From the holdings of the Department of Special Collections , Princeton University Library, Princeton, NJ

These documents are either in the public domain, or can only be used for educational and research purposes (Fair use) as per U.S. Copyright law (text below). By accessing this file, all users agree that their use falls in the public domain or within fair use as defined by the copyrightlaw. We do not charge any permission or usage fees for the publication of images of material inour collections, including those provided by Princeton University Library, via our website, catalog, or directly from staff.For more information regarding copyright, credit, and citation please review the information onthe Copyright, Credit and Citation Guidelines page of our website.
https://library.princeton.edu/special-collections/policies/copyright-credit-and-citation-guidelines. If you have any questions, including how to order higher-quality images from the collection, contact Special Collections Public Services via our Ask Us! Form. https://library.princeton.edu/special-collections/ask-us

---

Doctoral Dissertations, 2011-2026 Philosophy

2024

MORAL ABOUTNESS: INDEPENDENCE, OBJECTIVITY, AND DISAGREEMENT

Salinger, Elliot

Princeton, NJ : Princeton University

---

http://arks.princeton.edu/ark:/88435/dsp01f7623g97t

U.S. Copyright law: The copyright law of the United States (Title 17, §108, United States Code) governs the making of photocopies or other reproductions of copyrighted material. Under certain conditions specified in the law, libraries and archives are authorized to furnish a photocopy or other reproduction. One of these specified conditions is that the photocopy or reproduction is not to be used for any purpose other than private study, scholarship, or research. If a user makes a request for, or later uses, a photocopy or reproduction for purposes in excess of fair use that user may be liable for copyright infringement. This institution reserves the right to refuse to accept a copying order if, in its judgment, fulfillment of the order would involve a violation of copyright law. https://www.copyright.gov/title17/title17.pdf

MORAL ABOUTNESS: INDEPENDENCE, OBJECTIVITY, AND DISAGREEMENT

Elliot Moshe Salinger

A DISSERTATION

PRESENTED TO THE FACULTY

OF PRINCETON UNIVERSITY

IN CANDIDACY FOR THE DEGREE

OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE BY

THE DEPARTMENT OF PHILOSOPHY

Advisers: Gideon Rosen, Michael Smith

September 2024

© Copyright by Elliot Moshe Salinger, 2024. All rights reserved.

Dissertation Abstract

Metaethical realists standardly object that metaethical anti-realism violates the *independence* of normative ethics from metaethics, fails to respect the *objectivity* of morals, and cannot account for moral *disagreement*. This dissertation analyzes the notions of independence (**ch. 1**), objectivity (**ch. 2**), and disagreement (**ch. 3**) in terms of *subject matter* and explores anti-realist responses to the objections on this basis.

Chapter 1 addresses the objection that expressivism is revisionary of ordinary moral thought. The standard response to this objection is best understood as follows: expressivism is independent of normative ethics iff non-cognitivism about moral judgment does not rule out any plan about what to blame for. Although this response secures the *taxonomic* independence of expressivism from normative ethics, it does not secure its *moral* independence.

Chapter 2 develops an account of objectivity as not being about the subject (the relevant thinker or speaker or judge). Something is about a subject matter *S* in the relevant, *restricted* sense iff it *foregrounds S*. So a domain remains objective in the relevant sense even if the subject is in the *background*. Our concern for restricted objectivity stems from our concern for *moral aboutness*: not foregrounding *oneself* when morally significant interests are at stake. The account classifies as objective not only expressivism, but also anthropocentric forms of cognitivism and perhaps even error-theoretic cognitivism.

Chapter 3 develops an account of (dis)agreement in terms of subject matter or questions. Whenever A and B (dis)agree, there is some question *S* that their (dis)agreement is *about.* A and B agree about *S* when they accept the same answer to *S*; they disagree about *S* when they accept incompatible answers to *S*. To accept an answer to *S* is to have a committing attitude toward a

content that answers *S*, and answers are incompatible when non-jointly-realizable relative to *S*. This account countenances doxastic and practical disagreement, but not disagreement in attitude. It also sheds light on solutions to the problem of normative disagreement: the expressivist and contextualist, but not the relativist, have accounts on which normative disagreement is *genuine* and so on which the normative is objective.

Acknowledgements

I owe a tremendous debt of gratitude to the Department of Philosophy, which has been my intellectual home almost uninterruptedly for the past decade. I wish to thank, again, the many members of the faculty who introduced me to analytic philosophy and helped prepare me for graduate school. Among them, I thank especially Hans Halvorson, my undergraduate senior thesis adviser, for his intellectual mentorship and encouragement.

I am particularly grateful to my dissertation advisers, Gideon Rosen and Michael Smith, and to Sarah McGrath, who although not formally a dissertation adviser was very much a third adviser to me. Gideon, Michael, and Sarah helped give shape to my inchoate ideas and form them into their most philosophically interesting versions. I am grateful for every conversation and round of comments, and for the patience and warmth that always attended them.

Thanks to Mark Johnston and Una Stojnić for important philosophical conversations about my work; to Michael, Johann Frick, and Tom Kelly for guidance in teaching; and to Liz Harman for professional advice and support when I was a UCHV Graduate Prize Fellow in AY 2022–2023. I also thank the administrators in the Department of Philosophy and the UCHV for all their logistical help and support, as well as my peers in the graduate program for philosophical conversation, advice, and camaraderie.

Thanks to my parents Amy and Jay Salinger for their continued love and support in every way, and thanks to my parents-in-law Dr. David Ricklan and Dr. Deborah Mehl for their many years of love and support. Thank you, Sarah, for everything. And thank you, Claire, my study buddy for these last several months, who has been a source of joy and light in dark times.

I express my final acknowledgment through silence. "לך דמיה תהילה."

Contents

Dissertation Abstract **3**

Acknowledgements **5**

Introduction **7**

Chapter 1: Expressivism and Moral Independence **14**

Chapter 2: Moral Objectivity as Moral Aboutness **43**

Chapter 3: Genuine Disagreement **75**

References **107**

Introduction

This dissertation concerns a set of classic, and related, problems faced by versions of metaethical anti-realism. Metaethical realists standardly object that anti-realism 1) violates the *independence* of normative ethics from metaethics, 2) fails to respect the *objectivity* of morals, and 3) cannot properly account for moral *disagreement*. I develop accounts of the notions of independence, objectivity, and disagreement which take into account their connection to *subject matter* or *aboutness*; and I deploy these accounts to defend versions of anti-realism against the classic objections.

Chapter 1 defends expressivism by showing that, at least given expressivist assumptions about the nature of moral judgement, the view is independent of normative ethics in the sense that it has no bearing on normative ethical subject matter. Chapter 2 proposes a novel, aboutness-theoretic analysis of objectivity to argue for the objectivity of a strikingly wide range of anti-realist views in metaethics. Finally, Chapter 3 develops a question-relative account of disagreement to argue that the expressivist has, the contextualist might have, and the relativist probably lacks, an account on which normative disagreement is genuine.

Chapter 1, "**Expressivism and Moral Independence**," (Published as Salinger 2024) addresses the objection that owing to its commitment to non-cognitivism about moral judgment—the view that moral judgment is desire-like rather than belief-like—expressivism is revisionary of ordinary moral thought. The standard response to this objection is to say that since the expressivist's theoretical commitments about the nature of moral judgment are *independent* of normative ethics, the view cannot be revisionary of normative ethics. In this chapter I clarify

the objection and evaluate the standard response by exploring several senses of independence that expressivism might enjoy from normative ethics.

Although the claim of non-cognitivism about moral judgment is standardly thought by its defenders to be a metaphysical claim and therefore not a moral claim, in the present context this cannot simply be assumed. Accordingly, I develop a taxonomy to adjudicate this question. A natural way to develop a philosophical taxonomy is by appealing to the notion of *subject matter*, what a sentence or proposition is *about*. I develop my taxonomy within a generalized version of the theory of subject matter due to David Lewis, on which subject matters are (indirect) questions. On this theory, some content *c* is about a subject matter *S* iff *c* answers *S*. Within this theory, we can say that *c* is independent of *S* iff *c* has no bearing on *S*, which is to say that that there is no answer to *S* that *c* rules out. So expressivism will be independent of normative ethics in the defined sense if non-cognitivism about moral judgment does not rule out any answer to the question constituting normative ethical subject matter.

I then provide an expressivist theory of normative ethical subject matter. Following Allan Gibbard, I understand this subject matter to be *practical*, answering the question of *what to do*. Moral subject matter narrowly construed will be some smaller subject matter included in the larger subject matter of *what to do*, such as *what to blame for*. To assess whether non-cognitivism about moral judgment has a bearing on this subject matter, I generalize Lewis's theory to accommodate practical questions and contents and develop a theory of when descriptive content rules out practical content, or *plans*. Given this account, I argue that expressivism has no bearing on normative ethical subject matter because (funny cases aside) the non-cognitivist thesis does not rule out any blaming plans. In so arguing, I distinguish between two notions often conflated: the *presuppositions* of a plan—what must be the case for it to be metaphysically possible to carry

out a plan—and the *circumstances* a plan is for—the situation in which a plan becomes actionable.

Although this chapter essentially vindicates the standard response to the original objection, it also highlights a major limitation thereof. This is that even for the expressivist, the *taxonomic* independence of non-cognitivism about moral judgment from normative ethics does not guarantee its *moral* independence. That is, showing that non-cognitivism about moral judgment is not a moral claim is not itself to show that it is not a morally *relevant* claim, a claim whose truth or falsity might have implications within some normative ethical theory. That is a first-order moral question rather than a taxonomic one, and so it can only be resolved on first-order moral grounds.

Chapter 2, "**Moral Objectivity as Moral Aboutness**," develops a novel, aboutness-theoretic account of objectivity and subjectivity. On this account, subjectivity consists in being about the subject (subjectivity as *subject-aboutness*)—where 'the subject' denotes the relevant thinker or speaker or judge—and the opposing notion of objectivity consists in being about some object other than the subject (objectivity as *non-subject-aboutness*, or *object-aboutness*). To know whether some theory of moral thought, moral language, or moral metaphysics is subjective, we need to know what it is for a thought, sentence, or fact to be about the subject. Theories of aboutness may be *restricted* or *unrestricted*. To see this distinction, consider a painting of a harvest scene set against lakes and rolling hills. If on an unrestricted conception the painting is about everything it depicts—the harvest scene as well as the lakes and hills—on a restricted conception the painting is only about the harvest scene in its *foreground*, not the lakes and hills in its *background*. Since in its restricted sense something is about a subject matter *S*

only if it foregrounds *S*, a domain is subjective given the restricted sense of aboutness only if it foregrounds the relevant thinker or speaker or judge.

Turning from a general discussion of objectivity to the question of moral objectivity specifically, I argue that the sort of objectivity that the metaethical anti-realist hopes to secure is one on which moral thought, moral discourse, and moral metaphysics are not about the subject in the *restricted* sense, the sense in which a domain is objective despite the subject's presence in the background. I argue for this claim in part by arguing that our concern for securing the objectivity of morals has a basis in concerns familiar from moral theory and ultimately from everyday moral life, what I call our concern for *moral aboutness*. This is that we are concerned about who is centered in situations where morally significant interests are at stake, and we are especially concerned that people not make such situations *about themselves*.

Accordingly, I propose the following substantive criteria for when a domain *D* of thought, language, or facts foregrounds the subject: *D*-thought foregrounds the subject if the subject features in *D*-thought's *aim of inquiry*, *D*-language foregrounds the subject if the subject features in a *question D*-language answers by default, and the *D*-facts foreground subject if they are *partly grounded* in facts about the subject. Each defining foreground notion corresponds to an opposing *background* notion: mental or linguistic content that is *not-at-issue*—i.e. that is irrelevant to the aim of inquiry or default question—and *background conditions* on grounding relations. Given the restrictive conception of aboutness, an anti-realist metaethical theory will count as objective even if, according to the theory, the subject features in the not-at-issue content of moral thought or language or in the background conditions of the moral facts.

These substantive criteria allow us to distinguish objective theories of moral thought, moral language, and moral metaphysics from subjective theories. My account of objectivity

places a striking range of views on the objective side of this divide, including not only the expressivism of Blackburn and Gibbard, but also anthropocentric forms of cognitivism such as constructivism, and perhaps even error-theoretic cognitivist views. Traditional anti-realist accounts of objectivity—developed mainly by expressivists—are best understood as providing non-cognitive analyses of the rhetoric of objectivity, or else of what it is to believe in objectivity. But since my theory is a constitutive one, it can make the stronger claim that expressivist interpretations of moral thought and moral language simply are forms of objectivism.

Finally, I address the objection that my account renders judgments of taste objective. I argue that thought about what is tasty is derivative of a subject's experience with the thing judged tasty, and so has a degree of subjective involvement that moral thought lacks.

Chapter 3, "**Genuine Disagreement**," develops an original account of disagreement and explores its implications. It has gone largely unnoticed that agreement and disagreement bear the following conceptual connection to subject matters or *questions*: whenever A and B agree or disagree, there is some question *S* that their agreement or disagreement is *about.* I propose the following analyses: A and B agree about *S* when they accept the *same* answer to *S*; they disagree about *S* when they accept *incompatible* answers to *S*. To accept an answer to *S* is to have a committing attitude toward a content that answers *S* (belief is a committing attitude, supposition is not), and two answers to *S* are incompatible when they are non-jointly-realizable relative to *S*.

In doxastic disagreement, where A believes that *p* and B believes that ~*p*, A and B accept incompatible answers to the question *whether* p. This is because A and B have committing attitudes—beliefs—with contents—*p* and ~*p*—that are not jointly realizable relative to the question *whether* p. Now, just as beliefs are committing attitudes with descriptive or

propositional contents, intentions (or a state suitably related to intention) are committing attitudes with practical—or *planning*—contents. In practical disagreement, subjects accept incompatible answers to some practical question by having intentions whose contents are not jointly realizable relative to *what to do*, as when I intend for us to spend the evening at the cinema and you intend for us to spend it rather at the symphony.

The orthodox view is that disagreement is always doxastic; heterodoxy admits disagreement in *interest* or *attitude*. My account countenances both doxastic disagreement and practical disagreement, but not mere disagreement in attitude. On my account, disagreement requires divergent attitudes to concern the same question; it is not sufficient merely that they diverge.

The shared question requirement also distinguishes my account from those on which A and B disagree whenever it would be incoherent for a single subject to have the attitudes of both A and B. On this sort of account, subjects with contrary preferences or credences always disagree; on my account, however, the matter is more complex. In having contrary preferences or credences, subjects do not disagree about *what to prefer* or *what credence to have*, since these are practical questions answered by plans about what to prefer or about what credence to have rather than by preferences or credences themselves. They might disagree about *what is preferable* or *what is credible*, though this seems rather like doxastic disagreement about the instantiation of a normative property. I conclude that preference-involving disagreement parallels practical disagreement, and credence-involving disagreement is importantly linked to subjects' judgments about the likelihood of *p* given some body of evidence (even if these judgments are *sui generis* and not reducible to beliefs).

Finally, I discuss the implications of my account for the problem of normative disagreement. This is the problem, faced by metanormative views on which normative language does not straightforwardly express beliefs about normative reality, of explaining why subjects with contrary normative views thereby disagree. I clarify this problem by distinguishing *superficial* from *genuine* disagreement—linguistic disagreement is genuine when it is backed in the right way by mental-state disagreement; otherwise it is superficial—and I evaluate extant solutions to the problem proposed by expressivists, contextualists, and relativists. I conclude that the expressivist has, the contextualist might have, and the relativist probably lacks, an account on which normative disagreement is genuine. Finally, I discuss the conceptual connection of disagreement to objectivity. In my preferred formulation, a domain *D* is objective just in case it admits of genuine disagreement. If we accept this dictum, we can say that the expressivist has, the contextualist might have, and the relativist probably lacks, an account on which the normative can be objective.

Chapter 1: Expressivism and Moral Independence

1. Introduction

Metaethical expressivists maintain that moral language expresses non-cognitive attitudes: intentions or desires or states of approval and disapproval. Expressivism faces the perennial objection that its commitment to non-cognitivism about moral judgment renders the view revisionary of our ordinary moral thought. The standard response to this objection is to say that since the expressivist's theoretical commitments about the nature of moral judgment are *independent* of normative ethics, the view cannot be revisionary of normative ethics.

This essay seeks to evaluate the standard response to the perennial objection by exploring several senses of independence that expressivism might enjoy from normative ethics. Although the non-cognitivist claim is standardly assumed to be non-moral, in the present context this cannot simply be assumed. Accordingly, I develop a taxonomy to adjudicate this question. I argue that, at least by the expressivist's own lights, normative ethics is not dependent on non-cognitivism about moral judgment in a way that might render that claim itself normative ethical. The argument will require us to formulate a theory of *moral subject matter* according to the expressivist, a matter of independent interest that contributes to our understanding of the view generally.

Although this discussion will essentially vindicate the standard response to the perennial objection, it will also highlight a major limitation thereof. This is that even for the expressivist, the *taxonomic* independence of non-cognitivism about moral judgment from normative ethics does not guarantee its *moral independence* from normative ethics; that is, showing that non-cognitivism about moral judgment is not a moral claim is not itself to show that it is not a

morally *relevant* claim, a claim whose truth or falsity might have implications within some normative ethical theory. This question, usually discussed under the heading of *objectivity*, is ultimately first-order moral rather than taxonomic, and so can only be resolved on first-order moral grounds.

2. The perennial objection and the standard response

I begin by presenting one version of the perennial objection, pressed by Ronald Dworkin (1996; 2011).[1] Dworkin's argument concerns the expressivist's project of *accommodation*.[2] This is the project of providing non-cognitive interpretations not only of simple moral sentences such as 'stealing is wrong,' but also of fragments of our moral discourse that might have seemed available only to a metaethical realist, such as 'stealing would be wrong even if I didn't

[1] See (Blanshard 1949) for an early version of this sort of objection, which is often put as the claim that expressivism amounts to a form of *subjectivism* that we have good first-order grounds to reject. (See **n3**.) In any case, these *normative* objections to expressivism should be distinguished from those arguing that expressivism is a form of subjectivism by appealing to considerations from the philosophy of language. The classic statement of this sort of objection is due to (F. Jackson and Pettit 1998); for some responses see (M. Smith and Stoljar 2003; F. Jackson and Pettit 2003; Dreier 2004a; Suikkanen 2009; Schroeder 2014).

[2] The term is Blackburn's (1993a, 153). This way of understanding the upshot of Dworkin's objections roughly follows the discussion in Gibbard (2003, chap. 9), although that discussion concerns (Dworkin, 1996) exclusively—(Dworkin, 2011) not having been published—and Dworkin's two discussions vary in several ways. (I will not, in the text, address these subtleties.) Expressivists such as Blackburn and Gibbard who take on the project of accommodation are known as *quasi-realists*.

disapprove of it.' For example, if 'stealing is wrong' expresses disapproval of stealing, then 'stealing would be wrong even if I didn't disapprove of it' might express rejection of conditionalizing one's disapproval on one's own sensibility.[3] Dworkin argues that the project of accommodation is in principle unable to succeed, since it faces the following dilemma: if the expressivist cannot accommodate the entirety of ordinary moral discourse, the view fails; but if the expressivist can, then the view will no longer be a form of anti-realism. Inevitably, Dworkin thinks, whatever the expressivist says to distinguish the view from realism will be revisionary of ordinary moral thought.[4]

The standard response to the objection is this: The expressivist's core theoretical commitment is non-cognitivism about moral judgment, the view that moral thought consists (ultimately) in having a conative or affective mental state rather than in a(n ordinary,

[3] This is roughly Blackburn's (1993b, 172–74) suggestion. The interpretation of such counterfactual conditionals is central to Blackburn's argument that his view does not posit a problematic dependence of the moral facts on our attitudes, one facet of morality's objectivity. See (Cassam 1986, 450–51; Köhler 2014; Berker 2020) for a discussion of this strategy. I return to a discussion of objectivity and its relation to independence in **§6**.

[4] I should note that in *Justice for Hedgehogs* (Dworkin 2011, chap. 3), this version of the critique is not, strictly speaking, leveled against quasi-realist expressivism, but rather against earlier versions of non-cognitivism such as the views of Ayer (1946) and Stevenson (1937). In the (2011) treatment, Dworkin reads Gibbard and Blackburn (at least in their later work) as grounding their anti-realism in a distinction between two language games: that of 'morality' and that of 'philosophy.' As stated, this is implausible—it reads Gibbard and Blackburn more as fictionalists than as expressivists—so I will not address it here.

representational) doxastic state.[5] Since the realist denies non-cognitivism about moral judgment, expressivism is distinct from realism.[6] But since non-cognitivism about moral judgment is not a moral claim, it is not revisionary of ordinary moral thought. Michael Smith offers this response to Dworkin's version of the objection, explaining that the cognitivism that the expressivist denies is a "a distinctively *philosophical* thesis, ultimately, a metaphysical thesis about the constitution of moral belief," (M. Smith 2010, 519) where in context categorizing a claim as philosophical is a way of categorizing that claim as non-moral.[7]

---

[5] Early forms of non-cognitivism (Stevenson 1937; Ayer 1946) held that moral sentences are not truth-apt and therefore cannot amount to knowledge, and for this reason Blackburn eschews the label for his own view. But Blackburn's and Gibbard's views may still be called non-cognitive in the non-pejorative, technical sense given in the text. The parenthesized elements highlight that the expressivist can—and does—maintain that there are moral beliefs, but that these beliefs are extraordinary rather than "prosaic" (Gibbard's term), derivative of non-cognitive states and not representational.

[6] This briefly summarizes Dreier's (2004b) "'explanation' explanation" solution to the problem of creeping minimalism, which is just the problem of distinguishing a fully successful expressivism from realism. Dreier's solution is that for the expressivist, the explanation of facts of the form [*S* judges that *X* is *N*], where '*N*' is some normative property, makes no mention of *N*-ness or having *N* among the explanantia, whereas the realist must cite among the explanantia something involving *N*. What the expressivist *will* cite in the explanation of [*S* judges that *X* is *N*], rather than anything about *N*, is the fact that *S* has some non-cognitive attitude concerning *X*. See also (Fine 2001; O'Leary-Hawthorne and Price 1996; Gibbard 2003, chap. 9) for related thoughts.

[7] Earlier, Smith appeals to the philosophical/moral distinction when defending the coherence of the error theory, writing that the premises of the error theorist's argument are "all distinctively *philosophical* claims and not moral

Expressivists have thought that their analysis of moral judgment is not itself a moral thesis because it makes a claim not about what has *N*, where *N* is some normative property, but rather about what it is for someone to judge that something has *N*. This move, which Jamie Dreier has called the "expressivist sidestep," (Dreier 2015, 283) features prominently in Gibbard's classic argument for expressivism (1990, chap. 1). There Gibbard argues that no analysis of normative concepts such as rationality adequately captures the role they play in our language and thought. For it seems that two parties can agree about all the straightforwardly descriptive facts and yet *genuinely disagree* about what is rational—that is, their difference counts as disagreement, and they can disagree without talking past one another and without conceptual confusion. So, even if we believe that an action is rational if and only if it (e.g.) maximizes expected utility, it seems mistaken to *analyze* the concept of rationality in terms of the concept of maximizing expected utility.[8] If so, then the only illuminating conceptual analysis in the vicinity, Gibbard concludes, must be one not of what it *is* for a thing to be rational, but rather of what it is to *judge* a thing rational.

