

Penultimate draft.
Please consult the official version in Analysis.

Critical Notice

Having it Both Ways: Hybrid Theories and Modern Metaethics,
GUY FLETCHER AND MICHAEL RIDGE (eds.),
Oxford University Press USA, 2014, xvi + 302 pp.

ALEX SILK

A dialogue between a saint and a sinner:

“This rich collection features cutting-edge work on an exciting recent approach in metaethics: hybrid theories of normative language and thought. The volume, which includes papers by many of hybridism’s most prominent advocates and critics, is essential for people working in metaethics. Anyone with interests in expressive language or practical thought more generally would benefit from it.”

“Pshaw. There is no category of reader for whom I would recommend this volume. If you are a metaethicist working on hybrid theories: either you participated in the conference from which the volume resulted, or you are already familiar with the contributors’ previous work (with some exceptions, there isn’t *that* much new). If you are a young researcher interested in getting introduced to the field: most of the papers assume familiarity with the authors’ previous work. And if you are a specialist in a more general related area, like philosophy of language or mind, and are interested in seeing applications to the case of normative language or thought: just don’t. One might have expected, for a field birthed out of developments in philosophy of language on non-truth-conditional aspects of meaning, to find references to such developments. Ditto for work in psychology (think: primacy of affect, dual-process model, etc.). Yet seek, and you shall not find. In principle, hybridism opens up the potential for exciting advancements in our more general understanding of (e.g.) linguistic expressives, implicature, projective content, and the nature and structure of motivation, among other things. Such advancements may need to await future research.”

We will return to this “tale of two assessments” (§4). To foreshadow: neither is Faithful and True. Meanwhile, a mean between two extremes...

Hybrid metaethical theories claim (i) that “moral claims express *both* belief-like *and* desire-like mental states,” or (ii) that “moral judgments are constituted by both belief-like and desire-like components” (ix). The strategy is to appeal to the belief-like component to avoid pure expressivism’s putative problems with embedding and logic (Frege-Geach), and to appeal to the desire-like component to avoid pure cognitivism’s putative problems capturing the connection between moral judgment and

motivation. On the language side, many theories appeal to pejoratives as a precedent for expressions having both truth-conditional and attitudinally expressive aspects of meaning. Examining the aptness of this analogy will provide a natural foray into various core issues in the volume. I will focus primarily on Mark Schroeder's and Guy Fletcher's contributions. I close with brief remarks on the remaining papers.

1. *Pejoratives and projection*

Following Schroeder in his chapter, let's consider 'cheesehead' as our candidate pejorative (274).¹ Uttering (1) conventionally (i) describes Mark as being from Wisconsin, and (ii) expresses some sort of negative attitude toward people from Wisconsin (or toward Mark merely because he is from Wisconsin).

(1) Mark is a cheesehead. (281)

Pejoratives can also embed in complex constructions and figure in valid arguments, as in (2).

(2) a. If Mark cannot drive in traffic, then Mark is a cheesehead.
b. Mark cannot drive in traffic.
c. ∴ Mark is a cheesehead. (280–81)

Pejoratives thus provide hybrid theorists with a license for optimism. For everyone needs an account of the dual descriptive+attitudinal meaning of pejoratives. The hope is then that we can apply our best semantics for pejoratives to paradigmatic normative terms in responding to the Frege-Geach problem.

This hope is short-lived. There are clear linguistic differences between pejoratives — and other types of linguistic expressives (honorifics, expressive attributive adjectives, certain discourse particles) — and words such as 'wrong'. The expressive content of pejoratives *projects*: it tends to be regarded as a commitment of the speaker even when the pejorative is embedded under entailment-canceling operators, as in Family-of-Sentences examples:

(3) a. Mark isn't a cheesehead.
b. Mark might be a cheesehead.
c. Is Mark a cheesehead?

Further, the expressive content needn't contribute to the expression's local context: it

¹Readers for whom 'cheesehead' is less well-attested may wish to substitute an alternative pejorative. All occurrences of pejoratives here are mentioned, not used.

needn't contribute to the conventional contents of embedded clauses, like the complements of belief-predicates.

- (4) [Context: We hate people from Wisconsin, though Steve has nothing against them. We want to know where Mark is from. You leave to ask Steve. When you get back, I ask you what he thinks. You say:]
Steve thinks Mark is a cheesehead.
- (5) Steve thinks that cheesehead Mark is a great guy.

The expressive content of 'cheesehead' in (4)–(5) isn't taken to characterize Steve's state of mind. It projects simply as a commitment of the speaker.

These projective properties of pejoratives are well-known.² Hence it is puzzling why pejoratives were thought to provide a promising model for normative terms generally. 'Wrong' contrasts with 'cheesehead' with respect to both of the above properties (which isn't to say that either moral terms or linguistic expressives form homogeneous linguistic classes; they don't). This is evident on various ways of specifying the expressive content of 'wrong'.

