

higher-order attitudes, Frege's abyss, and the truth in propositions

In nearly forty years' of work, Simon Blackburn has done more than anyone to expand our imaginations about the aspirations for broadly projectivist/expressivist theorizing in all areas of philosophy. I know that I am far from alone in that his work has often been a source of both inspiration and provocation for my own work. It might be tempting, in a volume of critical essays such as this, to pay tribute to Blackburn's special talent for destructive polemic, by seeking to take down that by which I've been most provoked over the years. But Blackburn's biting wit has both more wit and more bite than I could hope to emulate. So instead I'll try to emulate here what I've admired the most about Blackburn – the *constructive* vein of much of his work.

In that constructive spirit, I'm going to explore one of Blackburn's earlier forays into his many discussions of 'Frege's Abyss' – the notorious problem about embeddings and logic which confronts projectivist/expressivist theories. In chapter 6 of *Spreading the Word*, Blackburn spelled out one of the earliest and, I believe, still one of the most interesting, attempts to account for the meanings of logically complex sentences and thoughts in terms amenable to the projectivist – an approach that has since been labeled the *Higher-Order Attitude* theory. In this paper I'll review a problem that I've long believed to be fatal for any account broadly like that Blackburn offered in *Spreading the Word*, and explain why I've come to believe that this problem is not as clearly fatal as I once thought.

The solution that I'll offer differs from Blackburn's original sketch remarkably little, while turning out to be much more defensible than I would have believed only a few years ago. The line of defense that I'll advocate, moreover, is broadly Blackburnian in spirit, in that it is intimately linked to a broadly deflationist picture of truth. But it also departs from the letter of much of Blackburn's work on at least one important matter: its attitude toward propositions. I'll close by explaining why I've come to believe that Blackburn's has been the wrong attitude for a thoroughgoing projectivist to take toward propositions.

I.1 projectivism/expressivism and Frege's abyss

It is the stated aim of the projectivist/expressivist to replace metaphysical modes of theoretical explanation with psychological modes. According to the proponent of this theoretical paradigm, it only creates trouble to try to ask *what it is* for something to be necessary, or true, or wrong, and looking outward to find the answers to these questions is simply a mistake of *projecting* our attitudes onto the world. Instead, we should look inward, and ask instead what it is to *think* that something is necessary, or true, or wrong. To think that something is necessary, for example, may be to be ready to rely on it in a certain kind of counterfactual reasoning, to think that something is true may be to agree with it, and to think that something is wrong may be to disapprove of it. Answers like these don't require there to *be* any answer to the questions of what it is for something to be necessary, true, or wrong, and so they provide, amongst other benefits, the prospect of escaping from the trap of unanswerable metaphysical questions.

However, it is one thing to say what it is to think that something is wrong, and another to say what it is to think that it is wrong to punish those who do no wrong, or that everything that is wrong will eventually be done, if it hasn't been already, or even simply that if stealing is wrong, then so is killing. The famous Frege-Geach Problem is the problem of taking a projectivist-friendly account of what it is to have *simple* thoughts, such as that this-or-that is wrong (or necessary, or true), and convert it into a fully general story of what it is to have *arbitrarily complex* thoughts. Frege originally argued, and Geach affirmed, that this is precisely why we need to postulate *propositions* to be the objects of our thoughts – which projectivism seems, at least in its treatment of the thought that something is wrong, to be trying to do without.¹

A satisfactory projectivist solution to the Frege-Geach Problem should do many different things. It should be *constructive*, in that it allows us to predict, on the basis of knowing what it is to think that P, and knowing what it is to think that Q, what it is to think that if P, then Q – and similarly for other ways of forming complex sentences. It should also allow us to explain the same sorts of things that other accounts of complex thoughts allow us to explain – for example, why someone who thinks that P and thinks that if P, then Q but does not think that Q is under a certain recognizable sort of rational pressure to either come to think that Q or else either cease thinking that P or cease thinking that if P, then Q. There are other important dimensions to the problem, and roadblocks that can come up along the way, but these give us a good place from which to start.

In *The Language of Morals*, Hare gave reasons to be optimistic that Frege's Abyss could be crossed – since imperatives can be logically complex, and complex imperatives bear predictable logical relationships

¹ See especially Searle [1962] and Geach [1965] for early articulations of this problem.

to one another, there is good reason to think there will be no obstacle to the complexity of moral sentences, or to important and predictable logical relationships between such complex sentences, if moral sentences work in a way that is much like imperatives, Hare believed and argued.

In 1970, Hare again offered an even more direct response to Searle’s articulation of the problem. He argued that the fact that complex sentences don’t express the same attitudes or perform the same speech-acts as their parts is no more obstacle to compositional semantics for the projectivist, than the fact that complex sentences don’t have the same truth-conditions as their parts is, for a more conventional theorist. All that each theorist needs, Hare argued, is a *compositional* theory of the right kind. But Hare didn’t actually tell us how such a compositional theory would work, and his work left it a bit mysterious what such a thing might look like. It was only Blackburn’s work over the next two decades which finally gave us some insight not just into why there might be room for a projectivist way of crossing Frege’s Abyss, but what such a crossing might actually look like.² It is therefore to Blackburn’s seminal work on this project and the subsequent critical work that it provoked, that we owe much of our present understanding of the possible resources for solving this problem.³

I.2 the HOA approach in *spreading the word*

In *Spreading the Word*, Blackburn offered one of the clearest and most explicit strategies for solving the Frege-Geach Problem, focusing on the case of conditional sentences. He starts with the idea that thinking that if P, then Q is a matter of *endorsing* a certain way of being – that of thinking that P only if you think that Q. In order to make this idea more precise, he imagines a language E_{ex} with no evaluative predicates, but with operators ‘B!’ and ‘H!’ (for ‘Boo!’ and ‘Hooray!’) which “attach to descriptions of things to result in expressions of attitude” [1984, 193]. In order to allow us to *refer* to attitudes, and not just express them, this language includes vertical bars ‘|’, which function as an operator taking the expression of attitude as input, and returning a description of the attitude expressed. So, for example, ‘B! (lying)’ expresses the attitude of booing lying, and ‘|B! (lying)|’ is a term referring to this attitude.