Still, we might worry that the standard response essentially begs the question against the objection that expressivism is revisionary of our ordinary moral thought: why should we think that a metaphysical claim about the nature of moral judgment could not also be a moral claim? Opponents of expressivism are likely to maintain that it is precisely expressivism's commitment to non-cognitivism about moral judgment that marks the view's deviation from ordinary moral thought: that this commitment either involves or entails some problematic first-order content. To

---

claims" (M. Smith 2010, 514).

[8] This thought is closely connected to G.E. Moore's arguments against naturalism (Moore 1903, chap. 1). For further discussion see (F. Jackson 2008; Köhler 2012).

put the standard response to the objection on more solid footing, what is needed is a principled basis for the thesis that non-cognitivism about moral judgment is independent of any moral claim. Specifically, since this thesis is *taxonomic* in nature, what is needed is a set of criteria for classifying claims as 'moral' and as 'metaphysical'—or at least, as 'non-moral'—and these criteria must yield the result that non-cognitivism is a non-moral claim. I develop such a taxonomy in the following section.

3. Moral taxonomy

A natural way to develop a philosophical taxonomy is by appealing to the notion of *subject matter*, what a sentence or proposition is *about*. I develop my moral taxonomy within a generalized version of the theory of subject matter due to David Lewis (1988a).[9] On such a scheme, to call a claim or proposition moral is to say that it has a moral subject matter.[10]

[9] I am not the first to use Lewis's theory of subject matter to construct a taxonomy of moral and non-moral claims (see Humberstone 1996). Lewis himself (1988b) uses the theory to develop a taxonomy of observation-statements and non-observation-statements. There is a parallel between these taxonomic projects: many believed there to be an inference barrier between observational and theoretical claims, just as many believe there to be an inference barrier between 'is'- and 'ought'- claims. The latter principle is known as *Hume's Law*, which I discuss in greater detail in **§4.** I thank an anonymous referee for *Philosophy and Phenomenological Research* for raising this point.

[10] Some other recent moral taxonomies include the pragmatic-commitment proposal of (Dreier 2002) and the ground-theoretic proposal of (Maguire 2015). I should note that on both taxonomies, non-cognitivism about moral judgment comes out as a non-moral claim: for Dreier it is non-moral roughly because it is

Subject matters can be designated using simple noun phrases, as in <the seventeenth century> or <the number of stars>. But for our purposes it will be best to understand subject matters as *indirect questions*, such as <what occurred in the seventeenth century> or <how many stars there are>.[11] On this conception, I propose we understand the general notion of being about a subject matter as follows:

(Aboutness) content *c* is about some subject matter <*S*> iff *c* is an (exact)[12] answer to the question corresponding to <*S*>.

---

assertible no matter which moral theory one accepts, and for Maguire it is non-moral roughly because it is fully grounded in non-moral facts.

[11] Lewis himself notes that subject matters may be understood as questions (1988a, 162), though Lewis officially analyzes subject matters as equivalence relations on possible worlds, or equivalently as partitions of logical space. (I assume that any subject matter that can be denoted by a simple noun phrase can alternatively be denoted as an indirect question. Even a subject matter such as <Canada> can be denoted as an indirect question such as <how thing stand with respect to Canada>.)

[12] My notion of being an exact answer corresponds to Lewis's notion of being *wholly* rather than *partially* about, where the relevant sense of partial aboutness is the *part-of-content* sense (Lewis 1988b, 17). When *c* is equivalent to the conjunctive content $m \wedge n$, then *c* is in this sense partly about any subject matter that *m* is wholly about and partly about any subject matter that *n* is wholly about; for example, *grass is green and it's raining* is partly about <what color grass is> and partly about <whether it's raining>. The definitions that follow should all be understood as pertaining to exact answers in this sense. But since none of the substantive discussion will turn on whether an answer to some question is exact or not, I will not venture a precise definition of exactness for answers. (See Cross and Roelofsen 2022, sec. 2 for a discussion of how to construe the notion of an answer generally within the theory of questions.)

For example, *there are exactly one trillion stars* is about <how many stars there are> because *there are exactly one trillion stars* is an answer to the question of how many stars there are. Another notion we can define is *subject matter inclusion*:

(Inclusion) <*S*> includes <*T*> if and only if every answer to <*T*> is also an answer to <*S*>.

For example, <what occurred in the seventeenth century> includes <what occurred in 1688> because any answer to the question of what occurred in 1688 is also an answer to the question of what occurred in the seventeenth century. Two more notions we can define are some content's *having no bearing* on a subject matter and of two subject matters' being *orthogonal* to one another:

(Having no bearing) *c* has no bearing on <*S*> iff *c* does not rule out any answer to <*S*>.

(Orthogonality) <*S*> and <*T*> are orthogonal iff no answer to <*S*> rules out any answer to <*T*> (or vice versa).[13]

If <*S*> and <*T*> are orthogonal, then no *c* about <*S*> has a bearing on <*T*> and no *c* about <*T*> has a bearing on <*S*>.

The stronger notion of orthogonality and the weaker notion of having no bearing are each suitable to serve as the basis for an independence thesis. For our purposes, we should focus for now on the latter, weaker notion: because the objection pertains narrowly to *expressivism* inasmuch as it is committed to non-cognitivism, it should be enough to show that non-cognitivism has no bearing on normative ethical subject matter in order to show that

[13] Cf. Lewis's informal characterization of orthogonality: "two subject matters M_1 and M_2 are *orthogonal* iff, roughly, and way for M_1 to be is compatible with any way for M_2 to be" (1988a, 167).

expressivism is independent in the desired sense. This sort of independence is captured by the following principle:

> (INDEPENDENCE AS HAVING NO BEARING) *p* is independent of <*S*> if *p* has no bearing on <*S*>.

According to this principle, non-cognitivism about moral judgment is independent of normative ethics if it has no bearing on normative ethical subject matter. Of course, if the subject matter of non-cognitivism, namely <what is moral judgment>, were orthogonal to the subject matter of normative ethics, that would entail the weaker claim that non-cognitivism about moral judgment has no bearing on normative ethics. Indeed, this seems to be the claim of the standard response, which is that *no* claim about the nature of moral judgment—and so not just the claim that moral judgment is non-cognitive—has a bearing on normative ethics. But this claim is stronger than what is needed to vindicate expressivism's independence from normative ethics, on INDEPENDENCE AS HAVING NO BEARING.

3.1. Normative ethical subject matter

To discover whether expressivism is independent of normative ethics in the sense of INDEPENDENCE AS HAVING NO BEARING, we must ask whether expressivism rules out any answer to the question corresponding to the subject matter of normative ethics. So, what is the subject matter of normative ethics? We can construct an expressivist theory of normative ethical subject matter using materials from Gibbard's *Thinking How to Live* (2003). Consider the following passage:

> Questions of what we ought to do are questions of what to do. Finish your deliberation, conclude what to do, and you've concluded what you ought to do. These crude sayings will, of course, need qualification, but the distinctive claim of an expressivist is that dicta

like these, suitably worked out, account for the subject matter of ethics. (Gibbard 2003, 19)

Gibbard here is essentially explicit in identifying the subject matter of ethics with <what to do>, which is itself equivalent to the titular subject matter of <how to live>.[14] Moral subject matter narrowly construed—what I have been calling normative ethical subject matter—will then be some smaller subject matter included in this larger subject matter: perhaps <what to do> restricted to some class of situations (e.g. situations in which certain kinds of interests are at stake) or, following a suggestion in Gibbard's earlier work, <what to blame for>.[15] (In what follows I choose for concreteness to take <what to blame for> to be the subject matter of normative ethics, but nothing will depend on this choice.)

Until the past paragraph, our example subject matters had all been *descriptive*. But notice that subject matters such as <what to do>, <how to live>, and the smaller subject matters they include are not descriptive but *practical*. (It is because Lewis's theory does not address practical subject matters that we needed to generalize the theory.[16]) Descriptive questions are answered

[14] Gibbard's identification of 'what I ought to do' with 'what to do' generally is the basis of his (2003) theory of normative content. Gibbard claims that we can clearly distinguish between planning to φ on the one hand and believing that φ-ing has a *sui generis* property of to-be-doneness on the other, and he argues that judgments of what I ought to do must be identified with the former rather than the latter. I discuss this identification more in **n21**.

[15] (Gibbard 1990, 40–48). Or, following the amendment of (Schroeder 2008a), <what to be for blaming for>. For a discussion of how to analyze other moral notions in this framework, see (Darwall 2022).

[16] The classic logical study of questions (Hamblin 1958) similarly neglects non-descriptive subject matters. See (Balcerak Jackson 2019) for an extended discussion of practical questions.

with *propositional content* which, in an intensional framework, is given by sets of possible worlds.[17] What sort of content answers a practical question the way that propositional content answers a descriptive question? Again following Gibbard, let us take the basic notion of practical content to be a *plan*: a possibility for action.[18] If I am deliberating about *what to eat for dinner* in my actual situation, my deliberation will conclude with a plan: a plan to eat salmon, a plan not to eat pasta, etc. Modeled formally, this sort of content would demarcate an area not of *logical* space, but rather of *practical* space: the space of possible actions, ruling some out while not ruling out others.[19] Just as the practical analogue of a proposition is a plan, the practical analogue

[17] This identification yields as special cases the Lewisian definitions for being (wholly) about, inclusion, having a bearing, and orthogonality. For Lewis, *p* is *wholly about* <*S*> iff the truth of *p* supervenes on <*S*>, i.e. iff *p*'s truth value never shifts within any pair of worlds *w* and *v* such that *w* and *v* agree on all matters pertaining to <*S*> (written $w \sim_S v$). *p* is *partially about* <*S*> iff part of *p* is wholly about <*S*>. *S includes T* if and only if the equivalence relation $\sim_S$ entails the equivalence relation $\sim_T$. *p* has no bearing on <*S*> iff for every *q* wholly about <*S*>, there is a world *w* where *p* and *q* are both true. <*S*> and <*T*> are *orthogonal* if and only if for any two worlds *w* and *v*, there is another world *u* such that *u* is exactly like *w* with respect to <*S*> and exactly like *v* with respect to <*T*>.

[18] In Gibbard (1990) the basic notion of practical or normative content is that of a *norm*, but this distinction will not matter for our purposes. For more on normative content as planning content, see (Gibbard 2012, 18–21, 169–77). This conception has important precursors in the metaethics of R.M. Hare, who held that normative 'ought'-claims are those that entail some imperative (Hare 1952, 163–72).

[19] It might well be possible to reduce the notion of a plan to a set of possible worlds. Plans are most often denoted by infinitives (e.g. *what to do* and *how to live*), but there is no deep syntactic distinction between infinitival and other noun phrases. Indeed, we can even make sense of what it is to plan or decide *that p*, allowing us to plan or decide for others (Ayars 2022). But I am not committed to there being any deep

of a possible world—a complete description of ways things might be—is a *hyperplan*, a plan for what to do in any possible situation.[20]

With our theories of practical content and moral subject matter, we can see that expressivism will be independent of normative ethics in the sense of INDEPENDENCE AS HAVING NO BEARING iff non-cognitivism about moral judgment does not rule out any plan for what to blame for.[21]

---

distinction between planning content and propositional content; in the accounts of plans ruling each other out and of propositional content ruling out planning content that follow, plans will essentially be identified with the set of worlds in which the plan is carried out. (Actions, understood as properties of agents, however, do differ in a principled manner from propositional content. This is a point made often by Mark Schroeder; see his 2011; 2021a; 2021b, chap. 2). I thank Sarah-Jane Leslie for pressing me to clarify this.

[20] (Additionally, just as propositions may be thought of as sets of possible worlds, plans may be thought of as sets of hyperplans.) The most important work that the notion of a hyperplan does for Gibbard is in the formulation of a semantics for moral language. In standard possible world semantics, the meaning of a sentence is modeled by the set of possible worlds it rules out. In Gibbard's semantics, the meaning of a normative sentence is modeled by the ruled-out sets of *plan-laden* or *fact-prac* worlds, ordered pairs which combine a possible world with a hyperplan.

[21] One might worry: does it not stack the deck in favor of the expressivist to claim that the view escapes the perennial objection if expressivism is independent from normative-ethics-*as-conceived-by-the-expressivist*? Although my aim here is to clarify what the expressivist should say about independence rather than simply to defend the view against the charge of revisionism, it is certainly worthwhile to ask whether the expressivist might hope to convince an observer undecided on the truth of expressivism that the objection fails. In this context, it is worth pointing out that Dworkin himself seems to accept a

3.2. Ruling out practical content

How should we understand the notion of ruling out that figures in the definition of having no bearing for the case of practical content? Since the ruling out relation holds between contents, it must be distinguished from any related relation that holds between mental states.[22] It also

---

practical account of normative ethical subject matter. This can be seen from the following passage, where Dworkin distinguishes between ethics and morality:

> An ethical judgment makes a claim about what people should do to live well…. A moral judgment makes a claim about how people must treat other people. Moral and ethical questions are inescapable dimensions of the inescapable question of what to do. (Dworkin 2011, 25)

If ethics and morality are "inescapable dimensions of the inescapable question of *what to do*" (emphasis mine), then these seem to be practical. (One chapter earlier, Dworkin defines ethics as "the study of how to live well" (Dworkin 2011, 13), which is expressly practical. Admittedly, the case for attributing to Dworkin the claim that moral subject matter is practical is not ironclad, inasmuch as it depends on identifying 'how we must treat others' with 'how to treat others,' and at one point Dworkin explicitly *rejects* the Gibbardian identification of 'what I ought to do' with 'what to do' (Dworkin 2011, 434–35 [n22 to Ch. 3]). This rejection, however, is based on a confusion: Gibbard's analysis pertains to the 'ought' of deliberation in its broadest sense, but Dworkin clearly misunderstands Gibbard's dictum as pertaining to the narrowly moral 'ought.' (See Hawthorne 2002 for non-confused skepticism about Gibbard's identification.)

[22] In particular, the relation of ruling out should not be conflated with the relation that holds between two mental states just in case a person cannot rationally possess both mental states (in the same place at the same time). Such a relation is borne, for example, by the mental states of *believing that* p and *believing that* not-p. These states are not rationally co-tenable, but that is not the same thing as those states' ruling one another out.

cannot be conflated with the more specific relation ruling out amounts to in the special case of descriptive content. In that case, where the relevant sort of content is propositional, *p* rules out *q* just in case *p* and *q* cannot be true together, i.e. when their conjunction entails a contradiction. One plan rules out another plan not by entailing a contradiction strictly speaking, but by being *not jointly realizable*, where plans are jointly realizable just in case there is some possible world where they are both carried out. The maximally specific plan to φ and the maximally specific plan not to φ, for example, rule each other out because there is no possible world in which any agent can both φ and not φ.

When does propositional content rule out planning content? To fix ideas, consider combinations of beliefs with planning states that are not rationally co-tenable. It is irrational for me to plan to walk to the twentieth floor of my apartment building if I believe it only to have fourteen stories. Why is this irrational? Plausibly, it is irrational because the aim of a plan is to realize some possible state of the world, and if that state of the world is not possible given my beliefs, then by my own lights I cannot bring it about. So the irrationality of combining these beliefs and plans seems to be derivative, ultimately, on a relation of exclusion that their contents bear to one another, a relation like non-joint-realizability:

> (FACT-PLAN INCOMPATIBILITY) propositional content *p* rules out the plan to φ just in case if *p*, then φ would be unrealizable.

Although not quite identical with the ruling out relation discussed above,[23] I take it that FACT-PLAN INCOMPATIBILITY is suited to play the same role as the genuine ruling out relation in our general definition of having no bearing.

[23] FACT-PLAN INCOMPATIBILITY is plausibly distinct from the genuine ruling out relation for the following reason. Ruling out is a *symmetrical* relation: propositions rule each other out when they cannot

The propositions that seem capable of making a plan unrealizable are those that are incompatible with what I will call its *presuppositions*. The plan to walk to the building's twentieth story, for example presupposes that it has at least twenty stories, and the plan to grab coffee with my friend Josh presupposes that Josh exists. These plans are made unrealizable by any proposition incompatible with the building's having at least twenty stories, and by any proposition entailing that Josh does not exist, respectively.[24] Crucially, this seems to be the *only* way for propositional content to rule out planning content, such that if a plan's presuppositions are all met, then that plan cannot be ruled out by any proposition. If this is right, then expressivism will be independent of normative ethics if no plan about what to blame for presupposes the falsity of expressivism.

---

be jointly true and plans rule each other out when they cannot be jointly realized; but it seems strange to speak of a plan *ruling out* a proposition. I think this is because the notion of propositional content ruling out planning content is just a special case of the notion of an impossible plan, a plan that it is impossible to realize. Plans can be impossible in different ways, corresponding to difference senses of necessity: the plan to go to the store and not go to the store is logically impossible, the plan to travel faster than the speed of light is physically or nomologically impossible, and the plan to walk to the building's twentieth floor is impossible given the way things actually are. The idea that propositional content rules out planning content by constraining the space of possible plans in this way would explain the observation about presuppositionality I discuss in the text presently.

[24] There is also a sense in which my plan to φ presupposes that I am able to φ. I take it that this sense of presupposition is not, however, a relation that holds between contents, but rather a *rational precondition* of intention, one which obtains no matter the value of φ. I thank Alice van't Hoff for raising this concern.

There might well be funny cases of plans that presuppose the falsity of expressivism, such as the plan to blame Jones for accepting Blackburn's false theory of moral judgment. But these funny cases are beside the point, since they presuppose the falsity of expressivism in the same way that any proposition *p* is presupposed by the plan to blame Jones—or, for that matter, to praise or smile at Jones—for accepting the false proposition that *not-p*. However exactly the objector believes expressivism is revisionary, it does not seem that it could be through a relation that any arbitrary proposition might bear to a plan. Such funny cases aside, it is difficult to think of how *any* plan, including a plan for what to blame for, might presuppose the falsity of expressivism in the way that my plan to grab coffee with Josh presupposes his existence. And if there are indeed no such cases, then we may tentatively conclude that expressivism is independent from normative ethics in the sense of INDEPENDENCE AS HAVING NO BEARING.[25]

[25] What if the expressivist's claims about the semantics of moral language are themselves plans of some sort? This would be the case, given the expressivist's analysis of moral belief in terms of plans, if meaning were itself a normative notion. Gibbard himself defends such a view in his *Meaning and Normativity* (2012), on the grounds that claims about what a term or sentence means analytically entail 'ought'-claims. On this view, meaning-claims express plans for belief, plans for which sentences to accept under a range of epistemic circumstances (2012, chap. 8). Since plans for belief seem jointly realizable with a range of plans for what to blame for, the two subject matters appear to be largely orthogonal. So the expressivist's semantic commitments will still be independent from normative ethics even if the former is also understood to be normative and therefore practical. I thank an anonymous referee for *Philosophy and Phenomenological Research* for raising this challenge and suggesting this response.

4. Assessing INDEPENDENCE AS HAVING NO BEARING

To resist this tentative conclusion, the objector might propose the following counterexample to the claim that no plan for what to blame for is ruled out by the truth of expressivism:

(X) Stealing is wrong, but nothing would be wrong if expressivism were true.

Since (X) entails the falsity of expressivism, it seems a good candidate for a genuine, non-funny counterexample to our tentative conclusion. To see whether we have a counterexample, we must identify the plan(s) that (X) expresses. Supposing that moral subject matter consists in plans about what to blame for, (X) expresses the following plan:

(φ_X): to blame for stealing, and not to blame for anything if expressivism is true

Does expressivism rule out (φ_X)? It will be helpful here to distinguish between *pure* and *impure* plans. Our concept of a plan is the concept of some course of action available to an agent *in some situation*: planning to φ is really elliptical for planning to φ in some set of circumstances *C*. *Pure* plans, then, are always conditional, they are plans to φ in *C*. *Impure* plans are plans that are *actionable*, where an actionable plan is a plan for a set of circumstances that include the agent's actual circumstances. To accept some impure plan φ, then, is in part to have—or more precisely, to be rationally committed to having—a self-locating belief that the circumstances that φ are for include one's own.

We can see that (φ_X) is a conjunction of two sub-plans. The second sub-plan (not to blame for anything if expressivism is true) is clearly *pure*: it is for the circumstance in which expressivism is true. But the first sub-plan (to blame for stealing) is *impure*. (This much is conveyed by the 'but' conjoining the two clauses in (X), which indicates that the second sub-plan is for a counterfactual set of circumstances.) The first sub-plan, it seems clear, is a plan to blame

for stealing in worlds where expressivism is false. So an agent who accepts at least the first sub-plan of (φ_X) is rationally committed to believing that expressivism is false in her actual circumstances.

Here, then, is what we should say about the relation between expressivism and (φ_X). The pure planning content expressed by (X), the pure plan from which the impure plan (φ_X) is derived, is:

(φ_X)* to blame for stealing if expressivism is false, and not to blame for anything if expressivism is true

The pure plan (φ_X)* is clearly not ruled out by the truth of expressivism or its falsity. In fact, neither is the impure plan (φ_X), at least not in the sense of FACT-PLAN INCOMPATIBILITY: although *accepting* an impure plan rationally requires the planner to *believe* that the relevant circumstances obtain, that does not make it *impossible* for an impure plan to be carried out in worlds where those circumstances fail to obtain. If I accept (φ_X), for example, it is irrational, but not metaphysically impossible, for me to blame for stealing and for expressivism to be true; by contrast, it is metaphysically impossible to walk to the twentieth floor of a building that has only five stories. So (φ_X) fails to provide a counterexample to the claim of independence construed as INDEPENDENCE AS HAVING NO BEARING. (Should the objector argue that expressivism does rule out the impure plan (φ_X) in some relevant extended sense? I discuss this gambit in **§5**.)

Having shown the independence of expressivism in the sense of INDEPENDENCE AS HAVING NO BEARING from normative ethics, we can now make the following observation: the considerations we have adduced in favor of this conclusion seem to support the standard response's stronger claim that <what is moral judgment> is *orthogonal* to <what to blame for>. (Or at least, it seems to support the stronger conclusion that no plausible answer to <what is

moral judgment> rules out any non-funny answer to <what to blame for>.) Upon reflection, since this sort of independence seems to hold for a wide range of descriptive and practical subject matters, it seems that the independence of expressivism or <what is moral judgment> from <what to blame for> is merely an instance of a much more general phenomenon, which I call *Humean Independence*:

> (HUMEAN INDEPENDENCE) A practical subject matter <Φ> and a descriptive subject matter <S_p> are Humean independent iff no plan wholly about <Φ> presupposes any proposition wholly about <S_p>.

