

# Self-Consciousness and Reductive Functionalism

By Arvid Båve, Hamburg University

## *Abstract*

It is argued that, although George Bealer's influential "Self-Consciousness argument" refutes standard versions of reductive functionalism (RF), it fails to generalize in the way Bealer supposes. To wit, he presupposes that any version of (RF) must take the content of "pain" to be the property of being in pain (and so on), which is expressly rejected in independently motivated versions of Conceptual Role Semantics (CRS). Accordingly, there are independently motivated versions of (RF), incorporating (CRS), which avoid Bealer's main type of refutation. I focus particularly on one such theory, which takes concepts to be event types that are individuated by their psychological roles, which has the resources of responding to each of the more specific worries Bealer expresses.

Keywords: Self-Consciousness, Functionalism, Mind-Body Problem, Materialism, Conceptual Role Semantics, Inferentialism

## Introduction

Physicalistic reductive functionalism (RF) is to many the most promising materialistic theory of the mind. But there is one especially serious objection against it, the "Self-Consciousness argument", developed in detail by George Bealer.<sup>1</sup> In contrast to most other objections against (RF), Bealer's does not depend on accepting any premise uncongenial to functionalists (like the claim that there are intrinsic *qualia*) (cf. Bealer (2010: 141)). Finding an adequate response to this argument, which is the purpose of this paper, would thus be an important step toward consolidating materialism.

The response is that Bealer presupposes that functionalists must make certain assumptions about meaning and content that they may in fact reject. Indeed, adherents of CRS are committed to rejecting them, and thus the combination of RF with CRS avoids Bealer's main objection, as well, I argue, as his additional objections. (CRS) is of course contentious, but since Bealer's ambition is to show that all versions of (RF) fall afoul of his argument, its contentiousness is irrelevant. Since I agree with Bealer that his argument refutes standard

(RF), one can conclude from the case I make in this paper that adherents of (RF) must adopt (CRS).

I first introduce (RF) in its standard form (§1), Bealer's argument against it (§2), his discussion of non-standard varieties of (RF) (§3), and go on to argue that Bealer illegitimately presupposes that the contents of predicates are the properties they express (§4). I next present two versions of (CRS), which I submit are untouched by Bealer's arguments, Paul Horwich's Use-Theory of Meaning (§5) and my own account of concepts (§6). The rest of the paper focuses on the latter, arguing that it provides independent reasons for deeming irrelevant some (true) claims that Bealer stresses (§7), and is able to accommodate two principles about attitude-contexts that Bealer thinks doom (RF) (§8). I finally discuss questions of the reference of psychological concepts and how they relate to Bealer's argument, and finally summarize the paper (§9).

## **1. Functionalism: reductive and non-reductive**

Functionalism is best introduced as an improvement on its predecessors, behaviourism and the mind-body (type) identity theory. The identity theory identifies psychological properties with physical properties, claiming, e.g. that being in pain is the same as undergoing the firing of C-fibres (or some such physical property). Behaviourism, by contrast, identifies psychological properties with sensory-behavioural *dispositions* of individuals, rather than categorical, intrinsic physical properties. Both theories clash with widely shared intuitions: it seems clear that one can be in any given psychological state without being in any specific physical state, and without having any specific sensory-behavioural disposition (and *vice versa*).<sup>1</sup>

Functionalism has seemed to many to incorporate the insights of both predecessors while steering clear of their failings. According to the kind of reductive, physicalistic

functionalism at issue here, to have a psychological property, like being in pain, is not to have a specific physical property, but rather to have *some* physical property playing the appropriate (typically causal) role, e.g. the “pain-role”. A physical property playing the role of a psychological property is said to *realize* that property.

Most functionalists also think that when a physical property realizes a psychological property, the person's instantiating the former is *identical* to her instantiating the latter. But this is not *type* identity, but *token* identity. Finally, all theories I am here calling “functionalism” are versions of so-called “role-functionalism”, in contrast to “realizer-functionalism”. On the present terminology, the latter does not count as a kind of functionalism, but rather as a type-identity theory with a kind of “functionalist” view of the way the reference of psychological terms is fixed.

Playing the role of a given psychological property is a matter of standing in the right relationships to sensory inputs, behavioural outputs, and *other mental states*. This sounds like defining psychological properties in terms of other psychological properties, but this is merely an informal explanation. As we will see anon, it can be explicated clearly using so-called ramsification, by which we can turn a psychological theory implicitly defining all psychological predicates within it, into fully reductive, non-circular, explicit definitions of each psychological property.

The remaining exposition of functionalism will roughly follow Bealer's. The differences are motivated by considerations of clarity and do not substantially affect the discussion of his arguments. Reductive functionalism (RF) is defined as the view that each psychological predicate has a *reductive definition*. To obtain a reductive definition of a psychological property, one first considers an “original theory”, O, which is a long conjunction of open sentences, with the free variable “*x*”, ranging over thinking beings. Bealer actually takes the clauses of this theory to be closed sentences (see, e.g. (1997: 77) and

(2001: 286)). However, this results in too weak (hence, false) reductive definitions. Since Bealer's arguments do not hinge on this mistake, however, I will make nothing more of it.

The conjunction of these open sentences is meant to give a complete, implicit definition of all the mental terms it contains. Candidates for such “clauses” may be:

- (1)  $(x \text{ is in pain} \& C) \rightarrow x \text{ believes that } x \text{ is in pain.}$
- (2)  $(x \text{ believes that } [x \text{ does } y \rightarrow p] \& x \text{ desires that } p \& C') \rightarrow x \text{ does } y.$
- (3)  $(\text{it looks to } x \text{ as if } p \& C'') \rightarrow x \text{ will believe that } p.$

The dummy-conjuncts “ $C$ ”, and so on, stand in place of qualifying conditions that needs to be inserted into (1)-(3), such as (perhaps) “ $x$  considers whether  $x$  is in pain”, “no inhibiting factors are present”, etc.

Next, we replace all monadic mental predicates by an equivalent predicate containing a property designator, so that, e.g. “is in pain” is replaced by “has the property of being in pain”. Similarly, dyadic predicates are replaced by equivalent predicates containing relation designators, so that “believes” gets replaced by, “stands in the relation of believing to”. Then, we replace all designators of mental properties and relations with first-order variables, “ $r_1$ ”, ..., “ $r_n$ ”. For instance, all occurrences of “the relation of belief” might get replaced by “ $r_2$ ”. Let  $V$  be an instruction for making such replacements and let  $X^V$  be the result of making those replacements in the expression  $X$ . Thus,  $O^V$  is the result of replacing all mental predicates in  $O$  in accordance with  $V$ . Assume also that “is in pain” $^V = "r_1"$ .