Finally, Blackburn introduces “the semi-colon to denote the view that one attitude involves or is coupled with another” [1984, 194]. So ‘H! (|B! (lying)|; |B! (stealing)|)’ expresses hooraying – a kind of *endorsement* – of the state of combining the booing of stealing with the booing of lying. It is this kind of endorsement, Blackburn suggests, that plays the role of the conditional, ‘B! (lying) \rightarrow B! (stealing)’. I call

² This is not to say that there was not other, though less prominent, important work done during the same period.

³ In particular, Schueler [1987], Hale [1993], van Roojen [1996], and Unwin [1999].

such an account a *Higher-Order Attitude* theory, because it treats logically complex thoughts as higher-order attitudes toward the attitudes expressed by their parts.

E_{ex} (the language that Blackburn imagines) does not have any sentential connectives – it only has the term-combining connective ‘;’, the term-forming operator ‘| |’, and the expression of attitudes. But he imagines that it is somehow a more perspicuous language than English, in which moral predicates combine with arguments in order to express attitudes, and logically complex sentences really express endorsement of certain combinations of attitudes. Since E_{ex} is thought of by Blackburn as more perspicuous than English, it tells us something about a language like English, insofar as we can define up a mapping from sentences of English to sentences of this more perspicuous language. This mapping might say, for example, that ‘lying is wrong’ in English maps to ‘B! (lying)’ in E_{ex} , and that ‘if lying is wrong, then stealing is wrong’ in English maps to ‘H! (|B! (lying)|; |B! (stealing)|)’ in E_{ex} .

Before moving on, it is worth getting clear on which features of Blackburn’s *Spreading the Word* account are central, and which are peripheral. For example, once we see that the significance of his semantics turns on constructing such a mapping from the object language on which it is supposed to shed light, we can dispense with the idea that we need to invent a more perspicuous expressive language in the first place. Instead of mapping sentences of English to sentences of E_{ex} , we can simply map them directly to the states of mind which are supposed to be expressed by the sentences of E_{ex} . On this – more direct, I think – way of thinking of things, what our mapping does is to map ‘lying is wrong’ to the attitude of booing lying. We don’t need to invent a special sentence of an invented language that is stipulated to be expressive in function in order to do this; we just need to have a way of referring to the state of mind that is said to be expressed. This is a good thing, in fact, because it means, in contrast to Blackburn’s way of doing things, that the semantics can be fully explained even to someone who insists that she does not understand how expressive sentences work even in the artificial perspicuous expressive language E_{ex} .

I also find it a bit obscure what Blackburn originally meant by saying that he will use “the semi-colon to denote the view that one attitude involves or is coupled with another” [1984, 194]. It is not enough, to think that if P, then Q, to hooray thinking both – that sounds more like conjunction than a conditional. Nor does it help to say that the semi-colon denotes an “input/output function” [194]; that makes it sound dangerously like the right response to a conditional whose antecedent you believe is always to come to accept its consequent, but many times the right rational response is either to give up the antecedent, or to give up the conditional itself. It’s also a bit peculiar, having introduced both the attitude of hooraying and the attitude of booing, to identify accepting a conditional with *endorsing* a certain

connection between attitudes. On an intuitive reading, it makes sense to *endorse* each of two incompatible combinations of attitudes – “that’s swell – and so is that!”

As a result, I have an easier time understanding Blackburn’s *Spreading the Word* proposal in the negative mode. On this picture, the conditional is thought of as expressing a negative attitude – *booing* – toward the state of mind that consists in accepting its antecedent without accepting its consequent. Whereas it is tricky to understand exactly what Blackburn’s semi-colon means, this account only appeals to our independent understanding of conjunction and negation – of combination and lack. And whereas it is a bit unclear on Blackburn’s original account why someone who accepts the premises of a *modus ponens* argument without accepting its conclusion cannot reply, to the charge of irrationality, “sure, I’m all for combining booing lying with booing stealing, but I’m also all for combining booing lying with not booing stealing – it’s great shakes all around, if you just boo lying.” In contrast, when we put the picture in negative terms, someone who accepts the premises of a *modus ponens* argument without accepting its conclusion is booing the very state – combining booing lying with a failure to boo stealing – which she is in. Booing a state that you’re in seems more obviously problematic than hooraying a state that you’re not in.

As a result of these considerations, I find it more perspicuous to describe the sort of view we can find in chapter 6 of *Spreading the Word* in the following way. Rather than inventing a stipulatively expressive language, we will simply state our semantics in a restricted and regimented fragment of English. To keep things simple, we’ll abstract away from many of the syntactic features of natural language, and illustrate things by assuming the language for which we are giving a semantics has the syntactic structure of propositional logic. We’ll use brackets, ‘[]’ around a sentence of our object language as a metalanguage name for the state of mind expressed by that sentence, ‘&’ and ‘~’ as metalanguage abbreviations for ‘and’ and ‘not’, and ‘BOO(X)’ as a metalanguage *name* for the state of booing X. Finally, we’ll use ‘P’ and ‘Q’ to stand in for arbitrary sentences. This gives us a simple theory for a very small fragment of English:

simple HOA

[lying is wrong] = BOO(lying)
[stealing is wrong] = BOO(stealing)
[if P, then Q] = BOO([P]&~[Q])

In other words, this just says that ‘lying is wrong’ expresses the state of booing lying, ‘stealing is wrong’ expresses the state of booing stealing, and for arbitrary sentences ‘P’ and ‘Q’, ‘if P, then Q’ expresses the state of booing the state of being in the state expressed by ‘P’ but not being in the state expressed by ‘Q’.