The principle is so named because it seems closely related to two Humean doctrines of interest to metaethicists. The first doctrine is the view that there is a principled distinction between beliefs—mental states with propositional content—and desires—mental states with practical content: beliefs have a *mind-to-world* direction of fit, whereas desires have a *world-to-mind* direction of fit.[26] The difference in these two sorts of psychological states tracks the difference between the kind of contents they have: desires have practical contents, and (ordinary) beliefs have descriptive contents. The second doctrine is Hume's Law, roughly, that no sentence about what ought to be the case or what someone ought to do can be inferred from any set of purely descriptive sentences. It is a matter of controversy whether there exists a true and informative version of this principle, suitable to serve as a kind of constraint on metaethical theorizing. Whether or not this is so, the *expressivist* is well positioned to accept some version of this

[26] For classic discussions see (Anscombe 2000, sec. 32; M. Smith 1987).

principle and to provide a principled basis for its truth by appealing to the distinction between descriptive and practical content.[27]

5. INDEPENDENCE AS NON-CONTINGENCY

I have argued that non-cognitivism about moral judgment is independent of normative ethics in the sense of INDEPENDENCE AS HAVING NO BEARING, vindicating the standard response's claim that expressivism is not itself a moral thesis. If this result is just an instance of

[27] It is worth saying a bit more here about the project of formulating an adequate version of Hume's Law—most importantly, one immune from the classic counterexamples due to (Prior 1960)—and how it relates to my taxonomy. Some theorists have sought to vindicate the 'autonomy of ethics' by modeling semantic content with something like Gibbard's factual-normative/fact-prac worlds (see Humberstone 1996, 153–54; 2019, 1400). One recent attempt to do so is due to Daniel J. Singer (2015), who argues that if non-ethical content is understood as norm-invariant content, then the correctly formulated version of Hume's Law will be a theorem of any normative semantics relevantly similar to Gibbard's. In response, Jack Woods and Barry Maguire (2017) contend that to know whether norm-invariance correctly characterizes the descriptive, we need a theory of norms; and that any such theory will make metaethical assumptions likely to be rejected by philosophers unsympathetic to Hume's Law. Unlike Singer's defense of Hume's Law, I seek only to understand the relation between descriptive content and practical content—that is, normative content *as construed by the expressivist* and those who share the relevant expressivistic assumptions. It would be unwarranted to assume the practical conception of moral content in arguing for Hume's Law as a logical or semantic thesis, but this is not my aim. Skepticism has been voiced about whether the expressivist should be concerned about the autonomy of ethics (Behrens 2021, 13182); my discussion shows that, *au contraire*, the theoretical considerations motivating expressivism strongly suggest a principle very much like Hume's Law.

the phenomenon of HUMEAN INDEPENDENCE, then the independence of expressivism from normative ethics turns out to be almost trivial. That puts pressure on the objector to find some other sense of independence on which normative ethics turns out not to be independent of expressivism.

The discussion of impure plans might suggest to the objector the following alternative conception of independence:

(INDEPENDENCE AS NON-CONTINGENCY) A practical subject matter <*Φ*> is independent of *p* if no plan wholly about <*Φ*> is for the circumstance that *p* (*not-p*).

On this principle, the plan to blame only if expressivism is false renders <what to blame for> dependent on expressivism. Although it would be natural for the objector to retreat to INDEPENDENCE AS NON-CONTINGENCY at this point in the dialectic, it is unclear what significance this sort of independence has for the objection that expressivism has revisionary first-order consequences. For one, since the principle as formulated does not discriminate between plans that we accept and plans we do not accept, bizarre plans contingent on the truth or falsity of expressivism—e.g. the plan to blame for breathing only if expressivism is false—would suffice to spoil the independence of <what to blame for> from expressivism. This trivial sort of dependence seems irrelevant to the objection. And the principle remains too strong even when restricted to plans we *do* accept, since it would render normative ethics dependent on a wide range of propositions about subject matters which intuitively *do* seem independent of normative ethics. In particular, if the objector wanted a sense of dependence that might underlie the claim that non-cognitivism about moral judgment (and therefore expressivism) is itself a moral thesis, INDEPENDENCE AS NON-CONTINGENCY could not provide that sense.

Consider for example the plan to blame for investing in the fossil fuel industry. Agents who accept this plan accept it for worlds like the actual world, where the fossil fuel industry contributes to climate change and all its sequelae. They do not plan to blame for investing in the fossil fuel industry in counterfactual circumstances where, e.g., burning fossil fuels does not harm the environment. If so, then <what to blame for> will be *dependent* in the sense of INDEPENDENCE AS NON-CONTINGENCY on the climatological facts in the same way as it is dependent on expressivism, given the plan to blame for stealing only if expressivism is false. But it is not plausible that what makes expressivism revisionary of normative ethics is that normative ethics depends on the falsity of expressivism in *this* way, the way that normative ethics depends on the climatological facts; certainly, this would not make the climatological facts into normative ethical facts themselves.[28] So the objection to expressivism should not be formulated as the claim that expressivism is not independent of normative ethics in the sense of INDEPENDENCE AS NON-CONTINGENCY: even if the plans we accept do make blaming for stealing contingent on the truth

[28] Indeed, at least one purveyor of the objection, namely Dworkin, positively believes that expressivism is independent of domains like the domain of climatological facts. Dworkin professes a "radical" belief in "the metaphysical independence of value" (Dworkin 2011, 9) and avers that "morality is a distinct, independent dimension of our experience" (Dworkin 1996, 128). What's more, Dworkin seems to hold these views because of Hume's Law; in his words, "Hume's principle, properly understood, supports not skepticism about moral truth but rather the independence of morality as a separate department of knowledge with its own standards of inquiry and justification" (Dworkin 2011, 17, and see also pp.44–46). So it seems plausible to conclude that even for Dworkin, there is a good sense of moral independence guaranteed by INDEPENDENCE AS HAVING NO BEARING.

or falsity of expressivism, this contingency does not convert metaethical content into normative ethical content.[29]

6. Moral dependence and objectivity

To review: neither INDEPENDENCE AS HAVING NO BEARING nor INDEPENDENCE AS NON-CONTINGENCY captured a sense in which normative ethics might be *logically* or *formally* dependent on expressivism in a way that might render expressivism revisionary of normative ethics. INDEPENDENCE AS HAVING NO BEARING could not show any dependence of normative ethics on expressivism, and the sort of dependence captured by INDEPENDENCE AS NON-CONTINGENCY did not seem of the right sort to inform the perennial objection. But although the standard response to that objection maintained that expressivism could not have revisionary consequences because it is independent of normative ethics, it was never obvious that the charge of revisionism should be understood so narrowly. Perhaps there is a way to understand how the truth of expressivism might be revisionary of ordinary moral thought *without* needing to furnish a problematic relation of dependence that normative ethics might bear to expressivism. The suggestion I now explore on behalf of the objector is this: expressivism is revisionary because many of our normative ethical views are committed to the falsity of non-cognitivism about moral judgment and therefore to the falsity of expressivism.

[29] Notice that the expressivist's analysis of normative content as planning content played no role in this discussion. Even for a metaethicist who rejects this analysis, it is not plausible that metaethical content becomes moral just because there are possible, or even true, claims of the form 'if *M*, then *N*,' where *M* is metaethical and *N* is moral. See **n34** for more discussion.

This version of the objection is distinct from the version that invoked INDEPENDENCE AS NON-CONTINGENCY. That version of the objection held that contingency on the truth or falsity of expressivism itself constituted the way in which normative ethics might be dependent on expressivism, and it was found inadequate because that sort of dependence seemed innocuous. The current proposal, by contrast, *embraces* the contingency of the normative ethical views we accept on the falsity of expressivism, and concludes on that basis that expressivism must be false. We can say that expressivism would be revisionary, on this view, because some normative ethical claims we accept are *morally dependent* on the falsity of expressivism, allowing us to reason from our acceptance of those moral claims to the falsity of expressivism itself.[30]

Construing the complaint in terms of moral dependence affords a dialectical benefit to the objector, since nothing in the expressivist's argument for the formal or *taxonomic* claim that non-cognitivism about moral judgment is a non-moral thesis addresses this *first-order moral* concern. Yet I take it that this version of the objection, even if ultimately successful, marks a dialectical retreat for the opponent of expressivism. This is because the original, taxonomic version of the

[30] It should be noted that the objection does not beg any questions against the expressivist: the objection could be formulated in terms friendly to the expressivist, even as it seeks to show that the expressivist's conception of normative content as planning content is untenable given our first-order moral views. Put in such terms, the objection would be that since some of the plans we accept as actionable are for the circumstance that expressivism is false, we can reason directly from the actionability of our plans contingent on expressivism's falsity to a conclusion about the world we are in, namely the falsity of expressivism. Expressivism would be revisionary, on this view, because it would force us to abandon at least some of the impure plans that we accept as actionable.

objection was not beholden to any particular first-order moral views. But now, since the objection hinges on claims about the content of the *true* moral views, it is a partly moral one.

I now turn to a brief discussion of this version of the perennial objection. My aim here is not so much to argue against the moral version of this objection as it is to distinguish it from the taxonomic versions discussed earlier and to highlight its connection to the vexed question of *moral objectivity*.

A clear statement of this sort of first-order moral worry may be found in an early passage from *Justice for Hedgehogs*:

> Politics is coercive: we cannot stand up to our responsibility as governors or citizens unless we suppose that the moral and other principles on which we act or vote are objectively true…. So the old philosopher's question—Can moral judgments really be true?—is a foundational, inescapable, question in political morality. We cannot defend a theory of justice without also defending, as part of the same enterprise, a theory of moral objectivity. (Dworkin 2011, 8)

Dworkin here evinces a *moral* uneasiness with non-objectivist views, on the grounds that our political arrangements, which involve the marshalling of state power to enforce its laws, would be *morally wrong* if morality were not objective.[31]

[31] Dworkin's overarching concern with objectivity is clearer in (Dworkin 1996), though the notion there is undertheorized: Dworkin seems to take 'objective(ly)' as synonymous with 'real(ly),' and to the extent that we get a theory of objectivity, it is Blackburn's theory: "The claim that abortion is objectively wrong seems equivalent, that is, in ordinary discourse, to [the claim] that abortion would still be wrong even if no one thought it was" (Dworkin 1996, 98). See Blackburn's (1996) response to Dworkin for further discussion.

Dworkin does not himself parlay these thoughts into an objection directly; in context, the relation of morality to politics is portrayed more as *motivating* the relevance of the arguments to come against anti-realism than as *constituting* that argument itself. But others have pressed this sort of objection explicitly, by identifying specific actions we take to be morally permissible that, it is claimed, would be morally impermissible if morality were not objective. Jeremy Fantl (2006), for example, argues that violent interference could only be justified in order to prevent harm that is objectively wrong, and so we have reason to reject subjectivism if we think that violent interference in order to prevent harm is sometimes morally justified. In a similar vein, David Enoch (2011, chap. 2) argues that when two parties disagree about what to do concerning some matter of taste, it is morally impermissible to "stand [one's] ground" (2011, 17), i.e. to fail to come to a solution in which all parties' interests are respected; but since in the case of moral disagreement standing one's ground is permissible, we may conclude that moral judgments are not mere preferences or desires.[32] In the first instance this argument is meant to refute a view Enoch calls Caricaturized Subjectivism, and Enoch argues that it can be extended to refute expressivism, as well.

[32] Mark Kalderon (2005, chap. 1) has made a structurally similar argument for the opposite conclusion, arguing that unlike in the case of factual discourse, moral conversation exhibits what he calls *intransigence*: though these conversations concern the interests of others, participants do not seek more information in the face of disagreement. Kalderon maintains that the best explanation of this norm governing moral conversation is that moral disagreement is non-cognitive.

I take no stand here on whether these particular arguments are successful, either against expressivism or against the simple forms of subjectivism they most directly threaten.[33] What I do claim is that since these arguments proceed from the normative ethical views we (allegedly) accept, they are *moral* arguments, not taxonomic ones. This point has, I think, not been sufficiently appreciated, though it is crucial for understanding what these arguments come to and how they might be addressed.[34]

---

[33] See (Sinclair 2014) for an argument that Enoch's argument cannot be extended to refute expressivism, and see (Atiq 2016) for an argument that Enoch's argument does not even refute subjectivism.

[34] Fantl and Enoch each discuss the bearing their arguments have for the question of metaethics' normative neutrality, but this latter question seems importantly distinct from their main arguments against subjectivism. On Fantl's view, a response to (Dreier 2002), no metaethical position is normatively neutral because any metaethical view *M* rules out a combination of a moral view *not-N* together with a conditional of the form 'if *M*, then *N*.' Fantl (pp. 38–39) notices that his view entails that no proposition is morally neutral, and so fails to identify a sense of moral commitment lacked by, e.g., Newtonian mechanics, which Dreier (2002, 247) explicitly takes to be a desideratum of any such theory. But Fantl writes that since there are not many *true* conditionals of this form whose antecedents are (e.g.) physical rather than metaethical, the lack of neutrality of other domains with respect to morality is not of concern. This last claim seems false. There are many such true conditionals; recall the example of the wrongness of investing in fossil fuels (planning to blame for investing in fossil fuels) from **§5**. For another example, consider the triad: 'arsenic is poisonous,' 'if arsenic is poisonous, then it is wrong to put arsenic in someone's food,' and 'it is not wrong to put arsenic in someone's food,' where the conditional rendering the arsenic-facts non-morally neutral is also true. So Fantl's best argument against subjectivism is simply that there is a true moral claim conditional on the falsity of subjectivism; the question of the normative neutrality of metaethics is beside the point. Enoch's argument against Characterized Subjectivism is

Dworkin, Fantl, and Enoch all maintain that the moral views we accept are sensitive to the question of *moral objectivity*; it is the question of whether morality is objective that might have normative ethical implications. The notion of objectivity is something of a placeholder in these arguments, standing in for the underlying feature(s) of moral thought and language (or of the moral facts) to which our normative ethical views are ultimately sensitive, the underlying feature(s) that make(s) some metaethical view count as objectivist by respecting. Rather than

---

followed by a discussion (pp. 41–49) of the normative neutrality of metaethics. Enoch's preferred characterization of this neutrality is in terms of a notion like *conservative extension*; officially, the view is that metaethics is normatively neutral if and only if "there are good, discriminating arguments with a moral conclusion, and at least one indispensable metaethical premise" (p. 46), and Enoch notes that if his argument against Characterized Subjectivism is successful, then metaethics does not conservatively extend normative ethics in his sense. I take no issue with these claims, but I do disagree with Enoch's claim that being morally non-neutral is distinct from being (merely) morally relevant, on the grounds that what is "in dispute between those taking metaethics to be morally neutral and those denying this neutrality is about some internal connections, as it were, between metaethics and morality, or some non-contingent, at least morally necessary connections of this sort" (p. 48). But there is a morally necessary connection between, e.g., the dangerousness of arsenic and the wrongness of lacing someone's breakfast with arsenic, which seems to render the arsenic-facts a non-conservative extension of morality, as well. (If arsenic had the chemical properties of sugar it would not be wrong to "lace" someone's breakfast with it.) So I fail to see the distinction that Enoch wishes to draw. In any event, all Enoch's argument against subjectivism requires is the claim that the metaethical question of whether moral judgment (or, more precisely, moral disagreement) is cognitive or non-cognitive is morally relevant to the norms of joint deliberation and action.

adjudicating the debate over which senses of objectivity are available to the expressivist,[35] my aim here is to highlight that the debate over objectivity is essentially an instance of first-order moral inquiry: specifically, of identifying the moral views we do or should accept surrounding, *inter alia*, coercion, intervention, and joint action. If these first-order moral questions have been somewhat obscured by taxonomical debates, I hope they can now be seen more clearly.[36]

[35] See the citations in **n3**, as well as (Blackburn 1993a; 1998, chap. 9; 1999; Gibbard 1990, pt. III) for important classic discussions. For a recent discussion of Gibbard's views about objectivity and their background in the expressivist tradition, see (Railton 2022). One prominent candidate for the relevant kind of objectivity unavailable to the expressivist is metaphysical objectivity; this proposal, however, is not available to Dworkin, who mocks the view that moral facts are made true by moral entities ("morons") as a philosopher's fantasy (Dworkin 2011, 9).

[36] **Acknowledgement:** thanks to Gabriel Broughton and Aarthy Vaidyanathan for written comments on earlier versions of this chapter.

Chapter 2: Moral objectivity as moral aboutness

1. Introduction

According to many versions of metaethical realism, morality is clearly objective. According to such views, just as there is a way the world is in its physical respects, there is a way the world is in its moral respects; and how we think and speak about the world in its moral respects is not fundamentally different from how we think and speak about the world in its physical respects. Although metaethical anti-realists are committed to rejecting some part of this story, many have claimed that their views can also respect the objectivity of morals—most famously by providing anti-realist-friendly interpretations of counterfactuals such as 'stealing would be wrong even if I did not disapprove of it.' But even if the anti-realist can provide such interpretations, the original question of objectivity seems unresolved. This essay marshals a novel constitutive account of objectivity—*objectivity as object-aboutness*—to argue that a striking range of anti-realist metaethical theories count as objective in a sense even moral realists should care about. On the account I develop, whether some domain *D* is objective depends on its *subject matter*, what it is that *D* is *about*. In brief, the view is this: a metaethical view counts as objective if, on its analysis, moral thought is not about the thinker, moral language is not about the speaker, and the moral facts are not about the relevant judge(s).

I first provide aboutness-theoretic definitions of subjectivity and objectivity and distinguish between *restricted* and *unrestricted* senses of aboutness, which underlie restricted and unrestricted conceptions of objectivity. I then argue that metaethical anti-realists have hoped to secure the restricted sort of objectivity—objectivity given the restricted sense of aboutness—according to which a domain is objective so long as it does not *foreground* the subject (**§2**). I

then argue that our concern for moral objectivity in this sense has a basis in concerns familiar from moral theory and ultimately from everyday moral life, what I call our concern for *moral aboutness* (**§3**). I then provide substantive accounts of foregrounding the subject for the cases of thought, language, and metaphysics (**§4**). These substantive accounts in hand, I argue that a striking range of metaethical views count as forms of objectivism, including not only the expressivism of Blackburn and Gibbard, but also anthropocentric forms of cognitivism such as constructivism, and perhaps even error-theoretic cognitivist views (**§5**). I conclude by addressing the objection that my account fails to distinguish moral judgments from judgments of taste (**§6**).

2. Subjectivity, Objectivity, Subject Matter

The question of objectivity and subjectivity is a question about the status of some *domain D*: the question is whether *D* is objective or subjective. The notion of a domain is generic: a *mental* domain is a set of thoughts, a *linguistic* domain is a set of sentences, and a *metaphysical* domain is a set of facts. (When we ask whether some philosophical theory *T* of a domain *D* is a form of objectivism or subjectivism, we are asking whether *D* is objective or subjective given *T*'s interpretation of *D*.)

Let us begin with a definition of subjectivity, perhaps the more tractable notion of the two. The analysis I propose is this:

> *Subjectivity as subject-aboutness*. *D* is subjective when it is about, or its subject matter involves, the relevant subject.

Who the relevant subject is depends on the nature of the domain in question: in the case of thought the relevant subject is the *thinker*, in the case of language it is the *speaker*, and in the case of metaphysics it is a *judge* or *judges*—that is, the person(s) whose judgments play some

role in constituting the facts of the domain. (For example, if the color facts are partly determined by the perceptual seemings of standard perceivers, then the standard perceivers are the relevant judges in the case of the color facts.)

We may now leverage our definition of subjectivity into a corresponding definition of objectivity thanks to the fact that subjectivity and objectivity in their relevant senses are *privative opposites*: if a thing is apt to be objective but is not, then it is subjective, and *vice versa*.[1] So, if a domain is subjective when it is about the subject, it will count as objective just in case it is *not* about the subject: objectivity on this view just is *non-subject-aboutness*. Using the term 'object' to denote anything other than the subject, we can define the relevant conception of objectivity as follows:

> *Objectivity as object-aboutness*. *D* is objective when it is not about, and its subject matter does not involve, the relevant subject.

Metaethicists aim to provide theories of moral thought, moral language, and moral metaphysics; and we can ask the question of objectivity about each sort of domain. So, to discern whether some metaethical view counts as objective, we require substantive accounts of what aboutness in thought, in language, and in metaphysics amount to. But before we address substantive accounts of aboutness in these domains, we must distinguish between two senses of aboutness or subject matter: *restricted* and *unrestricted*. The distinction is best introduced by illustration. Consider Bruegel's *The Harvesters*, which depicts a harvest scene set against lakes and rolling hills. We can ask: what is *The Harvesters* about? In the unrestricted sense, the

[1] A potato is not apt to be objective or subjective in any sense, so it can be not objective without thereby counting as subjective, and *vice versa*. A jury is apt to be objective, but in a slightly different sense, whose opposite is not subjective but *biased*.

painting is about everything it depicts—the harvest scene as well as the lakes and hills. In the restricted sense, however, the painting is only about the harvest scene in its *foreground*, not the lakes and hills in its *background*. In general, the restricted conception of aboutness is *discriminating* relative to the unrestricted sense: *c* is about *S* in the unrestricted sense so long as *S* is somehow "on the scene" of *c*, but *c* is about *S* in the restricted sense only if *S* is in *c*'s foreground.

Each sense of aboutness may be understood as informing some conception of objectivity: objectivity given the unrestricted sense of aboutness and objectivity given the restricted sense of aboutness. Which is relevant to the question of objectivity?

Let us first consider the traditional metaphysical debate about objectivity. Here we want to know in general which domains of fact are about the world as it is in itself, independent of us; and which domains of fact are, perhaps very subtly, not so. The contrast here is between, for example, facts about mass and electrical charge (as we ordinarily take them to be) on the one hand, and facts about social constructions such as the law or monetary value on the other. For this question, the relevant conception of objectivity seems to be the unrestricted one. It seems that if for instance the color facts or facts about space and time implicated or involved a subject—in these cases, a perceiver—*in any way*, they would be unlike the facts about mass and electrical charge and so not objective in the relevant sense.[2]

[2] The relation between metaphysical objectivity and realism is disputed. For some writers, metaphysical objectivity is necessary for realism, and the question of objectivity is precisely what is at stake in realism debates across philosophy. Metaethicists have discussed whether the moral facts are objective in this sense, usually glossed as *stance-independence*, which Ronald Milo defined as "independen[ce] of how we are disposed to respond to the world in terms of our affective or volitional responses" (1995, 182).

But this cannot be the conception of objectivity that the metaethical anti-realist hopes to secure. This would be the view that the facts in question are just like the facts about mass and electrical charge, and this view is incompatible with anti-realism if anything is.[3] So, since anti-realists must concede that moral thought, language, and metaphysics are not unrestrictedly objective, the strongest claim to objectivity they can venture involves the restricted sense of aboutness: that moral thought, moral language, and moral metaphysics do not foreground the thinker, speaker, or judge.

It would then seem that given the strongest sort of objectivity the anti-realist can hope to secure, our concern that morality be objective is sourced in a concern to avoid foregrounding the subject. In the following section, I argue that the concern to avoid foregrounding the subject—in plain language, to avoid *making things about oneself*—is familiar from everyday moral life and

---

Authors who essentially accept this definition include Russ Shafer-Landau (2005, 15–16) and David Enoch (2011, 3)—who affirm morality's stance-independence—and Sharon Street (2010, 370–71)—who denies it. Others, such as Geoffrey Sayre-McCord (1988) and Gideon Rosen (1994), deny that objectivity is necessary for realism. (Rosen calls this conception *minimal realism* and questions whether the further notion of objectivity can be clearly characterized.) Although I will argue that moral anti-realism is compatible with restricted objectivity, I take no stand on whether unrestricted objectivity is necessary for realism in general, or for that matter whether the notion of unrestricted objectivity in metaphysics is ultimately a viable one. See also the discussion of "non-objectivism" below (**§5.4**).

[3] Gibbard distinguishes objectivity as "Platonism" from the senses of objectivity he seeks to secure for normative thought and language. On Platonism about normative judgments and the intuitionism it requires Gibbard writes, "If this is what anyone seriously believes, then I simply want to debunk it" (1990, 154).

moral theory, and so should be shared by metaethical realists and anti-realists alike. If successful, then, this argument will show that the sort of objectivity that metaethical anti-realists hope to secure speaks to concerns we all have, irrespective of our metaethical commitments.

3. Moral aboutness

The concern to avoid foregrounding the subject forms part of a larger concern for what I call *moral aboutness*, a concern for which aboutness-properties are instantiated in situations of moral significance. We think it matters morally what—and especially *who*—is and is not *centered* in such situations: who, for example, is the object of our attention or the focus of conversation. When it matters morally who some situation is supposed to be about, it is inappropriate to center any other object, and it is especially inappropriate to center *the subject*—that is, oneself. The claim then is that our theoretical concern to avoid foregrounding the subject is grounded ultimately in familiar first-order concerns having to do with the virtue of attending to the morally significant interests of others and the vice of self-centeredness. I now describe a few examples that illustrate our concern for moral aboutness.

