Inferential Expressivism and the Negation Problem

Luca Incurvati & Julian J. Schlöder

Department of Philosophy / ILLC

University of Amsterdam

Abstract

We develop a novel solution to the negation version of the Frege–Geach problem by taking up recent insights from the bilateral programme in logic. Bilateralists explain the meaning of negation in terms of a primitive *B-type* inconsistency involving the attitudes of assent and dissent. Some may demand an explanation of this inconsistency in simpler terms, but we argue that bilateralism's assumptions are no less explanatory than those of *A-type* semantics that only require a single primitive attitude, but must stipulate inconsistency elsewhere. Based on these insights, we develop a version of B-type expressivism called *inferential expressivism*. This is a novel semantic framework that characterises meanings by inferential roles that define which *attitudes* one can *infer* from the use of terms. We apply this framework to normative vocabulary, thereby solving the Frege–Geach problem generally and comprehensively. Our account includes a semantics for epistemic modals, thereby also explaining normative terms under epistemic modals.

1   Introduction

Expressivism about a class of linguistic terms holds that the meaning of these terms is to be explained by appealing to the mental *attitudes* that these terms are used to express. Historically, expressivism has been applied to normative terms by taking, say, *wrong* to express the

non-cognitive (desire-like) attitude of *disapproval*[1] (Hare 1952), but has more recently been extended to other terms (e.g. Humberstone 2000; Schnieder 2010; Yalcin 2012).

Expressivism has several attractive features. It replaces representational (and hence meta-physically loaded) explanations of meaning with non-representational ones, by linking meaning with our psychological lives. In addition, ethical expressivism has the capacity to explain disagreement phenomena and the motivating power of moral judgements.

No matter its area of application, expressivism must however contend with the *Frege–Geach embedding problem* (Frege 1919; Geach 1965; Schroeder 2008b). The goal of this paper is to use insights from expressivism about negation to develop a novel form of expressivism about normative terms. As we shall show, this novel form of expressivism has the resources to solve the Frege–Geach problem generally and comprehensively.

In its original version, the Frege–Geach problem challenges the expressivist to explain how a term that putatively expresses an attitude can embed in conditional antecedents. However, current debate centres on the *negation version* of the Frege–Geach problem (Unwin 1999; 2001; Schroeder 2008a). In brief, it goes as follows. The negation of a descriptive term appears to be different from the negation of a term treated in expressivist terms. To wit, the two normative claims *Murdering is wrong* and *Murdering is not wrong*, if receiving expressivist treatment, must be read as expressing *different attitudes* towards the *same content*. Thus, it seems that *not* here modifies attitudes. However, in the descriptive claims *Grass is green* and *Grass is not green*, the term *not* modifies the *content* of what is claimed. The expressivist is challenged to explain these facts without any *ad hoc* stipulations. The problem permeates semantic analysis on different levels; we discuss its different aspects in Section 2.

We propose to solve the negation problem by adopting the framework of *inferential expressivism*. Inferential expressivism is a general semantic framework that has its roots in the bilateral programme in logic (Rumfitt 2000; Incurvati and Schlöder 2017). Very briefly, expressivism explains terms by reference to the attitudes they are used to express, inferentialism explains terms by reference to the inferences they feature in, and inferential expressivism explains terms by reference to what attitudes one can infer from their use. The possibility of com-

---

[1]We stipulatively call the attitude expressed by *wrong* 'disapproval', but any other term would do.

bining expressivism with inferentialism is acknowledged (Chrisman 2008; Baker and Woods 2015; Tiefensee 2016), but we are the first to develop it systematically and give the formal details of an inferential expressivist account of normative language. Moreover, our inferential expressivism includes an account of epistemic modals (Incurvati and Schlöder 2019). Building on this prior work, our account also explains how epistemic modals interact with normative vocabulary, improving on Mark Schroeder's (2008a) analysis, which encounters difficulties with sentences like *Murdering might be wrong*.

Our inferential expressivism allows us to locate the difference between normative and descriptive vocabulary *entirely* at the level of lexical semantics. On our account, uttering a declarative sentence standardly expresses *assent* to a proposition, where the meaning of that proposition is given by inference rules. Descriptive terms have *attitude invariant* inferential semantics; e.g., the semantics of *bachelor* is such that from the expression of assent to ⌜*John is a bachelor*⌝ one can infer assent to ⌜*John is unmarried*⌝.[2] In contrast, normative terms have *attitude changing* inferential semantics; the semantics of *wrong* is such that from assent to ⌜*murdering is wrong*⌝ one can infer the different attitude of *disapproval* towards ⌜*murdering*⌝. This is the only difference needed to define inferential expressivist semantics for logical operators—in particular, negation—that behave as they should.

In the next section we describe the different aspects of the negation problem in detail. In Section 3, we introduce inferential expressivism, extend it to a semantics for normative vocabulary, and discuss how this solves the negation problem. In Section 4, we compare our approach to Schroeder's (2008a) and argue that we improve on his account by solving what he calls the *new new negation problem* concerning normative sentences involving epistemic modals. We conclude in Section 5 by locating inferential expressivism within the metaethical landscape.

---

[2]To distinguish sentences from what they express, we encase contents one can have attitudes (both cognitive and non-cognitive) towards in Gödel brackets; e.g. asserting *John is unmarried* expresses assent to the proposition ⌜*John is unmarried*⌝. Formulae and variables generally stand for such contents and are therefore not bracketed.

2 The negation problem

Expressivists claim that a sentence like *Murdering is wrong* (1a) is used to express the attitude of *disapproval* towards the act of murder. The expressivist problem with negation is to explain what is expressed by *Murdering is not wrong* (1b). There is one clear constraint on any such explanation: to utter both (1a) and (1b) sounds *inconsistent* and this ought to be explained.[3]

(1) a. Murdering is wrong.

    b. Murdering is not wrong.

    c. Not murdering is wrong.

    d. Not murdering is not wrong.

Which attitude is expressed by (1b)? It cannot be *disapproval of not murdering*, since that is what one expresses by uttering (1c). It is a mistake to equate (1b) with (1c), since (1c) is inconsistent with (1d), whereas (1b) is not. Moreover, (1b) cannot express the *absence* of *disapproval of murdering*, as this would wrongly conflate (2a) and (2b).

(2) a. Jane does not think that murdering is wrong. = Jane does not disapprove of murdering

    b. Jane thinks that murdering is not wrong.   = ?

    c. Jane thinks that murdering is wrong.   = Jane disapproves of murdering

    d. Jane thinks that not murdering is wrong.   = Jane disapproves of not murdering

Nick Unwin (1999; 2001) notes that the problem is really quite simple: in *Jane does (not) think that (not) murdering is (not) wrong* one can insert up to three negations, and no two choices should be equivocated. But the expressivist paraphrase *Jane does (not) disapprove of (not) murdering* only has space for two, so the expressivist requires an attitude other than disapproval to describe what attitude is expressed by (1b). One may—by stipulation—call this attitude *tolerance* and describe (1b) as expressing tolerance of murder. However, according to Schroeder (2008a), this leaves an explanatory gap. The expressivist needs to explain the inconsistency of *Murdering is wrong* and *Murdering is not wrong*. She may attempt to do so

---

[3]Here, 'inconsistent' is used pre-theoretically. We discuss the different relevant notions of inconsistency in due course.

by claiming that it is inconsistent to simultaneously tolerate and disapprove of the same act, in much the same way that it is inconsistent to hold the attitude of belief towards two inconsistent contents or the attitude of intent towards two incompatible goals.