Although this fragment is very small, it's easy to see how it can be extended, and it is easy to see that its compositional rule can be recursively applied. Although as we've seen this isn't exactly how Blackburn describes his own theory, I think the differences are immaterial for what will be important for us, and that my presentation simply makes it a bit easier to see what is going on.

Before getting to some of the worrisome features of this theory, let me emphasize two features that I think are particularly important. First, this account is *constructive*, in that it genuinely tells us how to determine what it is to think that if P, then Q, so long as we know what it is to think that P, and what it is to think that Q. It doesn't just tell us that there is such a state of mind, or describe it only in terms of other features that it is supposed to have. Rather, it picks it out in the same sort of way that we can pick out the state of mind of *liking Jill*, simply on the basis of putting the attitude of *liking* together with a possible object of that liking – *Jill*.

The second important feature of Blackburn's Higher-Order Attitudes account is that it seeks to genuinely *explain* why it is that someone who both thinks that P and thinks that if P, then Q, is under a certain recognizable sort of rational pressure to either come to think that Q or else either cease thinking that P, or thinking that if P, then Q. According to Blackburn as I've described him, this is because to think that if P, then Q is to boo thinking that P without thinking that Q. So if you think that P, think that if P, then Q, and don't think that Q, then in virtue of thinking that P and not thinking that Q, you are in the very state of mind that you boo, in virtue of thinking that if P, then Q. So you boo your own state of mind.

Since it seems rationally problematic to boo your own state of mind, this puts you under rational pressure to get out of this situation – which you can do either by coming to think that Q, or by either ceasing to think that P or to think that if P, then Q. What is important about this feature of Blackburn's account is that it offers a *real explanation* of why this rational pressure exists – it doesn't just say that the state of mind of thinking that if P, then Q needs to be the right kind in order to give rise to this sort of rational pressure.

2.1 the van rooijen problem

It is possible to raise a number of different problems for Blackburn's *Spreading the Word*-era account as I've described it. For example, it diagnoses the kind of irrationality involved in accepting the premises of a *modus ponens* argument as a special case of *moral akrasia* – booing your own mental state is, by the terms of this very account, just thinking that what you are doing is wrong. It is also very unclear how to pair it with

accounts of ‘not’ and ‘and’ which preserve the relationship we should expect between the conditional and the material conditional.⁴ These, however, are not particularly deep problems for Blackburn’s account, and it is only to be expected that a groundbreaking account will go astray over matters of detail, especially where these details are not central to the main points it was intended to illustrate. So I don’t believe that many of these problems reflect negatively either on the spirit of Blackburn’s central proposal, or on the scope of his accomplishment. There is one important problem, however, which appears on its face to be a perfectly general problem for *any* Higher-Order Attitude account. I call it the van Roojen Problem, because I learned it from Mark van Roojen’s 1996 article, ‘Expressivism and Irrationality’, which originally introduced it.

The allegation at the heart of the van Roojen Problem, as I understand it, is that any theory that is, like Blackburn’s, a *Higher-Order Attitude* theory will *overgenerate* predictions about validity and inconsistency. To see the problem, recall that according to the simple HOA account, to think that something is wrong is just to *boo* it – or, as I have been saying, to *disapprove* of it. Yet according to the simple HOA account, to think that if P, then Q is just to boo thinking that P without thinking that Q. Consequently, the simple HOA account predicts that the thought that if P then Q is the very same thought as the thought that it is wrong to think that P without thinking that Q. Since the simple HOA account explains the inferential properties of sentences solely in terms of the properties of the corresponding states of mind, and since his account identifies these two states of mind, it predicts that ‘if P then Q’ and ‘it is wrong to think that P without thinking that Q’ *will have the same inferential properties*. This prediction is bad.

It is important to note that as we’ve stated it so far, this problem turns on an idiosyncratic feature of the the simple HOA account – the fact that on according to it, conditionals express the attitude of *booing*. As we saw, Blackburn himself said in chapter 6 of *Spreading the Word* that conditionals express the attitude of *hooraying*. But this difference is immaterial. For just as booing is also expressed by the predicate ‘wrong’, hooraying may be expressed by some other predicate. At some points Blackburn suggests that this other predicate may be ‘good’ or ‘great’.⁵ But then Blackburn’s account falls into the same problem. Now, instead of predicting that ‘if P, then Q’ has the same inferential properties as ‘it is wrong to think that P without thinking that Q’, he will predict that it has the same inferential properties as ‘it is good to combine thinking that P with thinking that Q’. This prediction is equally bad.

⁴ See chapter 6 of my *Noncognitivism in Ethics*.

⁵ In ‘Attitudes and Contents’, Blackburn actually introduces the ‘Hooray’ operator by explicit comparison to ‘great’: “Normally, if I make plain to what I feel, say about the Bears, I will most probably do so using a sentence with an ‘expressive’ predicate: ‘the Bears are great!’... Suppose we spoke an ‘emotivist’ language, in which expressions of attitude wore this function on their faces. We would not have the predicative form, to keep such expressions in the indicative mood, but an ejaculatory mood, correspond to that of ‘Hooray for the Bears.’” [1988, p187].

It's no surprise that Blackburn's *Spreading the Word* account faces this problem. Since he had already introduced the attitude of *booraying*, it was convenient for Blackburn, when he introduced the higher-order attitude treatment of the conditional, to appeal to the very attitude that he already had to work with, rather than bothering to introduce a new attitude to play this role in the account. But just because this was convenient for him to illustrate the view, it doesn't follow that this was an essential feature of the approach that he was offering. Fortunately, it is easy to revise Blackburn's view by holding that the higher-order attitude expressed by complex sentences such as conditionals is *different* from that expressed by moral words like 'good' and 'wrong'.