For an example from everyday moral life, imagine a conversation between two friends, Annie and Bertie. Annie is distraught, having just been rejected from another graduate program and uncertain whether she will be accepted anywhere she'd like to go. Looking to vent, Annie calls Bertie, hoping that he will lend a sympathetic ear. But when Annie starts speaking about her disappointment and fear of not being accepted to graduate school, Bertie interjects by complaining about how frustrated *he* is by some relatively trivial matter ("I hear what you're saying about disappointment: my Amazon package was delayed today!"). Annie might well become irritated at Bertie's having shifted the topic of conversation, expressing her anger by

saying, "This conversation is not about you!" And we would hope that Bertie would respond contritely by saying something like, "You're right. I'm sorry for making this about myself."

In the imagined scenario, Annie wants her concerns to be centered in conversation, to be the object of a certain kind of attention and standing in the social space occupied by her and her interlocutor: in short, for the conversation to be about her. And, given the facts of the case, the desire to be centered assumes moral significance: Annie phoned Bertie hoping that a good venting session would help alleviate her psychic anguish, a hope Bertie extinguished by shifting the topic. When it is a matter of moral significance who some conversation is about, it is a violation of the moral norms governing conversation to shift the topic away from that person, and it is a particularly severe violation of these norms to shift the topic towards *oneself*.[4]

Now consider a slightly more theoretical example, having to do with *moral worth*. It is widely agreed that to receive moral credit for doing the right thing, one must do the right thing for the right reason. Performing the right action for some ulterior motive or otherwise "accidentally" does not confer moral worth; one must rather act out of an intrinsic desire to do what is right. Moral philosophers have disagreed about what kind of desire this is precisely,[5] but

[4] For a classic defense of the view that we can in general be held responsible for what we attend to and find salient, see (A. M. Smith 2005); see (Watzl 2022; Archer 2022) for recent discussion of the moral norms governing attention and the varieties of salience for which we might be held responsible.

[5] On one view, moral concern consists in non-instrumental concern about what is right *per se*; on another, moral concern consists in non-instrumental concern about the right-making features, those naturalistic features of the world that make an action right. Prominent defenders of the former view include (Herman 1981; Sliwa 2016; Johnson King 2020a; 2020b), and prominent defenders of the latter view include (Arpaly 2003; Markovits 2010; Arpaly and Schroeder 2013). Following (M. Smith 1994, 75), these views

everyone should agree that the following sort of motivation to do what is right fails to confer moral worth: one that is derivative of a concern for the *agent's own* righteousness and virtue. This sort of concern seems, to borrow a phrase of Mark Schroeder's in a related context, "objectionably self-regarding."[6] The virtuous person is concerned not about herself and her own righteousness or virtue, but rather about the morally significant features of the situation that call upon us to respond. For the self-righteous person, moral concern is concern about himself.[7]

Finally, consider *agent-centered restrictions*. Given a ranking of states of affairs from morally best to morally worst, maximizing consequentialism holds that is obligatory to bring about (some) best outcome and impermissible to bring about any other. Non-consequentialist theories may deny either of these consequentialist commitments. If there are *agent-centered prerogatives*, then it is sometimes *permissible* for an agent to fail to bring about the morally best

---

are often called the *de dicto* and *de re* views of moral worth respectively, owing to how each reads the phrase 'desire to do what is right.' For other ways of construing the relevant desire, see (Howard 2021; Johnson King 2022).

[6] In his defense of a version of the Humean theory of reasons, Schroeder (2007, chap. 2) argues that his view is not committed to the claim he calls *No Background Conditions*, that an agent's desiring to φ must always be part of an agent's reason to φ. Schroeder accuses this view of having an "objectionably self-regarding" conception of practical reasoning.

[7] Proponents of the *de re* view have sometimes accused the *de dicto* view of endorsing a sort of *fetishism*, alleging that it is not only not morally worthy but positively morally vicious to be motivated by concern for rightness itself. (See M. Smith 1994, 75 for a classic statement of this thought and Johnson King 2022, 264–65 for citations of others.) Perhaps those who have found this objection compelling believe that concern to do what is right *de dicto* just is concern to do what *I* ought to do, and that this concern inappropriately involves the subject.

outcome; and if there are *agent-centered restrictions*, then it is sometimes *impermissible* to bring about the best outcome. If there is an agent-centered restriction against killing, for example, then not only would it be impermissible to kill to prevent some other kind of bad, it would even be impermissible to kill in order to prevent two other killings. Although this verdict seems intuitively acceptable, it can seem puzzling upon reflection: if the wrongness of killing is explained by its disvalue, then why is it impermissible to prevent two killings even by committing one? Robert Nozick called such restrictions "irrational" (1974, 30), and Samuel Scheffler writes that they are "surround[ed]" by a "persistent air of paradox" (1994, 112).

After considering a litany of defenses of agent-centered restrictions and finding them all wanting, Scheffler ultimately proposes a debunking explanation of their intuitive appeal, having to do with the benefits the relatively well-off have by living in a society in which agent-centered restrictions are respected (Ibid, 113). But another debunking explanation might be this: the intuitions supporting agent-centered restrictions ultimately stem from a concern that one's own agency not be corrupted by performing some generally impermissible act, by acting with "dirty hands" in Michael Walzer's (1973) phrase. I take it that this is a debunking explanation of the intuition in favor of agent-centered restrictions, since it is simply inappropriate to value one's own moral hygiene as lexically prior to whatever other value one's action might bring about or disvalue it might prevent.[8] Even if it can be shown that agent-centered restrictions are not

[8] In his response to Phillipa Foot's (1983; 1985) defense of agent-centered restrictions, Scheffler notes that virtue-theoretical views might have more resources than deontological views to identify a desirable aim secured by observing agent-centered restrictions, specifically by appealing to the agent's own virtue. To this Scheffler responds, "this, it seems to me, rather glaringly fails to capture our actual sense of what is ordinarily wrong with these things" (146).

paradoxical and that accepting a normative ethical theory that includes them does not violate a norm of rationality, it must further be shown that the proposed justification for accepting such a theory does not rely on a morally objectionable conception of what is valuable.[9]

4. Aboutness in metaethics

Having seen the basis of our theoretical concern for not foregrounding the subject, we may now turn to substantive accounts of the foregrounding/backgrounding distinction invoked by the restricted sense of aboutness. By telling us when, in their relevant senses, moral thought is about the thinker, moral language is about the speaker, and moral metaphysics is about the relevant judge(s), these accounts provide substantive criteria for the restricted objectivity of moral thought, moral language, and moral metaphysics.

---

[9] According to *consequentializers*, any deontological theory may be translated into a consequentialist one (Dreier 1993; 2011; Louise 2004; Portmore 2007). To represent agent-centered restrictions, these theories rank options not in terms of *agent-neutral* value, but in terms of *agent-relative* value: for each agent, the option of killing one to prevent two killings ranks below the option of refraining from killing even when this fails to prevent two killings. We might object that this sort of ranking is, to borrow again Schroeder's phrase, objectionably self-regarding, since it seems to imply that what I morally ought to do is to maximize a sort of value that attaches specially to my own agency. Michael Smith claims that the consequentializer is not in fact committed to this problematic view. Smith suggests that instead of claiming that "there is value in the outcome of an agent's not violating anyone's rights himself," the consequentializer should rather claim that "the agent's not violating anyone's rights [is] a *condition* of the value of (some or all of) the outcomes of the actions he performs" (M. Smith 2009, 271 n1, italics his). Given the account I defend of restricted aboutness in metaphysics (**§4.3**), such a view would *not* be about the subject and so would evade the objection.

4.1. Aboutness in thought

D-thought foregrounds the subject if the *aim of* D-*inquiry* and D-*deliberation* involves the thinker.[10] If the aim of moral inquiry and deliberation is to answer the question, *what do I disapprove of* (under that description), then moral thought foregrounds the thinker. But if the aim of moral inquiry and deliberation is to discern the morally relevant features of a situation and decide what to do or what to disapprove of on this basis, then moral thought is objective.[11] And it remains objective even if the thinker's self-directed attitudes are present somewhere in moral inquiry and deliberation—as, for many anti-realist views, they must be—so long as they are *not-at-issue* rather than *at-issue* relative to the aim of moral inquiry and deliberation.[12] (But don't the

[10] Thanks to Derek Baker for suggesting this way of formulating the distinction between the foreground and background of thought.

[11] Of course, since the morally relevant features of a situation might include facts about the thinker, good moral deliberation might include thoughts about the thinker, such as *that I promised to help* or *that it is only in my power to help*. But there is no 'irreducibly' first-personal character to these thoughts, since the underlying facts can equally well be described third-personally.

[12] Compare Philip Pettit and Michael Smith's discussion of the relation between reasons and desire (Pettit and Smith 1990). According to the *strict background view of desire* they defend, although an agent always possesses some relevant desire when she is motivated to φ, the agent need not believe that her justifying reason for φ-ing is that it would satisfy one of her desires. In their terminology, although desire is always in the background of the agent's decision to φ, it need not be in the foreground of that decision. An important implication of this view is that "phenomenology offers no argument for cognitivism" about moral judgment (578–9). That is, the fact that the phenomenology of moral deliberation does not include reference to the agent's desires does not itself show that moral judgment does not consist in having some

questions of *what to do* and *what to disapprove of*, which are part of the aim of moral inquiry and deliberation, foreground the subject inasmuch as they are practical? I discuss this in **§5.3**.)

4.2. Aboutness in language

D-language foregrounds the subject if it answers by default questions involving the speaker. (For example, the sentence 'the sweater is blue' answers by default the questions *what color is the sweater* and *which things are blue*, but not the question *how many sweaters are there*) This criterion parallels that of *D*-thought: each defines foregrounding in terms of what is *at-issue* relative to some question towards which the domain is oriented.[13] If moral language answers by default questions such as *what does the speaker approve of*, then it foregrounds the subject. But if moral language, say the sentence '*x* is good (right),' answers by default questions

---

desire. Putting together this view with the definition of restricted objectivity yields the result that moral thought can be objective even if non-cognitivism about moral judgment is true. (See **n23** below.)

[13] The relevant sense of the at-issue/not-at-issue distinction for linguistic content follows what (Koev 2018) calls *Q-at-issueness*. Linguistic content is Q-at-issue if it helps answer a conversation's *question under discussion* (QUD), the question(s) around which a conversation is oriented; otherwise it is not-Q-at-issue. In the QUD framework developed by Craige Roberts (2012), an utterance satisfies the Gricean maxim of relevance just in case it helps answer the conversation's QUD. This is a theory in linguistic pragmatics rather than semantics, and it has been argued that the theory is ill-equipped to do genuine semantic work (Grindrod and Borg 2019). My criterion of when linguistic content foregrounds the subject is not purely semantic, since I do not assume that a given sentence is semantically associated with a small set of questions. But since the range of questions a sentence answers *by default* is partly a function of the lexical properties of the words that sentence comprises, my criterion is not merely pragmatic.

such as *what is* x *like* or *what is good (right)* then it is objective. And it remains objective even if in recherché contexts it can be used to answer questions about the speaker and even if information about the speaker is not-at-issue and therefore remains recoverable from moral assertions.

There is a sense in which the objectivity of moral language is a datum of the language we speak, which predicates moral properties of objects. But clearly this is not sufficient to establish objectivity in any interesting philosophical sense. Since a complete metaethical view provides at least a theory of moral thought or of the moral facts alongside a theory of moral language, we can say that a metaethical view is objective only if it has an objective theory of moral thought or of the moral facts. For any view failing this condition—such as those to be discussed in **§5.1** below—the surface objectivity of moral language belies a deeper subjectivity.

4.3. Aboutness in metaphysics

The *D*-facts foreground the subject if the relevant judge(s) features in the *grounds* of the *D*-facts. I take grounding to be an explanatory relation between facts: a set of facts Γ grounds a fact [p] if and only if the fact(s) in Γ metaphysically explain [p]; if Γ grounds [p], then [p] obtains *because* the facts in Γ obtain.[14] Here are some paradigmatic instances of grounding claims: [the Brandenburg gate is in Berlin or Frankfurt] is grounded in [the Brandenburg gate is in Berlin], and [x is a square] is grounded in [x is a rectangle] and [x's sides all have the same length]. In general, disjunctive facts are grounded in their true disjuncts, and conjunctive facts are grounded in their conjuncts taken together. In this last case, each conjunct is a *partial* ground

[14] For introductions to this notion and the work it might do in metaphysics, see (Rosen 2010; Fine 2012; Raven 2015).

of the relevant conjunction. Sometimes, when [*q*] is a partial ground of [*p*], there is some other fact [*r*] relevant to the full grounding explanation of [*p*], but which seems to play a different role from the partial ground [*q*]. In some such cases, what [*r*] seems to explain is not [*p*], but the explanatory fact [[*q*] explains [*p*]]. In such cases we may call [*r*] a *background condition* in the grounding explanation of [*p*].[15] The distinction between grounds and background conditions is an instance of a general distinction made in explanatory contexts, including those of causal and

[15] A good example of this in the metaphysical grounding case is accidental universal generalizations, such as [everyone in Bob's family was born in January]. Suppose that Bob's family = {*b*, *c*, *d*}. This universal generalization is not fully explained by the following particular facts alone: [*b* was born in January], [*c* was born in January], and [*d* was born in January]. What fully explains this generalization is these particular facts together with a general fact, namely that [*b*, *c*, and *d* are all the members of Bob's family]. Yet to many it seems that this *totality* fact explains not the universal generalization—at least, not in the same way as do the particular facts—but rather explains the explanatory fact that the instances of the universal generalization ground that generalization. Similarly on Ralf Bader's view (ms, 11–14), a universal generalization [all *F*s are *G*s] is grounded in all the particular facts [F_1 is *G*], [F_2 is *G*], … with the background condition that there are no other *F*s. Stephen Yablo has a brief discussion of universal generalizations in his *Aboutness* (Yablo 2014, 63–65). He first considers regimenting the totality fact as a truth-maker-maker rather than as a truthmaker, but this would have the result that a sentence's truthmakers are not guaranteed to necessitate its truth. Yablo therefore takes the truthmakers of a general fact to be its instances "*qua* everything." For Yablo a sentence's subject matter is the union of the set of its semantic truthmakers with the set of its semantic falsemakers. So, since on neither proposal is the totality fact a truthmaker of the generalization, it follows that on neither proposal is it part of the generalization's subject matter. (I prefer to conduct our metaphysical inquiry in terms of ground rather than semantic truthmaking, but the difference will not matter for our purposes.)

normative explanations, between direct explainers on the one hand and meta-explainers or background conditions on the other.[16]

If the moral facts are partly grounded in the facts about some judge's attitudes, facts such as [*S* disapproves of *φ*-ing], then the moral facts foreground the subject. But if the moral facts are partly grounded only in what we intuitively think of as the right/wrong- or good/bad-making features of some action or state of affairs, then the moral facts are objective. And they remain objective even if facts about some judge's attitudes appear in the background conditions of these explanations.

Many have been skeptical that a distinction between direct explainers on the one hand and background conditions on the other can be sustained.[17] I remain officially neutral on the

[16] In the literature on causation, some authors have sought to distinguish causes from background conditions: the former produce the relevant effects, whereas the latter enable or disable this production. This view has recently been defended by Bradford Skow (2018, chap. 2). Skow's view is that background conditions—which are *states*—are reasons why some cause is a reason for its effect—which are both *events*. A related distinction is Jonathan Schaffer's (2017b; 2017a) distinction between grounds or causes ("sources") on the one hand and grounding or causal laws ("links") on the other. Similarly, some have sought to distinguish normative grounds—what we intuitively think of as the right/wrong- or good/bad-making features of some action or state of affairs—from enablers and disablers, strengtheners and attenuators, and normative principles. Champions of this view include Jonathan Dancy (2004, chap. 3), Bader (2016; 2017), and Shlomit Wygoda Cohen (2020; 2021). Wygoda Cohen has also applied this distinction to the question of objectivity; see **n30** below.

[17] David Lewis called this distinction, in the causal case, "invidious discrimination" (1973, 559). (For discussion of this distinction in the causal case generally, see Gallow 2022, sec. 1.2.3.) In the normative case, Gideon Rosen has defended the view that moral laws feature as partial grounds of the particular

question. My claim is conditional: if facts about a subject's attitudes feature only in the background conditions of the grounding explanations of the moral facts, their objectivity is not spoiled. And this conditional claim is compatible with there being no background conditions, only partial grounds. If the distinction between grounds and background conditions cannot be sustained, then views whose objectivity relies on it would indeed collapse into versions of subjectivism. (I discuss such views in **§5.4** below.)

5. The varieties of restricted subjectivism and objectivism

These substantive accounts of foregrounding the subject in hand, we may now ask which metaethical views count as forms of restricted subjectivism and which as forms of restricted objectivism.

5.1. Subjectivism

Consider a language that lacked predicates such as 'right,' 'wrong,' 'good,' and 'bad.' Suppose further that non-cognitivism about moral judgment is true—that moral judgment consists ultimately in having some conative or affective mental state. Speakers of this language plausibly could make moral judgments—they could have the states of approval and disapproval

---

moral facts. (See Rosen 2017b; 2017c; and especially 2017a, where he (pp. 283–84) explicitly addresses Bader's view. This view has also been defended by David Enoch; see Enoch 2019.) Kit Fine (2020, 161 n33) registers the following complaint about Yablo's view of universal truths, "I myself am somewhat unhappy with truthmakers carrying presuppositions in this way. Truthmakers are in the world; and it seems to me that presupposition properly belongs to language, not to the world." Rosen (2010, 120–21), for his part, views the totality fact as another partial ground.

or what have you—but they could only voice these states by describing their own psychologies, by saying 'I disapprove of stealing,' for example. Let us call this language *radically* subjectivist. Given our criterion from **§4.2**, we can say that a metaethical view according to which the radically subjectivist language is more faithful to the nature of moral judgment than a language that predicates moral properties of objects is a form of subjectivism. For the radically subjectivist language would be more perspicuous than a predicative language only if moral judgment were a way of venting our own non-cognitive mental states and only if there were no moral facts to describe. Paradigmatic cases of such views are simple speaker subjectivism—the view that '*x* is right' is true at a context *C* if and only if 'I approve of *x*' is true at *C*—and the classical emotivism of A.J. Ayer (1946) and C.L. Stevenson (1937) with which simple speaker subjectivism is often contrasted.[18] Other subjectivist views are those on which the facts about what is right or good are indexed or relativized to the subject or some group containing the subject, as in relativist forms of a dispositional theory of value.[19]

[18] For Ayer, moral language is defective; and for Stevenson, it is deceptive. Ayer writes that, setting aside the expressing/reporting distinction, "our theory of ethics might fairly be said to be radically subjectivist" (109) and concludes that "ethical philosophy consists simply in saying that ethical concepts are pseudo-concepts are therefore unanalysable" (112). Stevenson explains that having a predicate 'good' in the language to express one's desires—what I would call the surface objectivity of moral language—"enables one to make changes in a much more subtle, less fully conscious way. Note that the ethical sentence centres the hearer's attention not on his interests, but on the object of interest, and thereby facilitates suggestion" (26).

[19] On this sort of view, what is valuable is fixed by what some group *G* is disposed to value under conditions *C*. On David Lewis's (1989) view, group *G* is constituted by *us*. Lewis accordingly

5.2. Quasi-realist expressivism

We now turn to objectivist views, beginning with the version of metaethical anti-realism most likely to count as objective: quasi-realist expressivism. In arguing for the objectivity of quasi-realist expressivism, it will be instructive first to consider expressivist theories of moral objectivity.

The classic expressivist theory of objectivity as mind-independence was developed by Simon Blackburn (1993b). Blackburn's basic strategy is to regiment the claims of dependence to be denied as counterfactuals (e.g. 'if I did not disapprove of stealing, stealing would not be wrong') and to hear those counterfactuals as making some first-order claim. This allows the expressivist to provide a non-cognitive analysis of the relevant subjectivist claim—an account of which state of mind it expresses—as well as a non-cognitive analysis of its negation. The idea is roughly that the subjectivist claim expresses conditioning moral judgment on one's own occurrent sensibility, and its negation expresses a refusal to do so.[20]

---

distinguishes between an *absolute* version of the view, on which *G* is constituted by humanity; and *relative* versions of the view, on which *G* is constituted by the speaker and those relevantly similar. These relativist views are subjectivist in the sense at issue because according to them, what is valuable just is what is valuable *for me*. As Lewis writes, "for a speaker S to call something a value is to call it a value for the population of S and those like him; which means that S and those like him are all disposed, under ideal conditions, to value it" (126). (Lewis himself prefers what he calls the *wait-and-see* version, on which speakers make the strongest claim they can get away with.)

[20] The underlying idea is that the wrongness of stealing is a matter of its naturalistic profile, rather than the fact that we disapprove of it. As Blackburn writes: "'Even if we had approved of it or enjoyed it or desired to do it, bear-baiting would still have been wrong' can sound like a second-order realist

Blackburn's theory, like other expressivist theories of objectivity, suffers from an important dialectical limitation: its provides an account not of moral objectivity itself, but rather of what it is to believe in moral objectivity.[21] Here expressivists follow their general

commitment directly in opposition to projectivism. But in fact, on the construal of indirect contexts that I offer, it comes out as a perfectly sensible first-order commitment to the effect that it is not our enjoyments or approvals to which you should look in discovering whether bear-baiting is wrong (it is at least mainly the effect on the bear)" (Blackburn 1993b, 172–75). Selim Berker (2020) has argued that Blackburn would do better to analyze the problematic claim of dependence as a grounding claim rather than as a counterfactual. On this view, what the quasi-realist should be concerned to deny is sentences such as 'stealing is wrong because I disapprove of it,' sentences that claim that facts about our attitudes are the normative grounds of the moral facts. Berker's basic proposal, on behalf of the expressivist, is that these sentences express basing one's disapproval of stealing on one's own sensibility. (Berker himself expresses reservations about this proposal, since it cannot provide a non-circular analysis of 'I disapprove of stealing because it is wrong.')

[21] Gibbard (1990) argues for three forms of normative objectivity: that our normative views have "validity independent of acceptance," (164) that they represent "demands of rationality" rather than "idiosyncratic existential commitments," (155) and that they have "authority" for others and not only for ourselves (155). Gibbard's accounts of these senses of objectivity share the limitation of Blackburn's. The first sense of objectivity corresponds to Blackburn's notion of mind-independence, and Gibbard essentially adopts Blackburn's account. (See also Gibbard 2003, 183.) Gibbard cashes out the distinction between "demands of rationality" and "idiosyncratic existential commitments" in terms of ways one can *accept* a norm, not in terms of features of the norms themselves. ("To accept a norm as a requirement of rationality… is to accept it along with higher order norms that require its acceptance. To treat it as an existential commitment is to is to accept it along with higher order norms that permit it, but that permit

methodology—what Jamie Dreier has called the "sidestep[ping]" strategy (Dreier 2015, 283)—of providing an analysis not of our normative concepts directly, but of our normative beliefs and the sentences that express them. But in response to an opponent who doubts that anti-realist views in metaethics can be objective, the sidestepping strategy of analyzing moral objectivity seems evasive: the expressivist may be able to make sense of what it is to *believe* that morality is objective, but what we had wanted to know was whether, if expressivism is true, moral thought and moral language *are* objective.[22]

The restricted conception of objectivity as object-aboutness is essentially the constitutive analogue of the quasi-realist's theory of objectivity: it takes the quasi-realist's theory of what it is to believe in moral objectivity—roughly for one's moral judgment to be based on the morally relevant features of the situation rather than on one's own sensibility—and turns it into a theory of what moral objectivity consists in. In other words, the quasi-realist's analysis of the idiom of dependence shows that on its interpretation, moral thought is not about the subject—what is in the foreground of moral thought is the morally relevant naturalistic features of the relevant

---

accepting at least one incompatible alternative" (1990, 169)). And normative authority comes to *demanding* that others accept the norms you accept; authority pertains to a way I express the norms I accept rather than to a feature of the norms themselves. ("What is it, then, to claim authority for a judgment? […] What I do is to exert a conversational pressure. In effect I demand acceptance of what I am saying" (172).)