First, consider projection in Family-of-Sentences examples. Treating the expressed negative attitude as directed toward the subject of the predicate generates incorrect predictions. Utterances of (6) wouldn't typically be regarded as committing the speaker to (something like) disapproval of killing.

- (6) a. Killing isn't wrong.
b. Killing might be wrong.
c. Is killing wrong?

Treating the expressed attitude as a general attitude toward things, whatever they are, that satisfy the (alleged) descriptive, truth-conditional content of 'wrong' — a common move among hybrid theorists³ — won't work either. The following aren't infelicitous:

- (7) [Context: Nihilist says:]
Killing isn't wrong, since nothing is right or wrong.
- (8) [Context: Budding moral skeptic says:]

²See, e.g., POTTS 2005, 2007, MCCREADY 2010, TONHAUSER ET AL. 2013, and references therein. Note that a content's having local effect under attitude verbs is compatible with its also projecting (e.g., with the contents of factive attitude complements).

³See SCHROEDER 2009 for discussion of the relevant arguments.

I'm not sure if anything is right or wrong. But killing might be wrong. If anything is wrong, it's that.

(9) [Context: Budding moral skeptic says:]

I'm not sure if anything is right or wrong. But if anything is wrong, it's killing. Is killing wrong?

It's hard to hear analogous examples with pejoratives as failing to commit the speaker to the relevant negative attitude.

Second, it's hard to hear a belief ascription with 'wrong' as committing anyone other than the subject to the negative attitude, whether particular or general. Utterances of (10) are generally anomalous.

(10) #That whole community thinks killing is wrong, but they're all for it.

The (alleged) expressive content of 'wrong' systematically contributes to local content and fails to project in attitude contexts.

There are additional contrasts. For instance, the expressive content of pejoratives is typically infelicitous as the answer to a question, and continues to project when what is at-issue is the appropriateness of the relevant attitude.

(11) A: Is Mark to be despised merely because he's from Wisconsin?
B: #Well, he isn't a cheesehead.

B's denial in (11) cannot be used to communicate a negative answer to A's question. By contrast, B's response in (12) is perfectly felicitous (stiltedness aside).

(12) A: Is stealing to be disapproved of merely because it doesn't maximize happiness?
B: Well, it isn't wrong.

The contrasting behavior in attitude ascriptions of pejoratives and moral terms was pressed in Schroeder's 2009 comprehensive critical discussion of hybridism. Yet the extent of the linguistic differences don't seem to have been adequately appreciated. For instance, Daniel Boisvert's paper treats the expressive content of 'wrong' as "semantically and logically distinct" (44); sentence (13) is "rendered" (43) as (14), where F is the descriptive property picked out by 'wrong'.

(13) If insulting others is wrong, I won't insult others.

(14) If insulting others is F, then I won't insult others; down with F-things!

(43)

Boisvert cites various linguistic expressives as precedents to justify this move (44–45). The differences in projection properties go unnoticed.

David Copp is more circumspect concerning the aptness of the analogy to pejoratives. He explicitly qualifies that he is merely taking pejoratives as a “model” for expressions which conventionally both denote ordinary properties and express desire-like attitudes. Copp’s care in this regard is refreshing. Yet there is something unsatisfying when perhaps the best one can say about a volume’s treatment of a dialectically central issue is a case where the author *avoids* addressing it.

Schroeder’s paper — characteristically clear and systematic — is essentially a reply piece to his 2009 article. The conclusion there was that, in light of the sorts of projection differences under attitude verbs discussed above, modeling the meaning of moral terms on analogy with pejoratives is problematic. The current paper argues that a more promising model for (cognitivist) hybrid theories is ‘but’. On one common view, although the truth-conditions of (15) are that Shaq is huge and Shaq is agile, utterances of (15) also conventionally express that there is a contrast between (Shaq’s) being huge and (Shaq’s) being agile.

(15) Shaq is huge but agile. (274)

This contrastive implication arguably projects in Family-of-Sentences contexts; however, unlike the projective content of pejoratives, it systematically has local effect under attitude verbs. (I use ‘implication’ as neutral among entailment, presupposition, implicature, etc.)

(16) Sally thinks Shaq is huge but agile. (#But she sees no contrast between hugeness and agility.)