However, this still does not make the problem go away. So long as logically complex sentences express higher-order attitudes, and so long as those attitudes are expressed by *some* predicate or other, each logically complex sentence will have a counterpart, logically simple, sentence, which expresses the very same thought, and hence which has the very same inferential properties. And even if the attitude expressed by logically complex sentences isn't expressed by any *actual* predicate, the mere possibility that it *could be* predicts that there are *possible* logically simple sentences that have the same inferential properties as arbitrary logically complex sentences. This still sounds bad. And Blackburn has no obvious resources at his disposal to explain why it would be impossible for the higher-order attitude that is expressed by a conditional to also be expressed by a predicate.

In his original article, and following the previous influential discussions of Schueler [1987] and Hale [1993], van Roojen characterized his argument as showing that Blackburn's account showed the *wrong kind* of rational tension between accepting the premises of a valid argument and failing to accept its conclusion. Instead of showing that this tension was *logical*, he only showed that there was a broader kind of *pragmatic* tension in these cases. Schueler and Hale had already offered less articulate versions of this complaint, worrying that Blackburn had only shown that there was a *moral* failing, rather than a *logical* failing, in having inconsistent views or failing to follow your premises to their conclusions. But I don't think that it matters whether Blackburn's account establishes the right kind of rational tension. What the van Roojen Problem shows, is that even if it establishes the *right kind* of rational tension, it establishes *too much* of it. There is no valid argument from 'P' and 'it is wrong to think that P without thinking that Q' to 'Q'.

Following early critical discussions by Schueler and Hale, Blackburn took seriously the challenge of trying to say why the kind of inconsistency his account explained was a kind of *logical* inconsistency. In 'Attitudes and Contents', in particular, he did a lot to try to show that the kind of inconsistency his account diagnosed had much in common with familiar ideas about logical inconsistency. But nothing he

said really addressed the heart of the van Roojen Problem. Suppose, following Blackburn in ‘Attitudes and Contents’, that we postulate a higher-order attitude called being ‘tied to a tree’, and replace our compositional rule for ‘if P, then Q’ with the following:

$$[\text{if } P, \text{ then } Q] = \text{TREE-TIED}([\neg P],[Q])$$

According to this rule, to think that if P, then Q is to be in a tree-tied state between thinking that $\neg P$ and thinking that Q. Though we still need a general account of ‘ \neg ’ in order to fully evaluate this proposal (something that Blackburn hinted at, but didn’t fully provide, in ‘Attitudes and Contents’⁶), we know enough to see that the van Roojen Problem still applies. Let ‘is F to’ be a binary predicate that expresses the attitude of being tree-tied between its arguments. Then this account predicts that ‘thinking that $\neg P$ is F to thinking that Q’ will have the same inferential properties as ‘if P, then Q’. This is the same old problem.

Because it can be generalized in this way, the van Roojen Problem looks like a highly general obstacle to Higher-Order Attitude theories – an obstacle which I long believed to be clearly fatal. Consequently, I was not surprised to read and interpret subsequent work on the Frege-Geach Problem from the late 1980’s and later as dispensing with Blackburn’s Higher-Order Attitude approach altogether.⁷ So although I knew that the view described in chapter 6 of *Spreading the Word* was really responsible for allowing all of the rest of us to see, for the very first time, what a Projectivist/Expressivist attempt to cross Frege’s Abyss might even look like, and instrumental in helping the rest of us to understand what sort of standards it might make sense to hold it to, for a long time I congratulated myself that at least this much was clear: it was clearly not on the right track.⁸

⁶ ‘Attitudes and Contents’ relies on one idea about the negations of atomic moral sentences, and a different idea about the negations of complex sentences: “What of other truth-functional contexts? An important feature of inference using propositional calculus embeddings is that they can all be represented by the normal forms of conjunction and disjunction” [1988, p192]. His idea appears to be to first convert any sentence into disjunctive normal form, representing it as a disjunction of conjunctions of atomic sentences and their negations. Actually, if we are to have a mixed language with both normative and descriptive predicates, he will also need a third account of the negations of descriptive atomic sentences. It would unfortunately sidetrack us here to explore how each of these three ideas might be assimilated to a uniform recursive compositional rule for ‘not’.

⁷ In particular, I interpreted Gibbard [1990] and [2003] as offering a very different kind of view from Higher-Order Attitude theories. See chapter 7 of *Noncognitivism in Ethics* for the non-HOA interpretation I gave to Gibbard, and Wedgwood [2010] for a contrary view. It now appears clear (Gibbard [forthcoming]) that Gibbard does intend his view to be given an HOA interpretation.

⁸ Compare especially Schroeder [2010 a].

2.2 finding the right predicate

Though for a long time I believed that the van Roojen Problem was a decisive obstacle to any Higher-Order Attitude Theory, no matter its features, I have become convinced that this was a mistake, and that with the right tools – and given certain important commitments – there is exactly one reasonable solution to this problem. Finding it simply turns on locating the right predicate to express the higher-order attitude.

What makes the van Roojen Problem so highly *general*, as we saw, is that no matter what attitude it holds to be involved in complex thoughts, it will be hard-pressed to avoid the conclusion that that very same attitude could – at least possibly – be expressed by a simple predicate. And if it were, then there would be atomic sentences with the very same inferential properties as complex ones. And that, in the abstract, is what seems false. At least, it seemed false to me for some time. However, it is possible that this seems false only because we haven't been sufficiently imaginative in thinking about what the requisite atomic sentences will look like. Though this looks problematic for familiar transparently normative predicates like 'wrong', 'good', and 'irrational', perhaps we haven't widened our horizons sufficiently. That is the idea my solution to the van Roojen Problem exploits.

