Direct Reference, Psychological Explanation, and Frege Cases

SUSAN SCHNEIDER

Abstract: In this essay I defend a theory of psychological explanation that is based on the joint commitment to direct reference and computationalism. I offer a new solution to the problem of Frege Cases. Frege Cases involve agents who are unaware that certain expressions co-refer (e.g. that 'Cicero' and 'Tully' co-refer), where such knowledge is relevant to the success of their behavior, leading to cases in which the agents fail to behave as the intentional laws predict. It is generally agreed that Frege Cases are a major problem, if not *the* major problem, that this sort of theory faces. In this essay, I hope to show that the theory can surmount the Frege Cases.

There has recently been much philosophical interest in the view that psychological explanation should be wide, that is, the view that psychological kinds fail to supervene on the intrinsic states of the individual. Theories of wide content claim that content is individuated externally; theories of *broad* content, (as I am using the expression), in addition to this, take the basic semantic properties of thoughts to be denotation and truth.[1] In the domain of psychological explanation, those sympathetic to Russellianism, a currently popular account of attitude ascription, have claimed that intentional laws are sensitive to broad contents. This position is generally viewed as controversial because it rejects the standard view that intentional laws are supposed to be sensitive an agent's mode of presentation (MOP), or way of conceiving of things.

More specifically, according to Russellianism, the proposition expressed by the sentence, 'Cicero is Tully', is an entity that consists in the relation of identity, the man, Cicero, and the man, Tully. (Further, the sentence, 'Tully is Tully' also

I would like to thank Jerry Fodor for his comments on earlier drafts of this paper and for many interesting conversations. I would also like to thank the audiences at the Pacific APA in Seattle, Washington, '*Cognitive Science in the New Millenium*' in Long Beach, California, and '*Intentionality*' in Miscolc, Hungary (all spring, 2002), especially Ned Block, David Chalmers, Jeffrey King, Ruth Millikan, Jesse Prinz and Philip Robbins. And I am grateful to David Braun, Jon Cohen, Frances Egan, Mike Huemer, Gualtierro Puccininni, Ted Sider, and two anonymous reviewers at this journal for their helpful written comments and to Peter Hanks for presenting helpful comments on this paper at the Pacific APA.

Address for correspondence: Department of Philosophy, Moravian College, 1200 Main Street, Bethlehem, 18018 PA, USA
Email: susan@moravian.edu

[1] Here I'm following Fodor's usage (Fodor, 1994, p. 7). There is another usage in which 'broad content' is taken as synonymous with 'wide content'.

Mind & Language, Vol. 20 No. 4 September 2005, pp. 423–447.

expresses the same proposition). Russellians hold that 'believes' names a relation to a Russellian proposition. So they hold the surprising view that anyone who believes that Tully is Tully also believes that Cicero is Tully.[2] As a consequence of this, a Russellian about psychological explanation adopts the following claim about psychological explanation: (PE), beliefs differing only in containing different coreferring names, (e.g. 'Cicero'/'Tully'), are to be treated by an intentional psychology as being type-identical and are thereby subsumable under all the same intentional laws.[3]

Frege's puzzle about belief ascription is a well-known problem arising for the purely semantic version of Russellianism. Along similar lines, due to (PE), the Russellian about psychological explanation faces a related worry, the problem of 'Frege Cases'. To consider a well-known example of a Frege Case, consider Sophocles' Oedipus, who didn't realize that a woman he wanted to marry, 'Jocasta', happened to be his mother. So Oedipus has two distinct ways of representing the same person, and doesn't realize that they co-refer. This situation creates problems if psychological laws are based on Russellianism, for the laws are indifferent to these distinct ways. Indeed, Oedipus threatens to be a counter-example to the broad generalization:

> (M), *Ceteris paribus*, if people believe that they shouldn't marry Mother and they desire not to marry Mother, they will try to avoid marrying Mother.

Notice that Oedipus satisfies the antecedent of (M). However, Oedipus also fails to satisfy the consequent since, in virtue of his trying to marry Jocasta, it is true, according to a broad psychology, that he tries to marry Mother.[4]

In general, Frege Cases involve agents who are unaware that certain expressions co-refer, where such knowledge is relevant to the success of their behavior, leading to cases in which the agents fail to behave as the broad intentional laws predict. It is

---

[2] Russellianism has been defended by (inter alia), David Braun, Keith Donnellan, David Kaplan, Ruth Marcus, John Perry, Mark Richards, Bertrand Russell, Nathan Salmon, Scott Soames and Michael Thau. In addition to providing an account of proper names, Russellians typically extend their account to other expression types, for instance, the meaning of a simple predicate is a property or relation, rather than e.g. a sense.

