways for *x to be* P – that is, as the set of all ordered

pairs which have x as their first element, and any value within $[\![P]\!]^{\mathbf{D}}$ as their second element (i.e. all $\langle x, y\rangle$ such that $y \in [\![P]\!]^{\mathbf{D}}$). This is mathematically equivalent to the cross product $\{x\} \times [\![P]\!]^{\mathbf{D}}$. If we let **O** be the domain of all objects, we can rewrite the singleton $\{x\}$ as $[\![x]\!]^{\mathbf{O}}$. We can then express the propositional content of 'x is P' (notated $[\![x \text{ is P}]\!]$) as:

(105) $[\![x \text{ is P}]\!] \stackrel{\text{def}}{=} [\![x]\!]^{\mathbf{O}} \times [\![P]\!]^{\mathbf{D}}$

Note that $[\![x]\!]^{\mathbf{O}} \times [\![P]\!]^{\mathbf{D}}$ will necessarily be a region of the product space $\mathbf{O} \times \mathbf{D}$. What this means is that (**FGR**) formally models propositions as regions of object-predicate spaces; if $[\![x \text{ is P}]\!]$ is equivalent to $[\![x]\!]^{\mathbf{O}} \times [\![P]\!]^{\mathbf{D}}$, then $[\![x \text{ is P}]\!]$ corresponds to a region of $\mathbf{O} \times \mathbf{D}$.

While (105) captures the attribution of $[\![P]\!]^{\mathbf{D}}$ to x extensionally, we can streamline (**FGR**)'s representation of propositions by capitalizing on the fact that to attribute $[\![P]\!]^{\mathbf{D}}$ to x is to claim that x has a quality corresponding to some point in $[\![P]\!]^{\mathbf{D}}$ – and *a fortiori*, some point in **D**. Let this point be the output of a function $f^{\mathbf{D}}$, which maps objects $x \in \mathbf{O}$ to points in **D**; notationally: $f^{\mathbf{D}} : \mathbf{O} \mapsto \mathbf{D}$. For example, $f^{\mathbf{C}}$(the ball) yields the exact point in color space corresponding to the ball's color, while $f^{\mathbf{TS}}$(the air) returns the temperature, humidity, and wind speed of the air (in the domain of temperature sensation **TS**). Using $f^{\mathbf{D}}$, the propositional content of 'x is P' can now be modeled (more directly than before) by the claim that that x has a quality corresponding to some point in $[\![P]\!]^{\mathbf{D}}$:

(106) $[\![x \text{ is P}]\!] \stackrel{\text{def}}{=} f^{\mathbf{D}}(x) \in [\![P]\!]^{\mathbf{D}}$ [22]

This definition provides transparent truth-conditions: 'x is P' is true iff x has a quality corresponding to a point in **D** which is an element of the region of **D** corresponding to 'P'. So, 'the ball is red' is true iff the precise color of the ball counts as a way to be red, and 'sushi is delicious' is true just in case the exact tastiness of sushi (viz. a point in evaluative taste space) counts as a way to be delicious (i.e. is an element of the region in evaluative taste space corresponding to 'delicious').

[22] It is technically also required that $f^{\mathbf{D}}(x) \neq \emptyset$ (that x has a property which corresponds to a non-empty point in **D**), since for any $[\![P]\!]^{\mathbf{D}}$, it is trivially true that $\emptyset \in [\![P]\!]^{\mathbf{D}}$. For the sake of visual clarity I leave the requirement that $f^{\mathbf{D}}(x) \neq \emptyset$ unstated. Unless specified otherwise, it may be assumed.

With (**FGR**)'s machinery of domains, predicates, and propositions in place, I can now define actual representational mental states. An actual representational mental state – a "view" – is a proposition cognitively stored as true. To cognitively store a proposition as true is, in effect, to include that proposition in a special set (the finite set of propositions which we use to navigate the world and solve practical problems, and which we are constantly vetting and updating via experience and reason). We can therefore model the views of a person, S, as the finite set of propositions which S stores as true.[23] And since (**FGR**) models propositions as regions in object-predicate (**O** × **D**) space, (**FGR**) models S's views as a set of regions in object-predicate space – the finite set of regions S stores as true. Equivalently, (**FGR**) models any one of S's views as a specific region of some domain **D** attributed to some object *x* – an attribution which S stores as true.

Now, let the region in **D** corresponding to S's view about *x* be given by a function $v^{\mathbf{D}}$, which maps agents S ∈ **A** (the set of all agents) and objects *x* ∈ **O** to regions in **D**; notationally: $v^{\mathbf{D}} : \mathbf{A} \times \mathbf{O} \mapsto \mathscr{R}(\mathbf{D})$. For instance, $v^{\mathbf{C}}$(Mary, tomatoes) returns the region of color space that corresponds to Mary's view about the color of tomatoes. Likewise, $v^{\mathbf{ET}}$(John, chili) yields the region of evaluative taste space, **ET**, that corresponds to John's view about the tastiness of chili.

An important consequence of modeling both predicates and views about objects as regions in domains is that a person's views need not – and in general, will not – perfectly align with any basic predicate in their lexicon. For instance, I may fail to have a color concept which corresponds exactly to my view about the color of an object when its color is an unusual hue. Or, I may lack a temperature concept which corresponds exactly to my view about the temperature of water when it is hotter than warm but not quite hot. This mismatch is to be expected: given that the axes of domains tend to be continuous (e.g. temperature can vary infinitesimally), and that predicates and views about objects both correspond to discrete regions of domains, it will

[23] Since cognitive resources are finite, people generally do not store every proposition they are disposed to call "true." It is much more efficient to store only the *most specific* proposition which one can be confident in. For instance, if I observe six or seven birds in flight (and am uncertain as to which), the best use of my limited cognition is to store only the proposition that there are six or seven birds in flight. It would be wasteful for me to store the propositions that there are less than sixty birds in flight, that at least two birds are in flight, and so on. These other propositions can be derived (if the need arises) from the proposition that there are six or seven birds in flight.

only rarely occur that they perfectly overlap (i.e. that there is some predicate P in S's lexicon such that $v^{\mathbf{D}}(S, x) = [\![P]\!]^{\mathbf{D}}$).

Though also not guaranteed, it is far more common for a person's views about objects to be proper subsets of predicates in their lexicon (i.e. that there is some predicate P in S's lexicon such that $v^{\mathbf{D}}(S, x) \subset [\![P]\!]^{\mathbf{D}}$). At least in part, this is because views about objects (and especially those licensed by experience) tend to be more fine-grained than the lexical categories available to an agent; views about objects typically correspond to much smaller regions of domains than basic predicates do. For instance, when Mary observes a red ball, she typically observes a *specific shade* of red. Ideally, Mary would represent the ball's color as a single point in color space – but realistically, she mentally represents the ball's color as a relatively small region in the vicinity of that point (reflecting her marginal uncertainty). The important thing is that the region of color space Mary attributes to the ball is a region much smaller than $[\![\text{red}]\!]^{\mathbf{C}}$. Be that as it may, Mary represents the ball as a shade of red, and shades of red are still *red*. This state of affairs is described by the expression: $v^{\mathbf{C}}(\text{Mary, the ball}) \subset [\![\text{red}]\!]^{\mathbf{C}}$.

So far, I have said that actual representational mental states – views – are attributions of regions of domains to objects, stored as true. It is natural to wonder where this leaves *belief*. To believe a proposition is ordinarily supposed to be to take a particular mental stance towards that proposition: the stance that it is true. And since (**FGR**) defines views as a kind of stance towards propositions – the stance that stores them as true – it is reasonable to assume that beliefs and views come to the same thing. But this would be a mistake. For if we suppose that S believes that *x* is P just in case S has a view equivalent to [[*x* is P]], we are committed to the position that S believes that *x* is P iff the region (in **D**) corresponding to S's view about *x* is equivalent with the region picked out by 'P' (in **D**). In other words, we would be committed to the following account of belief:

(107) S believes that *x* is P $\leftrightarrow$ $v^{\mathbf{D}}(S, x) = [\![P]\!]^{\mathbf{D}}$

It is not hard to see what is wrong with this account of belief: identity between $v^{\mathbf{D}}(S, x)$ and $[\![P]\!]^{\mathbf{D}}$ is far too demanding a condition for belief. If (107) were true, then the region in **D** corresponding to S's view about *x* would have to be identical with the region picked out by 'P,' or else S would not count as believing that *x* is P. This would disqualify the vast majority of

mental states we intuitively take to be beliefs; most ordinary belief ascriptions would turn out to be literally false. To use my previous example, if Mary represents a ball as being a particular shade of red, then Mary would not count as believing the ball is red, since given $v^{\mathbf{C}}$(Mary, the ball) $\subset$ [[red]]$^{\mathbf{C}}$, it straightforwardly follows that $v^{\mathbf{C}}$(Mary, the ball) $\neq$ [[red]]$^{\mathbf{C}}$. Or, to use another example, if John represents a book as being five-hundred pages long, (107) requires that he not count as believing that the book is long unless he *only* counts five-hundred page books as long. If, instead, John counts any book in excess of three-hundred pages as long, then (107) disqualifies him from believing that the book is long, since [[five-hundred pages long]]$^{\mathbf{BL}}$ $\neq$ [[more than three-hundred pages long]]$^{\mathbf{BL}}$. In general, since most views about objects correspond to smaller regions of domains than lexicalized predicates, (107) dubiously entails that in general, people do not believe lexicalized propositions.