Schroeder (2008a), however, argues that this does not suffice. Any putative such explanation appears to entail that the expressivist is committed to there being two ways for attitude expressions to be inconsistent. First, *A-type* inconsistency, which arises from holding a single kind of *inconsistency-transmitting* attitude, such as belief, towards truth-conditionally incompatible contents (e.g. it is A-type inconsistent to believe both *p* and *not p*). Second, *B-type* inconsistency, which arises from holding different kinds of attitude that are in some sense mutually exclusive towards the *same* content. The supposed explanation of the inconsistency of (1a) and (1b)—that they express, respectively, disapproval and tolerance of the same act—is B-type. A-type inconsistencies arise in descriptive language: the propositional contents of (3a) and (3b) are incompatible and hence it is inconsistent to express belief in both.

(3) a. Grass is green.

   b. Grass is not green.

A-type inconsistency feels familiar, whereas B-type inconsistency may appear to be an artifact of expressivism. The expressivist cannot, Schroeder (2008a) contends, create the attitude of tolerance and its B-type properties by *fiat*. Rather, she needs to explain *why* there is an attitude that behaves like what she calls *tolerance*—by appealing to more fundamental concepts. To find these concepts or to explain away the complaint is the *negation problem*.

The expressivist may point to natural cases of B-type inconsistencies. For example, *believing p* and *wondering whether p* seem to be B-type inconsistent attitudes, regardless of whether expressivism about any particular class of terms is correct. Baker and Woods (2015) discuss further natural examples of B-type inconsistencies and conclude that they are no less familiar than A-type inconsistencies. May the expressivist, therefore, turn the negation problem on its head? We are inclined to take the inconsistency of *Grass is green* and *Grass is not green* to display the inconsistency-transmitting property of belief. The expressivist might claim that the very data constituting the negation problem show (i) that *Murdering is wrong* and *Murdering is not wrong* express different types of attitude, and (ii) that these attitudes are B-type incon-

sistent. Unfortunately for the expressivist, this will not suffice. At least the following problems remain.

The compositional problem. The contrast between (1) and (3) brings to light the *compositional* version of the negation problem. For these examples appear to show that expressivism requires two types of sentential negation: an A-type negation operating on *contents*, like the *not* in (3b), and a B-type negation operating on *attitudes*, like the *not* in (1b). This division however is not reflected in language—there are no *A-not* or *B-not* realised as distinct lexical items. Thus, one should give a uniform analysis of the particle *not* as a single compositional operator. It is unclear at this point how the expressivist could do this. It is also a mistake to take *not* to be *ambiguous* between two distinct operators not_A and not_B. In *Murdering is not wrong*, the B-reading is mandatory and the A-reading is not available. If an ambiguity is postulated, the unavailability of the A-type reading is still in need of explanation.

The logicality problem. We have observed that *Murdering is wrong* and *Murdering is not wrong* sound inconsistent, but we have been using the term 'inconsistency' quite loosely. But *Murdering is wrong* and *Murdering is not wrong* are inconsistent in a precise sense: they are *logically* inconsistent, meaning that it is sufficient to understand the meaning of *not* to recognise them as inconsistent; one may be entirely ignorant about the meaning of *murdering* or *wrong*. This is in contrast with *factual* inconsistencies, like *Grass is green and grass is blue*. To recognise these as inconsistent, one needs knowledge of the non-logical vocabulary featuring in them. However, the rough expressivist explanation of the relevant inconsistency rests on the recognition that *not* modifies the attitude expressed by *wrong* and that *not* does not modify attitude in every context. So it seems that some understanding of *wrong*—at least that it *is* an attitude expression—is required to recognise the inconsistency of *Murdering is wrong* and *Murdering is not wrong*, contravening logicality.

The explanatory problem. The expressivist has made the following assumptions: (i) there are B-type inconsistencies; (ii) for each attitude whose expression is linguistically realised (e.g. *wrong* realising the expression of *disapproval*) there is another attitude (e.g. *tolerance*) that

may not be linguistically realised; (iii) all these attitudes stand in appropriate logical relations to one another (e.g. disapproval and tolerance towards the same act are B-type inconsistent). Taken together, these assumptions commit the expressivist to the existence of a set of attitudes standing in a reasonably complex *web* of logical relations. The expressivist cannot just assume that this web exists, but must explain where it comes from.

An account that must contend with the explanatory problem is due to Terence Horgan and Mark Timmons (2006). They construct a language for *logically complex* attitude expressions, starting with two basic sentence-forming operators to express, respectively, descriptive is-commitment and normative ought-commitment. They then recursively define logical connectives on these operators, e.g. that for each operator there is another one that negates it. The meanings of these complex operators are given by their inferential role. However, Schroeder (2008a) incisively points out, commitments (or attitudes) cannot themselves be logically complex—only the language that Horgan and Timmons use to describe these commitments is logically complex. But a logically complex language might be overwrought and does not in itself guarantee that its expressions denote anything, so Horgan and Timmons have yet to explain why we should suppose the attitudes their language requires exist.

3 Inferential expressivism

3.1 A first pass

Inferential expressivism is an approach to meaning that combines aspects of expressivism and inferentialism. *Inferentialism* holds that the meaning of a term is given by its inferential role—that is, by when one can infer sentences containing the term, and what follows from such sentences. For example, inferentialists claim that the following *rules of inference* characterise the meaning of conjunction (Gentzen 1935; Prawitz 1965).

- From A and B infer $A \wedge B$.

- From $A \wedge B$ infer A.

- From $A \wedge B$ infer B.

We will write such inference rules in a natural deduction calculus. In this notation, the following rules for *conjunction introduction* and *conjunction elimination* give the meaning of $\wedge$.

$$(\wedge\text{I.}) \ \frac{A \quad B}{A \wedge B} \qquad (\wedge\text{E.}_1) \ \frac{A \wedge B}{A} \qquad (\wedge\text{E.}_2) \ \frac{A \wedge B}{B}$$

Expressivism and inferentialism share an intellectual heritage. Both are motivated by supposing that, in some sense, *meaning is use*. Expressivism holds that the expression of certain attitudes is central to how a term is used (Gibbard 2003). Inferentialists claim that inferential roles best characterise use (Dummett 1991; Brandom 1994). Inferential expressivism combines these views by specifying the meaning of an expression in terms of *the attitudes* one can *infer* from the use of sentences that contain that expression (Chrisman 2008; Incurvati and Schlöder 2019).

Let's start with the attitude of *assent*. Following Rumfitt (2000), write $+A$ for the expression of assent to the proposition A. Then, the inferential expressivist meaning for conjunction is given by the following rules:

$$(+\wedge\text{I.}) \ \frac{+A \quad +B}{+(A \wedge B)} \qquad (+\wedge\text{E.}_1) \ \frac{+(A \wedge B)}{+A} \qquad (+\wedge\text{E.}_2) \ \frac{+(A \wedge B)}{+B}$$

That is, from the expression of assent to $A \wedge B$, one can infer assent to A and assent to B (and *vice versa*). At first, this may seem an odd notion of inference. As Greg Restall (2005) notes, there is a mind-cluttering problem here (see Harman 1986). Since arbitrarily many propositions follow from any proposition A, inferential expressivism seems to entail that anyone who assents to a single proposition assents to an unbounded number of propositions. This is cognitively implausible. However, the proper understanding of *inferred assent* here is *social*: the meaning-conferring rules determine which attitudes one is *committed* to having (see also Dutilh Novaes 2015; Incurvati and Schlöder 2017). Someone who assents to A and to B need not assent to $A \wedge B$, since she may fail to draw the inference corresponding to $(+\wedge\text{I.})$. Nonetheless, she is *committed* to assenting to $A \wedge B$: if the inference is pointed out to her, she must assent to $A \wedge B$ or admit to a mistake.[4]

---

[4]This is not to say that a social interpretation is forced upon us. For instance, as Gideon Rosen pointed out to us, one might adopt a ground-theoretic interpretation whereby, e.g., assent to A and B grounds assent to $A \wedge B$. This interpretation too avoids the mind-cluttering problem and is also compatible with the account of normative vocabulary developed below.