[3] Perhaps the best-known proponent of (PE) is Jerry Fodor (Fodor, 1994). Fodor's discussion of these issues sometimes seems sympathetic to the hidden-indexical theory rather than Russellianism. I suspect that Fodor regards the hidden indexical theory as being compatible with (PE) because, for the purposes of doing scientific psychology, thoughts are typed by the proposition expressed by the that clauses, and, according to both hidden-indexical views and Russellianism, these are Russellian propositions.

[4] The Oedipus example is rather odd as 'mother' is not clearly a name. It is crucial to bear in mind that other, clearer, examples of Frege Cases can be provided (e.g. Cicero/Tully, woodchuck/groundhog). Unfortunately, the example is widely used in the literature on Frege Cases and certain objections to my view are not easily stated without it. To underscore that I have in mind a referential reading of 'mother' I will write 'Mother' (capitalized) throughout.

generally agreed that Frege Cases are a major problem, if not *the* major problem, that Russellianism about psychological explanation faces. And many find that this view is not faring well in responding to this problem; indeed, the literature on Frege Cases has generally been quite negative in its assessment of the capacity of Russellianism to surmount the Frege Cases (Aryo, 1996; Aydede, 1997 and 1998; Aydede and Robbins, 2001).[5] If these critics are correct then any theory of psychological explanation that is based on Russellianism is on poor footing indeed.

However, I believe that these criticisms are flawed; in this essay, I argue that Frege Cases are not genuine counterexamples to intentional laws because they can be included in the *ceteris paribus* clauses of the relevant laws. It is well-known that special science laws have *ceteris paribus* clauses. The presence of *ceteris paribus* clauses means that a case in which the antecedent of a special science law is satisfied, while the consequent is not, need not be a counterexample. Instead, the case may simply be a 'tolerable exception'—a situation in which the *ceteris paribus* conditions fail to hold.[6] Consider, for example, the generalization in economic theory that says that given a decrease in supply of an item and given no change in demand, the price of the item will go up. This generalization is taken to be *ceteris paribus*, having certain exceptions (e.g. it will not hold if there is price fixing, if the only supermarket that sales the item happens to close, etc.). Such situations are not taken to be counter-examples to the generalization, but are regarded as exceptions that are to be tolerated. For a case to be a counterexample, the antecedent must obtain, *and* the *ceteris paribus* condition must be met.

It is widely agreed that to include a given case in the ceteris paribus clause a justification must be provided for doing so. Critics have urged that no plausible justification has been provided when the given cases are Frege Cases. In light of this, the task of this paper is to provide the needed justification. I argue that including the Frege Cases in the *ceteris paribus* clauses is justified by a larger theoretical decision for intentional laws having a 'broad canonical form'. Intentional laws have a 'broad canonical form' when they are sensitive to mental events that are typed by their broad contents. Any decision to include the Frege Cases in the *ceteris paribus* clauses is the result of an overall assessment of the debate about which canonical form intentional laws should take, broad or narrow. While such a decision is a global affair, I will focus on the part of this theoretical decision that is internal to Frege Cases—the part that involves the issue of whether non-intentional explanation of the psychological difference between coreferring thoughts will suffice.

The literature on Frege Cases has been quite negative on this score. Three main lines of criticism have emerged. First, critics have argued that there are no grounds

[5] Criticisms of Russellianism and Frege cases have also appeared in the philosophy of language literature on belief ascription (e.g. Crimmons, 1992, pp. 32–34; Richard, 1990, p. 219).

[6] Herein, I'll be taking it for granted that there are special science laws and that further, they are *ceteris paribus* in nature, unlike the strict laws of microphysics. For debates on the general plausibility of *ceteris paribus* laws see (Schiffer, 1991; Fodor, 1991; Rey, 1995).

for treating Frege Cases as tolerable exceptions (Aryo, 1996; Aydede, 1997, 1998; Aydede and Robbins, 2001). Second, they have argued that even if the Frege Cases can be treated as tolerable exceptions, doing so leads to missed intentional-level predictions of certain thoughts and behaviors (Aydede, 1997, 1998; Aydede and Robbins, 2001). And third, they have argued that even assuming that the Frege Cases are tolerable exceptions and that there are no missed predictions, a broad theory will have to find a way to predict and explain Oedipus' behavior that is not intentional. And explanation of thought and behavior *must* be intentional (Aryo, 1996, Aydede, 1997, 1998; Aydede and Robbins, 2001). After exploring these criticisms in more detail, I will argue, contra these critics, that Frege Cases can be treated as tolerable exceptions. Further, there is no missed prediction of Frege Cases: a broad psychology does not fail to explain events in its laws that narrow psychology, on the other hand, captures (under narrow description). Finally, I will argue that there is no justification for believing that explanation of Frege Cases must be intentional in nature, as opposed to computational.