In order to respect the natural sense in which lexicalized belief ascriptions are often true, (**FGR**) jettisons the assumption that beliefs and views are identical – in effect, the assumption that to believe the proposition *p* is to store *p* as true. Instead, (**FGR**) models belief as a *derivative* representational state. In (**FGR**), views are primary, or basic representational states: S attributes a specific region in some domain **D** to some object *x* – typically, the smallest region in **D** which S can epistemically justify attributing to *x*. (**FGR**) then defines belief as a geometric relation borne between views and propositions. Specifically, S counts as believing the proposition that *x* is P iff the region in **D** corresponding to S's view about *x* is a subset of the region corresponding to the predicate P. Formally:

(108) S believes that *x* is P $\leftrightarrow$ $v^{\mathbf{D}}$(S, *x*) $\subseteq$ [[P]]$^{\mathbf{D}}$ [24]

A major consequence of modeling belief in this way is that a person, S, counts as believing every proposition which fully encloses (i.e. is a superset of) a region in object-predicate space that S stores as true. In other words, if S has a view about *x* in **D** (i.e. if S attributes a region in **D** to *x*), then S counts as believing every proposition which attributes a region in **D** to *x* that encloses S's view about *x* in **D**. This straightforwardly accommodates the

[24] As in footnote 151, we technically also require that $v^{\mathbf{D}}$(S, *x*) $\neq \emptyset$ (that S's view about *x* in **D** corresponds to a non-empty region of **D**), since for any [[P]]$^{\mathbf{D}}$, it is trivially true that $\emptyset \subseteq$ [[P]]$^{\mathbf{D}}$. But for the sake of visual clarity, I leave the requirement that $v^{\mathbf{D}}$(S, *x*) $\neq \emptyset$ unstated. Unless specified otherwise, it may be assumed.

flexibility of natural language belief reports. For instance, (108) explains why it is that if Mary mentally represents the color of a ball as a specific shade of red, it is perfectly natural to say that Mary believes that the ball is red (since $v^{\mathbf{C}}(\text{Mary, ball}) \subseteq [\![\text{red}]\!]^{\mathbf{D}}$), that the ball is not blue (since $v^{\mathbf{C}}(\text{Mary, ball}) \subseteq [\![\text{not blue}]\!]^{\mathbf{C}}$), that the ball is a warm color (since $v^{\mathbf{C}}(\text{Mary, ball}) \subseteq [\![\text{warm}]\!]^{\mathbf{C}}$), and so on. This definition of belief permits (**FGR**) to accommodate indefinitely many belief ascriptions to people with finite cognitive resources.

Moreover, if to assert '*x* is P' is to present oneself as believing that *x* is P (as (**EA**) claims), it follows, given (108), that to assert '*x* is P' is to present oneself as having a view about *x* somewhere in the region $[\![\text{P}]\!]^{\mathbf{D}}$. And since views about objects are generally much more fine-grained than lexicalized predicates (i.e. generally correspond to much smaller regions), it is no wonder why asserting '*x* is P' so often leaves so much of our view about *x* (in the relevant domain) unexpressed: to present oneself as having a view about *x* somewhere in $[\![\text{P}]\!]^{\mathbf{D}}$ is often akin to communicating your exact location by asserting that you are somewhere in New York City.

By modeling belief in this way, (**FGR**) helps to explain why assertion so often involves triangulation and approximation: to express something approximating our views in assertion, we start with the coarse-grained predicates in our collective lexicon, and whittle down the regions they correspond to with the help of logical connectives and modifiers. For instance, if I represent the color of a navy blue object, but lack that color concept, I may nevertheless express something approximating my view by saying that the object is a dark shade of blue (i.e. that $f^{\mathbf{C}}(\text{object}) \in ([\![\text{dark}]\!]^{\mathbf{C}} \cap [\![\text{blue}]\!]^{\mathbf{C}})$). Or, if I represent the feeling of a silky surface, but lack any concept approximating the region $[\![\text{silky}]\!]^{\mathbf{PT}}$ (where **PT** is the domain of phenomenal touch), I may nevertheless articulate a triangulation of my view by asserting that the surface is extremely smooth and soft, but not fuzzy (i.e. that $f^{\mathbf{PT}}(\text{surface}) \in ([\![\text{extremely smooth}]\!]^{\mathbf{PT}} \cap [\![\text{soft}]\!]^{\mathbf{PT}} \cap [\![\text{not fuzzy}]\!]^{\mathbf{PT}})$)).

Moreover, together with the geometric framework of (**FGR**), (108) offers a natural account of graded belief (i.e. graded acceptance of propositions). For example, to *almost* (or *less than fully*) believe that *x* is P is for one's view about *x* to be *almost* (or *less than fully*) a subset of $[\![\text{P}]\!]^{\mathbf{D}}$. Likewise, to *barely* believe that *x* is P is to have a view about *x* that only narrowly falls

within $[\![P]\!]^{\mathbf{D}}$. In general, (**FGR**) models degree of belief in the proposition that *x* is P by the extent to which one's view about *x* is contained within the region $[\![P]\!]^{\mathbf{D}}$. As that containment increases – i.e. as an agent's view about *x* is increasingly refined or sharpened within $[\![P]\!]^{\mathbf{D}}$ – so too does that agent's acceptance of the proposition that *x* is P. Once an agent's view lies entirely within $[\![P]\!]^{\mathbf{D}}$, belief is no longer partial or graded: the agent simply believes that *x* is P.

Yet another puzzle of belief which (**FGR**) can help illuminate is the distinction between *de dicto* and *de re* belief. While the *de dicto* / *de re* distinction has been cashed out in different ways, the puzzle that motivates (**FGR**) is that the truth of belief ascriptions are sometimes sensitive to one's choice of co-designating terms, and sometimes not. For instance:[25]

(109) Homer believed that the Aegean Sea was blue.

(110) Alice Hill believes that she has a dozen hens.

It is widely accepted that Pre-Classical Greeks – and by extension, Homer – did not have any concept approximating *blue*.[26] The Ancient Greek words 'κυάνεος' (denoting dark blue) and 'γλαυκός' (denoting light blue) are generally thought to have entered the Ancient Greek lexicon around the 5th century BCE, well after Homer's time. So, there is a natural sense in which (109) is false, given that Homer did not have anything approximating the concept *blue*. And yet, we have every reason to think Homer and his contemporaries perceived the Aegean in roughly the same way we do, and mentally represented its hue in a manner that aligns with what we now call blue.[27] On this somewhat less accessible reading, (109) may well be true, even if Homer lacked the concept *blue*.

---

[25] *De dicto* belief is typically illustrated by contrasting pairs of belief ascriptions that differ only in the substitution of co-designating terms referring to the same *object(s)*. Here, I opt to illustrate the distinction between *de dicto* / *de re* belief by comparing belief ascriptions that differ only in the substitution of equivalent *concept(s)*.

[26] Homer consistently referred to the sea as "wine-dark" (οἶνοψ πόντος). Classicists and historians are still unsure as to whether 'wine-dark' denoted a color or was being used metaphorically. Homer did not describe the sea using any other color word.

[27] It is very unlikely that any strong version of the Sapir-Whorf hypothesis is true. If Pre-Classical Greeks could not perceive the region of color space we refer to with the word 'blue,' then for what reason would Greeks of the 5th century BCE have introduced 'κυάνεος' and 'γλαυκός' into the Ancient Greek lexicon? The most plausible explanation is that these terms were introduced to respect pre-existing distinctions already present in people's phenomenal experiences of color.

Compare (110). Suppose that Alice Hill is a medieval English peasant, and is not familiar with the numerical concept *dozen* (the term 'dozen' first appeared in English in the 13th century). Suppose also, however, that she is perfectly capable of counting and that each morning she tallies her hens and confirms there are twelve. In that circumstance, (110) seems a perfectly natural statement of what Alice believes; this time around, the most accessible reading of (110) is that it is true. By contrast, the reading in which (110) is false (on the grounds that Alice lacks the concept *dozen*) strikes me as pedantic and strained. (Indeed, it seems comparable to saying that 'Alice Hill cree que tiene doce gallinas' is false, just because Alice doesn't speak Spanish.)

The ambiguity in cases like these is generally explained in terms of how the terms *blue* and *dozen* are scoped within their respective belief operators. Supposedly disambiguated, the four readings are usually framed as follows:

(109/ *de dicto*)	Homer believed the proposition that the Aegean Sea was blue.
(109/ *de re*)	The color denoted by 'blue' is the color that Homer believed the Aegean Sea was.
(110/ *de dicto*)	Alice Hill believes the proposition that she has a dozen hens.
(110/ *de re*)	The number denoted by 'a dozen' is the number of hens that Alice Hill believes she has.

On the standard view, *de dicto* belief requires possession of the relevant concept (e.g. *blue*, *dozen*), while *de re* belief merely requires that the agent represents the object in the right way – even if they lack the specific term or concept we use to describe it.

But the standard view breaks down in cases where agents possess concepts that are only approximately equivalent to the ascriber's. It gets the Homer case right: Homer lacks any concept close to *blue*, so we hesitate to ascribe to Homer the *de dicto* belief that the Aegean Sea is blue. But the standard view stumbles in the Alice Hill case. Alice may not know what 'a dozen' refers to, but she clearly grasps the concept *twelve*, and that seems close enough for her to count as believing that she has a dozen hens (even on the *de dicto* reading). A parallel modification of the Homer case reinforces the point: if Homer possessed the concept *κυάνεος* – a concept close to but not identical with *blue* – then we would be far more inclined to say that Homer believed that the Aegean Sea was blue. What's missing in the standard account is a way to recognize when an agent possesses a concept close enough to the ascriber's predicate for *de*

dicto belief ascription to be appropriate. So it looks like scope alone doesn't do justice to our intuitions about belief ascriptions.

Here, (**FGR**) develops and formalizes a strand of thought pressed by Quine (1956): that what we call "belief" really splits into two closely related but distinct attitudes – one which requires possession of *approximate* concepts (viz. *de dicto* belief), and one which does not (viz. *de re* belief). Since (**FGR**) already treats belief as a derivative state – as a geometric relation between a predicate and an agent's view about an object – there is little cost to bifurcating the notion of belief in this way. For (**FGR**) does not posit two distinct, *sui generis* representational states that just happen to resemble each other; rather, it merely distinguishes between two closely related ways of having a view in a predicate region: one in which a person also possesses a concept that approximates that region, and one in which they do not.

In the framework of (**FGR**), we can formalize the set of concepts a person possesses (with respect to a domain **D**) as the set of regions in **D** that they store as concepts (or predicates). Let the regions in **D** corresponding to S's concepts be given by a function $c^{\mathbf{D}}$, which maps agents $S \in \mathbf{A}$ to sets of regions in **D**; notationally: $c^{\mathbf{D}} : \mathbf{A} \mapsto \mathscr{P}(\mathscr{P}(\mathbf{D}))$. On this notation, $c^{\mathbf{C}}$(Homer) yields the set of regions in color space **C** which correspond to Homer's color concepts, and $c^{\mathbb{N}}$(Alice Hill) returns the set of regions in the domain of countable numbers $\mathbb{N}$ which correspond to Alice Hill's countable number concepts.