[22] This move is standardly justified by claiming that the question of dependence is first-order—internal to moral discourse and practice—there is no meaningful external question of dependence (Blackburn 1993b, 172–75). But that claim seems especially suspect when we recall that the question of dependence is just one facet of the question of objectivity, and that does not seem to be a merely first-order question.

situation, not the attitudes of the thinker—and so the theory counts as a form of objectivism in the restricted sense about moral thought.[23] What about the objectivity of moral language and of the moral facts? For expressivists, moral thought is primary—moral language is understood as expressing mental states—so the fact that moral language answers by default questions involving the world rather than the subject is derivative of the nature of moral thought.[24] Similarly for the expressivist's theory of the moral facts. The sense in which there are moral facts for the expressivist is a minimal one, derivative of the grammatical properties of the declarative sentences we use to express our moral judgments. Since the expressivist's moral facts are ultimately an ontic shadow of moral thought, it is no surprise that the expressivist's theory of the moral facts is an objective one by our criterion: on this view, the moral facts are *normatively* grounded in the relevant naturalistic facts, and they are not *metaphysically* grounded at all.[25] So

[23] Pettit and Smith (1990, 579 n19) essentially say this about quasi-realism, calling it "a form of non-cognitivism which would not foreground desire." For an important dissent, see (Johnston 2001).

[24] As Blackburn writes: "The correct opinion about these things is not necessarily the one we happen to have, nor is our having an opinion or not the kind of thing which makes for correctness. The standards governing projection make [our current opinion] irrelevant, in the way that opinion is irrelevant to the wrongness of kicking dogs…. When I say that Hitler was evil… I am not talking about myself" (Blackburn 1984, 219). This quotation nicely highlights the connection between the sort of dependence Blackburn denies and subjectivism in the sense of subject-aboutness.

[25] See (Blackburn 1993b, 172–73). The idea here must be not that the moral facts are metaphysically brute, like facts about the positions of fundamental particles in space, but rather that they are inapt to be metaphysically grounded at all. For discussion of whether the quasi-realist can consistently deny the mind-dependence of the moral facts given the view's theory of moral judgment, see (Cassam 1986, 450–51; Street 2011; Köhler 2014; Enoch 2019, 21).

there is no sense in which our attitudes feature in the explanation of the moral facts for the expressivist.

5.3. Plan-expressivism

So far, our discussion of expressivism has concerned accounts like that of Blackburn, on which moral judgment consists in approval or disapproval. Other forms of expressivism, most prominently the *plan-expressivism* of Allan Gibbard in *Thinking How to Live* (2003), raises special issues. On Gibbard's view, normative judgments are practical or *plan-laden*: to make a normative judgment is to have an intention. We might therefore worry that the question answered by normative judgment—the question of *what to do*—involves the subject, and so that the plan-expressivist's views of moral thought and moral language are subjectivist by my criteria.

The key to responding to this worry is to fix on the idea that plans are for some agent. My plan to prepare salmon for dinner is a plan for myself, but we can also make sense of plans for other agents.[26] And it seems constitutive of moral plans—however these are conceived exactly—that they are plans for everyone.

This touches on a point of broader significance. Just as we can ask about the 'where' of the subject in some domain—whether the subject is foregrounded, backgrounded, or entirely absent—we can also ask about the 'who'—whether the subject is a particular individual or a

[26] The idea of planning for others has been developed by Alisabeth Ayars (2022), who argues that it provides a better account of third-personal normative judgments than Gibbard's own. A common context in which we make plans for others is cases of what Ayars calls *delegated decision-making*, as when I ask you to decide for me what I should order at a restaurant.

generic individual, standing in for all humans or all rational creatures or the like. There is a clear sense in which domains involving a particular subject are less objective than accounts involving a generic subject: the former are genuinely subjective, whereas the latter are intersubjective. So, if the aim of moral thought is to decide what we all should do, and moral language answers the question what (are we all) to do, then we can call the plan-expressivist's account of moral thought and moral language a form of intersubjectivism.

Of course, intersubjective accounts are not forms of objectivism in the unrestricted sense. Nor, according to our definition, are they forms of objectivism even in the restricted sense. This may well be the correct result for the question of metaphysical objectivity, in which we want to know which features of the world are independent of us—even where 'us' might include a class as large as all rational beings. But the classification of intersubjectivism as a kind of subjectivism seems like the wrong result for the question of moral objectivity. This is because our concern to avoid foregrounding the subject does not seem to include a concern to avoid foregrounding the subject among others. (Concern for humanity does not seem to entail an objectionable sort of self-concern, even though one is also a human.) In these cases, the sense in which the relevant domain is about the subject is merely interaliotic (to coin a phrase): it is about the subject only inter alios, inasmuch as the subject is included in the set containing everyone (in a sufficiently broad group, such as all humans or all rational beings). This indicates that moral thought, moral language, and the moral facts should count as problematically subjective only if they foreground a single subject (or an insufficiently broad group, such as my clan or me and my mates). Given this understanding, plan-expressivism will turn out not to be a kind of subjectivism.[27]

[27] In his discussion of "Socrates's Question" of how we ought to live, Bernard Williams (1985) argues against the idea that practical thought is subjective even in the first-personal case. He writes that unlike

5.4. "Non-objectivism"

Let us now consider a family of views sometimes called versions of "non-objectivism."[28] These views are *cognitivist* since they hold that moral judgment consists in belief, and they are *success theories*—that is, not error theories—since they do not hold that our moral beliefs are systematically in error. So according to these views, there is a domain of moral facts. But these facts are explained in some way by facts about us and our attitudes, such as facts about what we would all disapprove of under ideal conditions.[29] These sorts of views deny that the moral facts are objective given the unrestricted sense of aboutness, and so deny that the moral facts enjoy the sort of objectivity at issue in the debate about objectivity in metaphysics generally.

But many versions of non-objectivism—I retain the label for convenience—will in fact count as forms of objectivism given the restricted sense of aboutness, supposing one view about how the moral facts are metaphysically grounded. Even non-objectivist views should accept the supervenience of moral properties of objects or actions on their natural properties—the right/wrong or good/bad-making features of an action or state of affairs—and thus see an explanatory role for facts involving these natural properties to play in the metaphysical grounding explanations of the moral facts. Two facts will then be involved in the complete metaphysical explanation of a fact such as [act *a* is wrong] according to the non-objectivist: [act

---

the question "what shall I do," the question "how should one live" is "impersonal," (4) and that it is "*anybody's* question" inasmuch as it "can be put to anyone" (20).

[28] I borrow the label from (Joyce 2021).

[29] This family includes ideal observer theories such as (Firth 1952), idea advisor theories such as (M. Smith 1994), and various versions of constructivism. (See Sharon Street's (2010) for discussion and defense of constructivism.)

a is *N*], where *N* is some natural property, and a fact like [we all (would) disapprove of any action having *N*]. This view will count as a form of restricted objectivism if only [act *a* is *N*] explains [act *a* is wrong] as a partial ground, with [we all (would) disapprove of any action having *N*] featuring in the explanation only as background condition.[30]

[30] Wygoda Cohen (2021) also discusses the implications of the distinction between partial grounds and background conditions for the question of moral objectivity. She distinguishes between objectivity as *mind-independence*, according to which the attitude facts play *no* role in the grounding explanations of the relevant facts (*Strong Objectivism*), and objectivity as *mind-nongroundedness*, according to which the attitude facts do not feature in the partial grounds, but may play some other role in the relevant grounding explanations (*the Intermediate Position*). She argues that several forms of what Joyce calls non-objectivism should be understood as versions of the Intermediate Position—including David Wiggins's (1987) Sensible Subjectivism, the voluntarism of Pufendorf and Hobbes (as understood by Christine Korsgaard; see her 1996, 21–27), and Schroeder's (2007) Hypotheticalism about practical reasons—and that this solves some puzzles for these views. My account allows me to draw a stronger conclusion from Wygoda Cohen's argument than she herself does, namely that versions of the Intermediate Position simply do count as forms of objectivism, given the conception of objectivity as object-aboutness and the restricted sense of aboutness. Other authors who have considered such grounding structures, such as Christopher Frugé (2022) and Justin Morton (2022), have defended the contrary view that this sort of explanatory structure should be understood as a form of subjectivism. (Morton's discussion concerns moral mind-dependence; he holds that particular moral facts count as subjective when their grounding explanations are governed by moral laws themselves fully grounded in facts about a subject's attitudes. Frugé argues that the insights that have motivated subjectivism about personal value motivate a view on which facts about what an agent values explain the personal-value-grounding facts, facts of the form [[x is N] grounds [x is personally valuable for S].)

More precisely, this view will count as a form of restricted objectivism about the *particular* moral facts. The objectivity of *general* moral principles will depend on precisely how these are understood.[31] If the general moral principles take the form [if *x* is *N*, then that makes *x* *M*], where *N* is a natural property and *M* a moral property, then the non-objectivist might claim that these principles are *universal generalizations* or *summaries* of particular grounding facts. This will allow the non-objectivist to deny that the principles are subjective: because universal generalizations are grounded in their instances, and we are granting that the instances (i.e. the particular moral facts) are objective in the relevant sense, so too will be the general moral principles.[32]

[31] Wygoda Cohen's (2021) discussion concerns the grounding explanations of the "most fundamental moral facts," (186) those moral facts that explain other moral facts and are not themselves explained by any other moral facts—facts about the moral laws or the essences of the moral properties. I suspect, however, that the views she identifies as versions of the Intermediate Position hold that it is the grounding explanations of *particular* moral facts, not the explanatorily fundamental ones, in which facts about our attitudes might feature as background conditions rather than as partial grounds.

[32] A very similar view will, however, come out as a form of subjectivism about the general moral principles by my criterion. Suppose that the general moral principles involve predications of act-types, such as [breaking a promise is pro-tanto wrong.] The non-objectivist would then analyze the predicate 'is wrong' in terms of some predicate like 'would be disapproved of by group *G* in circumstances *C*.' Then the general moral principles would be subjective, since they would be fully grounded in facts about the act-types whose performance some group disapproves of.

5.5. Yablo-style moral fictionalism

The last view to be discussed provides an objective theory of moral language by exploiting not-at-issue linguistic content. To build the view, take the metaphysical thesis that there exists no non-natural object rightly called *the objective moral standard* (objective in the sense of being an object, i.e. an entity) along with the semantic thesis that no sentence of folk moral discourse is literally true because that discourse presupposes the existence of an objective moral standard. So far, the view is theoretically indistinguishable from the moral error theory, and in particular a presuppositional version of the moral error theory.[33]

The view I have in mind therefore counts as a form of *moral fictionalism*—not one that involves any notion of make-believe or pretense, but rather a kind of fictionalism as developed by Stephen Yablo.[34] On this sort of view, although the characteristic objects of *D*-discourse do not exist, that fact is *irrelevant* to the subject matter *S* we use *D*-sentences to convey information about. As a result, when we use *D*-sentences, we can *subtract* their problematic commitments to *D*'s characteristic objects, leaving as a remainder the information they convey about *S*. So, even if the sentences of the domain are not, strictly speaking, true, they still may convey true information about *S*.[35]

---

[33] The classic defense of the moral error theory is J.L. Mackie's (1977); for presuppositional versions see (Perl and Schroeder 2019; Salinger 2021; Perl 2022).

[34] For other versions of moral fictionalism, see (Joyce 2001; 2005; Nolan, Restall, and West 2005; Kalderon 2005).

[35] Yablo comes to understand this form of fictionalism as an application of *logical subtraction* (2014; 2017a; 2017b), the phenomenon by which we are heard as speaking about some subject matter *S*—the subject matter of interest in conversation—despite our speaking falsely about some other subject matter

The case of greatest interest to Yablo is mathematical discourse, which presupposes the existence of abstract objects. There the thought is that even if there are no abstracta, and so no numbers in a Platonic realm, mathematical discourse is still true about a subject matter orthogonal to the metaphysical subject matter *what resides in Platonic heaven*. This subject matter is: *what the world is like in its concrete respects*. So, though 'the number of chairs in this room is ten' would not be true were there no number ten to which the number of chairs in the room could be identical, still the sentence might still say something true about the chairs.

A Yablo-style moral fictionalism will be, at least in broad outline, formally similar: although moral language presupposes the existence of an objective moral standard, sentences about this non-natural entity are irrelevant to moral subject matter, where this has to do with the naturalistic right/wrong- and good/bad-making features of the world.[36] However exactly we might develop this Yablo-inspired version of moral fictionalism, what is of interest to us is its

---

*S** in doing so. We can also say that according to such a view, moral discourse exhibits what Yablo calls *non-catastrophic* presupposition failure (Yablo 2006). In cases of *catastrophic* presupposition failure (e.g. an utterance of "the Holy Roman Emperor is six feet tall"), a sentence's false presupposition renders it defective in the way characteristic of presupposition failure, either truth-valueless or trivially false. But in cases of non-catastrophic presupposition failure, my utterance does not count as defective despite its backgrounded commitment to something false.

[36] This thought can be made precise if we formulate this version of moral fictionalism using Daniel Hoek's (2018) theory of *conversational exculpature*, which can be used also to implement Yablo's version of mathematical fictionalism. In this framework, we can say that moral sentences are equivalent to some (possibly long and disjunctive) sentences using only naturalistic vocabulary, conditional on the presupposition that there is an objective moral standard.

mere possibility. For when coupled with an objectivist theory of moral thought, what results is a metaethical view that remains objective despite accepting the two constituent claims of the moral error theory.

6. Moral judgments and judgments of taste

Someone worried that my account of moral objectivity is too permissive might raise the following objection:

Consider judgments about what is tasty, funny, disgusting, and the like. Moral judgments are supposed to be distinct from such judgments of taste—a feature of moral thought and language that even anti-realists have sought to secure.[37] Yet my account threatens to render these judgments as objective. For our language for expressing judgments of taste is not radically subjective; it rather includes predicates of personal taste which we predicate of the objects we find tasty or funny or disgusting or whatever. And it seems clear that these predicates function to foreground the objects of which they are predicated rather than the subject or her tastes, and so count as objective in my sense.

So far, the objector has given us no reason to think that the objectivity of our language for expressing judgments of taste differs from moral language as understood by the subjectivist: although it is a datum of the language we speak that sentences involving predicates of personal taste are objective, a further philosophical question remains about whether our language

[37] In his (1990, chap. 8) discussion, Gibbard sometimes seems at times to treat the problem of moral objectivity as identical with the problem of distinguishing moral judgements from judgments of taste, inasmuch as matters of taste are mind-dependent—lacking "validity independent of acceptance" (164)—and are more like existential commitments (167).

perspicuously represents thought about what is tasty or the facts about what is tasty.[38] But the objector might also claim that thought about what is tasty and the tastiness facts should also count as objective on my account. If thought about what is tasty foregrounds not the subject but how things taste—whether the thing's flavor is salty or sweet or bitter or the like—and the tastiness facts are grounded in the flavor facts with the background condition that one enjoys the taste of the relevant flavor—then matters of taste will be fully objective on my account.

I concede that on such a view, thought about what is tasty and the facts about what is tasty would indeed be objective. But they would still be distinguished from the analogous moral domains inasmuch as they exhibit a higher degree of subjective involvement. This is because thought about what is tasty is derivative of one's experience with the flavor and quality of the tasty thing, and so beliefs about what is tasty effectively record one's own responses to a gustatory experience: how some flavor strikes me. Moral thought, on the other hand, does not express one's experience with the thing judged right or wrong or good or bad.

This difference is reflected in our language for expressing moral judgments and judgments of taste. Judgments of taste seem relative to a judge or standard of taste parameter,

[38] On John MacFarlane's semantic treatment of 'tasty,' if licorice is tasty to me at the actual world, then the counterfactual 'if there were no tasters, licorice wouldn't be tasty' comes out false. This is because the counterfactual shifts the value of the world parameter of the index, not the value of the standard of taste parameter, and the closest world without any tasters is one in which licorice has the flavor it does at the actual world (MacFarlane 2014, 165–66). MacFarlane takes this analysis to show that in at least some sense judgments of taste are mind-independent. On my view this analysis vindicates the objectivity of tastiness judgments in language, but leaves unanswered the deep philosophical question of objectivity.

made explicit in constructions such as 'licorice is tasty to John' and 'kibble is tasty for Fido.'[39] And relatedly, simple statements ascribing predicates of personal taste exhibit an "acquaintance inference": from my utterance of 'licorice is tasty' you can infer that I have first-hand knowledge of or experience with licorice's flavor.[40] Moral language is disanalogous to language expressing judgments of taste in both respects. For we simply cannot speak of what is wrong to or for some person in the way that we can about what is tasty or funny for them. The best interpretation of the former idiom is that stealing is wrong according to the standard that the person accepts, and that seems an intellectualized and therefore strained interpretation of what it is for a food to be tasty to someone. (We can also speak of what is 'bald' or 'tall' according to the standard that someone accepts owing to these predicates' vague conditions of application, but this is clearly a different matter from that of what is tasty or funny to someone.) And there is no moral analogue of the acquaintance inference: my utterance of 'stealing is wrong' does not license the inference that I have ever been the victim or perpetrator of a theft.[41] Since predications of personal taste

[39] For some important semantic proposals about the nature of this relativity to a judge or standard of taste parameter, see (Lasersohn 2005; Stephenson 2007; MacFarlane 2014, chap. 7). To be clear, it is not the mere relativity to some standard that I am suggesting makes tastiness judgments subjective relative to moral judgments, it is the fact that the standard to which these judgments are relative—a judge or a 'standard of taste' that formally models my gustatory experience—involves the subject.

[40] For discussion of the precise nature of this inference, see (Pearson 2013; Ninan 2014; Willer and Kennedy 2022; Ninan forthcoming).

[41] For some related discussion, see (Franzén 2020, 1111–12). Franzén argues that since moral predicates embed under 'find'-constructions, this provides linguistic evidence that moral judgment is non-cognitive. Even according to Franzén, however, moral 'find'-constructions, unlike the analogous aesthetic and taste

involve the subject in ways that moral predications do not, my account still provides resources to distinguish moral judgments from tastiness judgments.[42]

judgments, do not give rise to the acquaintance inference. (Francén 2023 suggests that perhaps moral 'find'-constructions do have an acquaintance requirement, though one that is very easily met.)

[42] **Acknowledgement:** A version of this chapter was presented at the 5th Frankfurt Metaethics Conference (FrankMeta.5) at the Frankfurt School of Finance and Management in June 2023. I thank the conference participants for their questions and suggestions.

Chapter 3: Genuine Disagreement

1. Introduction

When do subjects count as disagreeing? Call this the *question of disagreement*. According to orthodoxy, subjects disagree only when they hold beliefs that cannot be true together. Heterodoxy, following ideas developed by C.L. Stevenson (1937; 1944), admits "disagreement in attitude" alongside disagreement in belief. On this view, subjects can disagree not only by having contrary beliefs, but also by having contrary non-doxastic states, such as desires or preferences. On the novel *question-relative* account of disagreement I defend, both orthodoxy and heterodoxy are mistaken: disagreement is not limited to belief, but there is no mere disagreement in attitude. The account I offer differs also from conceptions of disagreement as a form of incoherence, which has implications for whether subjects can disagree in virtue of their preferences and credences. The answer to the question of disagreement is thought to have a bearing on the *problem of normative disagreement*. This is the problem, faced by metanormative views on which normative language does not straightforwardly express beliefs about normative reality, of explaining why subjects with contrary normative views thereby disagree. I clarify this problem by distinguishing *superficial* from *genuine* disagreement, and I evaluate extant solutions to the problem proposed by expressivists, contextualists, and relativists. Finally, I discuss the conceptual connection between (genuine) disagreement and objectivity.

2. The analysis of (dis)agreement

The question of disagreement is, in the first instance, a question pertaining to subjects. The question is: under what conditions do subjects disagree? In the case of *linguistic disagreement*,

subjects disagree in virtue of the sentences they utter; in the case of *mental-state disagreement*, subjects disagree in virtue of the mental states they possess.[1] (This allows us to speak derivatively of when sentences or mental states disagree with one another.) This essay will primarily discuss the question of mental-state disagreement, which has an *a priori* character; the question of whether two sentences disagree, by contrast, seems to depend on empirical semantic facts about a given natural language. (The relation between linguistic and mental disagreement will be crucial for the discussion in **§§6–7**).[2]

My analysis of agreement and disagreement incorporates a largely overlooked insight, namely that they bear a conceptual connection to subject matters conceived as *questions*.[3] Notice

[1] Cappelen and Hawthorne introduce a distinction between disagreement as a *state* and disagreement as an *activity* (2009, 60–61). Two thinkers can be in disagreement as a state, but not as an activity, if they have no interaction with each other. Speakers can also be in disagreement as an activity even when they are not in disagreement as a state just by asserting contradictory sentences, even if one speaker does not truly believe what she says. MacFarlane (2014, 120) takes disagreement as a state to be the more fundamental notion; if this is roughly the same as disagreement in mental states, then my view follows MacFarlane's.

[2] MacFarlane expresses skepticism that there is such a thing as "*real* disagreement," arguing instead that we do better to distinguish several senses of disagreement and ask which senses a given semantic theory has the resources to explain (2014, 119). Perhaps this assessment is correct with respect to adjudicating disputes among semantic theories. But if a notion of genuine disagreement seems to do further work in philosophy—as I argue it does in **§§6–7**—then we should do our best to provide an account of it.

[3] The identification of subject matters with questions follows (a generalized version of) the theory of subject matter developed by David Lewis (1988a), an instance of what Peter Hawke calls the "way-based" conception of aboutness (2018, 698). Brendan Balcerak Jackson (2019) discusses the relation of

that whenever A and B agree or disagree, there is always some question we can report their agreement or disagreement as being *about.* (Dis)agreement is therefore relative to a question in roughly the way that motion is relative to a reference-frame: just as upon analyzing our concept of motion we came to realize that what it is to be in motion is to be in motion *relative to* some reference frame; so too, I claim, analyzing our concept of (dis)agreement reveals that what it is to (dis)agree is to (dis)agree relative to some question.

Questions are standardly analyzed as the union of their possible *answers* (correct and incorrect).[4] For example, the question *Who was the first president of the United States?* can be analyzed as, {Socrates was the first president of the United States, Cleopatra…, George Washington…, Phillipa Foot…, etc.} for every person in the domain. We can therefore say that some content constitutes an answer to a question just in case it is a member of the set constituting the question. I propose the following analyses:

(Agreement) A and B *agree* about *S* when they accept the *same* answer to the question *S*.

(Disagreement) A and B *disagree* about *S* when they accept *incompatible* answers to the question *S*.

The analyses invoke two further ideas in need of clarification: what it is to accept an answer to a question, and when two answers to some question are incompatible. To accept an answer to a question *S* is to have a *committing attitude* toward a content that serves as an answer

---

disagreement to questions, though his discussion concerns the relativity of disagreement as an activity to the pragmatic notion of *questions under discussion*.