Here is how my discussion will proceed: first, since Frege Cases arise for a very particular sort of theoretical apparatus, namely, a direct reference theory applied to the domain of psychological explanation, it will be helpful to explain this approach to psychological explanation. In the following section I outline such a framework, which I call 'Broad Psychology'. Then, I shall unpack and defend my solution to the Frege Cases.

1. Framework

According to a psychological theory informed by the semantics of direct reference, intentional generalizations are broadly referential in nature, that is, they are sensitive to the broadly referential properties of a mental state, while being indifferent to the state's conceptual role (that is, the particular role that the thought plays in one's cognitive economy).[7] As a result of this view, thoughts treated as intentionally type-identical may nonetheless function very differently in one's cognitive economy, causing very different thoughts and behaviors. So there is a tension between the causal functioning of these states, on the one hand, and the representational nature of the states on the other. Intentionally or referentially, the states are type identical; but causally, they are very different in ways that are clearly of interest to any psychology. This tension between *the representational* and the *causal* is the fundamental problem regarding psychological explanation for a Russellian theory.

While the direct reference theorist cannot accommodate any causal differences between intentionally type identical thoughts by appealing to semantic differences,

[7] I say 'broadly referential' for a technical reason. The view is referential in the case of proper names, indexicals and demonstratives, however, in the case of predicates, the content of a predicate is a property, rather than an extension at a world. Herein, I'll gloss over this difference, as many do, in calling the view 'referential', and in saying laws subsume states by their referential properties.

there are other theoretical wares that he can employ to do so. In the context of providing a solution to Frege's puzzle about belief ascription the Russellian often appeals to a 'guise' or 'mode of presentation' of a thought, a mental particular (or property of mental particulars) that thinkers are said to have when they entertain a Russellian proposition.[8] While theories of belief ascription have not provided any detailed elaboration concerning the nature of guises, it is fair to say that such entities are intended to capture the conceptual role of the thought. One popular way of cashing-out 'guises' or 'modes of presentation' is to say that they are expressions in one's language of thought (LOT). A similar appeal to LOT could be made in the context of the Frege Cases.[9] Proponents of LOT claim that LOT states are tailored to explain the production of narrowly described thoughts and behaviors in intelligent systems. Because LOT distinguishes between coreferring thoughts in a non-semantic manner, (unlike Fregean senses), the direct reference theorist can appeal to them to distinguish coreferring concepts and to help solve the Frege Cases.

This means that Broad Psychology appeals to a theoretical apparatus with two different levels of explanation, the computational and the intentional/semantic.[10] More specifically, Broad Psychology is a two-tiered psychological theory according to which there is both:

(i) An intentional level of explanation that subsumes mental states by their broad contents and is indifferent to the conceptual role of the mental state.
(ii) A relatively lower level of computational laws that is indifferent to the broad content of the mental state but is sensitive to the computational role; that is, the role the mental state plays in computation.[11]

---

[8] A classic Russellian appeal to guises is found in Salmon, 1986.

[9] Indeed, Jerry Fodor has advocated a general project that weds the semantics of direct reference to a classical theory of computation according to which mental processes involve the manipulation of mental symbols according to rules.

[10] Of course the expression 'Broad' in the name 'Broad Psychology', refers to the intentional element of the theory. A more precise name for the theory would perhaps be something like 'Broad Intentional/Narrow Syntactic Psychology', but alas, this name is too long.

[11] More specifically, computational laws are sensitive to the computational role of a mental state in the following sense: the laws are sensitive to symbols that are individuated by their computational role. Computational role could be spelled out in various ways. Elsewhere, I have argued that a primitive LOT expression should be individuated by its *total* computational role; that is, a primitive symbol is individuated by all computational-level causal relations between the symbol and proximal inputs, other primitive expressions, and narrow behaviors. And more specifically, a LOT symbol can be defined by the role that it plays in a Ramsey sentence that includes a list of all the computational laws in which LOT expression figures. (For a defense see Schneider, 2003). A different, 'molecularist' way of doing things would be to isolate certain elements of a symbol's computational role, regarding only these as being type individuative. This view is akin to the molecularist theory of narrow content, except the issue is not meaning or content individuation, but the individuation of words in the LOT vocabulary. In Schneider, 2003 I reject this sort of view because I argue that classicism *requires* that symbols be individuated by total computational role.

