WHY THE FREGE-GEACH PROBLEM DOES NOT REFUTE EXPRESSIVISM

Oliver Hallich
Universität Duisburg-Essen/Deutschland

Resumo: É assumido com frequência que o chamado "argumento Frege-Geach" refuta o expressivismo, ou seja, a visão de que sentenças morais não são primariamente fatos situados, mas expressam atitudes ou emoções. Neste trabalho, tento refutar essa suposição e demonstrar que o argumento de Frege-Geach não representa uma ameaça séria para o expressivismo. Depois de uma apresentação inicial do argumento Frege-Geach (Seção 1), eu tento abrir o caminho para uma defesa do expressivismo por meio do delineamento do que o expressivista tem que fazer e não fazer, a fim de defender-se contra aquela posição (Seção 2). Esta etapa preliminar, eu argumento, fornece ao expressivista uma resposta adequada à objeção de Frege-Geach (Seção 3). Finalmente, eu discuto a estratégia de Blackburn de defesa do expressivismo contra essa objeção e tento mostrar que, embora de forma não intencional por Blackburn, o fracasso dessa estratégia dá apoio ao expressivismo (Seção 4).

Palavras-chave: argumento Frege-Geach, expressivismo, Blackburn, cognitivismo, metaética.

Abstract: It is often assumed that the so-called "Frege-Geach argument" refutes expressivism, i.e. the view that moral sentences do not primarily state facts, but express attitudes or emotions. In this paper, I attempt to rebut that assumption and to show that the Frege-Geach argument poses no serious threat to expressivism. After an initial presentation of the Frege-Geach argument (Section 1), I try to pave the way for a defence of expressivism by delineating what the expressivist has to do and not to do in order to defend himself against it (Section 2). This preliminary step, I argue, provides the expressivist with a convenient reply to the Frege-Geach objection (Section 3). Finally I discuss Blackburn's strategy of defending expressivism against this objection and try to show that, though in a manner unintended by Blackburn, the failure of this strategy lends support to expressivism (Section 4).

Keywords: Frege-Geach argument, expressivism, Blackburn, cognitivism, metaethics.

It is often assumed that the so-called "Frege-Geach argument" refutes expressivism, i.e. the view that moral sentences do not primarily state facts, but express attitudes or emotions. In this paper, I attempt to rebut that assumption and to show that the Frege-Geach argument poses no serious threat to expressivism. After an initial presentation of the Frege-Geach argument (Section 1), I try to pave the way for a defence of expressivism by delineating what the expressivist has to do and not to do in order to defend himself against it (Section 2). This preliminary step, I argue, provides the expressivist with a convenient reply to the Frege-Geach objection (Section 3). Finally I discuss Blackburn's strategy of defending expressivism against this objection and try to show that, though in a manner unintended by Blackburn, the failure of this strategy lends support to expressivism (Section 4).

1. The Frege-Geach Argument

The Frege-Geach argument is most forcefully formulated by Geach, who traces it back to Frege's *Begriffsschrift*, in his articles "Assertion" and "Ascriptivism".[1] At the bottom of it lies the claim that the expressivist analysis of evaluative terms such as "good" or "bad" does not adequately represent the meaning of these terms when they occur in unasserted contexts such as conditionals and disjunctions. Therefore expressivism, it is said, cannot account for the logical validity of arguments in which these terms occur as unasserted components of, for example, conditionals and disjunctions. The following modus ponens argument is a standard example:

(1) Lying is bad.
(2) If lying is bad, then getting your little brother to lie is bad.
(3) *Hence*, getting your little brother to lie is bad.

A firm intuition shared by all defenders of the Frege-Geach argument is that the validity of arguments like (1)–(3) cannot sensibly be called into question. Now according to expressivist theories, so the objection goes, "bad"

---

[1] See GEACH, 1960 , p. 223–24, and 1965, p. 463–64; for a   similar argument see SEARLE, 196 2, p. 425–26.

serves to express an attitude towards the type of action in question. So, given the truth of an expressivist theory, by uttering (1) the speaker would express his attitude of aversion towards lying. But in (2) "bad" *cannot* be so used, since assenting to the conditional is compatible with not assenting to its antecedent and its consequent, i.e. with denying that lying is bad and that getting your little brother to lie is bad. Similarly, we would assent to "If all dictators are nice guys, Hitler was a nice guy" without assenting to "All dictators are nice guys" or to "Hitler was a nice guy". So while a speaker who utters a simple sentence like (1), according to expressivism, thereby expresses an attitude of aversion towards lying, a speaker who utters a complex sentence like (2) does not thereby express an attitude of aversion towards lying. The expressivist, so the objector says, would therefore have to assume that the meaning of "bad" varies in (1) and (2).