The reductive definition of pain then reads,

$$(RDP) \quad x \text{ is in pain iff}_{\text{def}} \exists r_1 \dots r_n (r_1, \dots, r_n \text{ are physical} \& O^V \& x \text{ has } r_1),$$

and analogously for all other mental predicates in O. (To avoid trivializing the *definientia*, the physical properties must also be assumed to be *sufficiently natural*, but I will leave this implicit.)

We get non-reductive (or semi-reductive) definitions by omitting the claim that  $r_1, \dots, r_n$  are physical. On such definitions, although mental properties can be defined in non-mental *terms*, they are still not fully reductive (more on this anon). *Non-reductive functionalism* is the view that all psychological predicates have such non-reductive definitions.

Bealer thinks non-reductive functionalism is true and that it is the natural retreat position for functionalists in view of his argument. But, says Bealer, non-reductive functionalism is no materialistic theory, for it allows psychological properties to be ontologically first-order, whereas (RF) takes them to be merely second-order, defined in terms of the interactions between their first-order realizations.

That a kind of property is second-order does not mean that it is unreal. It means only that it can be reductively defined in terms of first-order properties, in the way described. If there are merely non-reductive definitions of psychological properties, we can no longer say that “one is in pain just in case one has some *first-order* (e.g. physical) property which plays the pain role”.

First-order properties, Bealer says, are those taken as metaphysically irreducible or primitive, and may thus be referred to in the *definientia* of reductive definitions of higher-order properties. According to physicalism, only physical properties are first-order, but on, e.g. ontologically phenomenalist views, they are second-order, and the first-order properties are rather something like sense-data. We could in principle imagine non-physicalistic versions of (RF), but I will throughout use “(RF)” to refer to a physicalistic variety of (RF), i.e. one on which all first-order properties are physical.

## **2. Bealer's refutation of standard (RF)**

Bealer's papers on this topic are not always easy to follow, and my presentation below is the result of a rather arduous attempt to weed out the essentials and formulate them as clearly and succinctly as possible. Where any worries arise, Bealer's original texts should be consulted. The presentation of the first part of his argument, with which I have no issues, will contain only a minimum of exegetic remarks, whereas the latter part, which I argue is flawed in various ways, will be presented with closer attention to the original texts. Large swathes of Bealer's discussion are not discussed at all, since they concern various possible responses with no connection to mine, or concern some version of functionalism that I am not defending here.

The crux of the matter, according to Bealer, is that the original theory from which reductive definitions are to be extracted, must contain clauses in which *psychological predicates occur within the scope of propositional attitude predicates*, as in the “self-intimation” principle for pain (1) above. All leading reductive functionalists who have stated a view on the matter agree that if the reductive definitions are to come out true, the original theory must contain such principles (see, e.g. David Lewis (1983: 103) and Sydney Shoemaker (1994: 59)). Bealer (2000: 507ff.) also describes convincing counter-examples to definitions drawn from an original theory lacking such principles.

In a nutshell, the problem is that an original theory which contains the self-intimation principle (1) will yield reductive definitions that entail that if  $x$  is in pain (and the further qualifying conditions  $C$  obtain), and if pain is realized in  $x$  by physical property  $P$  (e.g. having firing C-fibres), then  $x$  will believe that  $x$  has  $P$ . Note that this consequence might be acceptable to type-identity theorists or “realizer functionalists”, since they identify psychological properties with physical ones. They can therefore say that the consequence is harmless, as long as the belief-ascription is read *de re*. But since role-functionalists say that

physical realizer-properties and the psychological properties they realize are *distinct*, the consequence is unacceptable. Even if the property of having C-fibres firing realizes pain in someone, they should certainly not be expected to come to know that they have that physical property merely by introspection.

To see how the argument goes, assume that principle (1) is among the clauses of our original theory O. Then, (RDP) will be equivalent to:

$$(RDP') \quad \text{Pain}(x) \text{ iff } \exists r_1 \dots r_n (r_1, \dots, r_n \text{ are physical} \& O^V \& x \text{ has } r_1 \& \\ ((x \text{ has } r_1 \& C) \rightarrow x \text{ stands in } r_2 \text{ to the proposition that } x \text{ has } r_1)).$$

(For visibility, I have put the crucial conjunct, corresponding to principle (1), last.) Further, assuming that “belief”<sup>V</sup> = “ $r_2$ ”, the definition of belief will read,

$$(RDB) \quad x \text{ believes } y \text{ iff}_{\text{def}} \exists r_1 \dots r_n (r_1, \dots, r_n \text{ are physical} \& O^V \& x \text{ stands in } r_2 \text{ to } y).$$

Now assume that  $x$  is in pain and that  $C$  obtain. Given (RDP’), it follows that

$$(i) \quad \exists r_1 \dots r_n (r_1, \dots, r_n \text{ are physical} \& O^V \& x \text{ has } r_1 \& C \& \\ (x \text{ has } r_1 \& C) \rightarrow x \text{ stands in } r_2 \text{ to the proposition that } x \text{ has } r_1).$$

But (i) and (RDB) logically entail,

$$(ii) \quad \exists r_1 \dots r_n (r_1 \dots r_n \text{ are physical} \& O^V \& x \text{ believes the proposition that } x \text{ has } r_1).$$

By conditional proof, if  $x$  is in pain and the further conditions  $C$  obtain, then, for some physical property  $P$  which realizes pain in  $x$ ,  $x$  believes that  $x$  has  $P$ .

In other words, the theory entails that, for all  $x$ , if  $x$  is in pain and  $C$  obtain, then  $x$  will believe that  $x$  has a certain *physical* property, namely, the physical property realizing pain in  $x$ , which is unacceptable. This argument applies very generally, as long as the original theory contains clauses with psychological predicates within the scope of a propositional attitude predicate, and regardless of which qualifying conditions we insert into these clauses.

The closest to this formulation of the argument is in Bealer (1997: 78ff.). The difference is that the argument found there appeals to a propositional attitude of “self-conscious awareness”, rather than belief (hence, the title of his original paper). However, this variation makes no difference for present purposes. The difference between these attitudes that Bealer emphasizes is that the latter is supposed to be an attitude one can hold only toward contents about one’s psychological properties, and never about one’s physical properties. The arguments are partly analogous to that above, but their upshot is a contradiction with the assumption that self-conscious awareness can only be awareness that one has a psychological property.

### **3. Modifications of standard (RF)**

Standard (RF) comprises two key elements responsible for the problem discovered by Bealer, namely, (i) the claim that the original theory features principles like (1)-(3) above, with psychological predicates within the scope of ordinary propositional attitude verbs, and (ii) the above method of extracting definitions from the original theory, where all occurrences of each psychological predicate are replaced by the same predicate variable.