Due to this two-tiered nature, there is a sense in which the traditional wide/narrow dichotomy fails to characterize Broad Psychology: such a psychology is wide, drawing from extrinsic states of the individual. But it is also narrow, since the computational level only appeals to a system's intrinsic states. As a result it aspires to accommodate both intuitions behind the well-known Twin Earth thought experiment. According to Hilary Putnam's thought experiment, we are asked to consider two molecular duplicates who are on planets that only differ with respect to the fact that one has H_20 in it, and the other has a substance that looks, smells, and tastes like water but, in fact, has a different chemical composition. We are then asked whether the meaning of the duplicates' respective 'water' utterances differs. Many agreed with Putnam that the meanings seem to differ because the kinds differ. And many held that the twins' psychological states differed, with these differences in meaning. According to them, mental content is wide, being determined, at least in part, by features in the environment that are external to the individual.[12]

Although many found the externalist intuitions about mental kind individuation attractive, the externalist position seems to face a serious difficulty. For it seemed to come into conflict with certain explanatory demands on content. Namely, reasoning involves a succession of representations that succeed one another in virtue of their internal, causal connections. And theories of mental processing in cognitive science seem to abstract away from relations to the environment, focusing on internal computations.[13] Such considerations have led many philosophers to observe that there is an important sense in which the twins' concepts and attitudes are the same. Namely, their internal, narrow, psychological states are the same. However, at least at first blush, this observation runs contrary to the popular externalist reaction to the Twin Earth thought experiment outlined above, which held that mental content is externally determined. A major issue in philosophy of mind is how to find a theory of psychological kind individuation that is able to accommodate these two, seemingly opposing, intuitions, or at least offer an effective refutation of one of them.

Broad psychology aims to accommodate the gist of both of these intuitions. More specifically:

- Broad Psychology aims to accommodate the intuition that the twins' thoughts play the same role in their respective cognitive economies, by subsuming the twins in the same predictions at the computational level.
- Because content is taken to be broad, Broad Psychology aims to accommodate the externalist intuition that their thoughts have different contents.

[12] The Twin Earth thought experiment first appeared in Putnam's 'The Meaning of "Meaning"'. For more discussion see this paper and the other papers in (Pessin and Goldberg, 1996).

[13] For further discussion of this issue see Egan, 1995.

Indeed, this framework can provide both a wide and a narrow taxonomy of thoughts. Suppose that I utter, 'I need a bucket of strong coffee just about now'. When I do so, I have a thought to this effect; that is, there is a particular, dated event in my brain that is characterizable by my public language utterance. There are two explanatory levels at which the Broad Psychologist may subsume this quite accurate thought. First, there is the intentional or representational level of explanation, according to which the thought is merely being singled out at a very coarse level of grain, so that any thoughts having the same grammatical form that refer to the same state of affairs are type-identical. Let us call this mental kind 'the broad thought'. In addition, there is another explanatory level at which the theory subsumes the very same mental event. This level is insensitive to content, but instead, subsumes the thought with respect to its computational expression type. The mental kind demarcated in this way is the narrow thought. On this view, narrow thoughts don't have their contents essentially. (That is, narrow thoughts are not taxonomized with respect to sameness and difference of content.) We could also identify a broad and narrow notion of a concept, where concepts are taken as mental particulars (or alternately, properties of mental particulars) that are constituents of thoughts.[14]

This being said, I should underscore the openness of this two-tiered framework. Although I've appealed to LOT, it is beyond the scope of this paper to delve into the arguments for LOT, and in any case, the basic framework can be adopted by the connectionist, as computational states individuated by their role in a connectionist network can differentiate #Cicero# from #Tully# thoughts as well.[15] Further, just as the computational level is open to both LOT and connectionism, the intentional level can be fleshed out by whatever broad theory of reference that, at the end of the philosophical day, seems most plausible (e.g. informational semantics, the causal theory of reference, and so on). At this point in the game, any openness is likely a virtue. My present task will be accomplished if the reader is persuaded that the basic framework is, overall, an approach worthy of further consideration.