This gives rise to at least two problems. First, it is unclear how the expressivist can construe the meaning of "bad" if it is embedded in a conditional as in (2). If "bad" is not used to express an attitude of aversion, what meaning can be attached to it within the framework of a theory according to which terms like "bad" *are* primarily used to express attitudes? Second, if the meaning of "bad" varies depending on whether it is used in a conditional like (2) or in a simple sentence like (1), arguments like (1)–(3) become logically *invalid*. They then have to be regarded as involving the fallacy of equivocation, similar to "He is working at the bank; if he is working at the bank he must be working near the river; so he must be working near the river". If "bad" means something different in (1) and in (2), (3) does not follow. This blatantly contradicts our conviction that (1)–(3) *is* a valid argument. So the expressivist seems committed to denying the validity of a valid argument.

2. What the expressivist has to do to rebut the Frege-Geach argument

The nucleus of the Frege-Geach argument is the claim that the following three propositions are true and that expressivism is committed to denying the truth of at least one of them:

(i) In (1), "bad" is used in an evaluative sense.
(ii) In (1) and (2), "lying is bad" has the same meaning.

(iii) An argument like (1)–(3) is valid.

In reply to this objection the expressivist's task is to show that, applied to particular arguments like (1)–(3), the truth of each of these propositions is *compatible* with expressivism. He owes us an account of why, within an expressivist framework, nothing *precludes* each of these propositions from being true. The Frege-Geach argument would be successful as an argument against expressivism if it turned out that this were not the case, i.e. if, for example, it were impossible without inconsistency to defend expressivism while also claiming that "lying is bad" has the same meaning in (1) and (2). However, what the expressivist does *not* have to show in order to defend himself against the Frege-Geach objection is that, applied to a particular argument like (1)–(3), *all* of these propositions must be true *at the same time*. He need not accept (i), (ii) and (iii) *en bloc*. He is free to argue (as I will do) that if "bad" is used in an evaluative sense in (1), then "lying is bad" does not have the same meaning in (1) and (2) and that if "lying is bad" does not have the same meaning in (1) and (2), then (1)–(3) is not valid, in other words: that the truth of (i) precludes (ii) and (iii) from being true. This would also mean (by modus tollens) that if the argument is logically valid, then "lying is bad" has the same meaning in (1) and (2) and that if "lying is bad" has the same meaning in (1) and (2), then "bad" cannot be used evaluatively in (1), in other words: that if (iii) is true, then (ii) is true and that if (ii) is true, then (i) is false. The expressivist has to take it upon himself to show that, in accordance with the premises of his theory, *each* of these propositions can be true, not that *all* of them are true at the same time.

In order to precisely demarcate what the expressivist has to do and not to do to defend himself against the Frege-Geach objection, we should also note that the expressivist is *not* committed to endorsing any of the following three general theses:

(i') In (1), "bad" *must* be used in an evaluative sense.
(ii') In (1) and (2), "lying is bad" *must* have the same meaning.
(iii') An argument like (1)–(3) is *always* valid.

Critics of expressivism often fail to distinguish between this set of propositions and the first one ((i), (ii), (iii)). They therefore often tacitly assume that the expressivist does not only have to show that his theory is

compatible with (i), (ii) and (iii), but also that it is compatible with (i'), (ii') and (iii'). However, both sets of propositions are different. Someone who endorses the first set of propositions does not regard (i), (ii) and (iii) as *necessary* truths. He can allow for exceptions, i.e. for cases in which these propositions are false: for example, (ii) is evidently false if "lying" means "deceiving" in (1) but "being in a horizontal position" in (2). In contrast, someone who endorses the second set of propositions ((i'), (ii'), (iii')) cannot allow for these exceptions. He has to deny that the truth of (i'), (ii'), (iii') may vary depending on what meaning is attached to terms like "bad" in the specific context of arguments like (1)–(3). As one can, without contradiction, endorse one or more of the propositions of the first set without endorsing any of the propositions of the second set, an expressivist who accepts the challenge that (i), (ii) and (iii) must be shown to be compatible with expressivism does not have to subscribe to any of the propositions of the second set. So should it turn out that the expressivist's defence strategy against the Frege-Geach objection commits him to denying one or more of the second set of propositions (as indeed it will) he can react stoically: as he does not embrace any of these propositions, this need not disconcert him.