Schroeder is right that “there is nothing fundamentally strange about secondary contents affecting the core, ‘truth-conditional’ content of attitude verbs” (274). The semantic status of the contrastive implication associated with ‘but’ is contentious. But there are other examples that would fit the bill — e.g., the complement implication of factive attitude verbs (‘know’), the prejacent implication of ‘only’, or the prestate implication of change-of-state verbs (‘stop’).⁴

Schroeder outlines a “simple, flexible model for secondary contents” on which sentences are “assigned to a triple of semantic values: a core content, a primary set, and a secondary set” (277). The “core content” is the primary, truth-conditional

⁴See Tonhauser et al.’s (2013) ‘Class C’ of backgrounded projective contents.

content. The primary and secondary sets represent any “secondary contents,” i.e. projective implications; the primary set represents projective contents which have local effect under attitude verbs (cf ‘but’), and the secondary set represents projective contents which don’t (cf ‘cheesehead’). Schroeder is clear in describing the proposed framework as “simple” (277). There is nothing wrong with using a simple model if the aim is simply to show that a language in which certain projective contents systematically have local effect is *possible*. (Should we have thought otherwise?) A concern, though, is that Schroeder’s ensuing worry for (cognitivist) hybrid theories turns on the very features of the model that make it not just simple, but simplistic.

Briefly, Schroeder’s worry is as follows. The following argument is “inference-licensing”: accepting the premises commits one to accepting the conclusion.

- (17) a. What Caroline believes is true.
 b. What Caroline believes is that stealing is wrong.
 c. ∴ Stealing is wrong. (289)

Accepting the conclusion, according to hybridism, commits one to a certain desire-like attitude — call it ‘D’. But accepting premise 1 needn’t involve having any particular desire-like attitude. So, it must be premise 2 which commits one to D. This can only happen, in Schroeder’s framework, if D figures in the secondary set. So, for any moral term, one must stipulate that the desire-like members of primary sets also be in the secondary sets. This seems *ad hoc*.

Schroeder worry gets its bite from the sharp distinction between primary and secondary sets, i.e. projective contents which contribute to the truth-conditions of attitude ascriptions and projective contents which don’t. But it is hard to imagine that our best semantics for a natural language such as English will include such a distinction. Many expressions only sometimes have local effect under attitude verbs. Pejoratives and other expressives themselves are arguably in this camp:

- (18) I don’t have anything against people from Wisconsin. But Steve, who does, thinks Mark is the worst cheesehead he knows.
(cf. SCHLENKER 2003: ex. 109b)
- (19) My friend said Professor X gave him a bad grade again. I don’t have anything against X, but my friend is pissed. He thinks the damn guy (/the jerk) always favours long papers. (cf. HARRIS & POTTS 2009: ex. 16)

Though it’s contentious how to capture such cases,⁵ there are other uncontentious

⁵See, e.g., SCHLENKER 2003, AMARAL ET AL. 2007, POTTS 2007, HARRIS & POTTS 2009.

examples. Consider the following contrasting examples with ‘too’:

- (20) [Context: We’re cousins, out at a party.]
My parents think you’re home in bed. They think I’m in bed too. *(local effect)*
- (21) [Context: My parents don’t know anything about you. You say you’re home in bed. I say:]
My parents think I’m in bed too. *(no local effect)*

Which set(s) includes the projective implication of ‘too’ that there is a salient alternative true proposition? It cannot be the primary set, since this would incorrectly treat (21) as ascribing to my parents beliefs about your whereabouts. And it cannot be the secondary set, since this would incorrectly treat (20) as inconsistent with the common ground. In reply Schroeder might treat the secondary set as including projective contents that don’t *necessarily* have local effect, and put the projective implication with ‘too’ there. But that won’t do for Schroeder’s purposes, for it is by treating the elements of secondary sets as invariably projecting from attitude contexts that Schroeder captures the validity of (17)-style arguments with ‘cheesehead’ (289).

The upshot, I take it, is that we should never have encoded a binary primary/secondary-set distinction into the semantics. But without such a distinction, Schroeder’s ad hocness worry (at least as presented) seems unmotivated. There has been substantial recent empirical and technical work on the logic/semantics of projective contents.⁶ The issues are bewilderingly complex. Initiating a dialectic with Schroeder’s “simple, flexible framework for thinking about secondary contents” (277) seems like a step in the wrong direction.

What *should* the hybrid theorist say about arguments like (17)? General issues concerning propositional anaphora and local effects provide independent grounds for questioning the validity of the inference schema ‘What S believes is true’; ‘what S believes is P’: ‘P’ (*cf.* SILK 2014: §4.2.4). Take a Donnellan-style case: We are talking about whether Steve knows that Mark, who is standing across the room, is a philosopher. Steve thinks Mark is drinking a martini, but we know Mark is drinking something else in a martini glass. Does accepting the premises in (22) commit one to accepting the conclusion?

- (22) a. What Steve believes is true.
b. What Steve believes is that the man drinking a martini is a philosopher.