---

[14] On my view, a concept is constituted by its computational role and its broad content. Concepts are thus naturally separable into two dimensions, a conceptual role dimension and a referential one. A conceptual role dimension is a useful addition to a referential factor because it captures the plausible intuition that the nature of a concept is at least partly a matter of the internal role that the concept plays in a system's cognitive economy. Readers familiar with the recent debate on the nature of concepts in this journal may note that this view is, in a sense, a 'pragmatist' theory, as it appeals to functional (and in particular, computational) role, while also being compatible with concept atomism, the view that lexical concepts are semantically unstructured. (For extensive discussion of atomism and pragmatism see the February 2004 issue of this journal.)

[15] Connectionism is commonly regarded as being a theory that invokes mental representations—indeed, even Paul Churchland talks of mappings between referents and (certain) computational states that are individuated in terms of their role in a connectionist network (Churchland, 1995). But for a dissenting opinion see Ramsey, Stich and Garon, 1990.

Now let me turn to a defense of the two-tiered structure. First, one might question the motivation for a broad level, over and above a narrow one. Indeed, even the semantic externalist may do so. Suppose that semantic externalism is true and the semantic contents differ between the twins. Still, one could hold that a semantic theory should stay out of a theory of psychological explanation. For example, one could dispense with the semantic level altogether and just appeal to LOT syntax, or one could appeal to something along the lines of Brian Loar's notion of psychological content to do the needed work.[16] Prima facie, psychological content and LOT syntax are the sort of entities that seem to have a suitable level of grain for the purpose of explaining the causation of (narrowly described) thought and proximal behavior. Since they are candidates for filling this role, and broad contents are not, why not dispense with a broad level altogether? Because this question can be asked by a semantic externalist, the Broad Psychologist must provide an independent motivation for a broad level; that is, she must give a motivation over and above the intuition that the twins' semantic contents differ. For even if semantic externalism is correct, there is still the further question: why should semantic properties serve to individuate *psychological* kinds?

The needed motivation for a broad level arises from two sources. First, it is well known that people can represent the same state-of-affairs in different ways from each other. Such ways of representing things are notoriously idiosyncratic, and many are sceptical that philosophy will provide a plausible theory of narrow content that succeeds in abstracting away from idiosyncratic differences, arriving at a sense in which different people share the same concept. An appeal to a purely broad intentional level provides a dimension in which concepts are shared. Second, broad laws capture predictive uniformities between mental events that are distinct ways of representing the same referent. Given that people live in the same environment and have similar mental structures, people's behavior towards the referent tends to converge despite idiosyncratic differences in their ways of representing the world. On the other hand, if the intentional laws are sensitive to narrow content or LOT syntax, then any predictive uniformity in their referent-directed behaviors is lost.[17] As Ned Block explains:

> ... wide meaning may be more useful [than conceptual role] in one respect: to the extent that there are nomological relations between the world and what people think and do, wide meaning will allow predicting what they think and without information about how they see things. Suppose, for example, that people tend to avoid wide open spaces, no matter how they describe these spaces to themselves. Then knowing that Fred is choosing whether to go via an open space or a city street, one would be in a position to predict Fred's

[16] See Loar, 1996.

[17] The motivations stated in this paragraph have been noted by Block, 1994; Fodor, 1994, p. 51; Pylyshyn, 1986.

> choice, even though one does not know whether Fred describes the open space to himself as 'that' or 'Copley Square' (Block, 1994).

Now, I don't know that people really aim to avoid wide open spaces; nonetheless, Block's central point strikes me as apt. Other generalizations making the same point are available: e.g. one may think of gold as 'the stuff with the atomic number seventy-nine'; another may think of it as, 'the sort of jewelry Jane likes to wear'. Nonetheless, both parties satisfy many of the same gold-related generalizations. In general, it seems uncontroversial that systems having different ways of representing the same entity will frequently behave in similar ways because they are embedded in similar environments and because they make similar demands on these environments. But it remains unclear how this tendency toward similar thoughts and behaviors can be captured by the generalizations of a narrow psychology.[18]

The usual criticism of this point is that coreferentiality does not ensure that the thoughts will *always* be behaviorally equivalent; e.g. one can represent the man, Cicero, under the mode of presentation, [Tully], and be unaware that he is also Cicero. As noted, agents unaware of coreferentialities relevant to their behaviors threaten to be counterexamples to putative broad generalizations because they satisfy the antecedents of the generalizations, but fail to behave as those knowing the relevant identity would behave. This brings us full circle to the problem that I'm hoping to solve. So let me now turn to the Frege Cases.