3. How the expressivist can reply to the Frege-Geach argument

After these preliminaries, let's now look at each of the propositions (i), (ii) and (iii) in turn. As to (i), it is uncontroversial that its truth is compatible with expressivism. Of course, the expressivist will claim that "bad" can be used evaluatively in (1), i.e. that it can be used to express a "con-attitude" towards lying. Undoubtedly he will claim that this is the case in many contexts and probably in those most important for moral discourse. However, this does not commit him to the claim that "bad" *must* be so used in (1). In his seminal work *The Language of Morals*, Hare points out that, apart from the evaluative (or, as Hare prefers to say, prescriptive) use of words such as "good", "bad" or "ought", there is also a purely descriptive use.[2] When these terms are used descriptively, sentences containing them are purely factual. "X is bad", in one of these uses, means nothing more than "X has the property P". To be more precise, there are three ways of using these

---

[2] See HARE, 1952, p. 118–126; p. 165–172.

words descriptively. First, there is the inverted comma use of these terms by means of which a speaker does not "give his signature" to an evaluative judgement, but rather reports what other people regard as relevant criteria of goodness or badness (often with a tinge of irony). For example, "She is a good Christian", uttered by a confessed atheist, might simply mean "She complies with the standards of goodness which other members of a relevant group (namely Christians) accept". Second, there is a descriptive use in which sentences containing these words serve to report social facts. For example, although the claim "One ought not to steal" *can* be used prescriptively according to Hare, it can also be used to mean nothing more than "It is regarded as socially unacceptable to steal". Third, sentences containing these words might also be used descriptively to report psychological facts about the speaker. "It is bad to tell lies", though usually used prescriptively according to Hare, might also simply report the speaker's negative feelings about lying, roughly meaning something like "I, the speaker, have a feeling that I ought not to lie".

This distinction between the evaluative and the non-evaluative uses of terms like "bad" opens the door for a convenient reply to the objection that the expressivist has to deny (ii). The expressivist need not deny that "lying is bad" has the same meaning in (1) and (2), *given* that "bad" is used in one of the "secondary", i.e. purely descriptive uses just mentioned. (1) and (2) would then have to be interpreted as "Lying has the property P (e.g. the property of being disapproved of by most members of a society)", and as "If lying has the property P, getting one's little brother to lie has the property P", respectively. In this case, (3) can soundly be deduced from (1) and (2): from "Lying is disapproved of by most members of society" and "If lying is disapproved of by most members of society, getting one's little brother to lie is disapproved of by most members of society" follows "Getting one's little brother to lie is disapproved of by most members of society". This also shows that there is no need for the expressivist to deny the truth of (iii): an argument like (1)–(3) *can* be reconstructed as a logically valid argument within an expressivist framework. The expressivist will only add the qualification that (iii) holds true only for a specific interpretation of "bad", namely, the secondary, non-

evaluative use. He will combine the acceptance of (iii) with a rejection of (iii'), which he can do without inconsistency.[3]

Things look different if "bad" *is* used evaluatively in (1), i.e. if by (1) the speaker expresses his aversion towards lying. In this case, it still holds true that "bad" *cannot* be so used in (2), as one can assent to the conditional without assenting to its antecedent or consequent. So in this case "lying is bad" does *not* have the same meaning in (1) and (2). And this means that (3) does *not* follow from (1) and (2). This sounds counterintuitive, but, unlike the cognitivist, the expressivist does not have to uphold the validity of (1)–(3). Nothing prevents him from admitting that, if "bad" is used evaluatively in (1), to deduce (3) from (1) and (2) would indeed involve a fallacy of equivocation as long as he concedes that (1)–(3) is valid under a *different* interpretation of "bad". As long as one rejects (iii'), as the expressivist is free to do, there is nothing contradictory in claiming that the argument is only valid given *some* interpretations of the terms used and that given others, it is a fallacy.