Since the van Roojen Problem predicts that there will be an atomic sentence with the same inferential properties as the sentence, 'if P, then Q', it can be solved by locating such a sentence. If its predicate expresses the attitude employed by the Higher-Order Attitude theory, and its subject refers to the state of mind appealed to by the Higher-Order Attitude theory, then the proponent of the HOA theory may claim that though she does indeed have the commitments about validity predicted by van Roojen's argument, that is not a *problem*, because these commitments are in fact intuitively plausible.

Fortunately for the Higher-Order Attitude theorist, there is indeed an atomic sentence which plausibly shares the same inferential properties as 'if P, then Q': 'it is false that P and not Q'. In general, 'it is false that P' and 'not P' are inferentially equivalent. But 'not (P and not Q)' is just the definition of the material conditional. So at least on the material conditional interpretation, that 'if P, then Q' should have exactly the same inferential properties as 'it is false that P and not Q' is exactly what we should expect. But 'it is false that P and not Q' is plausibly interpreted as having an atomic structure – predicating 'false' of 'that P and not Q'. So it is not so strange, after all, that a logically complex sentence should turn out to have the same inferential properties as an atomic sentence.

Indeed, it is not so strange that *arbitrary* logically complex sentences should turn out to do so. For any logically complex sentence, 'P', is equivalent to 'it is false that not P'. So the prediction of the van

Roojen Problem that every complex sentence will be matched with an atomic sentence that shares its inferential properties is not so peculiar after all. We just needed to know where to look.

Of course, the van Roojen Problem didn't just predict that any HOA theory would be committed to holding that there are atomic sentences with the same inferential properties as arbitrary complex sentences; it also predicted that the predicates of these sentences would express the very same attitude appealed to in the theory, and their subjects would refer to the object of this attitude. So to adopt this solution, the Higher-Order Attitude theorist needs to do two things. She needs, first, to say that 'false' and 'not' express the very same attitude, merely being realized in syntactically different ways. And she needs, second, to say that 'that' clauses like 'that P and not Q' refer to states of mind – the kinds of things which the attitude expressed by 'false' and 'not' takes for its objects. I'll return to the significance of each of these assumptions shortly. But first, I'll sketch a simple HOA theory that incorporates each of these elements.

3.1 a toy theory

In this section for concreteness I'll sketch a toy HOA theory that incorporates each of the elements I've described. In this HOA theory, the higher-order attitude expressed by 'false' and 'not' is called 'disagreement', and abbreviated 'DIS'. I'll state the theory as a compositional semantics for a simple language \mathcal{L} ; the idea being that if it works for \mathcal{L} , then it can be generalized to work for a natural language like English. \mathcal{L} contains 1-place predicates 'true', 'false', 'wrong', and 'common', simple referring terms 'stealing', 'walking', and 'killing', and a complex term-forming operator 'that()', which forms terms out of sentences. Atomic sentences are formed by prefixing a predicate to a term enclosed in parentheses, and complex sentences are formed by the one-place connective '¬' and the two-place connectives '∧', '∨', and '→' in the usual way.

Because \mathcal{L} is an expressivist language, its semantics consists in a mapping [] from sentences 'P' to mental states [P]. As before, if M and N are mental states, then M&N is the state of being in both M and N. I'll also use 'α' to stand in for an arbitrary term and 'P' and 'Q' to stand in for arbitrary sentences. DIS(M) is the state consisting of bearing the attitude of disagreement toward mental state M, BE(x is common) is the state consisting of believing of x, that it is common, and BOO(x) is the state consisting of booing x. The semantics is as follows:

\mathcal{L}
terms
 'stealing' refers to stealing

‘walking’ refers to walking
‘killing’ refers to killing
‘that(P)’ refers to [P]

predicates

If ‘ α ’ refers to x , then $[\text{wrong}(\alpha)] = \text{BOO}(x)$
If ‘ α ’ refers to x , then $[\text{common}(\alpha)] = \text{BF}(x \text{ is common})$
If ‘ α ’ refers to x , then $[\text{false}(\alpha)] = \text{DIS}(x)$
If ‘ α ’ refers to x , then $[\text{true}(\alpha)] = \text{DIS}(\text{DIS}(x))$

connectives

$[\neg P] = \text{DIS}([P])$
 $[P \vee Q] = \text{DIS}(\text{DIS}([P]) \& \text{DIS}([Q]))$
 $[P \wedge Q] \equiv [\neg(\neg P \vee \neg Q)] = \text{DIS}(\text{DIS}(\text{DIS}(\text{DIS}([P])) \& \text{DIS}(\text{DIS}([Q])))$
 $[P \rightarrow Q] \equiv [\neg(P \wedge \neg Q)] = \text{DIS}(\text{DIS}(\text{DIS}(\text{DIS}(\text{DIS}([P])) \& \text{DIS}(\text{DIS}(\text{DIS}([Q])))))$

The semantics I’ve just sketched for \mathcal{L} is a thoroughlygoingly Higher-Order Attitude account.⁹ Every logically complex sentence expresses a higher-order attitude toward the states of mind expressed by its parts. As the van Roojen Problem predicts, this semantics predicts that there is an atomic sentence which expresses the very same state of mind as ‘if P, then Q’. It is ‘ $\text{false}(\text{that}(P \wedge \neg Q))$ ’. But as we’ve noted before, this does not look like a particularly problematic prediction, after all. So by taking the right view about ‘false’, and the right view about the referents of ‘that’ clauses, this semantics for \mathcal{L} illustrates how it is possible to endorse a Higher-Order Attitude account without the resulting commitments being transparently problematic.

The semantics sketched in this section is not finished with the Frege-Geach Problem. To provide a complete treatment of the logical relationships between sentences, it needs to be able to prove that the right rational relationships hold between states of disagreement to back up the inferential relationships between the corresponding sentences, and it must defend the assumptions required for this proof. This goes for any higher-order attitude theory. But what interests me in this paper is the shape of the response to the van Roojen Problem, and so I will move on.