Then, let us define a notion of geometric similarity, denoted '≈', such that for any two regions $[\![X]\!]^{\mathbf{D}}$ and $[\![Y]\!]^{\mathbf{D}}$ within the same domain, $[\![X]\!]^{\mathbf{D}} \approx [\![Y]\!]^{\mathbf{D}}$ if and only if $[\![X]\!]^{\mathbf{D}}$ and $[\![Y]\!]^{\mathbf{D}}$ significantly overlap.[28] We now have the tools to formally notate what it is for S to have a concept *approximating* P:

(111) S has a concept approximating P $\leftrightarrow \exists [\![R]\!]^{\mathbf{D}} \mathbf{[}\, ([\![R]\!]^{\mathbf{D}} \in c^{\mathbf{D}}(S)) \wedge ([\![R]\!]^{\mathbf{D}} \approx [\![P]\!]^{\mathbf{D}}) \mathbf{]}$

In other words, S has a concept approximating P just in case there is some region $[\![R]\!]^{\mathbf{D}}$, such that (i) $[\![R]\!]^{\mathbf{D}}$ corresponds to one of S's concepts in **D**, and (ii) $[\![R]\!]^{\mathbf{D}}$ significantly overlaps with $[\![P]\!]^{\mathbf{D}}$.

---

[28] The precise conditions under which $[\![X]\!]^{\mathbf{D}}$ counts as sufficiently overlapping with $[\![Y]\!]^{\mathbf{D}}$ to warrant $[\![X]\!]^{\mathbf{D}} \approx [\![Y]\!]^{\mathbf{D}}$ are left unspecified here, as they lie well beyond the scope of the present work.

Alice Hill thus counts as having a concept approximating *dozen*, since she has the concept *twelve* ($[\![\text{twelve}]\!]^{\mathbb{N}} \in c^{\mathbb{N}}(\text{Alice Hill})$), and the concepts *twelve* and *dozen* significantly overlap ($[\![\text{twelve}]\!]^{\mathbb{N}} \approx [\![\text{dozen}]\!]^{\mathbb{N}}$). By contrast, Homer does not count as having a concept approximating *blue*, provided that he had no color concept whose region in color space significantly overlaps with $[\![\text{blue}]\!]^{\mathbf{C}}$.

We now have everything we need to gloss the intuitive distinction between *de re* and *de dicto* belief:

(108*) S believes (*de re*) that *x* is P $\leftrightarrow v^{\mathbf{D}}(\text{S}, x) \subseteq [\![\text{P}]\!]^{\mathbf{D}}$

(112) S believes (*de dicto*) that *x* is P $\leftrightarrow v^{\mathbf{D}}(\text{S}, x) \subseteq [\![\text{P}]\!]^{\mathbf{D}} \wedge \exists [\![\text{R}]\!]^{\mathbf{D}}$ **[** $([\![\text{R}]\!]^{\mathbf{D}} \in c^{\mathbf{D}}(\text{S})) \wedge ([\![\text{R}]\!]^{\mathbf{D}} \approx [\![\text{P}]\!]^{\mathbf{D}})$ **]**

(**FGR**) renders *de re* belief the more fundamental notion of belief, and *de dicto* belief a more complex condition – one which requires *de re* belief, but also requires possession of a concept approximating the predicate used in the belief ascription.[29]

In review, (**FGR**) modifies and extends Gärdenfors' conceptual spaces framework by modeling propositional content as the attribution of a region in a domain (i.e. a geometric space structured by one or more quality dimensions) to an object. On this account, primary representational states – *views* – consist in the finite set of propositions agents store as true. Since (i) views about objects rarely align perfectly with lexicalized predicates (and generally correspond to smaller regions than those predicates), and (ii) ordinary belief ascriptions attribute belief in *lexicalized* propositions, it follows that believing a proposition *p* and having a view which stores *p* as true are distinct states (on threat of most intuitively true belief attributions coming out literally false). Belief that *x* is P is better understood as having a view about *x* somewhere within the region picked out by 'P.' (**FGR**) thus defines belief as a *derivative* state: the condition of having a view that stands in the right kind of geometric relation to a proposition. This structural account of belief straightforwardly explains a wide range of puzzles about belief,

[29] As mentioned in footnote 154, I have opted to illustrate *de dicto* belief by emphasizing its sensitivity to substitution of co-designating terms for the same *concept(s)*, rather than by emphasizing its sensitivity to substitution of co-designating names for the same *object(s)*. To account for the latter sensitivity, (112) would also have to require that S know *x* by a name which at least approximates the one used in the belief ascription.

including (a) why it is natural to attribute indefinitely many beliefs to agents despite their finite cognitive resources, (b) why assertion often feels like an exercise in approximating intended meaning, (c) why belief can come in degrees, and (d) how belief possession can depend on concept possession in some cases but not others – as in the distinction between *de dicto* and *de re* belief. With (**FGR**) in view, I have now stated all five constraints on my solution to the grammatological problem; the path is clear to belief expressivism.

5.1.7 Belief expressivism

In Sections 5.1.1 - 5.1.6, I independently motivated five theoretical constraints on a solution to the grammatological problem – the challenge of saying what it is that moral, ordinary descriptive, and taste judgments have in common, in virtue of which they share so many grammatical and logical properties. For quick reference, the five constraints are:

> *Expressive Assertion* (**EA**): To assert the judgment '*p*' is, at a minimum, to present oneself as having the mental state required to agree with the content *p* – typically, belief that *p*.
>
> *Propositional Taste Nondescriptivism* (**PTN**): Mature taste discourse is nondescriptive; ordinary language users do not intend for their assertions of taste judgments to describe (or report) objective facts. Be that as it may, taste judgments have propositional content that is truth-evaluable (albeit only relatively), and taste disagreements are propositional disagreements.
>
> *Phenomenal Indistinguishability* (**PI**): The license afforded to taste beliefs by taste appearances is phenomenally indistinguishable from the license afforded to most beliefs by appearances.
>
> *A posteriori Non-objectivity* (**APNO**): It is only possible to learn that a class of judgments is non-objective by *a posteriori* means. It is discovered over time – through reflection on our practices and experiences – that certain phenomenal appearances (i.e. "apparent qualities") do not correspond to objective properties.
>
> *Finite Geometric Representationality* (**FGR**): Primary (or basic) representational mental states – "views" – are finite in number, and rarely map cleanly onto propositions expressible using the predicates of a natural language. Views and propositions can be modeled as regions in geometric conceptual spaces (attributed to objects). Beliefs are secondary (or derivative) states: a person believes every proposition whose associated region fully contains one of their views. The exact relation between views and beliefs is given as:
>
> > S believes that *x* is P iff the region (in the contextually salient domain **D**) that models S's view about *x* is a subset of the region (in **D**) that models P.

The first three of these constraints – (**EA**), (**PTN**), and (**PI**) – jointly constitute the core of belief expressivism, within which the notion of nondescriptive belief becomes intelligible. To see how, let us first take stock of what is ordinarily supposed to be incoherent about nondescriptive belief. All can agree that belief is a stance towards propositions: the stance that they are true. And ordinarily, for a proposition *p* to be true is just for the objective world to be as *p* describes. From these two points, it follows that belief is a stance on how the objective world is – that to believe that *p* is true is to believe that *p* is *objectively* true. In other words, given the assumption that truth entails objectivity, it follows that belief is inherently descriptive, and that nondescriptive belief is incoherent.

Enter (**PTN**). Using taste discourse as a proving ground, and drawing on the resources of minimalist expressivism and genuine relativism, (**PTN**) makes conceptual space for truth without objectivity. More specifically, (**PTN**) advances a notion of truth relative to an individual (or: truth given one's constitution). So conceived, taste propositions are truth-evaluable but not objective (i.e. true for some, but neither necessarily true for everyone nor necessarily false for everyone). This in turn makes room for belief in a proposition's truth without commitment to its *objective* truth (i.e. nondescriptive belief). But this move raises two immediate questions: (i) What is the point of believing a proposition is true if one denies its objective truth? And (ii) what communicative purpose is served by asserting propositions that lack objective truth-conditions?

(**PI**) answers the first question. The point of believing a taste proposition is the same as with any belief: to form a stable representation of how things appear to be, *all-things-considered*. Consider illusions – for example, the Müller-Lyer illusion – which present deceptive, sometimes even phenomenally stable appearances. We recognize such appearances as illusory only by placing our credence in other, more reliable appearances (e.g. precise measurements). We come to believe that the Müller-Lyer lines are equal in length, despite how they visually appear, by weighing these countervailing appearances. When we deny that rollercoasters can be objectively fun, by contrast, no such countervailing appearances are involved. This is a key point of disanalogy. Instead, our denial of objectivity instead reflects the recognition that the all-things-considered appearances which license taste beliefs are irreducibly subjective: they vary across subjects and depend entirely on one's constitution (i.e. one's palette), such that there is no objective way for objects to really be as taste propositions say.

That answers why we believe nondescriptive (taste) propositions, but still leaves unaddressed why we bother *asserting* them. (**EA**) answers this second question: to assert any proposition *p* is, at a minimum, to present oneself as believing that *p*. Ordinarily, this expressive function is secondary (or supplemental) to assertion's primary function: the description of objective truths. But once speakers deny that the propositions they assert have objective truth-conditions, they relinquish that primary, descriptive function. The belief-expressive function remains, however, and in the absence of the descriptive function of assertion, the expressive function ascends to primacy.[30]

It may be objected that once the subjectivity of taste is made salient, and the descriptive function of assertion no longer applies, speakers would naturally abandon ordinary taste judgments in favor of more direct statements of preference – for example, by retreating to first-person formulations, as in 'I like rollercoasters.' Certainly, some speakers do adopt this strategy. However, such a move is not necessary. The expressive function of assertion is sufficient, by itself, to meet all the conversational goals of taste discourse. There is no need to retreat to first-person statements in order to communicate preferences effectively.

Altogether then, (**PTN**), (**PI**), and (**EA**) allow us to reconceive of belief as rooted not in objectivity but in stable, all-things-considered appearances. Once we allow (i) that propositions can be truth-evaluable without having objective truth-values (as (**PTN**) claims), (ii) that belief in non-objective propositions is licensed by appearances in a way phenomenally indistinguishable from how most beliefs are licensed (as (**PI**) claims), and (iii) that nondescriptive assertions retain the belief-expressive function of assertion (as (**EA**) holds), the notion of nondescriptive belief (and nondescriptive assertion) becomes not only coherent but theoretically well-motivated. What looked like a contradiction – belief (and assertion) without commitment to objective truth – turns out to rest on a failure to distinguish between objectivity and truth, between description and representation, and between assertion's descriptive and expressive functions.

But even with (**PTN**), (**PI**), and (**EA**) in place, further questions remain. If beliefs in non-objective propositions are licensed by appearances in a manner that is phenomenally

[30] Here, I take the *ordinary* arrangement to be as follows: the descriptive function of assertion is semantically primary, the belief-expressive function is secondary, and the countless pragmatic purposes an assertion may serve – parasitic on these first two – are semantically tertiary.

indistinguishable from how most beliefs are licensed, how do we ever come to recognize that some propositions lack objective truth? And once we do, why doesn't that recognition prompt us to abandon those beliefs? These are the challenges addressed by (**APNO**) (and by further elaboration on (**PI**)).