[4] For an overview, see (Cross and Roelofsen 2022). On this analysis, questions are semantic values like propositions, not linguistic items like sentences.

to *S*.[5] Belief is a paradigm of a committing attitude; wondering and supposing are paradigms of non-committing attitudes.[6] For example, one way to accept an answer to the question *who wrote* Principia Ethica*?* is to have a belief that G.E. Moore wrote *Principia Ethica*, because having this belief commits you to some content—namely *that G.E. Moore wrote* Principia Ethica —that is an answer to the question *who wrote* Principia Ethica*?*[7] (The nature of the commitment a subject has when she has a committing attitude depends, of course, on the attitude in question: some

[5] I complicate this definition slightly in **§7.3** below.

[6] Michael Ridge (2013, 46) asks why attitudes are not always in disagreement when their contents are incompatible; on my account, the answer is that only *committing* attitudes are apt to underwrite disagreement.

[7] The idea that attitudes may be understood as answers to questions follows the account developed by Pamela Hieronymi (2005) in order to solve the wrong kind of reason problem for fitting-attitude analysis of value. For Hieronymi, "a reason is a consideration that bears on a question," and right- and wrong-kind reasons are to be distinguished "by distinguishing between kinds of questions on which a consideration can bear" (438). In particular, the right kind of reasons for having some attitude *A* are what Hieronymi calls *constitutive* reasons (to be contrasted with *extrinsic* reasons), which are those reasons that "*bear or are taken to bear* on a question, the answering of which amounts to having the attitude" (449, emphasis in original). Hieronymi also connects the conception of attitudes as answers to questions with the notion of commitment. She writes, "whenever one has an attitude that can be formed or revised simply by settling for oneself a question or a set of questions… one is committed to an answer to the relevant question(s)," and for this reason calls such attitudes "commitment-constituted" (450–451). (In the case of belief, right-kind reasons for belief that *p* are those that bear on the question *whether* p, since answering the question of *whether* p amounts to believing that *p*.)

committing attitudes—such as beliefs—have a *mind-world* direction of fit, and others—such as intentions—have a *world-mind* direction of fit.)

Two answers to some question *S* are incompatible when the contents constituting those answers are not jointly realizable relative to *S*. Contents are non-jointly-realizable when there is no possible world at which they jointly obtain. In the simplest sort of case, two answers to *S* are incompatible because the relevant contents are contradictory, and there is no possible world at which both *p* and *~p* obtain. In more complex cases, although the contents constituting two answers are not *absolutely* non-jointly-realizable, they are non-jointly-realizable relative to *S* owing to some presupposition of *S*. For example, consider the question: *which individual wrote* Principia Ethica*?* Since this question presupposes that the work has a single author, *G.E. Moore wrote* Principia Ethica and *Bertrand Russell wrote* Principia Ethica are incompatible answers to this question even though the relevant propositions are not contradictories.[8]

[8] This complication can be eliminated by building more assumptions into our definition of an answer. On a stricter conception, answers to a question must be mutually exclusive, with the result that two answers will be incompatible whenever they are not identical. Notice also that given the relaxed conception of an answer it will be possible for two subjects each to have committing attitudes toward some answer to *S* but without thereby agreeing or disagreeing about *S*. Suppose that A believes that Russell wrote *Principia Mathematica* and B believes that Whitehead wrote *Principia Mathematica*. So far, they neither agree nor disagree about *who wrote* Principia Mathematica (so long as the question does not presuppose [falsely] that the work has only one author). In another sort of example, suppose A believes that the keys are in the drawer and B believes that they are either in the drawer or in the car. These beliefs are *merely* compatible—neither incompatible nor identical—so on my view A and B neither agree nor disagree about

3. Doxastic and practical disagreement

The clearest case of disagreement is *doxastic disagreement*, which obtains when A believes that *p* and B believes that ~*p*.[9] My account explains doxastic agreement as follows: when A believes that *p* and B believes that ~*p*, they accept incompatible answers to the question *whether* p. And this because A and B have committing attitudes—namely beliefs—toward contents—*p* and ~*p*—that are not jointly realizable relative to the question *whether* p.

Since beliefs are not the only kind of committing attitude, not all disagreement is doxastic disagreement on my account. Just as beliefs are committing attitudes with descriptive or propositional contents, intentions are committing attitudes with practical—or *planning*—contents. Planning content underlies my theory of *practical disagreement*: disagreement about *what to do*. A and B have a practical disagreement when they accept incompatible answers to some practical question. For example, one way to accept an answer to the question of *what to eat for dinner?* is to form an intention to eat salmon for dinner, because having this intention commits you to the plan *to eat salmon for dinner*, which is an answer to the question *what to eat for dinner?* Just as propositional contents are incompatible when they are not jointly realizable, so too with practical contents: the plan to φ and the plan to ψ are not jointly realizable just in case there is no possible world where both plans are carried out. Consider Stevenson's (1937) classic example of interlocutors who propose incompatible plans for how to spend the evening: one intends for the group to go to the cinema, and the other intends for them rather to go to the

---

the question *where the keys are*, even though they agree about whether the keys *might* be in the drawer and disagree about whether the keys *might* be in the car.

[9] Gibbard (2003, 73) calls the view that all disagreement is doxastic the "orthodox" view, others call it the "simple account" (Beddor 2019, 822) or the "simple view" (MacFarlane 2014, 121).

symphony. Since it is part of the case that going to the cinema and going to the symphony are mutually exclusive options for how the group will spend the evening, their plans are not jointly realizable. The interlocutors therefore accept incompatible answers to the question *how (uniquely) to spend the evening* and so disagree about *what to do*.

This example brings out an important feature of planning content, whose importance for the question of disagreement has been recently emphasized by Alisabeth Ayars (2022). This feature is that plans are *for* an "executive subject," who may be the planning subject but may also be some other agent.[10] Suppose A plans to φ and B plans not to φ. One might have thought: inasmuch as it is not possible both to φ and not to φ (in exactly the same circumstances), A and B disagree about <whether to φ>. But, as Ayars notes, if A's plan to φ is a plan for *A* to φ—or a plan *that A* φ-s—and B's plan not to φ is a plan for *B* not to φ—or a plan *that B not* φ—then these

[10] Ayars argues that distinguishing between the planner and the executive subject of a planning judgment allows a better account of the non-cognitive state expressed by third-personal normative judgments—judgments of the form 'A ought to φ'—than Gibbard's account, on which my judgment of 'A ought to φ' expresses a plan for what *I* am to do if *A* in *A*'s exact circumstances (See especially Gibbard 2003, 50–53). Ayars points out that my judgment that A ought to φ should disagree with your judgment that A ought not to φ, but my judgment that I ought to φ if A in A's exact circumstances does not clearly disagree with your judgment that you ought not to φ if A in A's exact circumstances. On Ayars's theory of "deciding for others," normative language expresses decisions, which settle the practical question of what to do for the executive agent of the decision. In cases of what Ayars calls "delegated decision making," Agent A can delegate her decision to B, such that B's decision is for A, i.e. settles the practical question: *what is A to do?* (Note that for Ayars, decision are not intentions, but rather the state that settles intentions. I leave this complication aside. See also **§7.1** below.)

plans may well be jointly realizable. When contrary plans are *for* the same agent, however, they will be incompatible, because it is metaphysically impossible for the same agent both to φ and not to φ (in exactly the same circumstances). In the case of joint deliberation about what to do, where the relevant agent is the *group* constituted by the conversational participants (as in Stevenson's classic example), contrary plans will yield disagreement because it is metaphysically impossible for the group, as a collective agent, both to φ and not to φ (in exactly the same circumstances).

On my account, it is not sufficient for disagreement that parties simply accept answers whose contents are non-jointly-realizable; those answers must rather concern the same question. The question-relativity of disagreement excludes a spurious kind of disagreement that makes trouble for accounts on which non-joint-realizability is by itself sufficient for practical disagreement. In this kind of case, although two subjects have intentions that cannot be carried out together, no practical disagreement seems to obtain. Suppose, for example, that A and B each intend to take the last piece of sushi from the sushi platter. Their intentions are not jointly realizable—there is no possible world in which each takes the last piece of sushi—but it seems wrong to say that A and B thereby disagree.[11] My account excludes this sort of case because there is no practical question to which A's and B's plans constitute incompatible answers. A's plan to take the last piece of sushi answers, from a first-personal perspective, the question *what*

[11] Ridge discusses a similar case: "Suppose two people want both want to rent the same flat. Their preferences have incompatible contents, but we would not find it plausible to say that they disagree. Indeed, it might be natural to infer that they agree that the flat is an excellent one (2013, 47–48)." Lewis calls these cases in which subjects "desire alike" (1989, 119) and also denies that they amount to disagreement.

(is A) to do; and B's plan to take the last piece of sushi answers, from a first-personal perspective, the practical question *what (is B) to do*.[12] There is a closely related practical question to which A and B might accept incompatible answers, namely the question *who (of A and B) is to take the last piece of sushi*. Since plans answering this question are for everyone apt to be picked out by the pronoun "who," in this case A and B, this is just another case of contrary intentions for the same group agent, analogous to Stevenson's cinema or symphony case, and so a familiar example of practical disagreement.

Notice that my discussion has nowhere appealed to a notion of disagreement in attitude, which is often thought to underwrite a notion of disagreement in intention or preference or other non-doxastic state. On my view, proponents of disagreement in attitude have, by eliding the question-relativity of disagreement, generalized incorrectly from the case of belief: it is incorrect to infer that just as subjects with contrary beliefs thereby disagree, so too for subjects with contrary intentions or preferences. For although A's belief that *p* and B's belief that not *p* will always yield disagreement, the same does not hold with A's intention to φ and B's intention not to φ: contrary intentions must be for the same agent in order to answer the same question and therefore to disagree. Because disagreement requires a shared question, mere divergence in

[12] Of course, A and B can each pose their deliberative questions using the first-personal pronoun. But they do not thereby disagree about *what (I am) to do*. Just as context-sensitive expressions must be resolved in order to identify the propositional content expressed by a sentence and the question it is about, so too in the practical case. For example, said by A, the sentence "I am hungry" expresses the proposition *that A is hungry*; and if A says "I am hungry" and B says "I am not hungry," they do not disagree about the bogus question *whether I am hungry*. (Recall that questions are contents, not linguistic items.)

attitude is insufficient for disagreement. In a slogan: there is never merely disagreement *in*; disagreement must always be *about*.[13]

The shared question requirement also distinguishes my account from those on which A and B disagree whenever it would be incoherent for a single subject to have the attitudes of both A and B. This conception of disagreement, which Alex Worsnip has called *disagreement as interpersonal incoherence*, has been defended by James Dreier to argue that subjects can disagree in virtue of having contrary preferences (Dreier 2009), and by Worsnip to argue (*inter alia*) that subjects can disagree in virtue of having different credences toward the same proposition (Worsnip 2019).[14] In the following two sections, I address whether and when contrary preferences and credences yield disagreement according to my view.

4. Preference

When A prefers that *p* and B prefers that ~*p*, what might they disagree about? A tempting response is: *what to prefer*. It is not clear, however, that different answers to the question *what to prefer* are incompatible in the way necessary for disagreement. Notice that the question of *what to prefer* is a *practical* question—in a broad sense of practical—inasmuch as it is answered by

[13] Balcerak Jackson (2019, 4 and n7) explicitly refrains from claiming that disagreement is always about some question in order to leave space for disagreement in attitude, including cases not characterizable as instances of practical disagreement as I have defined it (for instance where A likes *x* and B dislikes *x*).

[14] This conception has been criticized by MacFarlane (2014) under the label of "noncotenability" and by Beddor (2019) under the label of "*de se* conflict." Worsnip defends this conception as an account of what he calls "disagreement in the wide sense," to be distinguished from "disagreement in the narrow sense," which is doxastic disagreement.

plans for what to prefer rather than by preferences themselves. Answers to *what to prefer* will therefore be incompatible when the relevant *plans* are not jointly realizable, and this is not the same criterion as the non-joint-satisfiability of the relevant *preferences*. A's plan for A to prefer that *p* is, quite clearly, jointly realizable with B's plan for B to prefer that ~*p*: this is exactly the case before us.

To recover disagreement about what to prefer, two further conditions must be met. First, we must understand the question of *what to prefer* as including some *uniqueness condition*, such that any answer to that question would be non-jointly-realizable with any other distinct answer. This condition is necessary because although it is rationally impermissible for the same subject both to prefer that *p* and to prefer that ~*p*, it does seem metaphysically possible to have both preferences. (If in fact it is not metaphysically possible to have both preferences, then no additional uniqueness condition will be required.) So plans to have these preferences will not be incompatible in the sense of non-joint-realizability unless we incorporate something like a uniqueness condition. Second, as with all plans, plans for what to prefer must be for the same agent in order for them to be incompatible in the right way for disagreement—that is, incompatible relative to a single question.

Notice, however, that if subjects with contrary preferences are understood as disagreeing about *what (uniquely) to prefer*, interpersonal preference-involving disagreement will be importantly disanalogous to the classic cases of belief- and intention-involving disagreement. In doxastic disagreement—where A believes that *p* and B believes that ~*p*—A and B may well disagree about *what to believe*, but they also disagree about *whether p*. And in practical disagreement—where A intends for some agent C to *φ* and B intends for C not to *φ*—A and B may well disagree about *what to intend*, but they also disagree about *whether (C is) to* φ. In the

doxastic and practical cases, disagreement is generated in the first instance by the question-relative incompatibility of the *objects* of beliefs and intentions (together with their status as committing attitudes) rather than by any incompatibility in having the attitudes themselves.

So, what is the analogue of *whether* p and *whether to* φ for the case of preferences—the question corresponding to the object of the attitude rather than to the having of the attitude itself? To answer this question satisfactorily, we need an account of what preferences are. I will assume that, for our purposes, we can understand having a preference that *p* as roughly equivalent to having all-things-considered desire for *p* rather than for ~*p* in some situation.[15] These desires should be distinguished from mere *pro tanto* desires, which a subject may coherently have for both *p* and for ~*p*, and even for *p* rather than ~*p* and for ~*p* rather than for *p* in the same situation. For example, I might have some desire to spend my vacation touring a foreign country, but also have some desire rather to stay at home and catch up on sleep. Although these are incompatible alternatives, no irrationality is involved in having some desire for each. It is, however, clearly irrational to *prefer* both incompatible options. (Presumably, this fact about the coherence of having contrary desires is to be explained by the fact that mere *pro tanto* desire that *p* is not a *full* commitment to *p*, and this strongly suggests that subjects with contrary desires of this kind do not disagree.[16])

---

[15] I assume for simplicity's sake that since what is preferred (desired) is some way for the world to be, the formal object of a preference (and desire) is a proposition—though often one involving the preferring subject. The formal object of preference (and desire) can also be represented as a centered world; nothing in my discussion will turn on this choice.

[16] It might not even be incoherent both to desire that *p* and be averse to *p,* where aversion to *p* is something distinct from simply desiring that ~*p*. It does, however, seem incoherent to desire that *p* and be

In normal cases, a subject who desires that *p* both likes *p* and has a disposition to bring it about that *p*. If A likes *p* and B dislikes ~*p*, do they disagree? I claim they do not. This is because liking—and more generally, affection—is a kind of *experience* rather than a committing attitude. Liking *x* is just a matter of how *x* feels to some experiencer and involves no commitment to any content, be it either propositional or planning content.[17] Accordingly, I suggest we look at the conative rather than affective element of desire in order to identify the question that subjects with contrary preferences disagree about.[18] If preference involves a commitment to bringing about its object, then a subject who prefers that *p* has a committing attitude to the question: *whether to bring it about that* p, or *whether to realize that* p, or *whether it is to be that* p, or something of the

---

completely averse to *p*, which suggests that we understand desire as a committing attitude, not toward *p* but rather toward *p-in-a-way* or *p-to-some-extent*. But if being completely averse to *p* is just the conjunction of being averse to *p* and having no desire that *p*, then we can explain the incoherence of both desiring that *p* and being completely averse to *p* without admitting commitments toward these sorts of contents.

[17] See **§7.4** below.

[18] It might be thought that, if the conative element of desire (to be discussed presently) is regulated in normal or rational subjects by her likings, my claim that affections do not underwrite disagreement might prevent desires from underwriting disagreement entirely. This objection rests on the claim that if an attitude *A* is regulated in normal or rational agents by attitude *A**, and *A** does not underwrite disagreement, then neither does *A*. But this claim seems false. Consider the case of perception and belief. The beliefs of normal or rational agents are regulated by their perceptions. These, too, are experiences rather than committing attitudes—perceptual seemings are just a matter of how the world appears to some perceiver and involve no commitment to any content. But it is uncontroversial that subjects can disagree by having contrary belief attitudes.

sort. On this account, when A prefers that *p* and B prefers that ~*p*, they disagree in the same cases as when A intends to bring it about that *p* and B intends to bring it about that ~*p*:[19] in each case, it is a practical commitment that generates disagreement.[20]

I mention another proposal largely to set it aside. It might be claimed that subjects with contrary preferences disagree about an evaluative question such as *whether* p *is preferable* or *whether ought* p or even *whether* p *is good.*[21] In some cases where A and B have contrary preferences, they do seem to accept different answers to such a question. Suppose A prefers that the wicked suffer for their wickedness, whereas B prefers otherwise. The conflict between A and B seems well-captured by saying that they disagree about *whether the suffering of the wicked is preferable* or about *whether the wicked ought to suffer*. But it remains unclear how, according to this proposal, preference-involving disagreement is distinct from a species of doxastic disagreement about the instantiation of some normative property such as *being preferable* or *being as the world ought to be* or *being good*. Unless the proponent of this conception of desire

[19] Since contrary practical commitments must be for the same agent in order to disagree, we would need a theory of how a subject can form a preference *for* at least a group of agents including herself, if not for a third-party entirely, in order to have a theory of interpersonal preference-involving disagreement. For a theory of *de nobis* desires, see (Marques and García-Carpintero 2014).

[20] It is a consequence of my view that subjects with contrary preferences for some state of affairs outside the control of any agent do not thereby disagree. Suppose A prefers that the number of stars be even and B prefers that it be odd. Unless A and B have beliefs on this subject with contradictory contents (see the next paragraph in the text for discussion), they will not disagree simply by liking different states of affairs.

[21] This sort of claim would follow an anti-Humean theory of desire.

provides us with a theory on which desire's evaluative commitment is not reducible to normative belief, then this proposal will have failed to identify a sort of disagreement distinct from ordinary doxastic disagreement.

5. Credence

The nature of credence-involving disagreement depends, of course, on what credence is. There are three families of views about the metaphysical priority relations between credences and beliefs. On one view, beliefs are reducible to credences; on a second view, credences are reducible to beliefs; and on a third view, neither credences nor beliefs are reducible to the other.[22]

If credences are reducible to beliefs, then subjects with different credences will disagree whenever the relevant beliefs disagree. For example, if credences are to be reduced to beliefs about the likelihood of a proposition, then subjects with different credences toward the same proposition will disagree about *how likely/probable is* p.

If credences are not reducible to beliefs—as on the second and third views listed above—then the question of disagreement becomes more difficult. Since on my account disagreement requires committing attitudes, we should first ask whether credences so construed count as committing attitudes.[23] If A has a credence of .6 toward *p*, to what is A committed? After a

[22] For an overview of these positions, see (E. G. Jackson 2020).

[23] Perhaps credences do not count as committing attitudes because no credence toward *p* closes inquiry about whether *p*. Here and elsewhere in this essay, the question of whether some attitude *A* can underwrite disagreement turns on the correct account of the nature of *A*. Although this question is strictly prior to the question of disagreement, our intuitions about when subjects disagree may well provide useful data for theorizing about the nature of the attitudes. Thanks to Sarah Stroud for discussion here.

certain threshold *n*—whether the value of *n* is set invariantly or is sensitive to context—credence of *n* that *p* seems to amount to a commitment that *p*, even if this commitment is distinct from the commitment embodied in belief. But even below the threshold *n* credences do seem to count as committing attitudes, because it is always incoherent for a single subject to hold different credences toward the same proposition, and this fact needs an explanation. (In this way, credence resembles preference rather than desire.)

One option is to identify credence-involving disagreement as disagreement about *what credence to have*, such that credence-involving disagreement is a species of practical disagreement that obtains in virtue of subjects' having non-jointly-realizable plans. This will then require us to offer a theory of plans about what credence to have and the nature of planning for others to have a credence. Beyond whatever difficulties might arise here, it seems anyways implausible to classify disagreement involving credences—which have mind-to-world direction of fit—as a species of practical disagreement. We must then identify the precise mind-to-world commitment involved in credence in order to identify the question that credence-involving disagreement is about.

Perhaps a credence that *p* is a like a *best guess* or *estimate* about the truth of *p*, which takes into account a subject's evidence concerning *p* but is not reducible to a belief about this evidence. This seems like the right sort of account to underwrite mind-to-world-directed disagreement with credences. I take it that on this view, credences associate a given body of evidence E for *p* with a number measuring a best guess or estimate (relative to some degree of precision) of the truth of *p*. These best guesses or estimates are mutually exclusive: relative to the same evidence and degree of precision, an estimate of *n* that *p* is incompatible because distinct from any other estimate *n** that *p*. If credences are commitments to these objects, we can

understand why single subjects with different credences toward *p* seem incoherent and why different subjects with different credences toward *p* disagree.

It is worth noting that, even if credences are best understood not as a committing attitudes or if the commitment constitutive of credence is somehow inapt to generate disagreement, we would not be compelled to say that no disagreement at all obtains between subjects with contrary credences toward the same proposition. We might instead identify some belief about *p* that, at least in a rational person, attends credence toward *p*—even if this belief does not constitute that credence—in which to locate disagreement. If there is some such belief, then rational subjects with contrary credences will disagree whenever—and when, also because—those beliefs disagree. Presumably, this belief would concern the evidence for *p*: or, phrased as a question, *to what degree does the evidence support that* p?[24] On this proposal, speakers with differing credences toward *p* disagree inasmuch as they have contrary beliefs about the degree to which the evidence supports *p*.[25]

On both the last two proposals—the proposal of credences as best guesses or estimates and the proposal of credences as rationally requiring beliefs about the evidence—the question of whether two subjects disagree will depend subtly on the question of which body of evidence is

[24] Another suggestion might be *(exactly) how credible is* p. But just as in the case of preference, it is unsatisfactory to understand credence in terms of belief about the instantiation of some normative property—in this case the property of credibility rather than preferability.

[25] Something like this proposal is also available for explaining why A and B might disagree if A believes that *p* and B suspends judgment on the truth of *p*. If they are rational, these agents will disagree about the degree of evidential support for *p*, and in particular whether there is sufficient evidence to close deliberation about whether *p*.

relevant. I think this is the correct result, since it explains why credence-involving disagreement seems slightly more elusive than classic doxastic disagreement. If A believes that *p* and B believes that ~*p*, then there is a good sense in which A and B always disagree about whether *p*—even if A and B never enter into a conversation together about whether *p* and even if A and B do not know one another.[26] But if A only has credence of .3 that *p* and B has credence of .7 that *p* (and neither has a full belief about whether *p*), it seems slightly harder to say that they disagree if A and B occupy different contexts. If A and B occupy different contexts, their beliefs about 'the evidence' might not concern the same body of evidence, and subjects will only disagree about *the* evidence when they share the same body of evidence. If A's credence relates only to *A*'s evidence or A's belief about the evidence merely concerns A's own evidence and likewise for B, then they will not disagree. If, however, the credence of each or the evidential beliefs of each pertain rather to something like *all the available evidence*, then they will thereby disagree. And if we assume that in conversation about whether *p* all conversational partners have access to one another's evidence bearing on the truth of *p*, then subjects in conversation with differing credences or evidential beliefs will be guaranteed to disagree.[27]

Having set out the account of disagreement, I now explore its implications for the problem of normative disagreement and for the question of objectivity.