Considered as an argument against standard (RF), Bealer’s argument is, I think, conclusive. However, Bealer also takes pains to show that various attempts to modify the

standard theory so as to avoid the problem will fail too. This question is clearly more vague, and can scarcely be answered as decisively. We can distinguish two (mutually exhaustive) ways of modifying standard (RF): abandoning either (i) or (ii).

The first option is to modify the formulation of the original theory. In particular, one might replace its ordinary-English attitude-ascriptions by some other kind of attitude-ascription. The second option is to alter the method of extracting reductive definitions from the original theory. Bealer argues that the latter kind of strategy is unworkable. For instance, we clearly cannot avoid the problem by deciding that the embedded occurrences of “pain” not be replaced by a variable in the procedure of extracting a definition. For unless all occurrences of psychological predicates are replaced by variables, the resulting definitions will have psychological predicates in their right-hand sides, contrary to the reductive ambitions of (RF).

Could one just replace the embedded occurrences of “pain” and unembedded ones with different predicate variables? Bealer claims one could not, since doing so would be acceptable only if these different occurrences differed semantically. Bealer insists, however, that they do not, appealing to the principles,

(P) The proposition expressed by a sentence “ $x$  is in pain” is the same as the proposition referred to by the that-clause in, “ $x$  believes that  $x$  is in pain” (1997: 82).

(P') An atomic sentence “ $Fa$ ” expresses the proposition denoted by the “that”-clause “that  $Fa$ ”; in each case, the proposition is formed from that which is expressed by the predicate “ $F$ ” (*ibid.*).

These and similar principles play an important role in Bealer's argument, as we shall see, and I will throughout grant them. But it is a *non sequitur* to think that accepting them somehow debars one from stipulating that embedded and unembedded occurrences should be replaced by different variables when stating reductive definitions. These two matters are just unrelated. Still, I will not defend this modified version of (RF) here, since I think it fails for other reasons. Accordingly, the kind of counter-examples against Bealer's contentions I will provide are not versions of (RF) on which terms functioning the same way semantically get replaced by different variables in the process of ramsification. Thus, although I reject Bealer's constraint on ramsifications, the counter-examples I will present happen to satisfy it.

Let us turn to the other kind of strategy, on which the original theory does not feature ordinary attitude-ascriptions. One alternative way of referring to propositional attitudes is based on the idea of a “language of thought” (LOT). On this hypothesis, to have the belief that snow is white is to have, tokened in one’s “Belief Box”, a sentence in LOT meaning that snow is white (and similarly for other attitudes and contents). If the original theory contained such attitude ascriptions instead of ordinary ones, it may be thought, the resulting reductive definitions would avoid the conclusion of Bealer’s argument. For, on this combined view, we would not have troublesome clauses in the original theory like (1) above, but rather something like,

- (1')  $(x \text{ is in pain} \& C) \rightarrow x \text{ will have, tokened in } x\text{'s Belief Box, a sentence of the form}$   
“ $F(a)$ ”, where “ $a$ ” refers to  $x$  and “ $F$ ” means *pain*.

Actually, Bealer does not consider this particular candidate, but it will prove to be a useful example in what follows, since it is neutral about what contents are. It is clear that some kind

of content specifications of the LOT expressions are necessary, since otherwise the principle would not say what beliefs it is that are prompted by pains.

On some ways of specifying the content of the Mentalese predicate “*F*” in principles like (1'), Bealer's suspicions are clearly borne out, e.g.

- (a)  $\forall x(\text{“}F\text{” is true of }x \text{ just in case }x \text{ is in pain})$ ,
- (b) the content of “*F*” = the property of being in pain,
- (c) the content of “*F*” = the function  $f: w \rightarrow \{x: x \text{ is in pain at }w\}$ .

With (c), for instance, we could conclude, with an argument analogous to the rebuttal of standard (RF) above, that if  $x$  is in pain and the “prompting conditions”  $C$  obtain, then a sentence of the form “*F(a)*” will be tokened in  $x$ 's Belief Box, where “*a*” refers to  $x$  and the content of “*F*” = the function  $f: w \rightarrow \{x: x \text{ has }r_k \text{ at }w\}$ , where  $r_k$  is a physical property realizing pain in  $x$ , rather than pain. But this of course assigns the wrong meaning to “*F*”, and so has the consequence that when  $x$  is in pain and  $C$  obtain, a sentence is tokened in  $x$ 's Belief-Box which means that  $x$  has physical property  $P$  (for the relevant  $P$ ).

The following content-specification, inspired by a causal theory of reference, would succumb to a similar problem:

- (d) the content of “*F*” = the property  $R$  such that the subject's having  $R$  would in normal conditions cause “*F(i)*” to be tokened in  $y$ 's Belief Box (where “*i*” is the first-person pronoun)

(cf. Bealer (1997: 94ff.)). For, Bealer insists, this specification would be correct only if the variable “*R*” would have pain among its values. But then, it could not take part of a definition of pain meeting the constraint set by (RF) that its variables range only over first-order (physical) properties.

This raises the question whether there is any way we could say, or ensure, that the predicate in the tokened sentence means pain, but without using “pain”, or at least not use it in such a way that it gives rise to the problem Bealer discovered. An obvious idea is that some kind of conceptual role semantics (CRS) might do the job, since on this kind of view, the meaning of an expression is not identified with any worldly item like an ordinary property, object, extension, or intension, but rather with the psychological role the expression (or concept) plays. Indeed, what would be more natural for a functionalist than to take meanings and contents to be individuated the same way as psychological properties, namely, by the roles they play in our psychology? This combination of functionalism and (CRS) has indeed been advocated by numerous leading functionalists, including Sellars (1968: Ch. 6), Harman (1973, 1987), and Block (1986).<sup>2</sup>

Bealer deals with this kind of proposal very quickly and essentially the same way he deals with the causal identification of the content of the Mentalese predicate above. He considers the following proposal (changed to fit our example of pain):

If it is nomologically necessary that [“*F*”] behaves in sentences in the respective Boxes (Thinking Box, considering Box, Belief Box, etc.) in the way the predicate [“pain”] behaves in [the original theory], then the content of [“*F*”] =<sub>def</sub> [the property of being in pain] (1997: 99).

and goes on to object,

As it stands, the analysis is viciously circular; the [property of being in pain] is mentioned on the right-hand side of the definition. But this is one of the [properties] that, ultimately, we are trying to define (*ibid.*).

There are many problems with this line of reasoning, but there is one that stands out and which, I will argue, constitutes the main mistake in Bealer's overall argument.

#### **4. Bealer's contentious presuppositions**

Bealer clearly presupposes that any adequate specification of the content of a predicate will identify this content with the worldly property we would normally say is “the property expressed” by the predicate, so that the content of “pain” is the property of being in pain, and so on. This is presumably due to a presupposition that the content of a predicate *is* the property it expresses in this “disquotational” way (this is an integral part of the philosophical system he lays out in Bealer (1982)). Given these assumptions, his arguments go through without much further ado.