According to (**APNO**), we come to know a discourse is non-objective not by introspection or conceptual analysis, but *a posteriori*: by experiencing and reflecting on the discourse's anomalous features. In the case of taste, we eventually notice that the familiar hallmarks of objectivity fail to materialize: no ways to decisively settle taste disagreements are found, "mistaken" taste beliefs fail to trigger the kinds of natural repercussions characteristic of objective error, and no objective properties are discovered to be shared between all the objects with the same apparent taste qualities. By taking stock of anomalous features like these, we steadily realize that taste discourse is non-objective.

But then, why does the revelation that the content of a belief is non-objective not lead to the abandonment of that belief? The short answer is that the non-objectivity of a proposition is not a reason against believing it. Notwithstanding admissions of ignorance (and a few other edge cases), belief that p is generally abandoned in light of reasons to think that p is false (even if such reasons do not decisively settle that $\neg p$). And since recognizing that p is non-objective is not a reason to believe that p is false, it is not a reason to abandon belief that p. Case in point: suppose that Mary believes that rollercoasters are fun because that is how they appear to her, all-things-considered. She may well deny that rollercoasters are *objectively* fun – but until they begin to seem *not* fun, Mary has no reason to abandon her belief that rollercoasters are fun.

In short, (**EA**), (**PTN**), (**PI**), and (**APNO**) collectively explain how it is possible to believe and assert propositions without objective truth-conditions, how such propositions come to be recognized as non-objective, and why belief in them persists even after recognizing their non-objectivity. Taken together, these constraints yield a solution to the grammatological problem: moral, ordinary descriptive, and taste judgments share the grammatical and logical properties that they do because they all express belief in (and are used to assert) propositions. Some beliefs are commitments to the truth of objective propositions (e.g. ordinary descriptive beliefs), others to the truth of non-objective propositions (e.g. taste beliefs), but all are

commitments to the truth of propositions. Thus, what all judgments have in common is propositional content, and the shared grammatological properties of judgments can be accounted for by the structure and logic of propositions.

In one sense, then, (**EA**), (**PTN**), (**PI**), and (**APNO**) suffice to explain why all judgments exhibit the same grammatical and logical properties. But a residual challenge remains. Even granting that all judgments are propositional, why do objective and non-objective propositions *themselves* support the same compositional treatment – why do *they* exhibit the same grammatical and logical properties? At this point, the grammatological and compositional problems intersect; the challenge of explaining why descriptive and nondescriptive judgments share so many grammatological properties blurs with the challenge of giving a formal account of what propositions *are*, such that their grammatical and logical properties are indifferent to questions of objectivity. What's needed is not just a unifying story about belief, but a unifying model of propositional content – one that explains why objective and non-objective propositions give rise to the same logical structure without merely presupposing that they do. This is the work of the final component of belief expressivism: (**FGR**).

(**FGR**) offers a geometric model of both belief and propositional content. On this model, speakers are said to have *views* – commitments to the truths of *propositions*, themselves defined as attributions of regions (i.e. subsets of domains) to objects. Belief is understood as a geometric relation on having a view: to believe that *x* is P is for one's view about *x* (in the relevant domain) to fall entirely within the region that models the predicate 'P.' On this account, the difference between objective and non-objective propositions merely depends on the kinds of domains they are drawn from: objective propositions belong to domains whose values correspond to real properties (e.g. different colors, masses, and temperatures), while non-objective propositions belong to domains whose values do not correspond to real properties (e.g. different degrees of tastiness, painfulness, and beauty).[31] As per (**PI**), views about objects in either kind of domain

[31] Although the values of non-objective domains do not *correspond* to real properties, they do *supervene* on them. For instance, objects are more painful to touch as they become hotter or induce stronger electric currents (above thresholds determined by one's constitution, and within experiential limits). But the degree to which something is painful is not a disguised fact about its temperature or electrical output. Too many stimuli are painful for the appearance of pain to track any discrete physical property. What seems to be happening is that the mind invents the phenomenal appearance of pain, and projects it onto an object whenever it detects at least one stimulus that the nervous system is predisposed to mark as painful. While evolutionary debunking arguments should be used cautiously, the explanatory appeal here is tremendous:

are licensed in the same kind of way: by the qualities objects appear to have. And as per (**APNO**), that the values of some domains correspond to real properties and others not is something we can only discover through reflection and experience.

In this way, (**FGR**) answers the residual grammatological challenge. In conjunction with (**PI**) and (**APNO**), (**FGR**) explains why objective and non-objective propositions share the same grammatical and logical properties: both are modeled as regions of domains attributed to objects, and our representational architecture is indifferent to whether those domains correspond to real properties. Our representational system is functional from an early age – well before we understand what it is for a proposition (or a domain) to be objective or not. And so, objective and non-objective propositions inherit a common logical and grammatical structure from the beginning. Our representational architecture cannot help but to treat objective and non-objective propositions the same way (viz. as regions of domains attributed to objects): its purpose is to codify stable, all-things-considered appearances – not to carve reality at its joints. By unifying a model of belief with a model of content, and putting the last vestiges of the grammatological problem to rest, (**FGR**) completes the account promised by belief expressivism.

5.2 Compositionality revisited

Belief expressivism is a form of expressivism, and as such inherits the traditional challenges expressivists face – including the Frege-Geach problem. As I explained in Part I (*Chapter 1*), this problem is best understood as comprising two distinct challenges. The first is what I call the *grammatological problem*: the challenge of explaining why descriptive and (supposedly) nondescriptive judgments share so many grammatical and logical properties. The second is the *compositional problem*: the challenge of explaining why (supposedly) nondescriptive judgments have the grammatical and logical properties that they do – that is, how their meanings compose across logical constructions. Traditional expressivists, such as Hare and Blackburn, took the compositional problem to be primary, and strove to reconstruct the

---

we are predisposed to find a large number of objects painful not because there is any real property they all share, but because it is evolutionarily adaptive to do so. The advantage of one's nervous system projecting pain onto harmful objects is obvious. But even so, we must not confuse painfulness with harmfulness. Some harmful objects – like rooms filled with noble gases and plutonium dust – are not immediately painful. Other stimuli – like capsaicin in spicy food and second-hand embarrassment – are painful but largely harmless.

grammatical and logical properties of moral discourse in a manner distinct from the standard truth-conditional approach for ordinary descriptive judgments.[32] This is strategically unsound, however, since it guarantees that no adequate solution to the grammatological problem will be possible: if the grammatical and logical properties of descriptive and nondescriptive judgments arise for completely separate reasons, it becomes utterly mysterious why both forms of judgment share them.

I follow Horgan and Timmons (2000, 2006) (and to a lesser extent, Gibbard (2003)), who start by answering the grammatological problem. I presented my solution to the problem in the previous section: what all judgments have in common is propositional content, and the shared grammatological properties of judgments can be accounted for by the structure and logic of propositions. The difference between descriptive and nondescriptive judgments is sustained by the distinction between objective and non-objective propositions. Both are treated the same by our representational system because that system is concerned with modeling the ways that objects appear – not with adjudicating which all-things-considered appearances correspond to real properties and which merely supervene on them.

Given my answer to the grammatological problem, the compositional problem looks very different. Rather than needing to reconstruct the inferential properties of judgments without appealing to truth-conditional content, we simply need an account of what propositions are and how they compose – one that applies uniformly to both objective and non-objective propositions. In this sense, the compositional problem collapses into the general problem of propositional semantics. Any plausible account – possible worlds semantics, truthmaker semantics, or some other framework – will do, provided it treats objective and non-objective propositions *isomorphically*. That is the only constraint my view imposes.

I will not attempt a full reconstruction of propositional logic here; that is a task for logicians and semanticists. What matters for my purposes is that belief expressivism – or more specifically, (**FGR**) – provides a model of propositional content and belief capable of satisfying

[32] Hare was the first to sketch out what this strategy should involve (i.e. the provision of formulas for the meanings of logically complex nondescriptive judgments as a function of the meanings of their parts), whereas Blackburn was the first to actually attempt to execute it.

the relevant constraint. Although my solution to the grammatological problem obviates the need for a distinct compositional semantics for nondescriptive discourse, it is worth briefly illustrating how (**FGR**) addresses some of the traditional pressure points – in particular, predicate negation and conditional embedding.

5.2.1 Predicate negation

Although expressivists (and the emotivists before them) have long struggled to explain how the meanings of nondescriptive judgments compose across *all* logical constructions, the difficulty for expressivism posed by negation has gained a lot of attention in the two decades since the publication of Nicholas Unwin's "Quasi-Realism, Negation, and the Frege-Geach Problem" (1999). Unwin (1999) pointed out that there are three different ways to negate propositional attitude reports, such as 'Mary believes that lying is wrong':

(n1) It is not the case that Mary believes that lying is wrong.
(n2) Mary believes that lying is not wrong.
(n3) Mary believes that not lying is wrong.

Each way to negate the original sentence is different: (n1) is compatible with Mary having no opinion about the morality of lying, (n2) requires Mary to have an opinion about the morality of lying inconsistent with believing that it is wrong, and (n3) requires Mary to have an opinion about the morality of a different act – failing to lie. For my purposes, I will refer to (n1) as *sentence negation* (since it negates the entire sentence), (n2) as *predicate negation* (since it negates the predicate of the proposition the reported attitude is directed towards), and (n3) as *subject negation* (since it negates the subject of the proposition the reported attitude is directed towards).

That there are three different places to insert negation into propositional attitude reports is a problem for traditional, non-propositional forms of expressivism, since they typically deny that moral belief is ordinary belief directed towards moral propositions. Instead, non-propositional expressivists analyze moral belief as a fundamentally different attitude from ordinary belief: it is supposed to be an attitude (or family of attitudes) distinct from belief, directed towards *objects* (i.e. nouns and act types) rather than propositions. For example, a non-propositional form of expressivism may analyze the moral belief that lying is wrong as a unique attitude of

disapprobation – *disapproval*, say – directed towards lying itself. Notwithstanding the worry that the non-propositional expressivist will have to posit a new attitude for every moral predicate (e.g. good, bad, permissible, abhorrent, and so on), another problem with analyzing moral belief in this way, as Unwin (1999) pointed out, is that it removes logical structure. Compare:

(w) Mary believes that [***lying*** is ***wrong***].
(w*) Mary disapproves of [***lying***].