⁶See, e.g., SCHLENKER 2003, POTTS 2005, 2007, MCCREADY 2010, TONHAUSER ET AL. 2013, MURRAY 2014.

- c. \therefore The man drinking a martini is a philosopher.

The description ‘the man drinking a martini’ in (22b) presupposes, roughly, that a certain maximally salient man, Mark, is drinking a martini. The falsity of this presupposition, call it \mathcal{P} , can interfere with our judgments about whether the argument is acceptance-preserving. We must test instead for what we might call — adapting a notion from VON FINTEL 1999 — *Strawson-acceptance-preservation*: we check whether the argument is acceptance-preserving given that any projective contents of premises and conclusion are (consistently) satisfied. And indeed, accepting the premises in (22) along with \mathcal{P} would commit one to accepting the conclusion. But the same also holds with (17), according to the hybrid theorist: accepting the premises, along with the projective negative attitude associated with ‘wrong’ in (17b)–(17c), commits one to accepting the conclusion. Requiring that (17) be acceptance-preserving in the general non-Strawsonified sense would be question-begging.

For all the discussions in the volume of pejoratives and the potential advantages of going hybrid for responding to the Frege-Geach problem, there is a conspicuous absence of sustained, rigorous linguistic inquiry into the specific relation between the descriptive and expressive meaning components, how they are represented in the compositional semantics, etc. This is unfortunate given the rich empirical and formal research in linguistic semantics investigating precisely these sorts of issues (n. 6). Though the analogy with pejoratives may be problematic, perhaps one could maintain a hybrid theory associating moral predicates with projective general desire-like attitudes, where this projective content has obligatory local effect. The lack of projection in (7)–(9) could be attributed to the fact that the projective content cannot be taken for granted in the discourse. Compare: the existence implication typically associated with ‘the king of France’ likewise doesn’t project in (23) where the question under discussion is whether France has a king.

- (23) A: Does France have a king?
B: The king of France wasn’t at the royal gala. (So, maybe not.)

We need a more comprehensive investigation of projection phenomena with the varieties of moral terms — and not just with thin moral terms like ‘wrong’. Detailed comparison with various types of linguistic expressives and other classes of projective contents is required.

2. *Implicature*

It is common to treat the projective, attitudinal component of linguistic expressives as a (Pottsian) conventional implicature. However, Guy Fletcher's paper argues that the putative attitudinal implications associated with uses of moral terms pattern instead with (generalized) conversational implicatures — specifically, in being reinforceable, detachable, cancelable, and calculable. The alleged contrasts Fletcher draws between moral terms and expressives strike me as problematic. Many of the issues are more complex than Fletcher's discussion lets on.

Reinforceability: Fletcher offers the following contrast:

- (24) a. Chelsea has signed that kraut Ballack. #I have derogatory attitudes toward German people.
b. This *fucking* printer cost £200. #I am angry about the printer.
- (25) Waterboarding terror suspects is wrong. I'd resent anyone who did so and feel guilty if I did.

(181–82)

However, the continuations in (24) don't adequately express the attitudinal contents of 'kraut' and 'fucking'. (Compare: 'Everyone failed. #Also, someone failed.')

The examples in (26) seem more analogous with those in (25).

- (26) a. Chelsea has signed that kraut Ballack. I can't *stand* Germans.
b. This *fucking* printer cost £200. I hate it!

The second sentences, though perhaps redundant in some sense, arguably provide felicitous continuations of the attitudes expressed in the first. (Compare: 'This is an electron. It has negative charge.')

Nondetachability: Fletcher claims that whereas the truth-conditional content of (27a) can be conveyed without expressing the problematic attitude, as in (27b), any way of conveying the truth-conditional content of (28a) expresses a negative attitude, as reflected in (28b).

- (27) a. Chelsea has signed that kraut Ballack.
b. Chelsea has signed that German Ballack.
- (28) a. Waterboarding terror suspects is wrong.
b. It's wrong to waterboard terror suspects. (183)

It is puzzling why Fletcher assumes that candidate non-implicature-generating analogues of (28a) must also use a moral term. Suppose 'wrong' denotes the property

of failing to maximize overall happiness. Then (29) has the same truth-conditions as (28a) but needn't generate the attitudinal implicature.

(29) Waterboarding terror suspects fails to maximize overall happiness.

Cancelability: Fletcher claims that attitude expression with moral terms, unlike with expressives, is cancelable:

(30) Chelsea has signed that kraut Ballack. #I have nothing against German people. (187)

(31) You say that it's morally wrong and you're absolutely right about that. But I don't care about what's morally right and wrong; I just want to make as much money and have as much fun as possible. It's not illegal. So lets do it and keep it quiet. Anyone who passed up this kind of opportunity is a sucker. (189)

Though Fletcher acknowledges several complications in testing for cancelability, the issues are still more complex than the discussion lets on. It's well-known that modal expressions can be used "non-endorsingly," i.e. used without presenting the speaker as endorsing the relevant considerations that would verify the modal claim, as in (32).⁷

(32) Ernie has to be home by 10. Aren't his parents stupid? I'd stay out if I were him.