The advantages of inferential expressivism can be appreciated by considering the case of negation. Traditional expressivism takes the connection between a term receiving expressivist treatment and the attitude it expresses to be *direct*: the term itself expresses an attitude (Ayer 1936; Stevenson 1937). Thus, traditional expressivism about negation holds that *not* expresses an attitude; call it *dissent*. Inferential expressivism, by contrast, takes the relation between negation and the expression of dissent to be *inferential*. That is, *not* does not directly express dissent, but from the use of a sentence containing a negation as the main operator one can *infer* that the speaker dissents. Such accounts of negation have been developed within the *bilateral programme* in the philosophy of logic (Smiley 1996; Rumfitt 2000). We now argue that they can be regarded as a special case of inferential expressivism.

3.2 Bilateralism and inferential expressivism about negation

Gottlob Frege (1919) argued that rejecting a proposition is the same as asserting its negation. Towards a *reductio*—in what is known as the Frege–Geach argument—he supposes that *not* is a speech-act indicator for rejection. He then notes the following valid inference.

(4) a. If not p, then not q.

    b. Not p.
    ———————————
    c. Not q.

Frege points out that the *not p* in (4a) cannot be a rejection of p, since somebody uttering (4a) might assent to p. Thus, the *not* in (4a) must *modify* the proposition p (instead of indicating a speech act *about* it). Suppose we nevertheless insist that in unembedded contexts such as (4b) *not p* is a rejection of p. Then (4b) does not coincide with the antecedent of (4a), and so (4) cannot be validated by *modus ponens*. For the inference to be an instance of *modus ponens*, the *not* in (4b) must modify p. But then *not* does not indicate rejection.

As observed by Timothy Smiley (1996), while the Frege–Geach problem does show that *not* should not be analysed as indicating rejection, this does not mean that one cannot *explain* negation by *appealing* to the speech act of rejection. Smiley claims that there is an independent and non-embeddable way of indicating rejection. To wit, posing a polar question to oneself

and answering negatively performs a rejection. Similarly, answering positively performs an assertion.

(5) a. [Is it the case that] *p*? Yes!   *asserts p*

   b. [Is it the case that] *p*? No!   *rejects p*

Smiley notes that these speech acts can figure in inferences.He draws attention to inferences such as (6).

(6) a. Is it the case that, if *q*, then *p*? Yes!

   b. Is it the case that *p*? No!
   _____

   c. Is it the case that *q*? No!

One might claim that uttering *Is it the case that p? No!* is simply a different way of asserting *It is not the case that p* and analyse (6) as *modus tollens*. However, Smiley notes, one can recognise (6) as valid without theorising about the nature of negation. Roughly, someone uttering (6a,b) cannot assert *q* on pain of being inconsistent: asserting *q* would entail *p* by (6a) and the standard semantics of the conditional, which is inconsistent with the rejection of *p* in (6b). By virtue of being unable to (consistently) assert *q*, she is committed to rejecting *q* (6c). This reasoning did not make any mention of negation. Similarly, we may consider an inference similar to (4), but where we make explicit that the minor premiss and conclusion are rejections.

(7) a. Is it the case that, if not *p*, then not *q*? Yes!

   b. Is it the case that *p*? No!
   _____

   c. Is it the case that *q*? No!

The speech-act indicators (5) that Smiley postulates feature in inferences like (6) and (7), but they are non-embeddable—*if is it the case that p? No!, then...* borders on incomprehensible. Thus, to simply reduce *p? No!* to *not p* seems unsatisfactory due to the divergent embedding behaviour of these phrases (Rumfitt 2000). But if (7b) is not reducible to the antecedent of (7a), then—following the structure of the Frege–Geach argument—how can (7) be valid?

Standard propositional logic deals only in asserted content (Frege 1879). However, based on the above arguments, Smiley (1996) concludes that the best explanation of *all* the data— including (6) and (7)—distinguishes asserted and rejected contents. Logic should be *bilateral*,

i.e. defined in terms of both assertion and rejection. The bilateralist takes *No!* in response to a yes-or-no question to be an (unembeddable) speech-act indicator and *not* to be an (embeddable) operator suitably related to the speech act of rejection. Their relation should, among other things, explain the validity of inferences like (7). Earlier work by Ian Rumfitt (2000) and ourselves (Incurvati and Schlöder 2017) specifies these relations to be *inferential* and takes them to be *constitutive* (i.e. meaning-conferring) of the negation connective.

This can be seen as an inferential expressivist account of negation. First, we may suppose that speech acts are expressions of attitudes; call the attitude expressed by assertion *assent* and the attitude expressed by rejection *dissent*. (This should come as a natural step in the context of the Frege–Geach debate: Frege's arguments about speech acts are typically thought of as arguments about attitude expressions.) Following Smiley, we can therefore consider answers to self-posed polar questions as non-embeddable expressions of attitude.

(8) a. *p*? Yes! *asserts p* and thereby *expresses assent to p*

    b. *p*? No! *rejects p* and thereby *expresses dissent from p*

Then, Rumfitt (2000) proposes the following inferences to be (part of) the meaning of negation: (i) from assent to a negated proposition, one can infer dissent from the bare proposition; and (ii) from assent to a proposition, one can infer dissent from its negation. Again writing $+$ for assent and now $\ominus$ for dissent, we can state these inferences as follows.

$$(+\neg\text{E.}) \ \frac{+\neg A}{\ominus A} \qquad (\ominus\neg\text{I.}) \ \frac{+A}{\ominus\neg A}$$

According to inferential expressivism, these rules (partially) constitute the meaning of negation: they state which attitudes one can infer from attitudes towards negated propositions, and from which attitudes one can infer (other) attitudes about negated propositions. For instance, uttering *Grass is not green* expresses assent to the proposition $\neg green(g)$. The meaning of this proposition is given by the inferential roles of its constituents, in particular the role of $\neg$. Part of this meaning is that from assent to $\neg green(g)$ one can infer dissent from $green(g)$, as per $(+\neg\text{E.})$. This is in contrast with traditional expressivism about *not*, which holds that uttering *Grass is not green* would *directly* express dissent from $green(g)$.

The precise statement of the meaning-conferring rules for negation varies across the literature; compare, e.g., Rumfitt (2000), Francez (2014) and Incurvati and Schlöder (2017). We present our preferred rules below. Taking negation to have these rules, however, is not sufficient to explain all the data considered above. The bilateralist also needs *coordination principles* governing the interaction between assent and dissent. Following Smiley (1996) and Rumfitt (2000), we adopt the following bilateral coordination principles.

$$(\text{Rejection})\ \frac{+A \qquad \ominus A}{\bot} \qquad (\text{SR}_1)\ \frac{\begin{matrix}[+A]\\ \vdots\\ \bot\end{matrix}}{\ominus A} \qquad (\text{SR}_2)\ \frac{\begin{matrix}[\ominus A]\\ \vdots\\ \bot\end{matrix}}{+A}$$

These rules characterise assent and dissent as contrary attitudes: (Rejection) states that it is inconsistent to assent to and dissent from the same proposition;[5] the *Smileian reductio* principles (SR) state that whenever it is inconsistent to hold one attitude towards some content, one is committed to holding the other one.[6] Together with bilateral *modus ponens*—the rule for conditional elimination $(+\rightarrow\text{E.})$—this suffices to explain Smiley's inference (6): if (6a) $+(q \rightarrow p)$ and (6b) $\ominus p$, then the supposition $+q$ entails $\bot$ by *modus ponens* and (Rejection). Smileian *reductio*$_1$ then allows us to infer (6c) $\ominus q$.

$$(+\rightarrow\text{E.})\ \frac{+(A \rightarrow B) \qquad +A}{+B}$$

Bilateralism—as a form of inferential expressivism—meets the original Frege–Geach challenge head on. Bilateralists agree that attitude expressions do not embed: the attitude expressions in (8) are linguistically unembeddable and the signs $+$ and $\ominus$ *decorate* formula, so cannot embed either. Bilateralists also accept Frege's conclusion that *not* cannot be an attitude expression, since it embeds. But, they continue, this does not mean that the meaning of *not* cannot be explained by appealing to the attitude of dissent. The general semantic framework of inferential

---

[5]This does not mean that one *cannot* simultaneously assent to and dissent from the same proposition (for example, one can compartmentalise). Rather, two attitude expressions are inconsistent when they result in a communicative breakdown (Baker and Woods 2015: 414).