---

[26] Cappelen and Hawthorne would call this state disagreement, which is roughly equivalent to disagreement in mental states.

[27] Worsnip (249–250) objects to similar proposals because they sometimes fail to predict disagreement in cases where subjects who diverge in credence possess different evidence. But since I see no motivation to call such cases—in which subjects lack attitudes with incompatible contents—instances of disagreement, I am unbothered by these objections.

6. The problem of (genuine) normative disagreement

If I believe that the atomic number of oxygen is 8 and you believe it is not, we disagree. And intuitively, if I believe that stealing is wrong and you believe it is not, we disagree. For the *cognitivist realist* about normative judgment, this intuitive judgment is easy to vindicate. According to this view, to accept a normative judgment is to have a belief about how the world is in its normative respects, and normative sentences are true when they correctly describe how the world is in its normative respects. So, as analyzed by the cognitivist realist, 'stealing is wrong' straightforwardly disagrees with 'stealing is not wrong.' But for other views about the semantics of normative language, it is more difficult to explain why contrary normative opinions count as disagreeing.

According to *contextualism*, normative language is context-sensitive: what proposition is expressed by the assertion of a normative sentence depends on what code or body of norms is salient in the context of utterance. For example, according to the contextualist, a speaker's utterance of 'stealing is wrong' is true at a context if and only if stealing is prohibited by the salient code or body of norms in the context. But if (e.g.) A's normative judgments concern the content of the norms *A* accepts and B's normative judgments concern the content of the norms *B* accepts, then there is no guarantee that their contrary normative beliefs have incompatible contents, and so it is unclear whether A and B disagree. This is an old objection, already formulated by G.E. Moore (1912, 50–51) against simple speaker subjectivism, the proto-contextualist view that '*x* is right' just means 'I desire *x*.'[28] Moore noticed that this view fails to predict disagreement where, intuitively, there is: If I assert '*x* is right' and you assert '*x* is not right' we disagree, but if I assert 'I desire *x*' and you assert 'I do not desire *x*,' we do not disagree.

[28] Stevenson (1937: 15–16) attributed a version of this view (the "interest theory") to Hobbes and Hume.

Expressivism is supposed to solve the disagreement problem for simple speaker subjectivism by distinguishing between *reporting* one's attitudes and *expressing* them. According to expressivism, to accept a normative judgment is ultimately to have some *non*-doxastic mental state, such as a desire or state of (dis)approval, and normative language expresses these non-doxastic mental states. As developed by Mark Schroeder (2008b), the view is that the sentence 'stealing is wrong' bears the same relation of expression to (say) disapproval of stealing as does the sentence 'oxygen's atomic number is 8' to the belief that oxygen's atomic number is 8. Just as speakers with contrary beliefs about oxygen's atomic number disagree only by expressing them and not by reporting them—'I believe that oxygen's atomic number is 8' does not disagree with 'I believe oxygen's atomic number is not 8'—so too in the case of normative attitudes. But this story does not amount to a positive account explaining how non-doxastic disagreement is possible: let us grant that there is some necessary condition for disagreement unmet in a discourse where speakers report their contrary mental states rather than express them, but this does not itself entail that expressions of contrary non-doxastic mental states constitute disagreement.

Relativism is supposed to solve the disagreement problem by claiming that it is truth rather than content that is relative to a set of norms. According to relativism, normative sentences express a fixed content, which is true or false relative to some code or body of norms. Some authors use relativism to describe any view on which the truth and falsity of the target sentences is relative to something like a standard; John MacFarlane (2014), however, reserves the label for the view that the relevant standard is determined by the context of assessment.[29] So, when A

[29] MacFarlane calls the other sort of view, on which the relevant standard is determined by the context of utterance, *nonindexical contextualism*. (See MacFarlane 2014, 88–89 and especially n24.)

asserts ‘stealing is wrong’ and B asserts ‘stealing is not wrong,’ A affirms the same content that B denies, viz. *that stealing is wrong*. But since according to the relativist this content might be *true* relative to the norms A accepts and false relative to the norms B accepts, A and B may each speak truly relative to the standard that each accepts, and this marks a clear disanalogy with the case of oxygen’s atomic number.[30] We might therefore wonder whether the relativist’s sort of disagreement is sufficiently robust to capture the nature of normative disagreement.

It seems clear, then, that to solve the problem of normative disagreement, we need a constitutive account of disagreement, as Dreier (2009) has urged.

Before we address solutions, we should first distinguish different versions of the problem. Schroeder (2008b, 90–91) distinguishes between the *shallow* problem of disagreement and the *hard* problem of disagreement. The shallow problem is to explain why speakers with contrary normative views disagree at all. The hard problem is to explain why a speaker may reject a contrary normative view by saying, “that’s false.” According to Schroeder, the shallow problem, but not the hard problem, may be solved with a ‘disagreement in attitude’ account.

We might, however, be left unsatisfied even with a solution to Schroeder’s hard problem. If, for example, a speaker may felicitously respond “that’s false” to any utterance she rejects,[31] then whatever account of disagreement solves the shallow version of the problem will thereby

[30] The relativist’s disagreement is often called *faultless* for this reason. On MacFarlane’s version of relativism as assessment-sensitivity, although A and B might each speak truly relative to the standard each accepts, there is no single context of assessment from which A’s and B’s contrary views count as speaking truly, and so no one can assess both A and B as speaking truly. (MacFarlane (133–136), accordingly, criticizes the label of faultless disagreement.)

[31] See (Jorgensen Bolinger 2022, 22 n33).

solve the hard problem. We might be left unsatisfied because the point of the hard problem, it seemed, was to note that we may distinguish between kinds of disagreement: a shallow or superficial sort of disagreement on the one hand, and a deep or genuine sort of disagreement on the other. And since normative disagreements—just like disagreements about the atomic number of oxygen—seem genuine rather than superficial, any metanormative view that denies this is false, or at least revisionary.

To be fully satisfied, then, we need a solution to the problem of *genuine* normative disagreement. We will therefore need an account of what genuine disagreement consists in. A paradigm instance of superficial disagreement is one that is *merely verbal*. One way to describe this phenomenon is as follows: in cases of merely verbal disagreement about *S*, interlocutors' mental states agree about *S*, but this underlying agreement in mental states is rendered opaque to the interlocutors on account of the language they use to articulate their beliefs about *S*. This characterization suggests that linguistic disagreement is superficial when not backed in the right way by mental-state disagreement. If so, then for a metanormative view to vindicate the genuine nature of normative disagreement, it must not only provide a semantic account that predicts linguistic disagreement in the correct cases—unlike simple speaker subjectivism, which fails to meet even this desideratum—but must also provide an explanation of how this linguistic disagreement is backed in the right way by mental-state disagreement. I take it, then, that the insufficiency of solving Schroeder's hard problem for establishing that normative disagreement is genuine is at least partly explained by the fact that Schroeder's hard problem may be solved without having to say anything about mental-state disagreement.

When does linguistic disagreement count as backed in the right way by mental state disagreement for it to count as genuine? A strict condition might invoke the expression relation

as defined by Schroeder (2008b). On this account, a sentence *s* expresses mental state *m* when having *m* is a sufficient condition for the assertability, in the narrow sense of "semantic correctness," of an utterance of *s*. (Schroeder argues that this "assertability expressivism" is the sort that should be endorsed by the metaethical expressivist.) On this criterion, linguistic disagreement expresses mental-state disagreement if and only if the disagreeing sentences respectively express disagreeing mental states. But perhaps this condition is too strict, and we should count linguistic disagreement as genuine so long as it is semantically explained in some other way by mental-state disagreement. In what follows I will assume this looser criterion, though I will not try to make it any more precise.

7. Evaluating solutions to the problem of genuine normative disagreement

Having developed a constitutive theory of disagreement and explicated the distinction between superficial and genuine disagreement, I now briefly discuss the prospects of expressivist, contextualist, and relativist solutions to the problem of genuine normative disagreement.

7.1. Expressivism

Because the expressivist analyzes normative sentences as expressions of non-doxastic attitudes, the expressivist is well-positioned to vindicate genuine disagreement: all the expressivist needs to do is to provide an account of the mental states expressed by contrary normative sentences on which these mental states disagree. Many expressivist proposals, however, have not succeeded in identifying the correct non-doxastic attitudes. The classic disagreement in attitude account will not do, because disagreement is never merely *in* but is

rather always *about*. Nor will it do to take disagreement as primitive, as does Allan Gibbard in his defense of plan-expressivism (2003, chap. 4), since this account predicts disagreement even in cases where the plans in question are jointly-realizable, as when I plan for myself to φ and you plan for yourself not to φ. But if we modify Gibbard's plan-expressivist account suitably, then we can predict practical disagreement just when subjects' plans are non-jointly-realizable: this is precisely what Ayars's (2022) account does. So there is at least one expressivist account that solves the problem of genuine normative disagreement.[32]

7.2. Contextualism

Contextualists have offered various proposals for where to locate normative disagreement. On the family of proposals Stephen Finlay (2017, 191) calls the "*quasi-expressivist* account," normative disagreement is to be found in speakers' contrary non-doxastic attitudes, even if these attitudes are not semantically expressed by normative language.[33] On the more restrictive definition of genuine disagreement, it is not clear that such a strategy would yield genuine disagreement. And even on the less restrictive definition, disagreement in mental states cannot simply 'float free' from linguistic disagreement: there must be some semantic connection between the two in order for linguistic disagreement to count as backed in the right way by

[32] It is worth noting that the analysandum of Gibbard's and Ayars's theories is normative judgment broadly-construed, not moral judgments narrowly construed. Within the plan-expressivist framework, we may think of the latter as plans for everyone about a question like *what to blame for*, following Gibbard's basic idea in his (1990).

[33] The other family of proposals he calls the "metalinguistic account." (As will become clear, on my view the best versions of the metalinguistic account are in fact forms of the quasi-expressivist account.)

mental-state disagreement. Although older subjectivist accounts likely cannot meet even this looser standard,[34] more recent contextualist accounts likely can.

Representative of recent contextualist theories of normative language are those developed by Alex Silk (2016, chap. 5; 2017) and by Justin Khoo and Joshua Knobe (2018). Each formal framework includes within the conversational context a parameter for a salient body of norms, and each explains the phenomenon of discourse disagreement as resulting from interlocutors' attempts, with their normative utterances, to shift the value of this parameter in incompatible ways.[35] We can then note: if normative disagreement is about *what body of norms to accept*, then it turns out that the contextualist, like the expressivist, requires an account of practical disagreement to make sense even of linguistic normative disagreement.[36]

---

[34] Frank Jackson (2008) argued that since the subjectivist can simply adopt the expressivist's account of disagreement in attitude, considerations stemming from disagreement do not support expressivism over subjectivism. Sebastian Köhler (2012) responded to Jackson that the resulting subjectivist account would still not provide a view on which moral judgements themselves disagree.

[35] On Silk's proposal, disagreement obtains when speakers' utterances presuppose incompatible values for this parameter (2016, 124–32; 2017, 223–25). On Khoo and Knobe's proposal, disagreement obtains when speakers' assertions offer proposals to update the norm-parameter value that are each compatible with the value as it stands, but incompatible with one another.

[36] Balcerak Jackson (2019, 17–22) argues similarly concerning how contextualists about the semantics of predicates of personal taste such as 'beautiful' should understand disagreement. He argues that disagreement about (e.g.) what is beautiful in the case where interlocutors do not share the same standard of beauty is practical disagreement about what to *count* as beautiful and thus what standard of beauty to accept, and he argues that this explains the sense in which these disagreements are faultless.

Is this linguistic disagreement genuine? Finlay complains—using his label for what I have called the problem of genuine normative disagreement—that if the contextualist views of Silk and of Khoo and Knobe do not provide a theory of disagreement independent of a conversational context, then they "do not really address the traditional metaethical Problem of Lost Disagreement" (2017, 196). Now it is true that these theories are concerned in the first instance with discourse disagreement. But this is because contextualism is a theory of normative language, not of normative thought, and it is open to such views to adopt an account of norm-acceptance on which normative discourse disagreement is backed in the right way by mental-state disagreement to count as genuine.[37] So, by borrowing the expressivist's account of practical disagreement, the contextualist likely has the resources to vindicate the genuine nature of normative disagreement, as well.

[37] Khoo and Knobe themselves might not, however, be happy with the package I have offered them. They motivate their account over the view they call "factualist invariantism" as well as expressivism by noting that only their account can predict the possibility of "non-exclusionary moral disagreements" (110), cases in which it is *incorrect* to call the opinion with which one disagrees false. (Khoo and Knobe take non-exclusionary moral disagreement to be possible on the basis of experimental results.) This is not, strictly speaking, an argument against their semantic proposal, but rather a note about how to understand the motivation for their proposal and its relation to other accounts of moral semantics. (Khoo and Knobe themselves believe that expressivism is committed to exclusionary moral disagreement owing to the view's acceptance of minimalism about truth (129). But suitable versions of expressivism like that defended by Ayars possess a deeper explanation of the nature of normative disagreement.)

7.3. Relativism

Let us sharpen the worry that the relativist's sort of disagreement is insufficiently robust to capture the nature of normative disagreement. On my account, subjects disagree when they accept incompatible answers to some question, and to accept an answer (in the case of mental state disagreement) is to have a committing attitude with a content. The idea behind the view that subjects disagree in virtue of the committing attitudes they possess is that disagreement is ultimately to be explained in virtue of subject's commitments; and in normal cases, what a subject is committed to in having some attitude A with content c is determined wholly by c and the nature of A. In special cases, however, we require additional information to determine a subject's commitments. These are cases where the contents of attitudes are *incomplete*—that is, where (e.g.) what is believed is only true or false relative to the value of some further parameter, such as a possible world or a time. The relativist holds that beliefs in the target domain are of this kind, which is what allows their contents to be true or false only relative to some further parameter (MacFarlane 2014, 84–88). In such cases, what a subject is committed to in believing that p—in the sense of commitment relevant for disagreement on my account—is the *complete* content, which includes the appropriate values of those parameters relative to which the truth of the belief must be determined. Here is why: Belief that p is apt to underwrite disagreement, on my account, because belief that p is a commitment to the truth of p. But in cases where the truth of p is relative to some further parameter t, then we must make explicit the intended value of t in order to accurately represent what it is a subject is committed to in believing that p.

What determines which value of these parameters is relevant to determining a subject's commitments? Let us first consider the simple cases of relativity to possible worlds and times. Suppose A, in possible world w_A, believes that the most abundant chemical element is hydrogen;

and B, in possible world w_B, believes that the most abundant chemical element is carbon. Even if we do not want to say that A's and B's beliefs include as part of their contents anything about their respective worlds, it seems clear that their commitments pertain only to their respective worlds. And this explains why it seems odd to say that A and B disagree about *which chemical element is the most abundant*. Similarly, if A at t_A believes that Socrates is sitting and B at t_B believes Socrates is not sitting, there is little inclination to say that A and B disagree about *whether Socrates is sitting*. And this verdict about A's and B's commitments, and therefore about whether they disagree, depends only on the possible worlds and times occupied by A and B.

Given this understanding of commitment, we can see that if truth in normative discourse is relative to some parameter like a normative theory, then the relevant value of this parameter for determining a subject's commitments will always be the value accepted by the subject, not the value accepted by some third party like an assessor. Therefore, even if the (incomplete) contents of subjects' beliefs appear to be incompatible, and even if their complete contents would be incompatible if evaluated for truth relative to the same value for the relevant parameter, this does not suffice to secure disagreement on my account. It is certainly interesting and important that, if relativism is correct, our language licenses speakers to voice their contrary incomplete beliefs using expressions of disagreement. But this fact about natural language is simply not sufficient to establish genuine disagreement.[38]

[38] Relativists, of course, define disagreement in a way on which such cases do amount to disagreement. MacFarlane argues that assessment-sensitive domains exhibit a form of disagreement he calls *preclusion of joint accuracy*:

7.4. Genuine disagreement and objectivity

Disagreement seems to bear a conceptual connection to *objectivity*. In my preferred formulation, a domain *D* is objective just in case it admits of genuine disagreement: just in case contrary *D*-opinions count as genuinely disagreeing.[39] The fact that *D*-discourse exhibits surface

---

> **Preclusion of joint accuracy**. *The accuracy of my attitudes (as assessed from any context) precludes the accuracy of your attitude or speech act (as assessed from that same context).* (2014, 129)

"Licorice is tasty" disagrees with "licorice is not tasty" on this view because there is no context of assessment from which both sentences are accurate. Notice that this conception of disagreement makes the question of whether two subjects disagree itself assessment-sensitive, because accuracy is an assessment-sensitive matter.

[39] Claims in this neighborhood have been endorsed by Crispin Wright and by Michael Smith. According to Wright (1992), a discourse *D* is apt for *substantial truth* only if it satisfies the *Cognitive Command* constraint, roughly that disagreement about *D* not due to vagueness or lack of relevant evidence "will involve something which may properly be regarded as a cognitive shortcoming" (144). (Discourse about the comic, for example, does not satisfy this constraint, since comic disagreement may be explained by differences in senses of humor.) And according to Smith (1994), what it is for morality to be objective is for '*x* is right' to disagree with '*x* is not right,' for there to be one correct answer to the question whether *x* is right (see 1994, 6, 33, 86 for instances of this assumption). A similar idea is in play in David Enoch's argument for the objectivity of morals (2011, chap. 2). The argument is that in situations of interpersonal conflict stemming from non-cognitive disagreement, it is morally impermissible to 'stand one's ground,' to refuse to prescind from the preferences one actually has and view them as being on a par with the preferences of other interested parties. In otherwise analogous cases stemming from cognitive disagreement, by contrast, standing one's ground is morally permissible. Since standing one's ground is

objectivity and thus admits of linguistic disagreement does not suffice to establish real objectivity—this requires linguistic disagreement to be backed by mental state disagreement.[40]

A paradigm example of a domain in which linguistic disagreement is possible but which seems non-objective in the deep, philosophical sense is that of *taste*. As the old saying goes: *de gustibus non disputandum est* (matters of taste are not to be disputed). The claim is not that we *cannot* disagree by voicing our contrary tastes—after all, "*x* is tasty" does disagree with "*x* is not tasty"—but rather that we *ought not* dispute such matters given their non-objective nature. My account delivers the result that disagreement voiced by simple attributions of *predicates of personal taste* such as 'tasty,' 'funny,' and 'disgusting' is superficial rather than genuine. Contrary tastes do not disagree on my account, because to find something tasty is not to possess a committing attitude with a content, but is rather to be disposed to having a certain kind of experience, namely that of reacting to the object found tasty with gustatory pleasure. Likewise for 'funny' and 'disgusting': to find something funny is to be disposed to react to it with amusement, and to find something disgusting is to be disposed to react to it with disgust. So linguistic disagreement about what is tasty or funny or disgusting is not backed in the right way by mental-state disagreement to count as genuine.[41]

---

permissible in the case of moral disagreement, Enoch argues, moral disagreement is cognitive, and this captures a clear sense in which morality is objective (40–41).

[40] Early emotivists such as A.J. Ayer (1946) and Stevenson (1937) and error theorists such as J.L. Mackie (1977) did not deny the surface objectivity of moral language; they claimed rather that this surface objectivity belies the nonexistence of any moral reality to which it might correspond.

[41] But aren't there beliefs about what is tasty, which sentences of the form '*x* is tasty' express? I claim that a subject's beliefs about what is tasty are in a sense fully reducible to the subject's relevant experiences,

or at least to beliefs about these experiences. First, it is unclear that A's belief that *x* is tasty is distinct from A's belief that *x* is tasty to A, since beliefs about what is tasty might be essentially *de se* (Beddor 2019, 821–22). (Beddor argues, partially on this basis, that attributions of disagreement are licensed when, and only when, subjects' beliefs are reported such that their complements are incompatible. And this condition has an irreducible linguistic component, because "our intuitions about whether two beliefs disagree are sensitive to syntactic features of the sentences used to report them" (829). For example, A's utterance of 'licorice is tasty' reports the same belief as A's utterance of 'licorice is tasty to A,' but A's utterance of 'licorice is tasty' disagrees only with B's utterance of 'licorice is not tasty' and not with B's utterance of 'licorice is not tasty to B.') And although MacFarlane explicitly denies that A's belief that *x* is tasty has the same content as A's belief that *x* is tasty to A, matters are not so simple even by MacFarlane's own lights. For MacFarlane generally, what distinguishes assessment-sensitive relativism from a formally isomorphic expressivism about what is tasty are their views in the philosophy of mind: for the expressivist, beliefs in the target discourse are the same as the underlying non-doxastic states, whereas for the relativist these are distinct (2014, 172–74). But even granting MacFarlane's claim about how he and Gibbard differently view the relation between normative beliefs and intentions, there might be room to resist the non-identity claim about 'tasty'-beliefs owing to special features of predicates of personal taste. Even for the relativist, '*x* is (not) tasty' is assertable by A if and only if A has *first-hand knowledge* of *x*'s (non-)pleasing-to-A flavor, and similarly for other predicates of personal taste. (This is MacFarlane's assertion rule for 'tasty' **TP** (2014, 4).) If the first-hand-knowledge requirement is not merely an additional necessary condition for the assertability of tasty-predications but rather a reflection of some feature of the underlying beliefs they express, this would indicate that beliefs about what is tasty are closely linked to a subject's experiences in way that beliefs about (e.g.) what someone ought to do are not.

Given the link between genuine disagreement and objectivity, any metanormative account on which normative disagreement is genuine provides an account that vindicates the objectivity of moral language and thought. But the way in which the normative is objective need not be the way that chemistry is objective. Chemistry is objective because disagreeing chemical beliefs cannot be true together. But if normative disagreement is analyzed as a species of practical disagreement, then the normative is objective because disagreeing plans cannot be carried out together.[42] So the expressivist has, the contextualist might have, and the relativist probably lacks, an account on which the normative is objective.[43]

[42] Justin Clarke-Doane (2020, chap. 6) has recently described a form of objectivity without realism, the objectivity of "we can only do one thing" (164).

[43] **Acknowledgement:** Versions of this chapter were presented at the Exploring Expressivism in Ethics, Language, and Metaphysics Conference at York University in April 2023 and at the 6th Annual Chapel Hill Normativity Workshop at the University of North Carolina at Chapel Hill in March 2024. I thank the conference participants for their questions and suggestions.

References

Anscombe, G. E. M. 2000. *Intention*. 2nd ed. Cambridge, Mass: Harvard University Press.

Archer, Sophie. 2022. "Salience and What Matters." In *Salience*, edited by Sophie Archer, 113–29. London: Routledge. https://doi.org/10.4324/9781351202114-7.

Arpaly, Nomy. 2003. *Unprincipled Virtue: An Inquiry into Moral Agency*. New York: Oxford University Press.

Arpaly, Nomy, and Timothy Schroeder. 2013. *In Praise of Desire*. New York: Oxford University Press.

Atiq, Emad H. 2016. "How to Be Impartial as a Subjectivist." *Philosophical Studies* 173 (3): 757–79. https://doi.org/10.1007/s11098-015-0518-x.

Ayars, Alisabeth. 2022. "Deciding for Others: An Expressivist Theory of Normative Judgment." *Philosophy and Phenomenological Research* 105: 42–61. https://doi.org/10.1111/phpr.12800.

Ayer, A.J. 1946. *Language, Truth and Logic*. New York: Dover Publications.

Bader, Ralf. 2016. "Conditions, Modifiers, and Holism." In *Weighing Reasons*, edited by Errol Lord and Barry Maguire, 27–55. New York: Oxford University Press.