⁹ With respect to the connectives, it is very similar to the HOA account sketched in Wedgwood [2010]. Wedgwood’s clause for conjunction is somewhat different; whereas I’ve taken disjunction as basic and introduced the clause for conjunction that is the result of defining conjunction from de Morgan’s laws, Wedgwood gives parallel basic definitions for each that are designed to make de Morgan’s laws fall out of his account of logical consequence. Note that neither my account nor Wedgwood’s identifies believing a conjunction with believing both conjuncts.

3.2 the quasi-redundancy theory of ‘false’

As noted, the solution I’ve offered to the van Roojen Problem requires two parts: first, assuming that ‘false’ and ‘not’ express the very same attitude, and second, assuming that ‘that’ clauses denote states of mind. In part 4 I want to consider the significance of the assumption about ‘that’ clauses, but first it’s worth examining the Blackburnian spirit and independent appeal of this idea about falsity.

One important theme throughout much of Blackburn’s work is the idea that we need a deflationary treatment of truth, and that offering such a treatment is a key element in the project of defending a projectivist/expressivist view without committing to a problematic error theory. The idea about ‘false’ to which my solution has needed to appeal is a deflationary one. In fact, by identifying the meaning of ‘it is false that P’ with the meaning of ‘not P’, it has much in common with the so-called *redundancy* theory of truth – a paradigmatically simple deflationist theory of truth.

But there are two important respects in which my treatment of ‘false’ is better than the redundancy theory of truth. The ordinary redundancy theory of truth identifies the meaning of ‘it is true that P’ with that of ‘P’, and is exhausted by this identification. It faces two important problems, one of which arises from this identification, and the other of which arises from the fact that it is exhausted by this identification. The first problem is that even though ‘it is true that P’ and ‘P’ are inferentially equivalent, it certainly seems *possible* to think one without thinking the other. But if thinking that it is true that P and thinking that P are the very same thought, then this is not possible. So it looks like it would be a mistake to identify these two thoughts.

The account of falsity that my solution to the van Roojen Problem relies on offers a similar identification – identifying thinking that it is false that P with thinking that not P. Is this identification equally problematic? I don’t think so. It’s much less obvious that it is possible to think that not P without thinking that it is false that P, or to think that it is false that P without thinking that not P, than it is for the case of truth. This is less obvious, in part, because many of the ways that we have of forming negative sentences in natural language actually look like they need the notion of truth or falsity, to begin with. Although sentences of formal language may easily be formed with wide-scope negations, in natural language we often need to say things like ‘it is not the case that’, in order to scope ‘not’ over conjunction, disjunction, or a conditional. But what does ‘the case’ mean? It is just a synonym of ‘true’. So ‘it is not the case that P’ is just a way of saying that it is not true that P – i.e., that it is false that P. So it is much less obviously a mistake to identify the meanings of ‘it is false that P’ and ‘not P’, than to identify the meanings of ‘it is true that P’ and ‘P’. This looks like a mark in the favor of the foregoing account.

A second and deeper problem arises for the redundancy theory of truth on account of its being *exhausted* by the identification of the meaning of ‘it is true that P’ with that of ‘P’. This is because sentences like ‘it is true that P’ are not, in fact, the only or even the most useful sort of sentences involving the word ‘true’. On the contrary, the most useful sentences are ones that *quantify into* the argument-place of ‘true’ – sentences like ‘what she said is true’, ‘that’s true’, and ‘everything the Bible says is true’. But the redundancy theory of ‘true’ offers us no way of seeing what these sentences mean, because its identification only applies to sentences of the form, ‘it is true that P’.

In contrast, although the foregoing treatment of ‘false’ *predicts* that ‘it is false that P’ and ‘not P’ have the same meaning, it is not *exhausted* by this prediction. On the contrary, it offers a general account of the meaning of ‘false’ in *arbitrary* sentences – it expresses the attitude of disagreement toward its argument. For this reason, the account can readily deal with sentences like ‘what she said is true’, ‘that’s true’, and ‘everything the bible says is true’. All that this treatment requires, is the introduction of quantifiers and other referring terms into the language. Again, this appears to have some promise.

In contrast to many deflationist theories of truth, therefore, the deflationist theory that is forced on us by the solution to the van Roojen Problem does not merely assert that ‘true’ and ‘false’ play certain inferential roles, or that certain biconditionals are analytic of ‘true’ and ‘false’. In contrast to Paul Horwich’s [1990] claim that deflationist accounts cannot offer explanations of the T-schema, this account is a genuinely *explanatory* theory, but still recognizably deflationist, because the explanations it offers recognizably appeal to no inflationary notions or picture of what it is for something to be true or false. It succeeds in this because the explanations that it offers are in the projectivist/expressivist vein.

Because this theory has these attractions going for it in its own right, therefore, and because these attractions are so patently Blackburnian in spirit, I take it that this is a sign of promise for this solution to the van Roojen Problem. We must now turn, however, to the identification of the referents of ‘that’ clauses required by this solution. Is this defensible? Is it Blackburnian?

4.1 propositions: the blackburnian stance

In many of his writings, Blackburn has suggested that it is part of his project to ‘earn the right’ to the ‘propositional surface’. This talk presents a picture on which though it is legitimate to talk as if there are propositions, this is a way of talking that is not really employed when we properly do theory. At some points Blackburn talks as though proposition talk is just functional talk – the reflection of isomorphic relationships that hold among different sorts of things. For example, in *Ruling Passions* he writes, “it is the

isomorphism between propositional structures and necessary practical states that is the heart of things” [1998, p77].