But these assumptions would typically be rejected by adherents of (CRS). Consequently, the alleged (CRS) content-specifications he considers do not resemble the content-specifications that adherents of (CRS) would propose. This should have been clear from the outset: (CRS) does not claim that the content of an expression is the worldly property it expresses (so that the content of “pain” is the property of being in pain, and so on), but rather its psychological role. The notion of “psychological role” is not by itself (i.e. without explication) very precise, but looking at some more specific proposals bears the point out.

For instance, some adherents of (CRS) say that the content of an expression is a concept, which is an entity individuated by its possession conditions, which in turn are to the effect that it plays a certain conceptual (psychological, inferential, functional, etc.) role. On

this view, the content of “pain” cannot possibly be the property of being in pain, and likewise for “red” and the property of being red, and so on. In other words, propositions whose identity is to be determined by the psychological roles of their constituents cannot be Russellian, since it typically makes no sense to speak of the “psychological role” of worldly objects, properties, and relations. Such propositions thus have to be Fregean.

Bealer asserts repeatedly that the content of “pain” obviously is pain. This is true as far as it goes, but the question is whether the second occurrence of “pain” here refers to the property of being in pain. It might as well—as many have indeed suggested on independent grounds—be taken to refer to a concept, distinct from the property pain (see Davis (2003)). Perhaps Bealer just *means* by the word “content”, “the worldly property expressed”. The word “content” is a term of art, after all, so this would not be in breach of any linguistic norm. If so, however, it is still clear that this part of his argument will not be very convincing, since the “content-specification” will then be held irrelevant by anyone with a different view about contents, properly so-called.

Consider also content-specification (1'), which I claimed is neutral as to the nature of meanings. If Bealer's argument is to be applicable here, he must assume that the italicized occurrence “*pain*” functions the same way as ordinary occurrences of “pain”. But this is far from obvious. For instance, we could not very well say that “*happy*” and “*and*” function the same way in “means *happy*” and “means *and*” as they do in ordinary contexts (where they occur rather as an adjective and a connective, respectively). But these phrases should plausibly be given a uniform account. Note also that saying that the occurrences of “pain” in “He experiences pain” and “‘Schmertz’ means pain” function differently is consistent with principles (P) and (P') above.

A similar response can be given to Bealer's charge that an argument analogous to his original one could be used to show that his problem could not be avoided by recourse to the

idea that the belief prompted by pain could be described as involving the *concept of pain*. The relevant reformulation of (1) he considers is,

if  $x$  has the property of being in pain [...], then  $x$  will [believe] a predicative proposition formed from the-concept-of-being-the-property-of-being-in-pain, which is a concept of the property of being in pain (1997: 83).

He then concludes that “the property of pain” in “the concept of the property of pain” must be quantified into when ramsifying, thus allowing his now-familiar kind of *reductio*. Pointing out the error here is a mere repetition of the points just made: it has been independently argued that expressions following “the concept (of)” do not function as they ordinarily do, and to claim this is consistent with (P) and (P').

The points made above reveal two further lacunae in Bealer’s argument. Suppose that the occurrence of “*pain*” in (1') does not work like ordinary occurrences of “*pain*”. Then, it need not be replaced by the same variable as those latter occurrences. Could it simply be left as it is? If taken to refer to a mental entity, then this would conflict with the aims of (RF). But the claim that it refers to a mental entity requires further support that Bealer does not provide.

Bealer would very likely reply that even if “*pain*” here does not refer to a mental entity, a predicate’s meaning *pain* is something that must be elucidated, and once it is, the result will run into problems analogous to those already discussed. But the claim that the content specification in (1') needs to be further elucidated at all *also* requires further support that Bealer does not provide. It is not an uncommon view, after all, that meaning-facts, like a predicate’s meaning *pain*, and so on, are *sui generis* and irreducible. If this view is correct, then we *could not* give more informative content specifications than those offered in (1'). I will not make anything more of this observation here, however, but rather grant that content-specifications must be elucidated in the original theory.

## **5. Counter-example 1: Horwich's Use-Theory of Meaning**

By a “counter-example” to Bealer’s contention about (RF), I mean a version of (RF) not subject to his objections. One such counter-example is a version of (RF) combined with Paul Horwich’s Use-Theory of Meaning (see his (2005: Ch. 2)). This theory is especially suitable for present purposes since Horwich, as we will see, accepts a version of the LOT hypothesis. On Horwich’s theory, the semantic axioms for primitive expressions take the following form:

- (M) The fact that *e* means *m* is constituted by the fact that the explanatorily basic fact about *e*’s use is that (competent) speakers tend to accept sentences *s*, *s'*, ... containing *e* (categorically, or under certain specified circumstances).

Here, we could replace “*x* accepts *s*” by, “*s* is tokened in *x*’s Belief Box” (Horwich indeed treats them as interchangeable—see his (2005: 40f.)). The idea is that the functionalist’s original theory could consist of the kind of attitude-attributions found in (1’), where the meaning-claims (““*F*” means *pain*”, etc.) are accompanied with specifications of the form (M), e.g.

- (F) The fact that “*F*” means *pain* is constituted by the fact that the explanatorily basic fact about its use is that competent speakers’ disposition to accept *s*, *s'*, ...

Bealer’s objections apply to the resulting version of (RF) only if “*pain*” functions like ordinary occurrences of “*pain*”, which has yet to be shown (and which is denied by Horwich on independent grounds—see his (2005: 33f.)).

This is the kind of content-specification that Bealer should have considered in making his case against (RF). If he had, I submit, he would not have been able to show it unworkable, at least not by any obvious modifications of his arguments against other variants of (RF). However, Bealer's case against (RF) is complex and many-pronged, involving *inter alia* the principles (P) and (P'), the question whether psychological predicates are typed, and a discussion about self-attribution, so we should not be content merely to have mentioned a counter-example: further investigation is required to show that all of Bealer's points can be duly responded to by adherents of the theory.

I will now leave Horwich's Use-Theory and focus instead on another theory, which is a second, and rather different counter-example to Bealer's argument. My discussion about the various further points Bealer makes will reference this theory rather than Horwich's. I do this mainly because I favour this theory over Horwich's, but also because it will be illustrative to see how such different theories can both be shown to steer clear of the problems Bealer describes. I also think my favoured theory is able to handle some of these problems in a clearer way.

A final, important reason for finding an alternative to Horwich's theory is that the latter incorporates the LOT hypothesis, which entails that attitudes are realized as physical tokenings of sentences in some physical medium. But functionalism is typically credited with the benefit of allowing the possibility-in-principle of "disembodied minds", i.e. non-physical realizations of mental properties. So, as Bealer notes (1997: 91), functionalists cannot accept the LOT hypothesis without sacrificing this benefit.