For the sake of argument, let us grant that the work originally done by the predicate 'is wrong' can be accomplished by the attitude of disapproval. Even if we grant this, (w*) is missing a bit of logical structure originally in (w): whereas (w) reports an attitude directed towards an entity with two parts (viz. subject and predicate), (w*) reports an attitude directed towards an entity with only one part (viz. object).

This loss of logical structure gives rise to Unwin's (1999) negation problem: there are three meaningful places to insert a negation into (w) – corresponding to (n1), (n2), and (n3) – whereas there are just two ways to meaningfully negate (w*):

(n1*) It is not the case that Mary disapproves of lying.
(n2*) ???
(n3*) Mary disapproves of not lying.

There is no way to *predicate negate* (w*), since the predicate 'is wrong' has been dispensed with – its work supposedly carried on by the attitude of disapproval. We cannot fill in (n2*) with the sentence 'Mary does not disapprove of lying' because it is not equivalent with (n2); (n2) clearly requires Mary to have a moral opinion, whereas 'Mary does not disapprove of lying' does not. Just as (n2) requires Mary to have an opinion about the morality of lying inconsistent with believing that it is wrong, (n2*) should require Mary to have a moral attitude about lying inconsistent with disapproving of it. But we cannot fill in (n2*) with something like 'Mary *not-disapproves* of lying', on the grounds that this is not well formed.

Now, the non-propositional expressivist may try to stipulate a new, logically primitive attitude – *toleration*, say – to serve as the complement (or "attitude negation") of disapproval. But then, the non-propositional expressivist is obligated to say why we should think that there are two attitudes – disapproval and toleration – that are inconsistent towards the same object, and

just so happen to be interdefinable via predicate negation; they must offer some structural account as to what disapproval and toleration are, such that it is clear why anyone who disagrees with disapproving of *x* necessarily tolerates *x*. In the absence of such an account, the non-propositional expressivist has simply assembled a wish list of properties she hopes certain noncognitive mental states have, without offering any constructive account for them. As Schroeder (2010) explains:

> The problem with [a theory that simply stipulates that *disapproval* states disagree with *toleration* states] is that it is nonconstructive. To be true, it requires that there must be states of mind with the right disagreement properties, but it avoids saying anything about why there are such states of mind or why they have those disagreement properties and no others. (p.151)

Without a constructive account of non-propositional conative attitudes, the non-propositional expressivist's analysis of moral belief deprives it of the logical structure necessary to account for predicate negations like (n2*).

Horgan & Timmons (2009) attempted to sidestep the dilemma at the bottom of Unwin's (1999) negation problem by arguing that the negation of moral predicates (as in (n2)) should be analyzed as a form of *contrary-forming* predicate negation, rather than a standard form of *contradictory-forming* sentential negation. They claim that moral predicates are often trivalent (i.e. that in many cases, a moral predicate *m* has some contrary *n* such that (a) it is logically possible for *x* to be neither *m* nor *n*, but (b) logically impossible for *x* to be both *m* and *n*). Horgan & Timmons (2009) point out that in natural language English, appending a predicate with an "adverbial prefix such as 'un-', 'in-', or the like" often signals its contrary. (p.96) They offer the following examples:

> [S]ome experiences are *pleasant*, some are *unpleasant*, and some are neither; some arguments are *persuasive*, some are *unpersuasive*, and some lie in between; some ways of addressing a task are *effective*, some are *ineffective*, and some fall under neither category. Trivalent concepts are legion, and such examples are easily multiplied. (Horgan & Timmons, 2009, p.96)

In that light, Horgan & Timmons' (2009) proposed solution to Unwin's (1999) negation problem is that the predicate negation (n2) of propositional attitude reports should be understood as an instance of contrary-forming predicate negation. As Horgan & Timmons (2009) write,

> [Unwin and] Schroeder [are] mistaken to say concerning statements w, n1, n2, and n3 above, [that] "There is simply one place not enough for negations to go around." On the contrary, in

> addition to the two available places for (contradictory-forming) *sentential* negation, there is also an available place for contrary-forming *predicate* negation. (p.98)

To borrow their own example, suppose we prefix 'permissible' with 'not,' and take the result to indicate a form of contrary-forming negation. Then, if the expressivist invokes an attitude to stand in the place of the predicate 'is permissible' – *tolerance*, say – then they are free to define another attitude that is logically related to *tolerance* as its contrary, in order to stand in the place of 'not permissible' – *intolerance*, say.

There are at least three problems with Horgan & Timmons' (2009) proposed solution. First, the syntactic operation which supposedly gives rise to contrary-forming predicate negation – the prefixing of a predicate with 'un-', 'in-', 'im-' or the like – is only well-formed on occasion. It works well enough for some gradable adjectives like 'interesting,' 'pleasant,' and 'fun,' but fails for others, like 'boring,' 'cold,' and 'short.' 'Unboring,' 'uncold,' and 'unshort' are unnatural constructions, even though 'boring,' 'cold,' and 'short' have perfectly intelligible contraries in 'interesting,' 'hot,' and 'tall,' respectively. By contrast, prefixing *any* predicate in English with the word 'not' is sure to be well-formed: 'not good' and 'not horrible' are perfectly natural constructions, whereas 'ungood' and 'unhorrible' are not. This is some evidence that the prefixing of moral predicates with 'not' is unlikely to always be contrary-forming. It is also evidence that not all attitudes will have a well-formed contrary. For instance, while 'Mary enjoys sushi' is a perfectly natural construction, 'Mary unenjoys sushi' is unnatural.

Second, prefixing a predicate with 'un-', 'in-', or 'im-' or the like sometimes yields a complex predicate that is fully *contradictory* with the first rather than merely its *contrary*. Horgan & Timmons' (2009) own example in 'permissible' demonstrates the point: impermissibility is fully *contradictory* with permissibility (not to be confused with an act's being *merely* permissible, or optional). It is not possible for an act to be neither permissible nor impermissible; there is no space in the deontic domain left over after ruling out its permissibility *and* its impermissibility. Other examples of the same phenomenon include 'opened' / 'unopened,' 'mortal' / 'immortal,' 'possible' / 'impossible,' 'available' / 'unavailable', and 'ambiguous' / 'unambiguous.' In all of these cases, the predicate and its prefixed twin jointly exhaust the space of possibility in their respective domains. This holds true, even if there are more fine-grained ways to carve regions of each relevant domain (e.g. 'completely open,' 'lived

for less than sixty years,' '40% likely,' and so on). These cases demonstrate that the syntactic operation of prefixing a predicate with 'un-', 'in-', or 'im-' or the like is not always contrary-forming.

Third and most pressingly, the prefixing of a moral predicate with 'not' is not always an indication of contrary-forming negation. Consider the attitude report 'Jones believes that corporal punishment is very good.' The predicate negation (n2) of this sentence is 'Jones believes that corporal punishment is *not* very good.' But if we interpret 'not very good' as an indication of contrary-forming negation, then we risk equivocating between 'not very good' and 'very bad' (assuming that 'very bad' is the most salient contrary of 'very good').

Intuitively, the meanings of 'not very good' and 'very bad' are quite different: for Jones to believe that corporal punishment is *not very good* is for him to think corporal punishment has a moral status besides a fairly high degree of moral goodness. This is compatible with Jones thinking that corporal punishment is deeply wrong, but so too with Jones thinking that corporal punishment is morally neutral (i.e. neither good nor bad), and with Jones thinking that corporal punishment is moderately good. (It may even be compatible with Jones thinking that corporal punishment is *exceptionally* good.) By contrast, for Jones to believe that corporal punishment is *very bad* is incompatible with Jones thinking that corporal punishment is moderately bad, morally neutral, moderately good, and so on. But since we intuitively understand 'Jones believes that corporal punishment is not very good' to be compatible with Jones thinking that corporal punishment is morally neutral or moderately bad, this demonstrates that we do not interpret 'not very good' to pick out some salient contrary of 'very good' (i.e. 'very bad'). And this suffices to show that prefixing a moral predicate with 'not' does not always indicate contrary-forming negation.

In sum, prefixing a moral predicate with 'not' does not seem to reliably indicate a form of *contrary-forming* negation (in contrast with standard, *contradictory-forming* negation). Moral predicates appear to behave just like all other predicates under negation. Horgan & Timmons' (2009) solution to the negation problem falls short. And in the absence of a constructive account of why anyone who disapproves of lying necessarily disagrees with anyone who tolerates lying – non-propositional analyses of moral belief lack the logical structure necessary to account for the

meanings of sentences produced via predicate negation (e.g. 'Mary believes that lying is *not* wrong' (n2)).

Belief expressivism avoids Unwin's (1999) problem of negation altogether, because it does not deprive moral belief of its status as a propositional attitude. More specifically, belief expressivism does not reduce moral belief to a unique non-propositional attitude (or family of such attitudes), and consequently, does not dispense with the logical structure inherent to ordinary belief. Moral belief – like all belief – is an attitude towards propositions: it is the attitude that they are true. But where I differ from descriptivists is in sustaining the possibility of *nondescriptive beliefs*. And unlike non-propositional expressivists, I do not suppose that "nondescriptive belief" is really a fundamentally different kind of mental state that just so happens to go by the same name (viz. 'belief'); nondescriptive beliefs are genuine, propositional beliefs. What I deny is that all genuine beliefs are inherently descriptive; I deny that truth entails objectivity, and hence that genuine beliefs are automatically descriptive in nature.

My formal account of propositional content and belief – (**FGR**) – ultimately analyzes predicate negation (n2) in the standard way: as a kind of contradictory-forming sentential negation. (**FGR**) models any given predicate P as a region of a contextually-salient domain **D**, notating that region as $[\![P]\!]^{\mathbf{D}}$. The negation of P picks out the complement of $[\![P]\!]^{\mathbf{D}}$ within **D**: all the space within **D** not enclosed by $[\![P]\!]^{\mathbf{D}}$. For instance, 'not red' picks out the rest of **C** (color space) not enclosed by $[\![red]\!]^{\mathbf{C}}$. In general, we may define the region picked out by the negation of any predicate P – notated $[\![not\ P]\!]^{\mathbf{D}}$ – as follows:

(113) $[\![not\ P]\!]^{\mathbf{D}} \stackrel{def}{=} \mathbf{D} \setminus [\![P]\!]^{\mathbf{D}}$

Given my account of belief (where for S to believe that *x* is P is for S to have a view about *x* that corresponds to a region in **D** that is a subset of $[\![P]\!]^{\mathbf{D}}$), to believe that *x* is *not red* is to have a view about *x* that corresponds to a region in **C** that is fully enclosed by $[\![not\ red]\!]^{\mathbf{C}}$ – that is, by the region specified by $\mathbf{C} \setminus [\![red]\!]^{\mathbf{C}}$. Hence, on the analysis I propose, if Mary mentally represents the color of a ball as a shade of aquamarine, she counts as believing that the ball is *not* red (since $[\![aquamarine]\!]^{\mathbf{C}} \subset \mathbf{C} \setminus [\![red]\!]^{\mathbf{C}}$).