[6]Smiley (1996) and Rumfitt (2000) do not consider these principles to be in need of justification, but linguistic and conceptual evidence for their validity is given in Francez (2014) and Incurvati and Schlöder (2017).

expressivism makes clear what kind of meanings bilateralists are proposing for negation: they explain *not* by stating which attitudes one can infer from its use.

However, the bilateralist strategy to explain negation raises a question. Bilateralists seem to uncritically take as basic two attitudes and stipulate them to stand in a B-type inconsistency relation to one another (as per the bilateral coordination principles). Does bilateralism therefore fall prey to the negation version of the Frege–Geach problem?

3.3 Bilateralism and the negation problem

Bilateralists take it as primitive that it is inconsistent to simultaneously assent to and dissent from the same proposition. Some may object that this already falls prey to the negation problem: the bilateralist owes an explanation of *why* these attitudes are inconsistent. But this objection is unfair.

It should be uncontroversial that one needs to take *some* basic type of inconsistency as primitive. Schroeder (2008a) posits as basic (i) the truth-conditional inconsistency between a proposition and its negation; and (ii) that some attitudes (such as belief and intent) are *inconsistency-transmitting*. The A-type inconsistency of *believing that p and believing that not p* follows from *p and not p* being truth-conditionally inconsistent and from this inconsistency being *transmitted* to the level of belief. The bilateralist, by contrast, posits as basic (i) the B-type inconsistency of assent and dissent; and (ii) that negation is explained by its inferential relation to the expression of dissent. This explains the supposedly A-type inconsistency of *believing that p and believing that not p* by noting that someone who believes that *p* must *assent to p*, and someone who believes that *not p* must *dissent from p*—and it is inconsistent to both assent and dissent from *p*.

Abstracting from there, we can say that there are *A-type expressivists* like Schroeder who posit A-type inconsistency as basic and attempt to reduce apparent B-type inconsistencies to A-type.[7] The bilateralists, by contrast, take B-type inconsistency as basic and attempt to reduce

---

[7]Schroeder (2008a) explains the apparent B-type inconsistency between *Murdering is wrong* and *Murdering is not wrong* as the A-type inconsistency between *being for blaming for murdering* and *being for not blaming for murdering*, stipulating that *being for* is an inconsistency-transmitting attitude. We discuss this further in Section

apparent A-type inconsistencies to B-type. Both need to take some kind of inconsistency as basic and use it to explain the other one.

The A-type expressivist might demand an explanation of the bilateralist of *why* assent and dissent are inconsistent—supposedly, asking the bilateralist to provide some fundamental property of the mind that would entail this inconsistency. But the bilateralist might equally demand an explanation of why some attitudes are inconsistency-transmitting—asking the A-type expressivist to name the underlying reason that explains why *belief* transmits inconsistencies, but *wonder* does not. So in terms of explanatory credentials, A-type expressivism and bilateralism are on equal footing.

To be sure, the negation problem we discussed in Section 2 is not thereby solved. Both A-type expressivists and bilateralists still need to explain compositionality and logicality. In addition, the bilateralist has not yet provided a semantics for *wrong*. At this stage, we are merely defending the starting point of bilateralism—that there are two attitudes we may call *assent* and *dissent* that stand in a B-type inconsistency relation—against the charge of being unexplanatory.

Also, we do not claim that the whole of the explanatory challenge is unfair. Expressivists like Horgan and Timmons (2006) who have logical operators such as negation *generating* new operators for attitude expressions are fairly challenged with explaining what these attitudes are and why they have the required logical properties. But the bilateralist does not do anything like this. Horgan and Timmons claim that a logically complex sentence expresses a logically complex attitude; the meaning of this attitude, they continue, is defined by its inferential role. The bilateralists agree with Schroeder (2008a) that there are no logically complex attitudes. They state the meanings of attitude expressions as inferential roles, but the complexity lies with the contents of those attitudes: bilateralists deal in *simple* attitudes (assent or dissent), held *towards* logically complex *propositions*. Furthermore, the two basic attitudes of bilateralism and their logical properties (given by the bilateral coordination principles) are linguistically observable (instead of being generated by logical vocabulary). While Horgan and Timmons (and others) do have an explanatory problem, the bilateralists do not.

4.

14

Schroeder nevertheless considers A-type expressivism superior, since 'B-type inconsistency is not something that expressivists can take for granted, because there are no good examples of it' (2008a: 48). However, there *are* natural examples of attitudes that stand in a relation of B-type inconsistency, e.g. *belief* and *wonder*, or in other logical relations, e.g. *certainty* entails *belief*. Such attitudes and their logical relations are linguistically observable, and in terms of explanation are on a par with the observation that certain attitudes (like *belief*) are inconsistency-transmitting (Baker and Woods 2015).

In sum, either there is no explanatory problem for the bilateralist or there is an analogous problem for the A-type expressivist. However, there is a serious explanatory challenge for expressivists that (i) take as basic attitudes and logical properties that are not observable, but are needed to make expressivist semantics work, or (ii) take *both* A-type and B-type inconsistency as basic. The bilateralist does neither. Both assent and dissent, as well as their logical properties, are linguistically observable by examining the speech acts of assertion and rejection. And A-type inconsistency is explained in terms of B-type inconsistency.

The bilateralist still needs to address the aspects of the negation problem related to compositionality and logicality and to include normative vocabulary in her account. To this end, we show how bilateral semantics can be extended to a *multilateral* semantics by taking further linguistically observable attitudes as primitive. We first present the inferential expressivist account of the epistemic modal *might*. This analysis of epistemic modals will be helpful in understanding the inferential expressivist account of *wrong* and in solving the *new new negation problem* in Section 4.

3.4   Inferential expressivism beyond negation

According to Smiley (1996), posing a polar question to oneself and answering it is one way of realising attitude expressions. He restricts attention to the answers *Yes!* and *No!*, but as we note in Incurvati and Schlöder 2019, there is a third possible answer: *Perhaps!*.[8] We call the speech act performed by answering *Perhaps!* to a polar question *weak assertion* and the attitude it

---

[8]There is linguistic evidence for *perhaps* being a speech-act indicator similar to *Yes!* and *No!* (Incurvati and Schlöder 2019). Notably, *perhaps* does not embed in e.g. conditional antecedents (#*if perhaps it is raining,...*).

expresses *weak assent*. As before, we mark assent with $+$, dissent with $\ominus$ and now weak assent with $\oplus$.