The phenomenon generalizes to other categories of expressions:

(33) This cat food is really tasty. Fluffy eats it right up. (cf. STEPHENSON 2007)

Recent discussions of expressives provide similar examples with epithets (see n. 5):

(34) I've heard similar arguments before: if companies like Symantec or McAfee make antivirus applications for the Mac, then Macs must truly be vulnerable somehow, somewhere. Steve Jobs and the rest of the Apple cronies must be lying.
[www.digitaltrends.com/features/antivirus-programs-for-mac-snake-oil-or-public-service] (in HARRIS & POTTS 2009)

⁷The distinction between endorsing and non-endorsing uses has been noted under various labels in many areas (see SILK 2014, 2015 for discussion and references).

It's contentious how precisely to analyze such examples — e.g., in terms of syntactic/semantic binding, anaphora, pragmatically mediated 'perspective shifting', etc. What is important here is simply that similar apparent "cancelation" phenomena can be observed with both expressives and moral terms. (Whether it is best understood as a kind of cancelation is another matter.)

Calculability: Fletcher argues that "attitude expression fits a conversational implicature model, as there seem to be numerous ways one could calculate an implicature (that the speaker has the relevant negative desire-like attitude) from their asserting that something is wrong" (191). For instance, "a speaker's engaging in discussion about what is required, permitted, prohibited, or recommended by some standard makes it likely that she cares about conformity to that standard" (191). One might have hoped for a more rigorous derivation of the implicature from the specific truth-conditions, specific maxims, and general features of contexts of use, but the basic idea is straightforward enough. The problem — as in the discussion of nondetachability above — is that there are potential complications depending on one's views about the truth-conditional content. Suppose 'wrong' denotes the property of failing to maximize overall happiness. One cannot in general calculate that the speaker has the relevant negative attitude from the truth-conditions of an utterance of 'Killing is wrong', i.e. that killing isn't happiness-maximizing. Indeed that is the one of the lessons of Open-Question-style considerations. Compare: one cannot in general calculate that the speaker has the relevant negative attitude from the truth-conditions of (1), i.e. that Mark is from Wisconsin.

So, depending on one's views about the truth-conditional contribution of moral terms, the respective attitudinal implicatures of uses of moral terms and pejoratives arguably pattern analogously with respect to reenforceability, nondetachability, cancelability, and calculability. One might think on this basis that, far from being a "clear nonstarter" (xiv), conventional implicature views of moral attitude expression are looking pretty good. But there are other important contrasts (§1).

Fletcher concludes with several objections to generalized conversational implicature views, ultimately favouring (what he regards as) a non-implicaturist, non-hybrid alternative. It is hard to see why his arguments wouldn't overgeneralize to show that no expressions give rise to generalized conversational implicatures. For instance, Fletcher argues that conversational implicature views cannot capture how eavesdroppers infer the desire-like attitudes of speakers who are talking to themselves: "There is no reason to expect people to adhere to the cooperative principle when, by their lights, they are not communicating with anyone.... More importantly... [the speaker] has no beliefs about what her hearer can work out. She believes that she is alone and may not believe anything about [the eavesdropper's]

competencies. Thus implicature cannot be the means by which [the eavesdropper] learns of [the speaker's] desire-like attitudes" (193). Nothing in this argument turns on features specific to uses of moral terms.

3. *The remaining papers*

I will close with brief remarks on the remaining papers, followed by a summary conclusion.

Ridge: Michael Ridge's paper defends a cognitivist hybrid theory of 'rational'. The cognitivist core is provided by an instrumentalist conception of rationality: 'S's ϕ -ing is rationally permissible' means that it is compatible with S's "adherence to those norms that an agent must be taken to adhere, for the most part, and absent some special explanation, in order for the idea that the agent is rational (*in the capacity sense*) to be intelligible" (9). (What if there isn't a unique such set of norms? What is the upshot of defining non-gradable expressions like '(im)permissible', 'required', etc. in terms of gradable notions of intelligibility? Why think the instrumentalist conception has become conventionalized in the lexical semantics of 'rational'?) The "hybrid 'twist'" is provided by a generalized conversational implicature: in deliberative contexts,

Intuitively, in saying that the option in question would be irrational without any further caveats or commentary, I am thereby *advising* my interlocutor against choosing that option. However, it would be insincere for me to advise you to do something if I did not think it was the thing to do... So insofar as I am being a cooperative interlocutor, one can derive from standard Gricean maxims of relevance and the like that I do indeed believe that my interlocutor ought to act accordingly. This in turn can explain why such an utterance expresses a proattitude [sic]... Insofar as I am myself rational... I will be motivated to perform [sic] such an action if I find myself in his or her circumstances. (17)

Further details concerning how the attitudinal implicature is calculated from the specific truth-conditions and maxims are not provided.