It is simply question-begging to claim, as Geach[4] and, more recently, Sinnott-Armstrong[5] do, that the argument is valid and that *therefore* "lying is bad" must mean the same in (1) and (2). This merely stipulates what can justly be called into question, namely the validity of the argument. What Geach and Sinnott-Armstrong should have said, more cautiously, is that *if* the argument is valid, then "lying is bad" must have the same meaning in (1) and (2) – which implies (via modus tollens) that if it does not have the same meaning, the argument is not valid. There is no need to dispute this.

So leaving the purely descriptive uses of "bad" aside and assuming that "bad" is used in an evaluative sense in (1), the expressivist, or so I argue, should not hesitate to "bite the bullet" and to accept the view that (1)–(3) is not logically valid.[6] To illustrate this, let's look at the following example. Imagine somebody talking about the two parts of the film "Lord of the Rings" by using the word "fantastic". This word might be used as a mere

---

[3] A defence of expressivism along these lines is developed by Eriksson (2009, p. 20–21) who claims that the validity of the modus ponens argument is preserved by the descriptive meaning of the relevant ethical terms. A similar strategy is adopted by Ridge (2006).

[4] GEACH, 1965, p. 463.

[5] SINNOTT-ARMSTRONG, 2000, p. 679–80.

[6] This option is considered, but rejected by van Roojen (1996, p. 330–31).

exclamation of appraisal and enthusiasm (meaning something like "great!" or "terrific!"). But it might also be used in a purely descriptive sense, meaning roughly something like "containing elements of fantasy". Let us indicate the first use by using exclamation marks ("fantastic!!") and the second by simply writing "fantastic" without exclamation marks. The fallacy involved in (1)–(3) is then analogous to the fallacy involved in the following argument:

(1') "Lord of the Rings", Part 1, is fantastic!!
(2') If "Lord of the Rings", Part 1, is fantastic, then "Lord of the Rings", Part 2, is fantastic.
(3') "Lord of the Rings", Part 2, is fantastic!!

(2') rests on the descriptive meaning of "fantastic": it means that if the first part of "Lord of the Rings" contains elements of fantasy, then the second does so as well. In (1') and (3'), in contrast, the use of "fantastic" is expressive. Now imagine someone argues: if one assents to (1') and (2') one is, by pain of logical inconsistency, committed to also assenting to (3'), for evidently the argument is a valid modus ponens argument. So given that (2') is true, someone who likes the first part of "Lord of the Rings" is, by pain of logical inconsistency, committed to also liking the second part. This is of course nonsensical. There is no logical mistake involved in enjoying the first part of "Lord of the Rings" and finding the second horrific. But, appearances to the contrary notwithstanding, it is nonsensical in the same manner to claim that if (2) is true, someone who disapproves of lying is, by pain of logical inconsistency, committed to also disapproving of getting one's little brother to lie. There is nothing logically wrong with someone disapproving of lying without disapproving of getting one's little brother to lie if we see that the meaning of "bad" in (1) and (2) differs in the way that the meaning of "fantastic" differs in (1') and (2').

At this point it might be tempting to object that (1')–(3') is not analogous to (1)–(3) because "fantastic" differs from "bad" insofar as the latter is not ambiguous. But according to expressivism "bad" *is* ambiguous in the same manner as "fantastic". The fact that "bad" can appear in unasserted contexts such as (2) shows exactly this. The argument is very simple: if we take a conditional like (2) at face value, i.e. as a coupling of singular sentences each of which is truth-apt, we *must* assume that "bad" is used descriptively in (2). So if we also assume that it is used expressively in (1),

there is no way of denying that "bad" differs in meaning in (1) and (2), just as "fantastic" differs in meaning in (1') and (2').