## **6. Counter-example 2: (CRS) with an event-type conception of concepts**

The theory I will now present has been defended and described in some detail in unpublished manuscripts, but I can here only summarize its main theses and motivations. I take every

concept to belong to some syntactic category or other: it can be an individual concept, predicative (monadic, dyadic, and so on), propositional, propositional-operator, and so on. Propositional concepts/contents are simply propositions. Further, I use words in boldface to form singular terms referring to simple concepts, so that “Socrates” expresses the individual concept **Socrates**, and so on (note that this transformation results in a felicitous expression only if the original word is semantically primitive).

On the theory I will now expound, concepts are *psychological event types* (see Davis (2003) for a detailed defence of this view). Concepts can be either simple or complex. To undergo (a token of) such an event type is to *entertain* the relevant concept. Thus, entertaining  $x =$  undergoing (a token of)  $x$ . As against this, one might object that concepts cannot be event types for one can undergo (tokens of) event types but one cannot undergo concepts. However, it has been argued convincingly by Benjamin Schnieder (2006) that this type of argument is fallacious. Roughly, the problem is that the negation in the premise is supported only on a metalinguistic reading, on which, however, the argument is invalid.

A complex concept is an event type of *conjoining* the concepts immediately involved in the complex concept in question, where conjoining is a multigrade relation whose first *relatum* is an agent, the second is an incomplete concept, and the following *relata* are concepts saturating the incomplete concept. This is an instance of the now-popular “act-type theory of propositions”.<sup>3</sup> To designate complex concepts, I will use “ $C$ ” as a functor, defined,

(CD)  $C(y_1, \dots, y_n)$  = the event type of conjoining  $y_1, \dots, y_n$  (in that order).

In accordance with the above gloss on what conjoining is like, one can think heuristically of  $C(y_1, \dots, y_n)$  as the event type of saturating  $y_1$  with  $y_2, \dots, y_n$  (in that order). It can be seen that the following equivalence holds:

$x$  undergoes (i.e. entertains)  $C(y_1, \dots, y_n)$  iff  $x$  conjoins  $y_1, \dots, y_n$  (in that order).

I have argued elsewhere that this theory allows us to capture many intuitive ideas about concepts, and withstands both Benacerraf worries and charges of chauvinism. The following identities will further illustrate this theory, and show how these concept-designators relate to our ordinary “that”-clauses:

- (I1) (The proposition) that Socrates is wise =  $C(\mathbf{wise}, \mathbf{Socrates})$ ,
- (I2) (The proposition) that Socrates is not wise =  $C(\mathbf{not}, C(\mathbf{wise}, \mathbf{Socrates}))$ ,
- (I3) (The proposition) that Mary loves John =  $C(\mathbf{love}, \mathbf{Mary}, \mathbf{John})$ ,
- (I4) (The proposition) that Mary believes that Socrates is wise =  
 $= C(\mathbf{believe}, \mathbf{Mary}, C(\mathbf{that}, C(\mathbf{wise}, \mathbf{Socrates})))$ .

In (I4), **that** is taken as a nominalizer-concept on propositions, i.e. a concept that can be conjoined with a proposition to yield an individual concept referring to that proposition. Such proposition-nominalization concepts are expressed by “that”-clauses.

The (CRS)-part of the theory borrows elements from Christopher Peacocke’s (1992) theory but also diverges in certain important ways. It proposes individuations of concepts in terms of their inferential roles, but takes a concept’s playing a certain role to be, not the condition of possessing it (as in Peacocke’s theory), but of *entertaining* it. Since concepts are taken to be psychological event types, and since undergoing such an event type is the same as entertaining the relevant concept, these concept individuations come out as instances of the more general fact that event types can be individuated by the conditions of undergoing them (or their tokens). This attractive feature is not easily had on Peacocke’s view.

On this view, a concept like **and** is individuated by a claim of the form,

- ( $\&$ )  $x$  undergoes (entertains) **and** iff, for some event-type  $e$ ,  $e$  plays inferential role  $R$  and  $x$  undergoes  $e$ .

The sentence, “ $e$  plays inferential role  $R$ ” can of course be explicated in many ways. For illustration, I here present a simple dispositionalist proposal:

(R) For all propositions,  $p, q$ :

- (i)  $(x \text{ believes } p \text{ and } q \& \dots) > x \text{ believes } C(e, p, q) \&$
- (ii)  $(x \text{ believes } C(e, p, q) \& \dots) > x \text{ believes } p \&$
- (iii)  $(x \text{ believes } C(e, p, q) \& \dots) > x \text{ believes } q,$

where “ $p$ ” and “ $q$ ” are *first-order* variables, “ $>$ ” is a counterfactual connective and “ $\dots$ ” are to be filled out with the necessary “prompting” conditions. These conditions should include  $x$ ’s entertaining the relevant propositions and the person’s not having any reason against making the inference. It is there also argued that the possibility of neural noise blocking the inference can be handled with a further complication, but for present purposes, we can rest content with (R).

The theory presented so far arguably would not count as a properly reductive theory if concept-designators like “**pain**” were left in the *definientia* of reductive definitions of pain, belief, etc. Concepts, after all, were taken here as *psychological* event-types. The conclusion to draw is that concept-designators, too, must be replaced by variables in the ramsification procedure. We can now obtain reductive definitions of concepts by replacing concept

designators like “**pain**” in  $O$  with variables ranging over physical event types, which in turn allows us to state reductive definitions of concepts like,

$x$  undergoes (i.e. entertains) **pain** iff<sub>def</sub>  $\exists r_1 \dots r_n (r_1, \dots, r_n \text{ are physical} \ \& \ O^V \ \& \ x \text{ undergoes } r_3)$ ,

where “ $r_3$ ” replaces “**pain**”. Although I think this expansion of (RF) is promising, whether it is ultimately workable is a question that goes far beyond the case Bealer has made, and need therefore not be further discussed here.

Let us call the theory of concepts presented above  $(CRS^+)$ . Note that  $(CRS^+)$  entails that concepts can be fully individuated without any claim to the effect that they refer to so and so, or have such and such an intension, etc. This is crucial for showing that  $(CRS^+)$  does not succumb to Bealer’s refutation (more on which in Section 7).

If we combine (RF) with  $(CRS^+)$ , the self-intimation principle for pain in the original theory will contain the concept-designator “**pain**”, rather than an occurrence of “pain” within the scope of “believe”. It could read, for instance,

$$(1'') \quad (x \text{ is in pain} \ \& \ C) \rightarrow \text{for some } de \ se \text{ concept } c, x \text{ will believe } C(\mathbf{pain}, c).$$

Bealer’s *reductio* cannot take off from (1’’), since there is no occurrence of “pain” within the scope of an attitude verb here. This point generalizes beyond the specific wording of (1’’). Thus any worries about the other aspects of (1’’), e.g. about “*de se* concepts” can be ignored, as long as there is *some* adequate way of referring to the proposition.