Boisvert: Daniel Boisvert argues that we should reject truth-conditional semantics and develop hybridism in the context of a "success-conditional semantics" (SCS), which associates sentences with the conditions under which they would be "successful" (true, complied with, or sincere, depending on the sentence type). It would be unfortunate to abandon everything which truth-conditional semantics has taught us in light of cursory dismissals such as these:

[N]either nondeclaratives nor sentences containing them are truth evaluable; consequently sentences of neither type have truth conditions; consequently constructing a truth theory to help explain our understanding of these types of sentences appears bound to fail. (23)

SCS can, while truth-conditional semantics cannot, explain the intuitive validity and invalidity of a variety of different kinds of logical transitions, including those across moods. (37)

It's as if the decades of research developing broadly truth-conditional semantics for interrogatives and imperatives never happened (Hamblin, Karttunen, Groenendijk & Stokhof, Ginzburg, Roberts, Han, Portner, etc.).

Copp: David Copp defends his cognitivist hybrid theory, developed in previous work, against two objections: first, that it cannot capture the element of endorsement in moral thought; second, that it cannot capture the “inference-licensing property” of valid arguments (roughly, that accepting the premises commits one to accepting the conclusion). On the latter point, Copp argues against SCHROEDER 2009 that the inference-licensing property isn't in fact a general feature of valid arguments. Copp offers counterexamples with pejoratives: for any pejorative ‘P*’, the argument schema in (35) is semantically valid, in the sense that any world where the premise is true is a world where the conclusion is true; but accepting the premise doesn't commit one to the negative attitude projecting in the conclusion.

- (35) a. Everything Anna thinks is true.
b. ∴ If Anna thinks Brenda is P*, Brenda is P*. (64–65)

A worry is that we have independent evidence — linguistic evidence concerning differences in projection and local effects — that the analogous arguments with moral predicates *are* acceptance-preserving. However, there are other strategies for replying to Schroeder's (2009) argument (see §1).

Hay: Many hybrid theories treat all uses of moral terms as expressing the same “general desire,” as to do what is morally right/wrong/etc., whatever it is. Ryan Hay argues that treating general desires as attitudes “toward types rather than propositions involving universals” (i.e., universal generalizations; 76) better captures the relation between agents' beliefs and desires and hence between moral judgment and motivation. Although Hay objects to several “interpretations involving the scope of the agent's desire and the universal quantifier” (83), one option he doesn't seem to consider is this: For each world w in the agent S 's doxastic alternatives Dox_S (worlds compatible with S 's beliefs), let D_w be the ordinary descriptive property denoted by

(say) ‘wrong’ in w , and let $R_{S,w}$ be the set of most desired worlds w' given the relevant facts in w . We might represent the “Generality Requirement” that S desires not to do whatever is morally wrong via the following structural constraint on S’s attitudes: for every $w \in Dox_S$, every $w' \in R_{S,w}$ is such that S doesn’t perform any acts which are D_w in w' . If S believes that stealing is wrong (“Belief Requirement”), then for every $w \in Dox_S$, stealing is D_w in w . So, given the Generality Requirement, for every $w \in Dox_S$, every $w' \in R_{S,w}$ is such that S doesn’t steal in w' , i.e. that S desires not to steal (“Resultant Requirement”). This would seem to satisfy Hay’s three Requirements, and reflect how general desires can interact with beliefs so as to lead to motivation in specific moral judgments.

Tresan: Jon Tresan defends a “diachronic” hybrid theory, according to which “when we use moral terms for moralizing, they have the realist [i.e., descriptive] meaning... But when we use them for metamoralizing they have an attitudinal or functional meaning along the lines posited by emotivists and antirealist expressivists. For instance, ‘they think dancing is morally wrong’ ascribes to ‘them’ a certain sort of negative orientation to dancing” (99). The view seems to be that moral terms have one meaning in asserted contexts, and a different meaning when embedded under attitude verbs. This kind of violation of compositionality and “semantic innocence” is typically regarded as a *reductio* of semantic theories.