2. The Frege Cases—a Solution

As a point of departure, note that *any* broad theorist who would like to solve the Frege Cases *must* include the Frege Cases in the *ceteris paribus* clauses of broad intentional laws. To see this first notice that if content is narrow, then there is no need to include the Frege Cases in the *ceteris paribus* clauses because the canonical form of the intentional laws is such that thoughts are subsumed by properties that are sensitive to differences between coreferring concepts. On the other hand, if the laws are broad then the canonical form of intentional laws glosses over differences

[18] An anonymous reviewer offered the following objection to my case for broad laws: scientific psychology is not really in the business of explaining behavior in terms of laws of any kind, much less laws about beliefs and desires. However, scientific psychology *is* concerned with laws. Although much of scientific psychology is concerned with explaining cognitive capacities, there is also much interest in discovering and confirming laws or effects. (For more discussion of these issues see Cummins, 2000, pp. 114–144.) Further, many of these laws or effects concern beliefs, desires and other attitudes. Such generalizations are employed throughout the field of social cognition, for example. And they are employed in computational psychology because they specify what the computational explanations are supposed to explain. Consider, e.g., computational explanations of the moon illusion, which seek to explain the generalization, 'people believe the moon looks larger when it's on the horizon'.

between coreferring concepts. In this case, the laws face putative counterexamples, so the Broad Psychologist *must* say that the Frege Cases are tolerable exceptions and should be included in the *ceteris paribus* clauses (otherwise the Frege Cases would be counterexamples and the theory would be false). So including Frege Cases in the *ceteris paribus* clauses is a consequence of a larger theoretical decision for a broad canonical form.[19]

Given that including Frege Cases in the *ceteris paribus* clauses is an inevitable result of the choice of a broad canonical form, to determine whether the cases are to be included in the clauses, we should ask: what determines which canonical form is the correct one? As far as I can tell, one's choice of a canonical form depends upon the following issues: the plausibility of the competing theories of narrow content, the plausibility of two-factor theories, and finally, whether Broad Psychology can give a satisfactory account of the computational difference between co-referring concepts. Of course, the litmus test for whether Broad Psychology can accomplish this latter task is whether there is a plausible computational account of the Frege Cases.

So it seems that justifying including Frege Cases in the *ceteris paribus* clauses is a global affair: if the best psychology should have a broad form, then we have motivation to include Frege Cases in the *ceteris paribus* clauses. While I cannot delve into all of the issues that determine the larger theoretical decision of which canonical form to choose, what I would like to do is focus on the part of the decision that involves the issue of whether computational explanation of Frege Cases will suffice.[20] In doing so I hope to persuade the reader of the following conditional thesis: assuming that the other issues listed above go in the direction of Broad Psychology (call this the 'almost all-things-considered judgment'), then, because computational explanation of Frege Cases adequately explains Frege Cases, the phenomenon of Frege Cases is unproblematic for the theory and it is justifiable to include them in the *ceteris paribus* clauses.

If I am correct, then any failed justification for including the Frege Cases in the *ceteris paribus* clauses would not emerge from problems *internal* to Frege Cases but from *external* issues involving one's choice of a canonical or proper form for intentional explanations. So I hope to put the Frege Cases on the theoretical back burner. Although limited, this is a significant result: to the best of my knowledge no one has attempted to respond to the variety of recent criticisms that computational explanation of Frege Cases will not suffice. Since Frege Cases are

[19] It is less well-known that Frege Cases arise for two-factor views as well. If one endorses a level at which thoughts are subsumed by their broad contents only, as many two-factor theorists do, there will be Frege Cases and the generalizations will face putative counterexamples. As I argue below, if Frege Cases arise, the theory must say that the Frege Cases are included in the *ceteris paribus* clauses. The solution I provide would be available to the two-factor theorist, although the details would differ because the component that distinguishes the co-referring names is a narrow content.

[20] However, I do offer a critical discussion of a few leading theories of narrow content at infra pp. 440–442.

considered to be a major problem for the theory, if not the major problem, proving that they don't speak against the theory eliminates a significant problem indeed.