Note that believing that *x* is *not* P (n2) is not the same as failing to believe that *x* is P (n1). According to (**FGR**), for S to *not* believe that *x* is P (n1) is for S to *not* have a view that is fully enclosed by $[\![P]\!]^{\mathbf{D}}$. This occurs if at least one of the following is true: (i) S has *no view* about *x* in **D**, (ii) S has a view about *x* in **D** that lies entirely outside of $[\![P]\!]^{\mathbf{D}}$, and (iii) S has a view about *x* in **D** that only partially overlaps with $[\![P]\!]^{\mathbf{D}}$. By contrast, S believes that *x* is *not* P (n2) just in case S has a view about *x* in **D** that lies entirely outside of $[\![P]\!]^{\mathbf{D}}$ (where this is just condition (ii) for sentence negation). Hence, if S believes that *x* is *not* P, it is logically necessary that S does not believe that *x* is P; the truth of (n2) guarantees the truth of (n1). The converse is not true: if either (i) or (iii) is true, then (n1) will be true but (n2) will not.

The structure of (**FGR**) thus enables a direct response to Unwin's (1999) negation problem. 'Mary believes that lying is not wrong' (n2) reports that Mary has a view about lying in the moral domain **M** that lies entirely outside of $[\![\text{wrong}]\!]^{\mathbf{M}}$; in the notation of (**FGR**), (n2) claims that $v^{\mathbf{M}}(\text{Mary, lying}) \subseteq \mathbf{M} \setminus [\![\text{wrong}]\!]^{\mathbf{M}}$. (**FGR**) thus analyzes each of (n1), (n2), and (n3) as:

(n1^{FGR}) $v^{\mathbf{M}}(\text{Mary, lying}) \not\subseteq [\![\text{wrong}]\!]^{\mathbf{M}} \ \vee \ v^{\mathbf{M}}(\text{Mary, lying}) = \emptyset$
(n2^{FGR}) $v^{\mathbf{M}}(\text{Mary, lying}) \subseteq \mathbf{M} \setminus [\![\text{wrong}]\!]^{\mathbf{M}} \ \wedge \ v^{\mathbf{M}}(\text{Mary, lying}) \neq \emptyset$
(n3^{FGR}) $v^{\mathbf{M}}(\text{Mary, not lying}) \subseteq [\![\text{wrong}]\!]^{\mathbf{M}} \ \wedge \ v^{\mathbf{M}}(\text{Mary, lying}) \neq \emptyset$

This treatment of negation shows that a nondescriptivist-friendly framework like (**FGR**) can handle compositional constructions like "not wrong" without collapsing into descriptivism or surrendering propositional structure. By modeling the contents of predicates as regions within conceptual domains, the negation of any given predicate P can be defined as its complement – all of the space in that domain *not* picked out by P. (**FGR**) can thus constructively distinguish predicate negation (n2) from both sentence negation (n1) and subject negation (n3). To *not* believe that *x* is P is to *lack* a view about *x* that is fully enclosed by $[\![P]\!]^{\mathbf{D}}$. To believe that *x* is *not* P is to have a view about *x* that is fully enclosed by $[\![\textit{not}\ P]\!]^{\mathbf{D}}$. And to believe that *not-x* is P is to have a view about *not-x* that is fully enclosed by $[\![P]\!]^{\mathbf{D}}$.

Belief expressivism thus resolves a traditional problem for expressivist accounts by leveraging the structure and logic of propositions via (**FGR**). The important theoretical work lies not in negation itself, but in rendering non-objective propositions intelligible – a challenge met earlier by belief expressivism's four other commitments. The expressive function of

nondescriptive assertion can thus be preserved without sacrificing semantic rigor: for when a speaker recognizes that the content they assert is non-objective, they relinquish the descriptive function of assertion, and the expressive function remains.

5.2.2 Conditional embedding

To many, the Frege-Geach problem is practically synonymous with the challenge of explaining the meanings of conditionals when (supposedly) nondescriptive judgments are embedded within them. The emotivists – Ayer (1936) and Stevenson (1944) chief among them – held that the meaning of any moral judgment is given by the noncognitive (i.e. non-propositional) attitude it is used to express. But since any moral judgment can be embedded in a conditional, where it is not used to perform the same speech-act as the original standalone judgment, it seems that the meaning of moral judgments is not exhausted by the emotivists' platitude. In particular, emotivists owe a non-truth-functional account of the inference properties of moral judgments.

For better or worse, Hare (1970) established an influential framework for how noncognitivists should respond: they should produce formulas for the meanings of logically complex sentences (e.g. conditionals) as a function of the attitudes which their (supposedly) noncognitive components are suited to express. Hare himself never supplied such formulas, but Blackburn (1984, 1993) took up the challenge. The problem with this entire strategy, I have argued, is that it dooms any solution to the grammatological problem: even if the expressivist succeeds in providing the right compositional rules, the result is that the logical and grammatical properties of moral and descriptive sentences arise for completely unrelated reasons. And that is simply too implausible to believe.

Any adequate solution to the grammatological problem must instead explain how the compositional behavior of moral sentences stems from a feature they share with both ordinary descriptive and taste judgments. My answer – which I call *belief expressivism* – is that the grammatological properties of all judgments derive from a shared expressive function: to assert a judgment *J* is to present oneself as having a belief with the same propositional, truth-evaluable content as *J*. In one sense, this is a concession: moral sentences are propositional, and thus not "noncognitive" in the traditional sense. But I deny that this gives up anything of real

consequence. The moral noncognitivist's core intuitions – (i) that there are no moral facts to render moral claims objectively true or false, (ii) that ordinary speakers do not intend to assert any such facts, and (iii) that the communicative function of moral assertion is to "express" (i.e. present oneself as having) a particular mental state – can still be preserved. Belief expressivism accommodates all three by distinguishing between objective and non-objective sentence contents, and between the descriptive and expressive functions of assertion.

As in the previous subsection, the account (**FGR**) offers of conditionals is not really the important contribution of belief expressivism. For the account on offer is a structural reformulation of propositional logic, akin in spirit to possible-worlds semantics. Geach's (1965) conditional embedding problem is supposed to show that moral judgments are cognitive – that is, propositional – and that is a conclusion I wholeheartedly accept (indeed, require for (**PTN**)). What matters is that (**FGR**) can provide an account of conditional embedding compatible with the distinction between objective and non-objective contents, and can do so in a way that allows moral, taste, and ordinary descriptive sentences to compose on equal semantic terms.

What is required of (**FGR**), then, is that (i) it be on a par with standard propositional accounts of the conditional, and (ii) accommodate the distinction between objective and non-objective contents without rendering that distinction obvious *a priori*. In light of the first requirement, I want to be clear that my goal in this section is not to resolve the many complications that have plagued propositional treatments of the natural language conditional(s) for the better part of a century. Any such attempt would take us too far afield. My aim is more modest: to briefly show that (**FGR**) can handle the material conditional ($\supset$), and that this suffices to place it in roughly the same position with respect to the conditional as standard accounts of propositional logic.

With expectations appropriately tempered, let us now see how (**FGR**) handles the material conditional. The truth-conditions of the material conditional are straightforward: $A \supset B$ iff $\neg A \vee B$. Given the previous subsection, we already know how to interpret the meaning of $\neg A$, given A: where A is 'x is P,' and its content in (**FGR**) is given by the region $[\![\mathrm{P}]\!]^{\mathbf{D}}$ attributed to the object x, $\neg A$ is the attribution of the complement region $\mathbf{D} \setminus [\![\mathrm{P}]\!]^{\mathbf{D}}$ to the same object. Recall

that the truth-conditions for [[x is P]] follow directly from its content: $f^{\mathbf{D}}(x) \in [[\text{P}]]^{\mathbf{D}}$. We may thus write the content of 'not A' as:

(114) $[[\neg A]] = [[x \text{ is not P}]] \stackrel{\text{def}}{=} f^{\mathbf{D}}(x) \in \mathbf{D} \setminus [[\text{P}]]^{\mathbf{D}}$

What remains to be seen is how to interpret the meaning of an inclusive disjunction of propositions. Let proposition A be 'x is P' and proposition B be 'y is Q'. What we need to know is how to interpret the semantic content of 'A or B' – that is, $[[A \vee B]]$ – in terms of the contents of its parts: [[x is P]] and [[y is Q]]. Preserving classical truth-conditions, $[[A \vee B]]$ is true iff [[x is P]] is true or [[y is Q]] is true. And since [[x is P]] is true in (**FGR**) iff $f^{\mathbf{D}}(x) \in [[\text{P}]]^{\mathbf{D}}$, and [[$y$ is Q]] is true iff $f^{\mathbf{E}}(y) \in [[\text{Q}]]^{\mathbf{E}}$, we may write:

(115) $[[A \vee B]] = [[(x \text{ is P}) \vee (y \text{ is Q})]] \stackrel{\text{def}}{=} f^{\mathbf{D}}(x) \in [[\text{P}]]^{\mathbf{D}} \ \vee \ f^{\mathbf{E}}(y) \in [[\text{Q}]]^{\mathbf{E}}$

In other words, (**FGR**) glosses 'x is P or y is Q' as the claim that either x has a quality in the region corresponding to the predicate 'P,' or y has a quality in the region corresponding to the predicate 'Q.' A disjunction of propositions is true just in case at least one of its disjuncts is.

With sentential negation and propositional disjunction in place, the truth-conditions for $A \supset B$ follow straightforwardly from those for $\neg A \vee B$:

(116) $[[A \supset B]] \stackrel{\text{def}}{=} [[\neg A \vee B]] = [[(x \text{ is not P}) \text{ or } (y \text{ is Q})]] = f^{\mathbf{D}}(x) \in \mathbf{D} \setminus [[\text{P}]]^{\mathbf{D}} \ \vee \ f^{\mathbf{E}}(y) \in [[\text{Q}]]^{\mathbf{E}}$

Thus, the material conditional *(x is P)* $\supset$ *(y is Q)* is true just in case either x has a quality in the complement of the region corresponding to 'P,' or y has a quality in the region corresponding to 'Q.'