(9) a. *p*? Yes!   *asserts p* and thereby *expresses assent to p* $(+p)$

    b. *p*? No!   *rejects p* and thereby *expresses dissent from p* $(\ominus p)$

    c. *p*? Perhaps! *weakly asserts p* and thereby *expresses weak assent from p* $(\oplus p)$

But how should we characterise these speech acts? Stalnaker (1978) famously argued that the essential effect of an assertion is a proposal to add a proposition to the common ground (or to accept such a proposal).[9] Within this framework, it is natural to take the essential effect of a rejection as that of marking a proposition as *not being accepted* in the common ground because the speaker dissents from it. This means that rejections can be *weak* (Dickie 2010; Incurvati and Schlöder 2017): asserting *not p* may be stronger than rejecting *p*, since there are circumstances where one may dissent from a proposition and its negation. The following is a natural example:

(10)   *Alice*: X or Y will win the election.

       Bob: Is it the case that X or Y will win? No! X or Y *or* Z will win.

Bob is here expressing dissent from ⌜*X or Y will win*⌝ by means of the particle *no*. But it would be a mistake to read Bob's rejection as expressing assent to ⌜*it is not the case that X or Y will win*⌝. For this would entail that he assents to ⌜*neither X nor Y will win*⌝ whereas he seems to be dissenting from that proposition as well—he leaves open who wins.

What about weak assertion? We take this speech act to have the essential effect of rejecting a negative (see Incurvati and Schlöder 2019 for details). Consider again (10). While dissent from ⌜*X or Y will win*⌝ is part of the content of Bob's utterance, dissent from ⌜*neither X nor Y will win*⌝ is merely *conveyed* via a pragmatic implicature: if Bob had not intended to dissent from ⌜*neither X nor Y will win*⌝, he would have explicitly assented to ⌜*Z will win*⌝. Indeed, this implicature can be cancelled.

---

[9]The widespread view that an assertion immediately results in the addition of a proposition to the common ground is not Stalnaker's. He is explicit that addition to the common ground is pending acceptance (1978: fn. 9). We concur.

(11) *Alice*: X or Y will win the election.

     Bob: Is it the case that X or Y will win? No! X or Y *or* Z will win. In fact, Z will.

As we argue in Incurvati and Schlöder (2019), *weak assertion* is the mechanism by which Bob can *explicitly* express dissent from ⌜*neither X nor Y will win*⌝ here. Compare (12), where this effect of uttering *perhaps*—that one is dissenting from ⌜*neither X nor Y will win*⌝—is non-cancellable.

(12) Is it the case that X or Y will win? Perhaps! #In fact, neither of them will.

Thus, weakly asserting *X or Y will win* by way of answering *Perhaps!* rejects the negative of ⌜*X or Y will win*⌝ (and nothing more). By generalisation, this explains the following inference rules for negation.

$$(\ominus\neg\text{I.})\ \frac{\oplus A}{\ominus\neg A} \qquad (\ominus\neg\text{E.})\ \frac{\ominus\neg A}{\oplus A}$$

$$(\oplus\neg\text{I.})\ \frac{\ominus A}{\oplus\neg A} \qquad (\oplus\neg\text{E.})\ \frac{\oplus\neg A}{\ominus A}$$

That is, if one dissents from a negative, one is committed to weakly assenting to the non-negated proposition and *vice versa*. And if one weakly assents to a negative, one is committed to dissenting from the non-negated proposition and *vice versa*. According to inferential expressivism, these rules are *meaning-conferring* for the negation operator.

As we show in Incurvati and Schlöder (2019), one can similarly give rules characterising the meaning of the epistemic modal *might* (formalised as $\Diamond$). We cannot go into detail here, but one can easily see that asserting *it might be that p* is inferentially equivalent to *p? Perhaps!* and that *perhaps* iterates with *might*. Thus, the following rules give the meaning of epistemic modals.

$$(+\Diamond\text{I.})\ \frac{\oplus A}{+\Diamond A} \qquad (+\Diamond\text{E.})\ \frac{+\Diamond A}{\oplus A}$$

$$(\oplus\Diamond\text{I.})\ \frac{\oplus A}{\oplus\Diamond A} \qquad (\oplus\Diamond\text{E.})\ \frac{\oplus\Diamond A}{\oplus A}$$

When we add the rules for epistemic modals, the following rules become derivable, which specify how negation interacts with assent. Later, these rules will be useful to characterise the meaning of *Murdering might be wrong* and how it differs from *Murdering is wrong*.

$$(+\neg \mathrm{I.}) \ \frac{\ominus\Diamond A}{+\neg A} \qquad (+\neg \mathrm{E.}) \ \frac{+\neg A}{\ominus\Diamond A}$$

Since our semantics is now *multilateral*—it takes more than two attitudes as basic—we need additional principles to coordinate assent and weak assent. These are the rules of (Assertion) and (Weak Inference).

$$(\text{Assertion}) \ \frac{+A}{\oplus A} \qquad (\text{Weak Inference}) \ \frac{\begin{array}{c}[+A]\\ \vdots\\ +B \qquad \oplus A\end{array}}{\oplus B}$$

if ($+\Diamond$E.) and ($\oplus\Diamond$E.) were not used to derive $+B$ and only $+$ed premisses occur in the derivation of $+B$.

The justification of (Assertion) is straightforward: if someone assents to A, she cannot also assent to *not A* (*pace* paraconsistent logics), so she is committed to dissenting from *not A*. The coordination principle (Weak Inference) is justified by the evident validity of inferences like (13):

(13) a. Is it the case that if A, then B? Yes!

 b. Is it the case that A? Perhaps!

---

 c. Is it the case that B? Perhaps!

Like Smiley's *modus tollens*-like inference (6), (13) is a valid scheme that is accounted for by coordinating the fundamental attitudes. We justify the proof-theoretic restrictions of (Weak Inference) elsewhere (Incurvati and Schlöder 2019); they are related to issues arising from weak attitudes interacting with conditionals in unexpected ways (Kolodny and MacFarlane 2010; Incurvati and Schlöder 2017).[10]

The inferential expressivist approach to negation and epistemic modality has a number of desirable properties. Write $\vdash^{EML}$ (Epistemic Multilateral Logic) for the proof theory containing all above inference rules. Then the account satisfies the following desiderata identified in the prior literature.[11]

---

[10]For the same reasons, this account of epistemic modals does not satisfy the conditional proof rule that from $+A$ entailing $+B$, one can infer $+(A \to B)$. Rather, one is only entitled to infer $+(A \to B)$ if the inference from $+A$ to $+B$ satisfies the restrictions on (Weak Inference).

[11]The first three desiderata may seem incompatible, because $+(A \wedge \Diamond \neg A) \to \bot$ classically entails $+\Diamond A \vdash +A$. However, Epistemic contradiction does not entail $+(A \wedge \Diamond \neg A) \to \bot$ because of the restrictions on the conditional proof rule.

- Classicality of assent: if $A \vdash B$ in classical logic, then $+A \vdash^{EML} +B$.

- Epistemic contradiction (Yalcin 2007; Mandelkern 2019): $+(A \wedge \Diamond \neg A) \vdash^{EML} \bot$.

- Non-factive *might* (Yalcin 2007): $+\Diamond A \nvdash^{EML} +A$.

- Rejection/dissent is weak (Dickie 2010): $\ominus A \nvdash^{EML} +\neg A$.

- Negation entails dissent (Rumfitt 2000): $+\neg A \vdash^{EML} \ominus A$ and $+A \vdash^{EML} \ominus \neg A$.

Epistemic Multilateral Logic is sound and complete with respect to an appropriate model theory and the calculus can be extended in a natural way to include quantifiers (Aloni *et al.* ms).