Blackburn’s idea that talk of moral propositions just reflects the fact of an isomorphic relationship is mirrored by his apparent concession that even the sentential connectives may not be strictly univocal, presumably because they do different things with moral and non-moral sentences. For example, at the close of ‘Attitudes and Contents’, he writes,

But it is unnecessary to claim that we make no jump at all. That would involve, for instance, defending the claim that negation is absolutely univocal as it occurs in $\sim H!p$ and in $\sim p$, and similarly for the other constants. *But this need not be claimed.* All we have is sufficient similarity of logical role to make the temptation to exploit *ordinary* propositional logic quite irresistible – and that is why we do. [197]

Here Blackburn seems to allow that the use of connectives like ‘not’ and ‘if...then’ is appropriate because of an isomorphism in inferential relationships, rather than because of strict univocality.

On this picture, ‘proposition’ is an honorific which we bestow on language which can embed under so-called ‘truth-conditional’ connectives, and of which we can ascribe the predicates ‘true’ and ‘false’. Bestowing this honorific, moreover, allows us to say things like ‘the proposition as to whether stealing is wrong has been put before the court’, allowing us to not only make sense of the embeddability of moral language, but of the very trappings of language that is sometimes used to distinguish ‘realist’ discourse.

But all of this, on the picture that Blackburn has long defended, is made possible because of the availability of a merely *deflationary* way of reading talk about propositions. On this picture, when we construct our theory that accounts for moral language and thought, we need employ no reference to moral propositions in order to do so. But part of what we explain, in offering this theory, is why it is okay to use talk about propositions in ordinary discourse – in other words, the word ‘proposition’ is among the words, like ‘good’, ‘wrong’, and ‘true’, whose use our theory legitimizes, though we do not use them in constructing the theory.

However, the solution I’ve been offering Blackburn to the van Roojen Problem for Higher-Order Attitudes accounts relies on a particular identification of the referents of ‘that’ clauses. And ‘that’ clauses have traditionally been characterized as referring to propositions. This means that my solution to the van Roojen Problem appears to rely on a particular theory about what propositions are. Not just a way of *talking* about moral propositions, but an out-and-out theory of what they are – something that Blackburn *seems* to have been trying to avoid.

It is no coincidence, moreover, that ‘that’ clauses have traditionally been characterized as referring to propositions. Propositions, after all, are theoretical entities that were originally introduced in terms of their theoretical role as the objects of the attitudes and (primary) bearers of truth and falsity. When Jack says what he believes but Jill disagrees with that, there is something that Jack says and believes, and is what Jill disagrees with – a common object of Jack’s saying, Jack’s believing, and Jill’s disagreement. Jack thinks it is true, and Jill thinks it is false, so the very thing Jack says and with which Jill disagrees needs to be the kind of thing that can be true or false, or at least of which truth and falsity are predicated. But these adjectives and verbs – ‘says’, ‘believes’, ‘true’, and ‘false’ – take ‘that’ clauses as among their arguments. So if propositions are the objects of these attitudes, they must be the referents of ‘that’ clauses. Certainly if ‘that’ clauses *do* have referents and ‘says’, ‘believes’ and ‘true’ are predicated of those referents, then there are propositions.

4.2 against the blackburnian stance: the theoretical roles of propositions

In contrast to Blackburn, I believe that the sophisticated proponent of a projectivist/expressivist picture not only need not, but should not, eschew appeal to propositions – moral and otherwise – in the construction of her theory. What she *does* need to do, is to offer a non-standard picture of what propositions *are*. Whereas the realist may hold that propositions mark out distinctions in some sort of metaphysical reality, the projectivist should insist that though *some* propositions may do this, others may not.

The reasons the projectivist should not eschew propositions are simple. They are all of the same reasons that propositions have always been postulated by theorists, in the first place. The explanatory roles that propositions play in theory cannot be played by merely deflationary talk about propositions, and so without real propositions to play these roles, the theorist who seeks to dispense with appeal to propositions, in the course of spelling out her theory, gets into a number of difficult problems.

So what explanatory roles do propositions play in theory? First and most simply, as the referents of ‘that’ clauses, they serve as the range for quantifiers into the argument places of ‘true’, ‘said’, and ‘believes’, explaining why when Jack says that P and it is true that P, it follows that there is something that Jack said and is true – namely, that P. Second and much more importantly, they explain why if it is possible to believe that P and possible to hope that Q, it is possible to believe that Q and possible to hope that P – and similarly for each of the other attitudes – wanting, wondering, preferring, being pleased that, being confident that, and more.

Without propositions to play the role of objects of the attitudes, the complicated projectivist task of explaining what it is to have moral beliefs replicates itself as the task of explaining what it is to have moral hopes, or other attitudes with moral contents. Similarly, the complicated Frege-Geach task of explaining what it is to have logically complex moral beliefs replicates itself in the tasks of explaining what it is to have logically complex attitudes of every other kind. Moreover, the force of these problems quickly explodes, given that just as we also need to be able to explain why moral and non-moral belief seem to have important enough features in common to merit both being called ‘beliefs’, we also need to explain why moral and non-moral hopes seem to have so much in common, and so on for all of the rest.¹⁰

Propositions also play the role of the arguments of modals, including epistemic modals like ‘might’ and ‘must’. In much of his work, Blackburn emphasizes the importance of distinguishing between claims like ‘if I fully reflect under full information, I will disapprove of stealing’ and ‘stealing is wrong’. But when he considers what it is to think that one might be in error about one’s own moral views, the only sense he is able to give to this thought, is to feed a non-moral content to ‘might’ – so that this is cashed out as the thought that one might change one’s mind, if one is to sufficiently reflect under full information. Since Blackburn would never countenance conflating ‘I think that P’ with ‘P’ outside the scope of ‘might’, however, it is hard to see why he should think this is a good idea inside the scope of ‘might’, and indeed Andy Egan [2007] has taken him to task for precisely this. This conflation could be easily avoided, if only we had moral propositions on which for ‘might’ to operate.¹¹

All of these explanatory roles for propositions require assuming, when we articulate our theory (and not just after we help ourselves to what it allows us to say) that there is a domain of things which play the role of being the arguments of attitude verbs and of which truth and falsity are predicated. But none of these explanatory roles for propositions require that these things cannot fit seamlessly into a projectivist/expressivist picture. The projectivist/expressivist simply needs a non-standard view about what propositions *are*. And the solution I’ve offered to the van Roojen Problem articulates such a view. It says that the proposition that P – the referent of ‘that P’ – is just the mental state expressed by ‘P’. Since the projectivist’s account of the mental state expressed by ‘P’ is not problematic in projectivist terms, it follows that the proposition itself is not problematic in projectivist terms, either.