This being said, let us turn to the details of my defense of computational explanation of Frege Cases. First off, let us note that in order for the Russellian to respond to the Frege Cases, she must defend the following claims:

(1) Frege Cases are tolerable exceptions, rather than counterexamples, to broad intentional *ceteris paribus* generalizations;
(2) Including Frege Cases in the *ceteris paribus* clauses does not result in Broad Psychology failing to predict Frege Cases. That is, Frege Cases can be predicted somewhere within the total theory (but not necessarily at the intentional level).
(3) There is no compelling philosophical argument that differences between co-referring concepts must have intentional, (as opposed to computational), explanation.

Intuitively, (1)–(3) are key components of any defense of the Frege Cases. (1) is crucial because it is generally agreed by those who believe in *ceteris paribus* laws that for a putative counterexample to be included in the *ceteris paribus* clause it must be shown to be a tolerable exception. But (1) is not sufficient to justify including Frege Cases in the clauses, even assuming that the almost all-things considered judgment is for Broad Psychology, for the following reason. If including Frege Cases in the *ceteris paribus* clauses leads Broad Psychology to fail to predict Oedipus' behavior, since greater scope is an advantage to a theory, then *ceteris paribus*, we have reason to take differences between co-referring names as being intentional in nature, rather than computational. In this case the canonical form will not be one that includes the Frege Cases in the *ceteris paribus* clauses. So (1) must be supplemented with (2).

But (2) only addresses the issue of predictive adequacy. Even if the theory is adequate in this respect, it may nonetheless fail to give satisfactory explanation of Frege Cases. For critics charge that Oedipus' behavior can only be rationalized or explained by giving intentional explanation. For many philosophers, the locus of doubt about the prospects of giving computational explanation of Frege Cases centers around (3).

I shall now proceed to argue for *(1)*. I will briefly outline why Oedipus is a tolerable exception. To avoid confusion, let me begin by underscoring that Frege Cases involve rational agents who generally are aware of co-referentialities relevant to the success of their actions, but, in certain cases, fail to be aware of a co-referentiality. Frege Cases are 'abnormal' only in the watered-down sense that they are a breakdown in the normal course of events—normally, they do not occur because agents tend to be aware that two expressions co-refer when it is important to the success of their behavior.

Further, in virtue of their failure to be aware that certain expressions co-refer, agents having Frege Cases fail to grasp the relevant propositions in matching ways. To see what is meant by 'matching ways', consider, again, the case in which Oedipus

is a putative counterexample to (M). Oedipus satisfies the antecedent in virtue of having beliefs and desires that employ the mode of presentation, #Mother#.[21] But, as a result of his ignorance of the coreferentiality, he represents Jocasta differently when he grasps the proposition figuring in the consequent; in this case he represents her under the mode of presentation, #Jocasta#. So he represents Jocasta in ways that do not match. As David Braun has suggested in a recent paper, this failure to represent Jocasta in matching ways is grounds for regarding Oedipus as a tolerable exception, rather than a counterexample, to (M) (Braun, 2001). All other things were not equal; although in general, people seek to avoid marrying their mothers, Oedipus was atypical in the sense that he was ignorant of a coreferentiality, and, in virtue of his ignorance, he represents her in mismatching ways. Including Oedipus in the *ceteris paribus* clause doesn't seem to take away from the usefulness of (M); after all, (M) embodies a generally accepted principle about human behavior and there is a large population that bears it out.[22]

(2) Now let me turn to the issue of whether Broad Psychology can predict Frege Cases somewhere within the total theory. I will identify a number of reasons why one may suspect that Broad Psychology cannot do so.

Aydede and Robbins have argued that:

> ... a narrow psychology ... can cover the occasional unsuccessful behavior, or accidentally successful behavior, to which Frege patients are prone. So a narrow psychology would have a wider scope, hence—*ceteris paribus*—greater

[21] I will designate MOPs by enclosing the relevant expression with the symbol '#', (e.g. #dog#).