To sum up, we have an inferential expressivist account of *not* and *might* that takes as basic the attitudes of *assent*, *dissent* and *weak assent*. The relation between the attitudes is characterised by the coordination principles, and the meaning of *not* and *might* is constituted by the above rules for their introduction and elimination. As argued above for bilateralism, our multilateral account is not subject to the explanatory part of the negation problem.[12] The fundamental attitudes we make use of are linguistically observable, as are the logical properties encoded as coordination principles. We now apply this framework to normative vocabulary.

3.5 Inferential expressivism about normative vocabulary

According to traditional expressivism, the semantic value of *wrong* is the expression of disapproval. According to inferential expressivism, the semantic value of *wrong* is that having the attitude of assent towards a proposition like ⌜*murdering is wrong*⌝ means that one can infer the attitude of disapproval towards ⌜*murdering*⌝. That is, as above for the terms *not* and *might*,

---

[12]This paper focuses on the Frege–Geach problem, but there are other problems related to the embedding of attitudes, such as the *Many Attitudes Problem* (Schroeder 2008a). What can the expressivist say about the mental state of someone who hopes, regrets or fears that murdering is wrong? We think that inferential expressivism has the resources to deal with this problem too. First, the inferential expressivist can explain mental states in terms of their inferential connection to its basic attitudes, for instance by explaining belief in terms of assent. Second, just as on the traditional picture, she can take the object of these mental states to be propositions, except that these are now understood as bearers of inference-potentials. We develop this strategy in ongoing work.

the meaning of *wrong* is given by how attitudes towards propositions containing *wrong* infer-entially interact with other attitudes towards different contents. In particular, the inferential expressivist meaning of *wrong* is given by the following rules of inference (writing $\mathscr{D}$ for the attitude of disapproval).

$$(+\text{WI.}) \ \frac{\mathscr{D}a}{+wrong(a)} \qquad (+\text{WE.}) \ \frac{+wrong(a)}{\mathscr{D}a}$$

This semantics provides enough places to insert a negation, since *wrong* is treated as an embed-dable predicate. The attitude expressed by uttering *Murdering is not wrong* is $+\neg wrong(m)$, the attitude expressed by *Not murdering is wrong* is $+wrong(\neg m)$.

What does $+\neg wrong(m)$ mean? The glib answer is that the meaning of this attitude expres-sion is given by the inference rules defining the meaning of $\neg$ and those defining the meaning of *wrong*.[13] This may be too glib to be satisfactory, so let us explore this inferential meaning.

First, $+\neg wrong(m)$ clearly entails $\ominus wrong(m)$, i.e. dissent from ⌜*murdering is wrong*⌝. This means that the speaker is committed to disagreeing with any expression of disapproval of ⌜*murdering*⌝. But dissent from ⌜*murdering is wrong*⌝ is already entailed by assent to ⌜*murdering might not be wrong*⌝ and there clearly is a difference between uttering *Murder-ing is not wrong* and *Murdering might not be wrong*. Indeed, $+\neg wrong(m)$ entails something stronger, namely $\ominus \Diamond wrong(m)$, i.e. dissent from ⌜*murder might be wrong*⌝. But in what sense exactly is this stronger than dissent from ⌜*murder is wrong*⌝? We can characterise $\ominus \Diamond wrong(m)$ attitude as expressing a special kind of dissent from ⌜*murdering is wrong*⌝. Namely, *stable* dis-sent.

The attitude of assent is *stable* in the sense that if one assents to a proposition p one cannot come to stop assenting to p by simply being provided additional information—it requires a change of mind to stop assenting (Restall 2005). But, in general, if one dissents from p or weakly assents to p, one may come to change these attitudes *merely* by receiving additional information (Incurvati and Schlöder 2017). The difference in stability can also be observed when considering the speech acts that are used to express these attitudes.

---

[13]This, to stress, differentiates inferential expressivism from Horgan and Timmons' (2006) proposal: they describe the meaning of *Murdering is not wrong* as a logically complex attitude; we describe ⌜*Murdering is not wrong*⌝ as a logically complex proposition towards which one may have attitudes.

(14)   *Alice*: X or Y will win the election.

   Bob: Is it the case that X or Y will win? No! X or Y *or* Z will win.

As argued above, the essential effects of our three speech acts are as follows: an assertion puts forward a proposal to accept a proposition into common ground; a rejection marks a proposition as not being accepted into common ground; and a weak assertion rejects a negative. Assertion is stable in the sense that once a proposition is in the common ground, it remains in it unless the common ground is revised. A rejection, however, is unstable in that a previously rejected proposition can become common ground simply because other, new information is added to the common ground and hence *without any revision of the common ground*.

(14) is an example of this phenomenon: Bob's utterance marks ⌜*X or Y will win*⌝ as unaccepted into the common ground. But if the proposition ⌜*X or Y or Z will win*⌝ asserted by Bob does become common ground and later Alice and Bob add ⌜*Z dropped out*⌝ to the common ground as well, ⌜*X or Y will win*⌝ will enter the common ground after all, but without any revision of the common ground. Hence, Bob's rejection of ⌜*X or Y will win*⌝ was unstable, as undoing it required no revision. The case for weak assertion being unstable is analogous.

Thus, our account has the resources to distinguish between saying *Murdering is not wrong* and *Murdering might not be wrong*. By uttering either sentence, the speaker expresses dissent from the proposition ⌜*murdering is wrong*⌝. However, uttering *Murdering is not wrong* expresses stable dissent (i.e. changing this attitude would be a change of mind), whereas uttering *Murdering might not be wrong* expresses unstable dissent—the speaker may come to express something else about *murdering* without changing their mind, simply by receiving additional information. The formal expression of stable dissent from a proposition p is $\ominus\Diamond p$, which is inferentially equivalent to $+\neg p$. Thus, the inferential meaning of *Murdering is not wrong* is the expression of stable dissent from ⌜*murdering is wrong*⌝.

Note that we did not introduce an auxiliary attitude (such as *tolerance*) to account for what is expressed by uttering *Murdering is not wrong*—we only need the three foundational attitudes of inferential expressivism together with the attitude of *disapproval*. If someone wishes to *call* the attitude expressed by $\ominus\Diamond wrong(m)$ by the name *tolerance of murder*, she can do so—this may be a useful and illuminating shorthand. But *only* a shorthand. The inferential expressivist

does not need to take the attitude of tolerance as a primitive.

In sum, inferential expressivism solves the negation problem for expressivism about *wrong*. The aspects of compositionality and logicality are accounted for by the fact that inferential expressivism posits the same negation operator for descriptive and normative language. The explanatory problem is avoided by not stipulating any auxiliary or complex attitudes, instead explaining the meaning of *Murdering is not wrong* by appealing to the joint inferential meaning of *not* and *wrong*.

The account straightforwardly generalises to further normative vocabulary like *right*. Supposedly, traditional expressivism takes *right* to be the expression of *approval*. Hence, according to inferential expressivism, the semantics of *right* is such that one can *infer* approval from assent to a sentence containing *right*. The meaning of *right* is given by the following rules (writing $\mathscr{A}$ for the attitude of approval).

$$(+R\text{I.}) \ \frac{\mathscr{A}\,a}{+right(a)} \qquad (+R\text{E.}) \ \frac{+right(a)}{\mathscr{A}\,a}$$

The meaning of *Murdering is not right* is then explained analogously to the meaning of *Murdering is not wrong*. Finally, to account for the factual inconsistency of *Murdering is wrong* and *Murdering is right*, it suffices to add the following axiom:

$$(\text{RW-contradiction}) \quad +\neg(right(a) \wedge wrong(a))$$

Note that (RW-contradiction) is not a coordination principle expressing fundamental properties of primitive attitudes, but an axiom expressing lexical knowledge about *right* and *wrong*. For the inconsistency of *Murdering is wrong* and *Murdering is right* is not logical, but relies on lexical knowledge.