¹⁰ For an example of an expressivist attempt to characterize what moral belief and non-moral belief have in common, see Horgan and Timmons [2006].

¹¹ For much more extensive and careful discussion of these and other explanatory roles for propositions, see Schroeder [forthcoming].

4.3 is this enough?

So far, I've been arguing that for independent reasons, even a projectivist/expressivist needs something to play the role of propositions, and needs a non-standard, projectivist-friendly, account of what these things are. Since my solution to the van Roojen Problem automatically commits itself to a projectivist-friendly answer to this question, I think this is again a sign of its promise, at least from the projectivist's point of view. A committed sophisticated projectivist *should*, I believe, get over Blackburn's hesitation about propositions, and attribute it to a hesitation about the kind of thing *other* theorists hold to play the theoretical role of propositions. (If Blackburn's hesitation about propositions is really so attributable, then this solution is still in the Blackburnian spirit after all.)

But before we get too excited about the prospects for my solution to the van Roojen Problem, it is worth asking whether the resulting account of the nature of propositions allows them to play their role *well*. And here I still have to express serious reservations.¹² If 'that' clauses denote the mental states expressed by their complements, then the relation expressed by the verb 'believes' will just be the instantiation relation. But hoping that P is not a matter of instantiating the mental state expressed by 'P'. It is a matter of being in some other state. So the HOA theorist who takes this task will have to treat all attitudes other than belief as attitudes *toward* beliefs. This gives us a very strong and I think surprising asymmetry between belief and other attitudes. It is striking that it should turn out that the hope that $\neg P$ is related to the hope that P in a very different way from that in which the belief that $\neg P$ is related to the belief that P.

However, it may be that Blackburn himself will not find this prediction so unsurprising. On page 70 of *Ruling Passions* he considers a few attitudes other than belief – wondering, denying, being undecided, and being certain. And he construes each of these as an attitude toward belief. So perhaps the view I've been describing is not that different from Blackburn's, after all.¹³

references

- Blackburn, Simon [1984]. *Spreading the Word*. Oxford: Oxford University Press.
_____[1988]. 'Attitudes and Contents.' Reprinted in Blackburn [1993], 182-197.
_____[1993]. *Essays in Quasi-Realism*. Oxford: Oxford University Press.
_____[1998]. *Ruling Passions*. Oxford: Oxford University Press.

¹² Compare Dunaway [2010].

¹³ Special thanks to Allan Gibbard, Scott Sturgeon, Mark van Roojen, Michael Smith, and Robert Johnson.

- Dunaway, William [2010]. 'Minimalist Semantics in Meta-Ethical Expressivism.' *Philosophical Studies* 151(3):
- Egan, Andy [2007]. 'Quasi-Realism and Fundamental Moral Error.' *Australasian Journal of Philosophy* 86(2): 205-219.
- Geach, P.T. [1965]. 'Assertion.' *Philosophical Review* 74: 449-465.
- Gibbard, Allan [1990]. *Wise Choices, Apt Feelings*. Cambridge: Harvard University Press.
- _____ [2003]. *Thinking How to Live*. Cambridge: Harvard University Press.
- _____ [unpublished]. *Thoughts and Oughts*. Unpublished book manuscript.
- Hale, Bob [1993]. 'Can There Be a Logic of Attitudes?' In Haldane and Wright, eds., *Reality, Representation, and Projection*. New York: Oxford University Press.
- Hare, R.M. [1952]. *The Language of Morals*. Oxford: Oxford University Press.
- _____ [1970]. 'Meaning and Speech Acts.' *The Philosophical Review* 79(1): 3-24.
- Horgan, Terry, and Mark Timmons [2006]. 'Cognitivist Expressivism.' In Horgan and Timmons, eds., *Metaethics After Moore*. Oxford: Oxford University Press.
- Horwich, Paul [1990]. *Truth*. Oxford: Oxford University Press.
- Schroeder, Mark [2008] *Being For: Evaluating the Semantic Program of Expressivism*. Oxford: Oxford University Press.
- _____ [2010a]. *Noncognitivism in Ethics*. New York: Routledge.
- _____ [2010b]. 'How to be an Expressivist About Truth.' In Cory Wright and Nikolaj Jang Peterson, eds., *New Waves in Truth*. New York: Palgrave MacMillan, 282-298.
- _____ [forthcoming]. 'Two Roles for Propositions: Cause for Divorce?' Forthcoming in *Noûs*.
- Schueler, G.F. [1988]. 'Modus Ponens and Moral Realism.' *Ethics* 98(3): 492-500.
- Searle, John [1962]. 'Meaning and Speech Acts.' *Philosophical Review* 71:423-432.
- van Roojen, Mark [1996]. 'Expressivism and Irrationality.' *The Philosophical Review* 105(3): 311-335.
- Unwin, Nicholas [1999]. 'Quasi-Realism, Negation and the Frege-Geach Problem.' *The Philosophical Quarterly* 49(196): 337-352.
- Wedgwood, Ralph [2010]. 'Schroeder on Expressivism: For or Against?' *Analysis* 70(1): 117-129.