4. Blackburn's alternative defence of expressivism

One often-discussed strategy of avoiding the concession that (1)–(3) does contain a fallacy of equivocation is presented by Blackburn in his attempt to defend expressivism against the Frege-Geach attack.[7] Blackburn interprets (2) in a way that preserves the evaluative meaning of "bad". At the same time he attempts to show that we do not have to charge arguments like (1)–(3) with a fallacy of equivocation. To do so, he proposes we regard the conditional surface structure of (2) as misleading. On a deeper level, according to Blackburn, (2) can be construed as expressing approval towards a coupling of attitudes. Using "H!" as a "hooray-operator", i.e. as indicating an attitude of approval; "B!" as a "boo-operator", i.e. as indicating an attitude of disapproval; square brackets as indicating that we are talking about attitudes of approval or disapproval; and the semi-colon to denote the coupling of different attitudes, Blackburn renders (1)–(3) as follows:

(1ex) B! (lying)
(2ex) H! [B! (lying); B! (getting one's little brother to lie)]
(3ex) B! (getting one's little brother to lie)

So Blackburn's reading of the conditional (2) is that by uttering it a speaker expresses a pro- attitude towards the coupling of a con-attitude towards lying with a con-attitude towards getting one's little brother to lie. Blackburn argues that someone who endorses the con-attitude towards lying is thereby committed to also endorsing a con-attitude towards getting one's little brother to lie, provided that he also has a pro-attitude towards the coupling of con-attitudes referred to in (2ex). Failing to do so would involve him in a *clash of attitudes*. He then has "a fractured sensibility which cannot itself be an object of approval"[8]. In other words, someone who combines the commitment to the premises with a lack of commitment to the conclusion

---

[7] BLACKBURN, 1984, p. 189–196.
[8] BLACKBURN, 1984, p. 195.

would fail to have a combination of attitudes of which he himself approves. This, Blackburn thinks, demonstrates that someone who endorses (1) and (2) without endorsing (3) makes a logical mistake.[9]

Against this attempt to defend expressivism, it has been argued that it fails to give an account of why (1)-(3) is *logically* valid. Someone who gets involved in a "clash of attitudes" shows that his system of attitudes is *incoherent* and he is likely to incur our *moral* disapprobation. But the incoherency involved in this is no logical inconsistency. Our disapprobation of a person who gets involved in a "clash of attitudes" is mediated by the acceptance of a prescription like "Do not adopt incoherent attitudes!", and this appeals to a moral principle, not to a logical principle.[10]

What then *is* wrong with somebody who, assenting to (1ex) and (2ex), does not assent to (3ex)? It is likely that we will accuse him of *insincerity*. Assenting to (3ex) is a sincerity condition for assenting to both (1ex) and (2ex). The mistake involved in assenting to (1ex) and (2ex) but not to (3ex) is analogous to the one involved in Moore's famous example "It is raining, but I do not believe it": there is nothing *logically* wrong in saying this, but if I do not believe that it is raining, I cannot sincerely assent to the proposition "It is raining".[11] Analogously, there is nothing logically wrong in not disapproving of getting one's little brother to lie while disapproving of lying and approving of the coupling of attitudes referred to in (2ex). However, one cannot *sincerely* claim to disapprove of lying and to approve of the coupling of con-attitudes referred to in (2ex) while *not* disapproving of getting one's little brother to lie. It is this insincerity which is likely to provoke our moral disapprobation. So Blackburn's argument from (1ex) to (3ex) corresponds to a contextual implication (in the sense defined by Moore) rather than to a logical deduction, and he can be suspected of having confounded these.[12]

So the standard objection against Blackburn's attempt to defend expressivism is compelling. The failing is moral, not logical. But in light of what has been said above, this should not disturb the expressivist in the least.

---

[9] This is not Blackburn's last word on the issue. For his later attempt to interpret the conditional in terms of a comittment-theoretic semantic and the flaws of this attempt, see Hale (2002).

[10] See WRIGHT, 1988, p. 33; SCHUELER, 1988, p. 499–500; HALE, 1993, p. 344; MILLER, 2003, p. 62–64; HOFFMANN, 2008, p. 224.

[11] See MOORE, 1993; see also VAN ROOJEN, 1996, p. 332–33.

[12] See MOORE, 1993, p. 210. I cannot here go into the details of the discussion of Moore's paradox. For the relation between contextual implication and sincerity conditions, see AUSTIN, 1962, p. 48–50.