———. 2017. "The Grounding Argument against Non-Reductive Moral Realism." In *Oxford Studies in Metaethics 12*, edited by Russ Shafer-Landau, 106–34. New York: Oxford University Press.

———. ms. "Conditional Grounding."

Balcerak Jackson, Brendan. 2019. "Essentially Practical Questions." *Analytic Philosophy* 60 (1): 1–26.

Beddor, Bob. 2019. “Subjective Disagreement.” *Noûs* 53 (4): 819–51. https://doi.org/10.1111/nous.12240.

Behrens, Singa. 2021. “No Normative Free Lunch: Relevance and the Autonomy of the Normative Domain.” *Synthese* 199 (5–6): 13163–86. https://doi.org/10.1007/s11229-021-03371-6.

Berker, Selim. 2020. “Quasi-Dependence.” In *Oxford Studies in Metaethics 15*, edited by Russ Shafer-Landau, 195–218. New York: Oxford University Press.

Blackburn, Simon. 1984. *Spreading the Word*. New York: Oxford University Press.

———. 1993a. “Errors and the Phenomenology of Value.” In *Essays in Quasi-Realism*, 149–65. Oxford University Press.

———. 1993b. “How to Be an Ethical Anti-Realist.” In *Essays in Quasi-Realism*, 166–81. New York: Oxford University Press.

———. 1996. “Response to Dworkin.” *Symposium at Brown University on “Objectivity and Truth: You’d Better Believe It”* (blog). 1996. https://www.brown.edu/Departments/Philosophy/bears/9611blac.html.

———. 1998. *Ruling Passions: A Theory of Practical Reasoning*. Reprinted. New York: Clarendon Press of Oxford University Press.

———. 1999. “Is Objective Moral Justification Possible on a Quasi-Realist Foundation?” *Inquiry* 42 (2): 213–27. https://doi.org/10.1080/002017499321552.

Blanshard, Brand. 1949. “The New Subjectivism in Ethics.” *Philosophy and Phenomenological Research* 9 (3): 504–11.

Cappelen, Herman, and John Hawthorne. 2009. *Relativism and Monadic Truth*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199560554.001.0001.

Cassam, Quassim. 1986. “Necessity and Externality.” *Mind* 95 (380): 446–64.

Clarke-Doane, Justin. 2020. *Morality and Mathematics*. New York: Oxford University Press.

Cross, Charles, and Floris Roelofsen. 2022. “Questions.” In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Summer 2022. https://plato.stanford.edu/archives/sum2022/entries/questions/.

Dancy, Jonathan. 2004. *Ethics without Principles*. New York: Oxford University Press.

Darwall, Stephen. 2022. “A Gibbardian Account of (Narrow) Moral Concepts.” In *Meaning, Decision, and Norms: Themes from the Work of Allan Gibbard*, edited by Billy Dunaway and David Plunkett, 109–24. Ann Arbor: Michigan Publishing.

Dreier, James. 1993. “Structures of Normative Theories.” *The Monist* 76 (1): 22–40.

———. 2002. “Meta-Ethics and Normative Commitment.” *Philosophical Issues* 12 (1): 241–63.

———. 2004a. “Lockean and Logical Truth Conditions.” *Analysis* 64 (1): 84–91.

———. 2004b. “Meta-Ethics and the Problem of Creeping Minimalism.” *Philosophical Perspectives*, no. 18: 23–44.

———. 2009. “Relativism (and Expressivism) and the Problem of Disagreement.” *Philosophical Perspectives* 23: 97–110.

———. 2011. “In Defense of Consequentializing.” In *Oxford Studies in Normative Ethics, Volume 1*, edited by Mark Timmons, 97–119. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199693269.001.0001.

———. 2015. “Explaining the Quasi-Real.” In *Oxford Studies in Metaethics, Volume 10*, edited by Russ Shafer-Landau. New York: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780198738695.001.0001.

Dworkin, Ronald. 1996. “Objectivity and Truth: You’d Better Believe It.” *Philosophy & Public Affairs* 25 (2): 87–139.

———. 2011. *Justice for Hedgehogs*. Cambridge, MA: Belknap Press of Harvard University Press.

Enoch, David. 2011. *Taking Morality Seriously*. New York: Oxford University Press.

———. 2019. “How Principles Ground.” In *Oxford Studies in Metaethics 14*, edited by Russ Shafer-Landau, 1–22. New York: Oxford University Press.

Fantl, Jeremy. 2006. “Is Metaethics Morally Neutral?” *Pacific Philosophical Quarterly* 87 (1): 24–44.

Fine, Kit. 2001. “The Question of Realism.” *Philosophers’ Imprint* 1 (1): 1–30.

———. 2012. “Guide to Ground.” In *Metaphysical Grounding: Understanding the Structure of Reality*, edited by Fabrice Correia and Benjamin Schnieder, 37–80. New York: Cambridge University Press.

———. 2020. “Yablo on Subject-Matter.” *Philosophical Studies* 177 (1): 129–71. https://doi.org/10.1007/s11098-018-1183-7.

Finlay, Stephen. 2017. “Disagreement Lost and Found.” In *Oxford Studies in Metaethics 12*, edited by Russ Shafer-Landau, 187–205. New York: Oxford University Press.

Firth, Roderick. 1952. “Ethical Absolutism and the Ideal Observer.” *Philosophy and Phenomenological Research* 12 (3): 317. https://doi.org/10.2307/2103988.

Foot, Philippa. 1983. “Utilitarianism and the Virtues.” *Proceedings and Addresses of the American Philosophical Association* 57 (2): 273. https://doi.org/10.2307/3131701.

———. 1985. “Utilitarianism and the Virtues.” *Mind* 94: 196–209. https://doi.org/10.2307/3131701.

Francén, Ragnar. 2023. “Finding Wrong.” *Mind* 132 (526): 493–504. https://doi.org/10.1093/mind/fzac041.

Franzén, Nils. 2020. “Evaluative Discourse and Affective States of Mind.” *Mind* 129 (516): 1095–1126. https://doi.org/10.1093/mind/fzz088.

Frugé, Christopher. 2022. “Structuring Wellbeing.” *Philosophy and Phenomenological Research* 105 (3): 564–80. https://doi.org/10.1111/phpr.12830.

Gallow, J. Dmitri. 2022. “The Metaphysics of Causation.” In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta and Uri Nodelman, Fall 2022. https://plato.stanford.edu/archives/fall2022/entries/causation-metaphysics/.

Gibbard, Allan. 1990. *Wise Choices, Apt Feelings: A Theory of Normative Judgment*. Cambridge, MA: Harvard University Press.

———. 2003. *Thinking How to Live*. Cambridge, MA: Harvard University Press.

———. 2012. *Meaning and Normativity*. 1st ed. New York: Oxford University Press.

Grindrod, Jumbly, and Emma Borg. 2019. “Questions Under Discussion and the Semantics/Pragmatics Divide.” *Philosophical Quarterly* 69 (275): 418–26. https://doi.org/10.1093/pq/pqy058.

Hamblin, C.L. 1958. “Questions.” *Australasian Journal of Philosophy* 36 (3): 159–68. https://doi.org/10.1080/00048405885200211.

Hare, R.M. 1952. *The Language of Morals*. New York: Oxford University Press.

Hawke, Peter. 2018. “Theories of Aboutness.” *Australasian Journal of Philosophy* 96 (4): 697–723. https://doi.org/10.1080/00048402.2017.1388826.

Hawthorne, John. 2002. “Practical Realism?” *Philosophy and Phenomenological Research* 64 (1): 169–78.

Herman, Barbara. 1981. “On the Value of Acting from the Motive of Duty.” *Philosophical Review* 90: 359–82.

Hieronymi, Pamela. 2005. “The Wrong Kind of Reason.” *Journal of Philosophy* 102 (9): 437–57.

Hoek, Daniel. 2018. “Conversational Exculpature.” *Philosophical Review* 127 (2): 151–96.

Howard, Nathan. 2021. “Moral Fetishism and a Third Desire for What’s Right.” *Journal of Ethics and Social Philosophy* 20 (3). https://doi.org/10.26556/jesp.v20i3.1249.

Humberstone, Lloyd. 1996. “A Study in Philosophical Taxonomy.” *Philosophical Studies* 83 (2): 121–69.

———. 2019. “Recent Thought on Is and Ought: Connections, Confluences and Rediscoveries.” *Journal of Applied Logics — IfCoLog Journal of Logics and Their Applications* 6 (7): 1373–1446.

Jackson, Elizabeth G. 2020. “The Relationship between Belief and Credence.” *Philosophy Compass* 15 (6): e12668. https://doi.org/10.1111/phc3.12668.

Jackson, Frank. 2008. “The Argument from the Persistence of Moral Disagreement.” In *Oxford Studies in Metaethics*, edited by Russ Shafer-Landau, 3:75–86. New York: Oxford University Press.

Jackson, Frank, and Philip Pettit. 1998. “A Problem for Expressivism.” *Analysis* 58 (4): 239–51.

———. 2003. “Locke, Expressivism, Conditionals.” *Analysis* 63 (1): 86–92.

Johnson King, Zoë A. 2020a. “Accidentally Doing the Right Thing.” *Philosophy and Phenomenological Research* 100 (1): 186–206. https://doi.org/10.1111/phpr.12535.

———. 2020b. “Praiseworthy Motivations.” *Noûs* 54 (2): 408–30. https://doi.org/10.1111/nous.12276.

———. 2022. “Deliberation and Moral Motivation.” In *Oxford Studies in Metaethics, Volume 17*, edited by Russ Shafer-Landau, 254–80. Oxford University Press. https://doi.org/10.1093/oso/9780192865601.003.0010.

Johnston, Mark. 2001. “The Authority of Affect.” *Philosophy and Phenomenological Research* 63: 181–214.

Jorgensen Bolinger, Renée. 2022. “Metalinguistic Negotiations in Moral Disagreement.” *Inquiry* 65 (3): 352–80. https://doi.org/10.1080/0020174X.2020.1850336.

Joyce, Richard. 2001. *The Myth of Morality*. New York: Cambridge University Press.

———. 2005. “Moral Fictionalism.” In *Fictionalism in Metaphysics*, edited by Mark Eli Kalderon, 287–313. New York: Oxford University Press.

———. 2021. “Moral Anti-Realism.” In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta and Uri Nodelman, (Winter 2022 Edition). https://plato.stanford.edu/archives/win2022/entries/moral-anti-realism/.

Kalderon, Mark Eli. 2005. *Moral Fictionalism*. New York: Oxford University Press.

Khoo, Justin, and Joshua Knobe. 2018. “Moral Disagreement and Moral Semantics.” *Noûs* 52 (1): 109–43. https://doi.org/10.1111/nous.12151.

Koev, Todor. 2018. “Notions of At-issueness.” *Language and Linguistics Compass* 12 (12). https://doi.org/10.1111/lnc3.12306.

Köhler, Sebastian. 2012. “Expressivism, Subjectivism and Moral Disagreement.” *Thought* 1: 71–78.

———. 2014. “Expressivism and Mind-Dependence: Distinct Existences.” *Journal of Moral Philosophy* 11 (6): 750–64. https://doi.org/10.1163/17455243-4681039.

Korsgaard, Christine M. 1996. “The Normative Question.” In *The Sources of Normativity*, edited by Onora O’Neill, 7–48. New York: Cambridge University Press.

Lasersohn, Peter. 2005. “Context Dependence, Disagreement, and Predicates of Personal Taste.” *Linguistics and Philosophy* 28 (6): 643–86. https://doi.org/10.1007/s10988-005-0596-x.

Lewis, David. 1973. “Causation.” *Journal of Philosophy* 70: 556–67.

———. 1988a. “Relevant Implication.” *Theoria* 54 (3): 161–74.

———. 1988b. “Statements Partly about Observation.” *Philosophical Papers* 17 (1): 1–31.

———. 1989. “Dispositional Theories of Value.” *Proceedings of the Aristotelian Society* Supplementary Volume 63: 113–37.

Louise, Jennie. 2004. “Relativity of Value and the Consequentialist Umbrella.” *Philosophical Quarterly* 54: 518–36.

MacFarlane, John. 2014. *Assessment Sensitivity: Relative Truth and Its Applications*. First edition. Context and Content. New York: Oxford University Press.

Mackie, J.L. 1977. *Ethics: Inventing Right and Wrong*. New York: Penguin Books.

Maguire, Barry. 2015. “Grounding the Autonomy of Ethics.” In *Oxford Studies in Metaethics 10*, edited by Russ Shafer-Landau, 188–215. New York: Oxford University Press.

Markovits, Julia. 2010. “Acting for the Right Reasons.” *Philosophical Review* 119 (2): 201–42.

Marques, Teresa, and Manuel García-Carpintero. 2014. “Disagreement about Taste: Commonality Presuppositions and Coordination.” *Australasian Journal of Philosophy* 92 (4): 701–23. https://doi.org/10.1080/00048402.2014.922592.

Milo, Ronald. 1995. “Contractarian Constructivism.” *Journal of Philosophy* 92 (4): 181–204.

Moore, G. E. 1912. *Ethics*. Edited by William H. Shaw. Oxford University Press. https://doi.org/10.1093/0199272018.001.0001.

Moore, G.E. 1903. *Principia Ethica*. Mineola, NY: Dover Publications.

Morton, Justin. 2022. "Normative Principles and the Nature of Mind-Dependence." *Philosophical Studies* 179 (4): 1153–76. https://doi.org/10.1007/s11098-021-01692-3.

Ninan, Dilip. 2014. "Taste Predicates and the Acquaintance Inference." *Proceedings of SALT* 24: 290–309.

———. forthcoming. "An Expressivist Theory of Taste Predicates." *Philosophers' Imprint*.

Nolan, Daniel, Greg Restall, and Caroline West. 2005. "Moral Fictionalism Versus the Rest." *Australasian Journal of Philosophy* 83 (3): 307–30.

Nozick, Robert. 1974. *Anarchy, State, and Utopia*. New York: Basic Books.

O'Leary-Hawthorne, John, and Huw Price. 1996. "How to Stand Up for Non-Cognitivists." *Australasian Journal of Philosophy* 74 (2): 275–93.

Pearson, Hazel. 2013. "A Judge-Free Semantics for Predicates of Personal Taste." *Journal of Semantics* 30 (1): 103–54. https://doi.org/10.1093/jos/ffs001.

Perl, Caleb. 2022. "Formulating Moral Error Theory." *Journal of Philosophy* 119 (5): 279–88. https://doi.org/10.5840/jphil2022119518.

Perl, Caleb, and Mark Schroeder. 2019. "Attributing Error Without Taking a Stand." *Philosophical Studies* 176 (6): 1453–71.

Pettit, Philip, and Michael Smith. 1990. "Backgrounding Desire." *Philosophical Review* 99 (4): 565–92.

Portmore, Douglas W. 2007. "Consequentializing Moral Theories." *Pacific Philosophical Quarterly* 88: 39–73.

Prior, Arthur. 1960. "The Autonomy of Ethics." *Australasian Journal of Philosophy* 38 (3): 199–206.

Railton, Peter. 2022. “Expressivism and Objectivity.” In *Meaning, Decision, and Norms: Themes from the Work of Allan Gibbard*, edited by Billy Dunaway and David Plunkett, 341–60. Ann Arbor: Michigan Publishing.

Raven, Michael J. 2015. “Ground.” *Philosophy Compass* 10 (5): 322–33. https://doi.org/10.1111/phc3.12220.

Ridge, Michael. 2013. “Disagreement.” *Philosophy and Phenomenological Research* 86 (1): 41–63.

Roberts, Craige. 2012. “Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics.” *Semantics and Pragmatics* 5. https://doi.org/10.3765/sp.5.6.

Rosen, Gideon. 1994. “Objectivity and Modern Idealism: What Is The Question.” In *Philosophy in Mind: The Place of Philosophy in the Study of Mind*, edited by Michaelis Michael and John O’Leary-Hawthorne, 277–319. Boston: Kluwer Academic Publishers.

———. 2010. “Metaphysical Dependence: Grounding and Reduction.” In *Modality: Metaphysics, Logic and Epistemology*, edited by Bob Hale and Aviv Hoffman, 109–35. New York: Oxford University Press.

———. 2017a. “Ground by Law.” *Philosophical Issues* 27 (1): 279–301.

———. 2017b. “Scanlon’s Modal Metaphysics.” *Canadian Journal of Philosophy* 47 (6): 856–76.

———. 2017c. “What Is a Moral Law?” In *Oxford Studies in Metaethics 12*, edited by Russ Shafer-Landau, 135–59. New York: Oxford University Press.

Salinger, Elliot. 2021. “Defending the Kratzerian Presuppositional Error Theory.” *Analysis* 81 (4): 701–9. https://doi.org/10.1093/analys/anab052.

———. 2024. “Expressivism and Moral Independence.” *Philosophy and Phenomenological Research* 108 (1): 136–52. https://doi.org/10.1111/phpr.12954.

Sayre-McCord, Geoffrey. 1988. “The Many Moral Realisms.” In *Essays on Moral Realism*, 1–23. Ithaca, NY: Cornell University Press.

Schaffer, Jonathan. 2017a. “The Ground Between the Gaps.” *Philosophers’ Imprint* 17 (11): 1–26.

———. 2017b. “Laws for Metaphysical Explanation.” *Philosophical Issues* 27 (1): 302–21. https://doi.org/10.1111/phis.12111.

Scheffler, Samuel. 1994. *The Rejection of Consequentialism: A Philosophical Investigation of the Considerations Underlying Rival Moral Conceptions*. New York: Oxford University Press. https://doi.org/10.1093/0198235119.001.0001.

Schroeder, Mark. 2007. *Slaves of the Passions*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199299508.001.0001.

———. 2008a. *Being For: Evaluating the Semantic Program of Expressivism*. New York: Oxford University Press.

———. 2008b. “Expression for Expressivists.” *Philosophy and Phenomenological Research* 76: 86–116.

———. 2011. “Ought, Agents, and Actions.” *Philosophical Review* 120 (1): 1–41.

———. 2014. “Does Expressivism Have Subjectivist Consequences?” *Philosophical Perspectives* 28: 278–90.

———. 2021a. “The Fundamental Reason for Reasons Fundamentalism.” *Philosophical Studies* 178 (10): 3107–27. https://doi.org/10.1007/s11098-020-01572-2.

———. 2021b. *Reasons First*. New York: Oxford University Press. https://doi.org/10.1093/oso/9780198868224.003.0002.

Shafer-Landau, Russ. 2005. *Moral Realism: A Defence*. New York: Clarendon Press of Oxford University Press.

Silk, Alex. 2016. *Discourse Contextualism: A Framework for Contextualist Semantics and Pragmatics*. First edition. Oxford: Oxford University Press.

———. 2017. "Normative Language in Context." In *Oxford Studies in Metaethics 12*, edited by Russ Shafer-Landau. Oxford University Press. https://doi.org/10.1093/oso/9780198805076.003.0009.

Sinclair, Neil. 2014. "On Standing One's Ground." *Analysis* 74 (3): 422–31. https://doi.org/10.1093/analys/anu068.

Singer, Daniel J. 2015. "Mind the Is-Ought Gap:" *Journal of Philosophy* 112 (4): 193–210. https://doi.org/10.5840/jphil2015112412.

Skow, Bradford. 2018. *Causation, Explanation, and the Metaphysics of Aspect*. 1st ed. Oxford University Press. https://doi.org/10.1093/oso/9780198826965.001.0001.

Sliwa, Paulina. 2016. "Moral Worth and Moral Knowledge." *Philosophy and Phenomenological Research* 93 (2): 393–418. https://doi.org/10.1111/phpr.12195.

Smith, Angela M. 2005. "Responsibility for Attitudes: Activity and Passivity in Mental Life." *Ethics* 115: 236–71.

Smith, Michael. 1987. "The Humean Theory of Motivation." *Mind* 96 (381): 36–61.

———. 1994. *The Moral Problem*. Malden, MA: Blackwell.

———. 2009. "Two Kinds of Consequentialism." *Philosophical Issues* 19: 257–72.

———. 2010. “Dworkin on External Skepticism.” *Boston University Law Review* 90 (2): 509–20.

Smith, Michael, and Daniel Stoljar. 2003. “Is There a Lockean Argument against Expressivism.” *Analysis* 63 (1): 76–86.

Stephenson, Tamina. 2007. “Judge Dependence, Epistemic Modals, and Predicates of Personal Taste.” *Linguistics and Philosophy* 30 (4): 487–525. https://doi.org/10.1007/s10988-008-9023-4.

Stevenson, Charles Leslie. 1937. “The Emotive Meaning of Ethical Terms.” *Mind* 46 (181): 14–31.

———. 1944. *Ethics and Language*. New Haven: Yale University Press.

Street, Sharon. 2010. “What Is Constructivism in Ethics and Metaethics.” *Philosophy Compass* 5 (5): 363–84.

———. 2011. “Mind-Independence without the Mystery: Why Quasi-Realists Can’t Have It Both Ways.” In *Oxford Studies in Metaethics 6*, edited by Russ Shafer-Landau, 1–32. New York: Oxford University Press.

Suikkanen, Jussi. 2009. “The Subjectivist Consequences of Expressivism.” *Pacific Philosophical Quarterly* 90: 364–87.

Walzer, Michael. 1973. “The Problem of Dirty Hands.” *Philosophy & Public Affairs* 2: 60–180.

Watzl, Sebastian. 2022. “The Ethics of Attention: An Argument and a Framework.” In *Salience*, edited by Sophie Archer, 89–112. London: Routledge. https://doi.org/10.4324/9781351202114-6.

Wiggins, David. 1987. “A Sensible Subjectivism?” In *Needs, Values, Truth: Essays in the Philosophy of Value*, 227–44. New York: Clarendon: Oxford University Press.

Willer, Malte, and Christopher Kennedy. 2022. “Assertion, Expression, Experience.” *Inquiry* 65 (7): 821–57. https://doi.org/10.1080/0020174X.2020.1850338.

Williams, Bernard. 1985. *Ethics and the Limits of Philosophy*. Cambridge, MA: Harvard University Press.

Woods, Jack, and Barry Maguire. 2017. “Model Theory, Hume’s Dictum, and the Priority of Ethical Theory.” *Ergo, an Open Access Journal of Philosophy* 4 (14). https://doi.org/10.3998/ergo.12405314.0004.014.

Worsnip, Alex. 2019. “Disagreement as Interpersonal Incoherence.” *Res Philosophica* 96 (2): 245–68. https://doi.org/10.11612/resphil.1771.

Wright, Crispin. 1992. *Truth and Objectivity*. Cambridge, MA: Harvard University Press.

Wygoda Cohen, Shlomit. 2020. “Not All Partial Grounds Partly Ground: Some Useful Distinctions in the Theory of Grounding.” *Philosophy and Phenomenological Research* 100 (1): 75–92. https://doi.org/10.1111/phpr.12524.

———. 2021. “Mind Independence versus Mind Nongroundedness: Two Kinds of Objectivism.” *Ethics* 132 (1): 180–203. https://doi.org/10.1086/715278.

Yablo, Stephen. 2006. “Non-Catastrophic Presupposition-Failure.” In *Things: Papers on Objects, Events, and Properties*, 269–95. New York: Oxford University Press.

———. 2014. *Aboutness*. Princeton: Princeton University Press.

———. 2017a. “If-Thenism.” *Australasian Philosophical Review* 1 (2): 115–32. https://doi.org/10.1080/24740500.2017.1346423.

———. 2017b. “Precis of Aboutness.” *Philosophical Studies* 174 (3): 771–77. https://doi.org/10.1007/s11098-016-0755-7.