What about “content specifications”? On the current theory, these will be concept individuations like (&). Concept individuations of psychological concepts (like **pain**) will

often refer to beliefs about psychological properties like belief. It is therefore necessary to say something about propositions in which such properties are ascribed. (I4) above identifies the syntactic structure of such propositions, and is most naturally coupled with a broadly “Fregean” view, on which **believe** is a regular, dyadic predicative concept, defined over individual concepts of people and individual concepts of propositions. The latter are event types of conjoining the “nominalizing” concept **that** with a proposition.

Now, individuations of psychological concepts like **pain** will be of the same form as (&), i.e.

(CP)       $x$  undergoes (entertains) **pain** iff, for some event-type  $e$ ,  $e$  plays inferential role  $R'$  and  $x$  undergoes  $e$ .

The specification of the role  $R'$ , further, should somehow correspond (at least loosely) to clauses in the original theory containing “pain”. For instance, loosely corresponding to the self-intimation principle (1), we may have, e.g.

(S)      For all individual concepts  $c$ , if  $x$  believes  $C(e, c)$  and ..., then  $x$  will believe  $C(\mathbf{believe}, c, C(\mathbf{that}, C(e, c)))$ .

The full specification of “ $e$  plays inferential role  $R'$ ” will be a conjunction of (S) and other clauses, which, when inserted in (CP), yield correct entertaining conditions of **pain**.

The resulting “content specifications” are unaffected by Bealer’s arguments, and for the same reason as for (1’): psychological predicates do not occur here within the scope of any attitude verbs. (CP) can thus safely occur within O, and no conclusion to the effect that

someone who is in pain will come to believe that she has some physical-realization property can be inferred.

Of course, the proposals made here are merely examples of many possible variants of (RF) congenial with (CRS<sup>+</sup>). (CRS<sup>+</sup>) itself could be varied in many ways compatibly with the present response to Bealer's argument. For instance, the event-type theory of concepts does not play any essential role in my response above (although it is significant that it differs from the LOT hypothesis in the way explained at the close of Section 4). It would have been possible to describe in more neutral terms what kind of theory is needed to avoid his objections, but such a description would have been more abstract and more difficult to grasp.

Here is a potential worry about individuating **pain**. In order to entertain **pain**, one must arguably have the disposition, *when in pain*, to believe C(**pain**, *c*) (for some *de se* concept *c*). This seems to invite Bealer's *reductio*, since the description contains "pain", which we tried hard to not use in concept individuations. But "pain" does not occur within the scope of any propositional attitude verb, which means that the unacceptable kind of consequence cannot be derived. We can indeed infer that for some physical property *P* (namely, the realizer in *x* of pain), when *x* has *P* and conditions *C* obtain, then *x* will believe C(**pain**, *c*) (for some *de se* concept *c* referring to *x*). But this is entirely unproblematic.

## 7. Excluding clauses allowing the *reductio* from the original theory

Obviously, **pain** refers to pain. But if this claim is added to the original theory, it will yield incorrect definitions, by assigning the wrong contents to the beliefs prompted by pains. Similarly for the claim that **pain** has a certain intension, etc. However, the conclusion to draw from this is simply that such claims are *not* to occur in the original theory. Not everything true must be part of the original theory. The "original theory" is just defined as "that theory from which reductive definitions of psychological predicates can be extracted by a certain method".

In §3, we saw that Bealer’s argument from (P) and (P') to rejecting certain kinds of reductive definitions was fallacious. A similar point applies here.

Of course, if it turns out *impossible* to individuate **pain** without mentioning that it refers to pain (etc.), then, and assuming in addition that concepts must be individuated, the current version of (RF) fails, for it would then not adequately describe what beliefs pains prompt. But then, also, (CRS<sup>+</sup>) would be false, since its central tenet is precisely that concepts can be individuated entirely by their psychological roles.

Of course, these psychological roles can be *wide*, that is, they may involve causal relations with external objects and properties, including the concepts’ referents. We have already seen that it is unproblematic to say that, in order to entertain **pain**, one must be disposed to believe C(**pain**, *c*), for some *de se* concept *c*, when in pain. But we run into problems only if O needs to *state that* the concept **pain** *refers to* pain. For every occurrence of “pain” must be ramsified, and thus we get the result that being in pain causes one to believe a proposition involving a concept referring to the physical realization of pain.

Thus, we are committed to saying that every concept can be fully individuated without claiming that it refers to so and so, has such and such an intension, etc. However, this does not mean that we are committed to denying that the individuation of a concept must play some part in explaining *why* it refers to a certain entity (or has a certain intension, etc.), perhaps along the lines of Peacocke’s “determination theory” (1992: 1.3, 5.3). But it does not follow that concept individuations must include these claims about what various concepts refer to.

One may worry that whereas (1) ensures that the beliefs prompted by pains are about pain, (1'') does not. (1'') may seem to leave open the possibility that someone’s belief, which is prompted by their pains is not about pain.<sup>4</sup> In response, one can reply simply that **pain** cannot fail to refer to pain. Although this seems true enough, it will evoke the follow-up question, how can we be sure that **pain**, *as characterized by O*, is guaranteed to refer to pain?

The answer to this, it seems to me, is given immediately by the account of pain and **pain**, together with the standard, Merrick-Lewis view about how reference relations are determined, namely, by way of “best fit” between one’s beliefs and worldly entities.

The notion of best fit is usually spelt out in terms of accepted sentences and worldly entities. Very roughly, if I accept sentences  $S_1(e)$ , ...,  $S_n(e)$ , then  $e$  refers to whatever best satisfies the open formulas,  $S_1(v)$ , ...,  $S_n(v)$  (where “ $e$ ” and “ $v$ ” can belong to other syntactic categories than names, e.g. predicates). I assume that something analogous can be said of concepts and propositions. Best fit is typically taken to be determined by some weighing of naturalness (“eligibility”) of the entity in question, and satisfaction, where important considerations about the latter include how many beliefs are satisfied and how “entrenched” or “central” the beliefs are.

(CP) states conditions of entertaining **pain** to the effect that one has certain belief dispositions, and these dispositions will match the clauses about pain in O in the way suggested above (see in particular the example (S) above). This match guarantees that anyone who can entertain **pain** will refer thereby to pain. For the clauses for pain in O, if correct, will say what pain itself is like, and the individuation of **pain** will say what one must be disposed to believe in order to be able to entertain **pain**. What best satisfies the latter will thus inevitably be pain. There is no reason to think that the principles of reference determination should be part of O, however, wherefore Bealer’s *reductio* cannot get off the ground.