It is, of course, well known that the material conditional ($\supset$) fails to capture many aspects of the natural language 'if A, B' construction.[33] For one thing, belief that $\neg A$ does not seem to automatically justify belief that if A, B. Moreover, it is rationally coherent to affirm $\neg A$ and yet deny that if A, B. For instance, it is perfectly coherent to accept 'Mary is not hungry' and to reject 'if Mary is hungry, she will eat a handful of sand.' In that light, I do not take the analysis

[33] See Edgington (1995) for a comprehensive survey of the difficulties encountered by treatments of 'if A, B.'

given in (116) for the material conditional to suffice to explain the meaning of 'if *A*, *B*'. But the provision of an adequate account of 'if *A*, *B*' is a particularly thorny problem for all theories of propositional content – not just nondescriptivist-friendly analyses like (**FGR**). I am content to leave the matter here: with a straightforward account of the material conditional, and without offering a satisfactory account of the 'if *A*, *B*' natural language construction (because the provision of such an analysis is far beyond the scope of the present project), (**FGR**) situates itself among standard propositional frameworks: it handles $\supset$ cleanly, while leaving open the more difficult question of how to capture the nuances of natural language conditionals.

Now, how does (**FGR**) manage to accommodate the distinction between objective and non-objective contents without building that distinction in *a priori*? It does so by treating the function $f^{\mathbf{D}}(x)$ – which returns the point in domain **D** corresponding to *x*'s quality in **D** – as metaphysically neutral. That is, $f^{\mathbf{D}}(x)$ specifies "how *x* is" with respect to **D**, regardless of whether **D** tracks objective features or not. To say that $f^{\mathbf{D}}(x)$ is metaphysically neutral is to say that the operative notion of "how things are" (that concerns our representational architecture) is indifferent to questions of objectivity: it includes "tasty" as a way sushi might be, and "green" as a way grass might be, without regard to which quality is *objective* and which not. Put differently, there is a perfectly good sense in which our representations aim to capture how objects *are* (as opposed to merely how they appear) – but this sense of "how objects are" is metaphysically undifferentiated; our representational system does not, from the outset, distinguish between those qualities which are objective, and those which are not.

This neutrality is what enables (**FGR**) to model both objective and non-objective judgments using a single formal framework. The difference between domains like color and taste, or mass and moral wrongness, is not one that our semantics must hardwire into the logic of truth or composition. It is, rather, a distinction that emerges through reflection on the epistemic features of a domain – its justificatory boundaries, the consequences for error, the nature and persistence of disagreement – exactly the route emphasized by (**APNO**). (**FGR**) preserves the logical and grammatical integrity of assertoric discourse across all domains, because the compositional rules it provides operate on contents that are structurally uniform, regardless of their metaphysical status. In this way, belief expressivism vindicates a propositional,

truth-evaluable treatment of all judgments, even as it allows for an *a posteriori* recognition that not all sentence contents have objective truth-conditions.

5.3 Belief expressivism, morality, and taste

The semantic promises of belief expressivism have now been fulfilled. The grammatological and compositional challenges have been met: the grammatical and logical properties of judgments are explained by the fact that all judgments express beliefs with propositional content. Some propositions, however, are non-objective: they do not correspond to objective ways the world might be, and they lack objective truth-conditions. Nevertheless, belief in non-objective contents is licensed by appearances in a manner phenomenally indistinguishable from how most beliefs are licensed. We form stable representations of how things appear to be, all-things-considered – whether those appearances reflect real properties, or are projections of the mind that merely supervene on real properties.

Moral judgments and taste judgments are especially ripe for treatment as non-objective in this sense. Setting aside the semantic questions – I have said enough in the foregoing sections – I turn now to a final and distinctive task: how do non-objective discourses differ from one another, if at all? In particular, how do moral and taste judgments differ in the beliefs they express?

In this final section, I defend the view that both moral and taste judgments are non-objective – that they lack *objective* truth-conditions, and are *primarily* asserted to present oneself as believing the content asserted.[34] This explains their shared grammatical and logical properties, but the similarities largely end there. Moral and taste beliefs differ in several important respects: they function differently in interpersonal discourse, they play different roles in practical reasoning, and they are typically licensed by different kinds of appearances. These differences can be accommodated within belief expressivism by distinguishing the belief-expressive function of assertion from the nature of the beliefs expressed. The resulting picture simultaneously supports a pluralist view of nondescriptive discourse, and resists the temptation to erect rigid boundaries between judgment classes.

---

[34] I use 'primary' here in the semantic sense. An assertion can of course be primarily motivated by any number of practical purposes, but my claim concerns what is *semantically* primary in assertion. See footnote 159 for further clarification.

To begin with, what reason is there to think that moral judgments are non-objective, assuming that they are propositional? Let us approach this question by reviewing why taste judgments are plausibly non-objective, and consider how far those same considerations extend to morality. An abridged set of reasons includes:

(i) although there is widespread agreement on many basic taste propositions (e.g. that ice cream is tasty), there are no established methods for decisively settling taste disagreements;

(ii) there are no *natural corrective repercussions* for believing "objectively false" taste judgments when they are aligned with one's own palette;

(iii) objects with no overlapping objective properties often share the same taste appearances (i.e. appear *as fun*, *as interesting*, etc.); and

(iv) taste disagreements typically seem *faultless*.

How far do these considerations reach into the moral domain? With respect to (i), the extension is straightforward. While many basic moral propositions enjoy broad consensus (e.g. that stealing is wrong), there appear to be no canonical ways to decisively settle moral disagreements. It is widely known that scientists do not concern themselves with settling moral questions. And while there are ethicists who claim to have resolved longstanding moral disputes, that any such resolution is possible remains deeply contested. And in any case, there are people who claim to have decisively settled longstanding taste disputes (e.g. that French cuisine is the best in the world).

The perspectival character of both kinds of judgment reappears at the level of applied practice. Culinary specialists, game designers, and artists (to name just a few professions) generally aim to produce things as tasty, fun, or beautiful as they can manage. But we generally understand these goals as efforts to maximize how things *appear* to an intended audience – whether that be "most people," "people like me," or some other group. Similarly, although humanitarians, activists, and policymakers generally aim to improve the lives of others inasmuch as possible, we understand their work to be the maximization of human flourishing *from within a particular framework* of what that involves – where the relevant framework is supplied by each agent's moral constitution.

As for (ii), a similar story can be told in the moral domain. There do not seem to be any *natural corrective repercussions* for "objectively false" moral beliefs, so long as those beliefs are

aligned with one's moral constitution. Suppose that Mary and Jones disagree as to whether it is good to sleep for nine hours each night. Mary finds that when she sleeps any less, she becomes irritable and easily distracted. By contrast, Jones finds that he is well-rested with six hours of sleep, and that any more is a slothful waste of daylight. Let us stipulate that each person's beliefs are well-aligned with their own moral constitution: Mary's life goes best with nine hours of sleep; Jones's, with six. In that case, it is difficult to conceive of any natural repercussion – any obvious cost or crisis – that would reveal which of them is objectively mistaken.

Suppose, for instance, that in his later years, Jones begins to suffer the effects of sleep deprivation and changes his view. Would this show that he was *always* objectively mistaken? Clearly not. It would merely show that Jones's constitution has changed. Jones's earlier belief was true *for him*, just as Mary's is true *for her*. That both can thrive with opposing beliefs, and without incurring any conceivable natural penalty, suggests that 'it is good to sleep for nine hours' is non-objective.[35]

Before turning to (iii), it is worth differentiating objective from non-objective error – for although I claim that moral propositions are non-objective (and thus, cannot be objectively mistaken), there is a clear sense in which moral beliefs can be mistaken. People frequently hold moral beliefs that are misaligned with their own constitution. In *Chapter 4*, I referred to this phenomenon as *substantive yet non-objective error*. If Sophia believes it is wrong to pursue a career in financial management – even in service of non-profits – there is a substantive, though not objective, sense in which she may be mistaken: Sophia may be mistaken about what kind of work is *wrong for her*. Moral belief is often informed by knowledge of which actions promote or undermine flourishing, but having a moral belief does not guarantee such knowledge. Even if moral propositions are non-objective, it simply does not follow that an action φ is wrong for S just in case S believes that φ is wrong. A person can be *objectively* mistaken about what is wrong for them, just as they can be objectively mistaken about what is fun, tasty, or interesting to them.

---

[35] One might object that the dispute between Mary and Jones can be dissolved by appealing to a higher level of generality: what is good for all is simply for each person to sleep however long promotes their health and well-being. But this maneuver either trivializes the issue or simply restates the relativist point. If it is meant to clarify what is good for all, it smuggles the contested notion of goodness into the very criterion that is supposed to explain it. And if it is taken as a substantive claim, it amounts to a concession that what is good for Mary can differ from what is good for Jones, without either being objectively mistaken.

And these personal mistakes help clarify how it is possible to be substantively, yet non-objectively, mistaken about what is wrong *simpliciter* – just as one can be substantively but non-objectively mistaken about what is fun *simpliciter*.[36]

It is also worth briefly pointing out that non-objectivity does not require widespread disagreement, either. On my view, moral propositions are true or false relative to individual constitutions – but most people have broadly similar constitutions. Indeed, moral judgments that enjoy widespread agreement (e.g. that stealing is wrong) reflect cases in which nearly everyone's constitution aligns. For instance, nearly everyone is disposed to lay claim to, and exercise exclusive control over, certain objects (viz. the objects they possess). Theft is thus experienced as a violation by nearly everyone, regardless of what they *believe*.

Finally, that a moral proposition is true for everyone (or nearly everyone) does not entail that it is *objectively* true. For one thing, the scope of 'everyone' is ambiguous. Suppose, for instance, that an aggressive strain of genetically engineered mosquitos escapes containment and threatens ecosystems and human life. The rapid eradication of this strain may be good *for everyone* – but not, of course, *for the mosquitos* themselves. More importantly, an objectively true proposition is true independently of anyone's constitution – true in virtue of how things *really are* – not merely a proposition that happens to be true for everyone (or nearly everyone) in some gerrymandered world. With these qualifications in hand, I return to (iii).