Finlay: Stephen Finlay’s paper defends a “quasi-expressivist” account of discourse disagreement with normative language. On Finlay’s “end-relational” contextualist semantics (developed extensively in other work), speakers uttering normative sentences ‘ ϕ ’ and ‘No, $\neg\phi$ ’ often assert consistent propositions about their respective information and/or ends. What explains the speakers’ disagreement are “conflicting preferences that are pragmatically expressed by asserting [these] logically consistent normative propositions” (125). In motivating his contextualist account Finlay notes that speakers don’t always evaluate normative claims with respect to their own current information/ends. While I agree that this raises a *prima facie* worry for relativism and (pure) expressivism, the fact that normative claims *sometimes* receive intuitively contextualist readings has been granted by all parties in the broader literature (*pace* Finlay’s suggestions on 130, 145).⁸ The question is at what level to explain how conversational factors are affecting interpretation — e.g., whether they are affecting the determination of content from character, or whether a relativist vs. contextualist pronoun figures in the syntax, among other options. The central challenges for contextualism remain — most pressingly, in my view, to explain the *systematic* contrasts between (e.g.) normative language and paradigm context-sensitive

⁸Cf. STEPHENSON 2007, MACFARLANE 2014, MOSS 2015.

expressions. (Why do speakers systematically assert “normative propositions” they don’t have a “fundamental interest” in (133)? Why do we systematically use normative terms, unlike paradigm context-sensitive expressions, to communicate information about the features of context on which their contents depend? Why...) The force of these challenges has been underappreciated among contextualists.⁹

Eriksson: John Eriksson’s paper considers how (hybrid) expressivism should understand meaning in general. The central move is to deny that expressivism needs to “explain the meaning of a sentence by appeal to its expressing a particular belief” or desire-like attitude (162; *cf.* 169); rather, meaning should be explained in terms of “more basic states of mind” — namely, a “thought” (160–62) in the case of descriptive meaning, and “entertaining a plan or determination of what to do” (165) in the case of nondescriptive meaning. This is to avoid “the traditional problem with expressivism” (155) that complex sentences needn’t express the attitudes that would be expressed by their constituent clauses considered on their own. If Eriksson’s point is that expressivists should treat the meanings of expressions in terms of their *contribution* to the attitude expressed, I would have thought expressivists have been doing this since HARE 1970.

Barker: Stephen Barker’s paper argues against conventional implicaturist hybrid theories. Barker argues that the conventional implicatures associated with “a range of locutions and constructions” (200) require pure expressivist treatments. (Barker focuses exclusively on ‘even.’) Since the required “pure cognitivist expressivist” framework can be applied to moral language, going hybrid is otiose. Barker’s paper fails to advance our understanding of conventional implicature. For instance, the initial sections examine whether conventional implicature can be “reduc[ed]” (205) to semantic presupposition, “secondary propositional content,” or pragmatic presupposition. An alleged feature of Barker’s positive view is that it “explains implicating’s presuppositional status” (209). It’s hard to know what to make of this: the consensus is that conventional implicatures are to be distinguished from presuppositions and other kinds of projective contents. On Barker’s positive view, what distinguishes assertion from (conventionally?) implicating is that “In an assertion the speaker’s purpose is to manifest a *defensive stance* with respect to [a mental state] P, and in an implicative act it is merely to manifest the state” (208). Not enough is said about “defending” vs. merely “manifesting” mental states to generate testable predictions. Natural questions include: What distinguishes assertions/implicatures from questions or imperatives? (Intuitively, my uttering ‘Will it stop?’ manifests

⁹See my 2014 for elaboration on these points and development of a contextualist theory which begins to address them.

(a disposition to defend) my mental state of wanting to know whether it will stop.) Why would “merely manifesting” a mental state lead to backgrounding or (the varieties of) projection phenomena?

Bar-On, Chrisman, & Sias (BCS): BCS’s paper develops Bar-On & Chrisman’s (2009) “ethical neo-expressivism,” on which ethical claims express propositions and motivational attitudes, but in two different senses of ‘express’: it is ethical sentences which semantically express (“s-express”) propositions (i.e., have propositions as their semantic contents), and uses of ethical sentences (i.e., utterances qua acts) which express (“a-express”) motivational attitudes. It isn’t clear to me how simply marking the s-/a-express distinction advances the dialectic. The expressivist strategy, I take it, is to explain the semantics, or conventional meaning, in terms of use — in BCS’s terms, to explain the meaning of a sentence *S* fundamentally in terms of the attitude conventionally a-expressed in using *S*, rather than in terms of any content s-expressed by *S*. The central move is to *do away with* treating a notion of s-expression as playing a fundamental explanatory role. Gibbard, for instance, still semantically associates sentences with (s-expressed) contents; he just denies that these (s-expressed) contents fundamentally explain the sentences’ semantic properties. Likewise, the expressivist agrees with BCS in treating having the relevant motivating attitude as a “*propriety condition* on making (genuinely) ethical claims” (240), but simply treats this condition as conventionalized.