Given the above claims about how the reference of mental concepts is determined, we can even *infer* (1) from (1’), thus ensuring that the kind of reference failure we were concerned about is impossible. The further, needed premise is the schema,

(RP) If  $x$  believes  $C(y, z)$ , where  $y$  is a monadic, predicative concept referring to the property of  $\Phi$ :ing, and  $z$  is an individual concept referring to  $a$ , then  $x$  believes that  $a \Phi$ :s.

Given the appropriate instance of (RP), we can infer (1) from (1''). Given that the instances of (RP) are necessarily true, this shows that it is not possible for someone to satisfy the antecedent of (1) without also believing that he is in pain. As with the other cases, if we include (RP) in O, then Bealer's *reductio* goes through, so these referential claims must again be excluded from O.

Bealer stresses repeatedly that we might, instead of (1) above, have a self-intimation principle involving the relation of self-attribution, e.g. the claim that if someone is in pain and  $C$  obtain, then he will self-attribute the property of being in pain (see (2000: 504), (2001: 288), and (2010: 146f.)). Having such a clause in the original theory would immediately allow the *reductio*. We should not deny that pains prompt the relevant self-ascriptions, but, as before, we can hold without strain (pending further argument) that such claims need not be part of the original theory. But can we really accept the self-ascription claim at all, given our commitment to saying that propositions involve concepts rather than properties? Yes, for we can say that to self-attribute a property  $P$  is to believe, for some predicative concept  $c$  that refers to  $P$  and some *de se* concept  $c'$ , the proposition  $C(c, c')$ .

## **8. Principles (P) and (P') and attitude-ascriptions**

Much of the plausibility of Bealer's line of thought relies on inferences between claims about the semantics of attitude-ascriptions, in particular, the principles (P) and (P'), and, on the other hand, the nature of attitudes and contents themselves. It has been independently argued that these are rather separate matters (Eaker (2013) and Yalcin (2014)). Here, I will argue

similarly that there are obvious and principled ways of rejecting these inferences on the basis of (CRS<sup>+</sup>).

The argument that will be pursued here does not involve rejecting (P) and (P'), but, on the contrary, shows how easily they can be accommodated on the most natural semantics of attitude-ascriptions within the framework of (SC). This semantics similarly accommodates the plausible view that Bealer stresses, that psychological contents are type-free.

It is perhaps natural to think the present account violates (P) and (P'). After all, if John believes that Mary believes something, then, according to this account, there is firstly a relation of belief involved, and secondly the concept **believe**, which is distinct from the relation. So, one might think that “believe” in an attitude-ascription like (B) below must function differently in violation of (P) and (P'). However, this simply does not follow, for (P) and (P') concern the semantics of attitude-*ascriptions* whereas (CRS<sup>+</sup>) concern the nature of attitudes and contents.

To see how (P) and (P') can be accommodated by our theory, note first that, on this theory, a sentence like,

(B) John believes that Mary believes that *p*,

expresses

(BP) C(**believe**, **John**, C(**that**, C(**believe**, **Mary**, C(**that**, *x*))))

(assuming that “*p*” expresses *x*). The sentence embedded under the first “believe” in (B), further, expresses the proposition C(**believe**, **Mary**, C(**that**, *x*)). Now, to accommodate (P) and (P'), this proposition must come out as the one denoted by the “that”-clause “that Mary

believes that  $p$ ”. This comes out as true on our account, given the following plausible assumptions:

- (a) A denoting expression denotes what is denoted by the individual concept it expresses.
- (b) For all propositions  $x$ ,  $C(\mathbf{that}, x)$  denotes  $x$ .

If “that”-clauses are not denoting expressions, then, of course, we must reject (b). But on this hypothesis, also, (P) and (P') fail, so this eventuality is dialectically irrelevant. I conclude that (P) and (P') can easily be accommodated on the basis of the most natural semantics of attitude-ascriptions given (SC) and (CRS<sup>+</sup>).

A further claim, in the same spirit as (P) and (P'), is,

(P'') “Believe” in (B) functions the same way semantically in both occurrences.

On an intuitive interpretation of (P''), it is easily accommodated by our semantics, since the latter holds that both occurrences of “believe” express the same concept **believe**. An obvious suggestion is that it is by expressing **believe** that “believe” in general contributes to which propositions are expressed by sentences in which it occurs. If we interpret “functions the same way semantically” simply as “contributes the same way to what proposition is expressed”, then (P'') follows. Bealer's views about the functioning of attitude-ascriptions, while true, thus pose no problem for (CRS<sup>+</sup>).

Let us now consider the charge that the above account leads to a type-theory of psychological contents (see Bealer (1997: § 2.2)). Again, that it does is a natural suspicion.

For on my account, if John believes that Mary believes something, then there is firstly a relation of belief involved, and secondly the concept **believe**, which is a distinct entity. Now suppose I believe that John believes that Mary believes something. Then what I believe must be something involving these two kinds of entities, the belief-relation and the distinct concept **believe**. So, one might think, by the first assumption, my believing cannot be the same as that ascribed to John, so we must have three distinct entities involved (and so on). Surely, we are off on a regress, and must posit a full hierarchy of belief-like entities, one for each embedding.

But, on the contrary, there are only two entities involved here, belief and **believe**, and belief contents merely involve the same concept **believe**, even if it is within the scope of some other occurrence of **believe**. Thus, on our semantics, psychological concepts are manifestly type-free. To think otherwise is again to mistake contents for the states of affairs they may represent, as in the fallacious reasoning just rehearsed.

This take on attitudes and concepts of attitudes will naturally raise further questions and worries, but I leave them out here, since they go far beyond Bealer's case against (RF).

## **9. Concluding remarks**

We found that although Bealer succeeds in refuting standard (RF), featuring an original theory with ordinary attitude-ascriptions, his case for thinking every modification of (RF) must fail is flawed. The most important failing is that he presupposes that any such theory must take the content of “pain” to be the property pain, and that in “the concept (of) pain”, “pain” must be taken to function as it ordinarily does. These assumptions, however, are explicitly rejected (and for obvious reasons) by adherents of (CRS), who instead take concepts to be entities individuated by their psychological role.

I then presented two pre-existing versions of (CRS), which are not vulnerable to Bealer's complaints, Horwich's Use-Theory and (CRS<sup>+</sup>), both of which are independently motivated. Further, it was shown that (CRS<sup>+</sup>) provides principled reasons for excluding certain principles that Bealer stresses from the original theory, which, if included there, would have made the reductive definitions come out false. (CRS<sup>+</sup>) was also found consistent with—indeed, as accommodating—principles (P) and (P'), which Bealer thought must contradict any version of (RF) of the kind defended here.