As with (i) and (ii), (iii) extends naturally from the taste domain into the moral domain. In the domain of taste, objects with no overlapping objective features can nonetheless present with the same taste appearances: science-fiction novels and brisk mountain hikes might both appear *as fun*, despite having little in common. Typically, some objective feature explains why objects appear similar: foods that appear *as salty* generally contain salt, and surfaces that appear

---

[36] Why do objective mistakes about what is, say, good *for oneself* entail substantive but non-objective mistakes about what is good *simpliciter*? On the belief expressivist view, belief that *x* is good *simpliciter* is licensed by appearances of *x* as good – appearances shaped by one's moral constitution, much like taste appearances are shaped by one's palate. A person can be objectively mistaken about what is good for them (i.e. misaligned with their constitution) by misidentifying which propositions are licensed by the sum total of appearances their constitution gives rise to. But *a fortiori*, this is to misidentify which propositions are (or would be) licensed by the sum total of appearances. Thus, objective confusion about what is good *for oneself* entails a kind of substantive – albeit non-objective – confusion about what is good *simpliciter*.

as rough tend to be microscopically uneven. But taste appearances do not correspond to any such real properties – and moral appearances even less so. Wildly dissimilar actions can all appear *as good*: donating to a global charity, taking time for daily leisure, and maintaining a disciplined exercise routine might all appear *as good*, though they have almost nothing in common.

One might hypothesize that all actions that appear *as good* nonetheless share some property – for instance, the promotion of desire-satisfaction. But the evidence resists such analyses. Child-rearing, for example, is widely reported as both one of the most morally significant and one of the most desire-frustrating activities. Likewise, learning to delay gratification is celebrated across moral systems. Similar counterexamples arise for other naturalist hypotheses. This reinforces the idea that moral appearances – like taste appearances – are not straightforwardly reducible to objective features, but instead reflect an evaluative projection of the mind. Just as taste appearances supervene on objective features but are fundamentally a product of one's palate, moral appearances supervene on objective features but are fundamentally a product of one's constitution.

Beginning with (iv), however, the analogy between taste discourse and moral discourse begins to break down. While disagreements in taste discourse typically seem faultless (e.g. between one party alleging that cilantro is tasty and the other that it is disgusting), moral disagreements rarely do. To be sure, some moral disagreements *do* seem faultless – as in the examples already discussed, such as whether it is bad to sleep for fewer than nine hours, or whether a particular career path is good. But many other moral disagreements do not seem faultless. For instance, disagreements over whether stealing is wrong, capital punishment is unjust, or abortion is permissible all seem to demand that at least one party is mistaken.

Cases like these pose a *prima facie* challenge to the claim that moral judgments are non-objective. But belief expressivism is not without resources to explain why many moral disagreements seem faultful. For one thing, if the parties to a moral disagreement have the same moral constitution, then at least one of them must be *substantively but non-objectively mistaken* (in the sense explained previously). And if most people have broadly similar moral constitutions, then most moral disagreements will seem faultful in precisely this way. This marks a key

divergence from the taste domain: while variations in palate are common, differences in moral constitution are comparatively rare.

A second potential explanation is that many moral disagreements are practical disagreements about *what to do*, rather than purely theoretical disagreements about an action's or person's moral status. Consider the difference between:

(117) Mary: We should sleep for nine hours tonight.
Jones: No, we shouldn't.

(118) Mary: It is good to sleep for nine hours each night.
Jones: No, it isn't.

While (118) is a theoretical disagreement about the moral status of sleeping in, (117) is a practical disagreement over collective action. And although such disagreements often go hand-in-hand, practical disagreements resolve differently: whereas theoretical disagreements ideally end with one party changing their mind, practical disagreements end when both parties settle on a plan of action – even if neither party changes their mind. For example:

(119) Mary: We should sleep for nine hours tonight.
Jones: No, we should sleep for six hours.
Mary: Fine – six it is. I still think nine would be better, but whatever.

In such cases, successful resolution requires agreement to a plan of action, not convergence of belief. And since only one proposal (or a compromise) can be enacted, it is a feature of practical disagreements that at least one party's proposal will be frustrated or compromised. While this is not the same as being mistaken, it helps explain why many moral disagreements *resist being set aside.* Moral disagreements that are practical in nature (e.g. whether to permit abortion) cannot simply be defused by agreeing to disagree – the disagreement only festers. To avoid indefinite deadlock, the parties to practical disagreements are rationally compelled to "see the disagreement through to its resolution," much like parties to a factual dispute. This practical demand to agree to a common course of action mimics a central feature of objectively faultful disagreements – namely, that they call for a final determination of who is mistaken. At a glance, the rational compulsion to resolve a dispute can easily be mistaken for its faultfulness. Combined with the possibility of substantive but non-objective error, this helps explain why many moral disagreements seem genuinely faultful – even if, on reflection, they are not.

In sum, the case for the non-objectivity of moral judgments closely parallels the case for the non-objectivity of taste judgments (though some qualifications apply):

(i) although many moral propositions enjoy broad agreement (e.g. that stealing is wrong), there are no established methods for decisively settling moral disagreements;

(ii) there are no *natural corrective repercussions* for believing "objectively false" moral judgments when they are aligned with one's constitution;

(iii) objects with no overlapping objective properties often share the same moral appearances (i.e. appear *as good*, or *as wrong*, etc.); and

(iv) although moral disagreements typically seem *faultful*, (a) some clearly do not, (b) some involve substantive but non-objective error, and (c) some are practical disputes which demand resolution, despite not being strictly faultful.

With the case for the non-objectivity of moral propositions complete, I will conclude by taking stock of the differences between moral and taste propositions. These differences will highlight that not all non-objective propositions are made equal, and invite pluralism about non-objective discourses.

I have already discussed a few of the differences between moral and taste propositions. First, differences in palate are more widespread than differences in moral constitution. As a result, moral disagreements are more likely than taste disagreements to involve substantive but non-objective error. Second, moral disagreements often spill over into practical disputes about a common course of action in ways that taste disagreements do not. Consequently, moral disagreements are more likely to *resist being put aside* – they often rationally compel disputants to reach a resolution (and not to agree to disagree) in ways that taste disagreements do not.

A third and closely related difference is that moral propositions are intrinsically motivating[37] in a way taste propositions are not. For instance, to accept that stealing is wrong is, by itself, to be at least somewhat motivated – to have a reason, if only a defeasible one – not to steal, and to discourage others from doing so. By contrast, one can believe that rollercoasters are

[37] I will not attempt to resolve the debate between moral internalism and externalism here, as it lies beyond the scope of this discussion. While I am broadly sympathetic to internalism – and regard it as a desirable feature for any account of moral belief to accommodate – if externalism turns out to be correct, then the gap between moral and taste discourse narrows further. In that case, accommodating moral discourse within a belief-expressivist framework becomes even more straightforward, given the account already in place for taste discourse.

fun, that sushi is tasty, or that exercise is painful, without encountering the slightest motivation to ride rollercoasters, eat sushi, or avoid exercise. To produce motivation from such beliefs, they must be paired with some further desire – to have fun, to eat something tasty, to avoid pain, and so on. And at least some of the time, we simply lack these desires.

It is worth briefly considering why it might be that moral beliefs are uniquely motivating. A familiar explanation appeals to their role as answers to practical questions. On this view, we ask ourselves: *what should I do?* – and moral propositions serve as candidate answers. The moral propositions one accepts – one's moral beliefs – thus bear directly on one's practical deliberation. And since settling questions about what to do is inherently action-guiding, and most beliefs do not directly bear on such questions, this helps explain why moral beliefs are intrinsically motivating, and why most other beliefs – taste beliefs included – are not.

Although I do not doubt that there are further differences between taste and moral discourse worth exploring, the fourth and final one I will consider here concerns the nature of the appearances that license moral and taste beliefs. First, a note on phenomenal appearances. In Section 5.1.3 (my defense of (**PI**)), I adopted a fairly expansive understanding of what counts as an appearance. It is not meant to refer narrowly to visual experiences, nor even to sensory experiences more broadly – though both certainly qualify. An appearance, as I use the term, is any *seeming* of which we are aware. This includes the kinds of truths we grasp by rational intuition – for instance, that *modus ponens* is valid, or that the law of the excluded middle holds. These are truths that simply *seem* true upon reflection. And as *seemings* of which we are aware, I include them in the broader pantheon of appearances.

With that in mind, it is worth observing that taste appearances are paradigmatically sensory in nature: they typically arise alongside experiences like eating chili, or riding a rollercoaster. Moral appearances, by contrast, are often of a peculiarly rational kind. We are able to apprehend that a wide range of actions – stealing, killing, betraying a friend – are wrong, even without directly witnessing them. In sharp relief, many taste beliefs seem to require firsthand sensory awareness. For instance, I cannot come to know that a meal is tasty just by thinking about it. It is doubtful whether I can even know it on someone else's authority – I *must* try it

myself. So, moral and taste appearances would seem to diverge: the former are prototypically rational, the latter prototypically sensory.

Still, while I'm happy to grant this distinction between paradigm cases, I do not think it tracks a deep or decisive divide between the two domains. For one thing, we often regard life experience as essential for genuine moral knowledge. It is widely accepted that one's moral beliefs are directly shaped by their *lived experience* – that someone who comes of age amid a ruinous civil war will simply have a different moral understanding than someone insulated from every conceivable hardship. So, it seems that experience plays a not-insignificant role in our moral life, after all.

From the other end, many taste appearances are amenable to, or even arise from, rational reflection. Consider how the appearance of something as *interesting* – a paradigmatic taste appearance – is often triggered by purely rational considerations. We reflect on some curious historical event, and it appears *remarkable*. We imagine possible game states in a strategy game, and they seem *fascinating*. We recall an old joke, and it immediately appears *funny*. We compose a line of poetry, and it seems *beautiful*. In what sense are these appearances sensory? And if these *are* sensory appearances, then I submit that moral appearances are sensory too.

We have seen, then, that one can go a long way in distinguishing the moral domain from the taste domain, even within the broader category of non-objective domains. Differences in moral constitution are less common and less pronounced than differences in palette. Moral disagreements frequently take on a practical dimension that taste disagreements lack. Moral beliefs seem to be intrinsically action-guiding in ways that taste beliefs are not. And the appearances that ground moral judgments are paradigmatically rational, while those underlying taste judgments are typically sensory.

Yet we have also seen that the boundaries between domains are often quite porous. Both moral and taste domains are evaluative in nature – and we have remarkably similar reasons to think that both are non-objective. Both allow for cases of substantive but non-objective error, even if such cases are more common in moral discourse. Life experience plays a colorful role in our moral life, and many taste appearances arise from rational reflection.

Finally, there is the matter of a shared evaluative vocabulary. I say of the meal I ate, that it is *good*; of the novel I read, that it is *good*; of my dog Sammy, that she is *good*; of the work of so many philosophers, that it is *good*; of my advisors' boundless generosity, that it is *good*; and of my wife's steadfast support, that it is *good*. I tell myself that these are different senses of the word – but I am quite unsure.