Schroeter & Schroeter (SS): SS argue that applying their general binding model of concepts to normative concepts opens up a way for pure cognitivism to capture the connection between normative judgment and motivation. Though being motivated to act in accordance with one’s normative judgments isn’t necessary for conceptual competence, having such motivational dispositions is crucial in constituting normal conceptual competence with our shared normative concepts and in fixing their reference. SS’s paper is refreshing. It draws fruitful connections with broader work on concept determination and individuation. One issue: SS argue that an adequate account of normative concepts must capture how one’s various token attitudes deploying a particular normative concept “strike the subject as pertaining de jure to the same [normative] topic” (252). A worry is that, absent an account of the nature of the attitude of “seeing as pertaining de jure to the same topic” in normative thoughts, SS’s account is neutral on questions of (non-)cognitivism or motivational internalism/externalism. For all SS argue, “keep[ing] track of sameness of normative topic among our thoughts” (251) might essentially involve being in a certain motivational state.

4. Conclusion

Let's return to our initial "tale of two assessments." As usual with these sorts of things, the truth is somewhere in the middle. On the one hand, hybridism highlights the potential for fruitful cross-disciplinary connections and new directions for future research. There *are* interesting questions raised by hybridism. Among them:

- What is the nature of fundamental semantic explanation? Should meaning be explained fundamentally in terms of contents/truth? Or states of mind?
- Do normative sentences have a descriptive component to their conventional meaning? Do normative uses of language conventionally express desire-like attitudes? If 'yes' to both, how are the dual descriptive+expressive aspects of meaning to be implemented in a compositional grammar?
- What are the similarities/differences among various linguistic expressives and paradigmatic normative expressions — e.g., concerning projection, local effects, discourse/information structure, felicity conditions?
- How are normative judgments responsive to ordinary features of the world and subjects' beliefs about them? How does the connection to motivation in normative judgment relate to affective coding in judgment more generally?

Substantial progress has been made on many of the more general issues. Yet precisely for this reason the volume is often frustrating. There is nearly no engagement with relevant broader research, and the issues are often more complex than the discussions let on. The reader is left thinking that there are exciting developments on the horizon, just not here. Don't get me wrong; cross-disciplinary work is hard, indeed daunting. But we can do better, metaethics. We must.

*University of Birmingham
ERI Building
Edgbaston, Birmingham
B15 2TT
UK
a.silk@bham.ac.uk*

References

- AMARAL, PATRICIA, CRAIG ROBERTS, & E. ALLYN SMITH. 2007. Review of *The logic of conventional implicatures* by Chris Potts. *Linguistics and Philosophy* 30, 707–749.

- VON FINTEL, KAI. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16, 97–148.
- HARE, R.M. 1970. Meaning and speech acts. *Philosophical Review* 79, 3–24.
- HARRIS, JESSE A. & CHRISTOPHER POTTS. 2009. Perspective-shifting with appositives and expressives. *Linguistics and Philosophy* 32, 523–552.
- MACFARLANE, JOHN. 2014. *Assessment sensitivity: Relative truth and its applications*. Oxford: Clarendon Press.
- MCCREADY, ERIC. 2010. Varieties of conventional implicature. *Semantics and Pragmatics* 3, 1–57.
- MOSS, SARAH. 2015. On the semantics and pragmatics of epistemic vocabulary. *Semantics and Pragmatics* 8, 1–81.
- MURRAY, SARAH E. 2014. Varieties of update. *Semantics and Pragmatics* 7, 1–53.
- POTTS, CHRISTOPHER. 2005. *The logic of conventional implicature*. Oxford: Oxford University Press.
- POTTS, CHRISTOPHER. 2007. The expressive dimension. *Theoretical Linguistics* 33, 165–197.
- SCHLENKER, PHILIPPE. 2003. A plea for monsters. *Linguistics and Philosophy* 26, 29–120.
- SCHROEDER, MARK. 2009. Hybrid expressivism: Virtues and vices. *Ethics* 119, 257–309.
- SILK, ALEX. 2014. Discourse contextualism: A framework for contextualist semantics and pragmatics. URL <http://goo.gl/aIDDmb>. Monograph, University of Birmingham. To appear, Oxford: Oxford University Press.
- SILK, ALEX. 2015. What normative terms mean and why it matters for ethical theory. In MARK TIMMONS (Ed.), *Oxford studies in normative ethics*, vol. 5. Oxford: Oxford University Press.
- STEPHENSON, TAMINA. 2007. Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy* 30, 487–525.
- TONHAUSER, JUDITH, DAVID BEAVER, CRAIG ROBERTS, & MANDY SIMONS. 2013. Towards a taxonomy of projective content. *Language* 89, 66–109.