I take this investigation to conclusively show that Bealer fails in his attempt to show that any version of (RF) must be unworkable. Still, many philosophers will remain unconvinced for the simple reason that the theory advanced here incorporates a version of (CRS), contradicting the popular view that meaning and concepts are essentially referential in character. This is not the place to further this debate. Suffice it, in closing, to note that if the reasoning in this paper is correct, there is yet another argument for (CRS) that will weigh heavily at least for materialists: that (CRS) allows a principled solution to a serious problem afflicting the most promising materialist theory of the mind.

## Acknowledgments

I want to thank the participants at the Forschungskolloquium at the Philosophy Department at Hamburg University for helpful comments, as well as the Alexander von Humboldt Foundation for a generous fellowship allowing me to write this paper.

---

<sup>1</sup> See his (1997, 2000, 2001, 2010). For critical commentary, see Tooley (1999) and McCullagh (2000). See also Shoemaker (2001), formerly a leading reductive functionalist, whom Bealer persuaded to abandon standard reductive functionalism. Although my claims connect in numerous ways with those of McCullagh and Tooley, I have chosen not to clutter

the text with references to their papers. It will be clear that my overall response, for all these connections, is very different from theirs.

<sup>2</sup> (CRS) has of course met with many objections over the years, but this is not the place to discuss them in any detail. Perhaps the best-known objection is due to Fodor and LePore, and charges (CRS) with being committed to accepting either an untenable analytic-synthetic distinction or an equally untenable holism (e.g. Fodor and LePore (1991, 1992)). Good overviews of the debate and important discussions about (CRS) include Davis (2003: §16.6), Greenberg and Harman (2008), and Whiting (2015), and Båve (2015) offers a full bibliography over (CRS).

<sup>3</sup> See, e.g. Jubien (2001), King *et al.* (2013), Hanks (2011, 2015), Soames (2010, 2015), Moltmann and Textor (forthcoming).

<sup>4</sup> Thanks to an anonymous reviewer at *Philosophical Quarterly* for this objection.

## References

- Båve, A. (2015), ‘Conceptual Role Semantics’, in D. Pritchard (ed.), *Oxford Bibliographies Online: Philosophy*, Oxford, Oxford University Press.
- Bealer, G. (1982), *Quality and Concept*, Oxford, Oxford University Press.
- Bealer, G. (1997), ‘Self-Consciousness’, *The Philosophical Review*, 106, 69-117.
- Bealer, G. (2000), ‘Fregean Equivocation and Ramsification on Sparse Theories: Response to McCullagh’, *Mind & Language*, 15, 500-10.
- Bealer, G. (2001), ‘The Self-Consciousness Argument: Why Tooley’s Criticisms Fail’, *Philosophical Studies*, 105, 281-307.
- Bealer, G. (2010), “The Self-Consciousness Argument: Functionalism and the Corruption of Content”, in G. Bealer and R. C. Koons (eds.), *The Waning of Materialism*, 137-58. Oxford, Oxford University Press.

- Block, N. (1986), ‘Advertisement for a Semantics for Psychology’, *Midwest Studies in Philosophy*, 10, 615-78.
- Davis, W. (2003), *Meaning, Expression, and Thought*, Cambridge University Press, Cambridge.
- Eaker, E. (2013), ‘Keeping Attitude Metaphysics out of Attitude Ascription Semantics (and Vice Versa)’, in Nottelmann, N. (ed.), *New Essays on Belief: Constitution, Content, and Structure*, 166-87. Basingstoke, Palgrave Macmillan.
- Fodor, J. and E. LePore (1991), ‘Why Meaning (Probably) Isn’t Conceptual Role’, *Mind and Language*, 6, 328-43.
- Fodor, J. and E. LePore (1992), *Holism: A Shopper’s Guide*, Oxford, Blackwell.
- Greenberg, Mark and Gilbert Harman (2008), ‘Conceptual Role Semantics’, in E. Lepore and B. Smith (eds.), *The Oxford Handbook of Philosophy of Language*, 296-322. Oxford, Oxford University Press.
- Hanks, P. (2011), ‘Structured Propositions as Types’, *Mind*, 120, 11-52.
- Hanks, P. (2015), *Propositional Content*, Oxford, Oxford University Press.
- Harman, G. (1973), *Thought*, Princeton, Princeton University Press.
- Harman, G. (1987), ‘(Nonsolipsistic) Conceptual Role Semantics’, in E. LePore (ed.), *New Directions in Semantics*, 55-81. London, Academic Press.
- Horwich, P. (2005), *Reflections on Meaning*, Oxford, Oxford University Press.
- Jubien, M. (2001), ‘Propositions and the Objects of Thought’, *Philosophical Studies*, 104, 47-62.
- King, J., S. Soames, and J. Speaks (eds.) (2013), *New Thinking About Propositions*, Oxford, Oxford University Press.
- Lewis, D. (1983), ‘An Argument for the Identity Theory’, in *Philosophical Papers, Volume I*, New York, Oxford University Press, 99-107.

- McCullagh, M. (2000), 'Functionalism and Self-Consciousness', *Mind & Language*, 15, 481-99.
- Moltmann, F. and M. Textor (eds.) (forthcoming), *Act-Based Conceptions of Propositional Content: Contemporary and Historical Perspectives*, Oxford, Oxford University Press.
- Peacocke, C. (1992), *A Study of Concepts*, Cambridge, MA., MIT Press.
- Schnieder, B. (2006), 'By Leibniz's Law: Remarks on a Fallacy', *The Philosophical Quarterly*, 56, 39-54.
- Sellars, W. (1968), *Science and Metaphysics*, London, Routledge & Kegan Paul.
- Shoemaker, S. (1994), 'The Mind-Body Problem', in R. Warner and T. Szubka (eds.), *The Mind-Body Problem: A Guide to the Current Debate*, 55-60. Oxford, Basil Blackwell.
- Shoemaker, S. (2001), 'Realization and Mental Causation', in C. Gillet and B. Loewer (eds.), *Physicalism and Its Discontents*, 74-98. Cambridge, Cambridge University Press.
- Soames, S. (2010), *What is Meaning?*, Princeton University Press, Princeton, NJ.
- Soames, S. (2015), *Rethinking Language, Mind, and Meaning*, Princeton, Princeton university Press.
- Tooley, M. (1999), 'Functional Concepts, Referentially Opaque Contexts, Causal Relations, and the Definition of Theoretical Terms', *Philosophical Studies*, 105, 251-79.
- Whiting, D. (2015), 'Conceptual Role Semantics', in J. Fieser and B. Dowden (eds.), *The Internet Encyclopedia of Philosophy* (<http://www.iep.utm.edu>), ISSN 2161-0002.
- Yalcin, S. (2014), 'Semantics and Metasemantics in the Context of Generative Grammar', in A. Burgess and B. Sherman (eds.), *Metasemantics*, 17-46. Oxford, Oxford University Press.