4 The new and new new negation problems

4.1 Schroeder on the negation problem

Schroeder (2008a) defends an A-type account of *wrong*. Instead of analysing the attitude expressed by *wrong* as *disapproval*, he analyses it as *being for blaming for*, stipulating *being for*

as his only primitive attitude. This allows him to insert a negation in the right place: *Murdering is not wrong* is analysed as *being for* ⌜*not blaming for murdering*⌝. Then Schroeder claims that *being for* is an inconsistency-transmitting attitude (like belief), thereby accounting for all relevant data. This is remarkably elegant, but Schroeder himself notes some issues.

Notably, he acknowledges the *new negation problem*, a revenge version of the compositionality problem. It goes as follows. To give a uniform semantics—including both descriptive and normative vocabulary—that is grounded in the basic attitude of *being for*, Schroeder analyses descriptive claims as *being for proceeding as if* (short, *being for pai*) expressions, e.g. *Grass is green* expresses the attitude *being for* ⌜*pai Grass is green*⌝. But then *Murdering is wrong* expresses *being for* ⌜*blaming for murdering*⌝ and *Grass is green* expresses *being for* ⌜*pai Grass is green*⌝. Thus, if *Murdering is not wrong* expresses *being for* ⌜*not blaming for murdering*⌝ and the compositional semantics of *not* is uniform (as it should), the meaning of *Grass is not green* would be *being for* ⌜*not pai Grass is green*⌝. But this is a mistake: the attitude expressed by *Grass is not green* should be *being for* ⌜*pai Grass is not green*⌝.

In essence, the negation problem is that for normative language there is one space for a negation missing. The new negation problem is that for descriptive language, Schroeder has one space *too many* to put a negation.

To solve the new negation problem, Schroeder introduces the notion of a *biforcated* attitude. A biforcation is a pair of two *being for* attitudes such that the *major attitude* entails the *minor* one.[14] We write biforcations as tuples and treat the left hand-side of a tuple to be the major attitude. What exactly the minor attitude is is determined by the lexical semantics of the relevant predicates; for descriptive vocabulary, one obtains the minor attitude from the major one by inserting two negations so that, e.g., the biforcated attitude expressed by *Grass is green* is ⟨*being for* ⌜*pai Grass is green*⌝, *being for* ⌜*not pai Grass is not green*⌝⟩, where *being for* ⌜*pai Grass is green*⌝ is the major attitude (Schroeder 2008a: 98).

Using such biforcations, Schroeder defines the meaning of negation as follows: the negation of a biforcation inserts in both the minor and major attitude a *not* between *being for* and the embedded action term and then *switches* major and minor attitude. The following is Schroeder's

---

[14]The term 'biforcated' is Schroeder's, a pun on 'bifurcated' and 'being for'.

compositional semantics for *not* (2008a: 106).

$$\llbracket not \rrbracket (\langle being\ for\ a,\ being\ for\ b\ \rangle) = \langle being\ for\ \neg b,\ being\ for\ \neg a \rangle$$

Thus, if a descriptive sentence *A* expresses the biforcation ⟨*being for* ⌜*pai A*⌝, *being for* ⌜¬ *pai* ¬ *A*⌝⟩, then *not A* expresses ⟨*being for* ⌜¬¬ *pai* ¬*A*⌝, *being for* ⌜¬ *pai A*⌝⟩. Deleting double negations, one obtains the desired major attitude *being for* ⌜*pai* ¬*A*⌝.

However, Schroeder now needs to insist that for *normative* propositions, the minor and major attitudes in a biforcation are identical, lest he run into the original negation problem again. For descriptive language, by contrast, the minor attitude is strictly weaker than the major one. This is not an *ad hoc* move, but a sound strategy: by encoding a difference between the attitudes expressed by descriptive and normative terms on the level of lexical semantics, he can maintain uniform compositional semantics for negation while solving both the negation problem and the new negation problem. As a matter of fact, we agree that lexical semantics is the (only) right place to locate this difference.

Unfortunately, this move leads to the *new new negation problem* (Schroeder 2008a: 113). Schroeder justifies his introduction of minor attitudes by appealing to a notion called *disacceptance*, which consists in not agreeing with a proposition without agreeing to its negation. Formally, on Schroeder's account, to disaccept *A* is to accept the minor attitude in the biforcation expressed by *not A*. But since normative vocabulary sentences express biforcations with identical major and minor attitudes, the distinction between disacceptance and acceptance of a negative collapses for normative sentences. This entails that one cannot withhold judgement about an act *X* by disaccepting both ⌜*X is wrong*⌝ and ⌜*X is not wrong*⌝. So the only way to withhold judgement is to *not express* an attitude. Schroeder notes that this is not an *outright* bad result, concluding that it might be 'a bullet to bite' (2008a: 115).

However, this bullet becomes harder to bite when also considering epistemic *might*. Schroeder (2008a: 181) suggests the following semantics for epistemic modals: *might* deletes the major attitude and replaces it with the minor attitude.

$$\llbracket might \rrbracket (\langle being\ for\ a,\ being\ for\ b \rangle) = \langle being\ for\ b,\ being\ for\ b \rangle$$

But given Schroeder's solution to the new negation problem, this means that *Murdering might be wrong* expresses the same attitude as *Murdering is wrong*, since *Murdering is wrong* expresses a biforcation with identical minor and major attitudes. Schroeder (2008a: 182) concludes that either his semantics for *might* or his identification of minor and major attitudes for normative vocabulary must be revised. But no suitable such revision is in sight. The semantics for *might* has compelling features, as Schroeder himself notes. To preserve these, and to also make sense of normative vocabulary under epistemic modality, Schroeder would likely need to revise the foundations of the biforcated attitude framework. However, it is unclear whether this can be done so as to solve both the original negation problem and the new negation problem.

4.2 Inferential expressivism and the new new negation problem

According to inferential expressivism, to express assent to *Grass is green* is to propose the addition of $green(g)$ to the common ground. Since the common ground records the tacit assumptions in a conversation, this can be roughly paraphrased as *being for* ⌜*pai green*(g)⌝. Thus, *being for pai* is analogous to our *assent* and Schroeder's minor attitudes are analogous to our *weak assent*. His notion of *disacceptance* similarly corresponds to our *dissent*. Likewise, to express assent to ⌜*murdering is wrong*⌝ is to propose the addition of $wrong(m)$ to the common ground, which is to modify the common ground so as to include disapproval of murdering. This can be roughly paraphrased as *being for* ⌜*blaming for murdering*⌝.

Thus, inferential expressivism, as developed in Section 3, can be examined in Schroeder's terms. Doing so shows no trace of the new and new new negation problems. The inferential expressivist account differs in that it sees no fundamental difference in how its basic attitudes operate on normative *vis-á-vis* descriptive vocabulary: we do not collapse assent with weak assent anywhere, whereas Schroeder collapses major and minor attitudes towards normative content. But how does inferential expressivism distinguish between normative and descriptive vocabulary if it does not, like Schroeder, locate their difference in the minor attitudes?

Inferential expressivism locates this difference in whether meaning-conferring inference rules *change attitude*. The semantics of descriptive vocabulary is given by rules which do not change attitude, such as *from assent to X, infer assent to Y* (e.g. the meaning of *bachelor* in-

cludes that from assent to ⌜*John is a bachelor*⌝ one may infer assent to ⌜*John is unmarried*⌝).
The semantics of normative vocabulary is given by rules which *do* change attitude. For instance, *wrong* switches between assent and disapproval. This is why if one negates *wrong*, it may appear that one negates an attitude (namely, disapproval, the attitude *wrong* switches with), whereas this is not so when negating a descriptive term. Like Schroeder, we locate the difference between normative and descriptive vocabulary at the level of lexical semantics: in the way the inference rules constituting lexical meaning characterise the interaction of a term with attitudes. This allows us to have uniform semantics for operators such as negation, without to collapsing any of our notions.