Quite on the contrary, Blackburn's failure to reconstruct (1)–(3) as logically valid ought to be regarded as *good* news for expressivism because the expressivist should regard it as his aim to show that if "bad" is used evaluatively in (1), (3) does not follow from (1) and (2). If what has been said above is correct, expressivism is committed to *denying* that (1)–(3) is valid (unless "bad" is used in a purely descriptive sense in (1) *and* (2)). Blackburn's failure to save expressivism from the objection that it exposes (1)–(3) to the charge of a fallacy of equivocation, albeit in a way not intended by Blackburn, gives support to the assumption that there is nothing which hinders the expressivist from holding on to this view: as things stand, attempts to pull the expressivist to the cognitivist camp fail. The failure of Blackburn's argument demonstrates that even with an expressivist interpretation of (1)–(3), it cannot be construed as a logically valid argument.

Conclusion

The upshot of my argument is that, leaving the purely descriptive uses of terms like "bad" aside, the expressivist need not and ought not spend his energy on attempting to show that (1)–(3) is valid. Quite on the contrary, he should welcome any persuasive account of why, given that "bad" is used evaluatively in (1), (3) does *not* follow from (1) and (2). Expressivists tend to make unnecessary concessions to their cognitivist opponents by accepting the constraint that, within an expressivist framework, (1)–(3) must be reinterpreted so as to free the argument from the charge of a fallacy of equivocation. But the key to defending expressivism lies in refusing to accept this constraint. The Frege-Geach problem poses a serious threat to expressivism only as long as the expressivist obediently accepts cognitivist metaethical assumptions. But he need not do this. Rather than trying to integrate cognitivist intuitions into his theory, he should reject them from the start. If he does that, it turns out that the Frege-Geach argument does not constitute a serious problem for expressivism.

References

AUSTIN, J. *How to Do Things With Words*. 2. Ed. by J.O. Urmson / M. Sbisà. Harvard: HUP, 1975.

BLACKBURN, S. *Spreading the Word. Groundings in the Philosophy of Language*. Oxford: OUP, 1984.

ERIKSSON, J. "Homage to Hare: Ecumenism and the Frege-Geach Problem". In: *Ethics*, 120, 2009, p. 8-35.

GEACH, P. "Ascriptivism". In: *Philosophical Review*, 69, 1960, p. 221-225.

_____ . "Assertion". In: *Philosophical Review*, 74, 1965, p. 449-465.

HALE, B. "Can There Be a Logic of Attitudes?". In: *Reality, Representation, and Projection*. Haldane, J., Wright, C. (ed.). Oxford: OUP, 1993, p. 337-363.

_____. "Can Arboreal Knotwork Help Blackburn out of Frege's Abyss?". In: *Philosophy and Phenomenological Research*, 65, 2002, p. 144-149.

HARE, R. *The Language of Morals*. Oxford: OUP, 1952.

HOFFMANN, M. *Kohärenzbegriffe in der Ethik*. Berlin/New York: De Gruyter, 2008.

MILLER, A. *An Introduction to Contemporary Metaethics*. Cambridge: Polity Press, 2003.

MOORE, G. E. "Moore's Paradox". In: MOORE, G. E. *Selected Writings*. Ed. by Th. Baldwin. Cambridge: Routledge, 1993, p. 207-212.

RIDGE, M. "Ecumenical Expressivism: Finessing Frege". In: *Ethics* 116, 2006, p. 302-336.

SCHUELER, G. F. "Modus Ponens and Moral Realism". In: *Ethics*, 98, 1988, p. 492-500.

SEARLE, J. "Meaning and Speech Acts". In: *Philosophical Review*, 71, 1962, p. 423-432.

SINNOTT-ARMSTRONG, W. "Expressivism and Embedding". In: *Philosophy and Phenomenological Research*, 61, 2000, p. 677-693.

VAN ROOJEN, M. "Expressivism and Irrationality". In: *The Philosophical Review*, 105, 1996, p. 311-335.

WRIGHT, C. "Realism, antirealism, irrealism, quasi-realism". In: *Realism and Antirealism*. Edited by P. French, Th. Uehling, H. Wettstein (*Midwest Studies in Philosophy* 12). Minneapolis: University of Minnesota Press, 1988, p. 25-49.