5 Conclusion

Bilateralism is best understood as a form of inferential expressivism about negation: the meaning of negation is given in terms of its *inferential* connections to the *expression* of certain attitudes. Bilateralists make use of two distinct attitudes, assent and dissent, *coordinated* to stand in particular relations. Smiley (1996) argued that this move solves the original version of the Frege–Geach problem. We have demonstrated that this move also solves its negation version. In particular, bilateralism is not subject to the explanatory aspect of the negation problem, since its foundational assumptions are on a par with A-type expressivists' assumption that certain attitudes transmit the inconsistency of truth and falsity.

Previous work of ours has extended inferential expressivism to a *multilateral* semantics which also accounts for the meaning of epistemic modals. In this paper, we have further expanded inferential expressivism so as to encompass moral vocabulary such as *wrong*. The meaning of *wrong* is given in terms of its inferential connections with the attitude of disapproval. This is a non-representational explanation of meaning, which does not appeal to the existence of moral facts. Nonetheless, inferential expressivism about moral vocabulary is *compatible* with moral realism. Although the combination of these two views might hold little appeal, we take it as an advantage of our semantics that it remains neutral about the existence of moral facts.

The following is an oft-told story about the metaethical landscape. Moral cognitivism takes

moral sentences to express beliefs but, unlike expressivism, has difficulty in explaining the motivating power of moral judgments. Traditional ethical expressivism takes moral sentences to express non-cognitive attitudes, but, unlike cognitivism, falls prey to the Frege-Geach problem. Hybrid approaches (see, e.g., Copp 2001; Boisvert 2008; Ridge 2014) attempt to have 'the best of both worlds' by taking moral sentences to express beliefs *and* non-cognitive attitudes. However, such approaches, in the end, fall prey to the standard problems for cognitivism or expressivism, and fail to properly explain the connection between the two alleged components of moral judgments.

According to inferential expressivism about normative vocabulary, the connection between assent to ⌜*murdering is wrong*⌝ and the attitude of disapproval is clearly articulated by the rules of inference for *wrong*. This approach solves the original version of the Frege-Geach problem as well as the negation version and iterations thereof. It therefore improves on Schroeder's (2008a) A-type theory, which assigns the same meaning to *Murdering is wrong* and *Murdering might be wrong*. We submit that inferential expressivism, like traditional expressivism, also has the resources to explain the motivating power of moral judgments, since on this view assenting to ⌜*murdering is wrong*⌝ commits the speaker to having the attitude of disapproval towards ⌜*murdering*⌝. If this claim can be made out, inferential expressivism might well be the best of both worlds.[15]

References

Aloni, M., Incurvati, L., and Schlöder, J. J. (ms). 'Epistemic modals in hypothetical reasoning', manuscript.

Ayer, A. J. (1936). *Language, Truth and Logic* (London: Gollancz).

Baker, D. and Woods, J. (2015). 'How expressivists can and should explain inconsistency', *Ethics* 125: 391–424.

Boisvert, D. (2008). 'Expressive-assertivism', *Pacific Philosophical Quarterly* 89: 169–203.

Brandom, R. (1994). *Making it Explicit* (Cambridge, MA: Harvard University Press).

Chrisman, M. (2008). 'Expressivism, inferentialism, and saving the debate', *Philosophy and Phenomenological Research* 77: 334–358.

Copp, D. (2001). 'Realist-expressivism: A neglected option for moral realism', *Social Philosophy and Policy* 18: 1–43.

Dickie, I. (2010). 'Negation, anti-realism, and the denial defence', *Philosophical Studies* 150: 161–185.

Dummett, M. (1991). *The Logical Basis of Metaphysics* (Cambridge, MA: Harvard University Press).

Dutilh Novaes, C. (2015). 'A dialogical, multi-agent account of the normativity of logic', *Dialectica* 69: 587–609.

Francez, N. (2014). 'Bilateralism in proof-theoretic semantics', *Journal of Philosophical Logic* 43: 239–259.

Frege, G. (1879). *Begriffsschrift: Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens* (Halle: Louis Nebert).

Frege, G. (1919). 'Die Verneinung: Eine logische Untersuchung', *Beiträge zur Philosophie des deutschen Idealismus* 1: 143–157.

Geach, P. T. (1965). 'Assertion', *Philosophical Review* 74: 449–465.

Gentzen, G. (1935). 'Untersuchungen über das logische Schließen I', *Mathematische Zeitschrift* 39: 176–210.

Gibbard, A. (2003). *Thinking How to Live* (Cambridge, MA: Harvard University Press).

Hare, R. M. (1952). *The Language of Morals* (Oxford: Clarendon Press).

Harman, G. (1986). *Change in View* (Cambridge, MA: MIT Press).

Horgan, T. and Timmons, M. (2006). 'Cognitivist expressivism', in T. Horgan and M. Timmons (eds.), *Metaethics after Moore* (Oxford: Oxford University Press), 255–298.

Humberstone, L. (2000). 'The revival of rejective negation', *Journal of Philosophical Logic* 29: 331–381.

Incurvati, L. and Schlöder, J. J. (2017). 'Weak rejection', *Australasian Journal of Philosophy* 95: 741–760.

Incurvati, L. and Schlöder, J. J. (2019). 'Weak assertion', *Philosophical Quarterly* 69: 741–770.

Kolodny, N. and MacFarlane, J. (2010). 'Ifs and oughts', *The Journal of Philosophy* 107: 115–143.

Mandelkern, M. (2019). 'Bounded modality', *Philosophical Review* 128: 1–61.

Prawitz, D. (1965). *Natural Deduction. A Proof-Theoretical Study* (Uppsala: Almqvist & Wicksell).

Restall, G. (2005). 'Multiple conclusions', in P. Hájek, L. Valdés-Villanueva, and D. Westerståhl (eds.), *Logic, Methodology and Philosophy of Science: Proceedings of the Twelfth International Congress* (London: Kings College Publications), 189–205.

Ridge, M. (2014). *Impassioned Belief* (Oxford: Oxford University Press).

Rumfitt, I. (2000). '"Yes" and "No"', *Mind* 109: 781–823.

Schnieder, B. (2010). 'Expressivism concerning epistemic modals', *Philosophical Quarterly* 60: 601–615.

Schroeder, M. (2008a). *Being For: Evaluating the Semantic Program of Expressivism* (Oxford: Oxford University Press).

Schroeder, M. (2008b). 'What is the Frege-Geach problem?', *Philosophy Compass* 3: 703–720.

Smiley, T. (1996). 'Rejection', *Analysis* 56: 1–9.

Stalnaker, R. (1978). 'Assertion', in P. Cole (ed.), *Syntax and Semantics*, vol. 9 (New York: Academic Press), 315–332.

Stevenson, C. L. (1937). 'The emotive meaning of ethical terms', *Mind* 46: 14–31.

Tiefensee, C. (2016). 'Inferentialist metaethics, bifurcations and ontological commitment', *Philosophical Studies* 173: 2437–2459.

Unwin, N. (1999). 'Quasi-realism, negation and the Frege-Geach problem', *Philosophical Quarterly* 49: 337–352.

Unwin, N. (2001). 'Norms and negation: A problem for Gibbard's logic', *Philosophical Quarterly* 51: 60–75.

Yalcin, S. (2007). 'Epistemic modals', *Mind* 116: 983–1026.

Yalcin, S. (2012). 'Bayesian expressivism', *Proceedings of the Aristotelian Society* 112: 123–160.