[22] An anonymous reviewer asks if treating Frege cases as tolerable exceptions doesn't result in an intentional theory that is a 'notational variant' of saying that the mental states covered by intentional laws have fine-grained propositional contents. To see why this doesn't occur consider (D): 'if people believe danger should be avoided, and believe they are in danger, then *ceteris paribus*, they will try to avoid danger'. On the standard construal of narrow intentional laws, if (D) is a narrow intentional law, then the proposition expressed by the relevant that clauses is individuated by the cognitive content of the thought, where such has been construed in various ways, most notably as a Fregean sense. When propositional contents are taxonomized in these ways, all those who satisfy a given law must conceive of the state-of-affairs refered to by the relevant that clauses in the same way, e.g. in the context of (D) all individuals must think of danger, avoidance, and so on, under the same mode of presentation. This is clearly not a notational variant of laws having a broad canonical form. Consider (D); if (D) is a broad intentional law, then mental events are subsumed in the law by their referential properties, not their cognitive contents. So one person who satisfies (D) is free to conceive of danger (avoidance, etc.) under a different MOP from another person who satisfies (D). Perhaps the confusion arises from the matching ways requirement in the *ceteris paribus* clauses of broad laws; I suppose that it may sound Fregean in spirit, as it requires that the MOPs match up. It should be underscored however, that the matching ways requirement merely requires that *within* a system that satisfies a given intentional law, the system must conceive of the relevant referents in ways that match up. It does not require a match in MOPs *across* systems the way that the Fregean or neo-Fregean does. Further, the required match up is not one of propositional content, but one of computational state-type.

> explanatory and predictive power. And that surely suggests its superiority to the broad alternative (Aydede and Robbins, p. 2001).

To restate their claim in terms of the Oedipus example, they claim that narrow psychology, but not a broad psychology, can cover Oedipus' unsuccessful attempt to not marry Mother. For as we've already noted, Oedipus fails to satisfy the following prediction:

> (M) *Ceteris paribus*, if Oedipus desires that he not marry Mother/Jocasta, and Oedipus believes that not marrying Mother/Jocasta is the only way to bring this about, then he will try not to marry Mother/Jocasta.

(I write 'Mother/Jocasta' rather than simply 'Mother' to indicate that I have in mind a referential reading of 'Mother'.) And here we seem to arrive at a problem for the broad theory: when the theory includes Oedipus' thought, *try not to marry Mother/Jocasta*, in the *ceteris paribus* clause, according to the critics, it will not cover Oedipus' unsuccessful behavior that is narrowly described as his trying not to marry Mother. The theory will fail to do so because the intentional level is insensitive to this thought. However, the critics continue, there is still a sense in which Oedipus tries not to marry Mother, even though he fails. But a broad theory fails to predict that Oedipus will try not to marry Mother. Hence, it seems that narrow psychology can predict behavior which the Russellian view cannot.

Of course I agree with Aydede and Robbins that the following principle is true: *ceteris paribus*, greater scope is an advantage to a theory. But it is simply not true that Broad Psychology fails to cover, 'the occasional unsuccessful behavior, or accidentally successful behavior, to which Frege patients are prone'. On the contrary, I will argue that *including Frege Cases in the clauses does not result in missed psychological explanation of Frege Cases*. This claim may seem surprising; after all, it is uncontroversial that Broad Psychology includes certain of Oedipus' thoughts in the *ceteris paribus* clauses. To prove my claim, let me begin by distinguishing the following points:

(i). Frege Cases are included in the *ceteris paribus* clauses of broad intentional laws.

(ii). Broad Psychology fails to explain thoughts and behaviors that narrow psychology, on the other hand, does explain (when the events are described narrowly): namely, Broad Psychology fails to explain the events that are included in the *ceteris paribus* clauses.

Of course, if (ii) is correct then chalk one up for narrow psychology. I believe that many are assuming that (i), together with some reasonable premises, entails that (ii) is true. But such an argument is not valid. I will argue that (i) is true while (ii) is false. Consider again the putative law (M); the critics urge that certain explanation of behavior is missed that narrow psychology does not miss, in not subsuming Oedipus under (M). Their concern is that at time t, before Oedipus

meets Jocasta, Oedipus instances (M). After he tries to marry Jocasta he fails to do so. So after t, explanation is unavailable, leading to a missed intentional generalization that narrow content does not miss since there is a sense in which Oedipus still tries not to marry Mother.

But this line of reasoning is flawed. Oedipus still instances (M) after (t). Intuitively, when Oedipus believes that he should try not to marry Mother he does so because he holds, like the rest of us, the moral prescription that one shouldn't marry Mother. The fact that he tries to marry Jocasta does not imply that he stops having this belief. Broad Psychology can reflect this intuitive picture in the following way. When Oedipus believes the moral prescription he has the mode of presentation, #try not to marry Mother#, and this belief is intentionally described as a *try not to marry Mother/Jocasta* thought. So at least until he meets Jocasta he satisfies the broad version of (M) like the rest of us do.

But what about after t? Certainly the broad theory does not *require* that after t Oedipus drop his belief in the prescription simply because he now has a *try to marry Jocasta/Mother* thought as well. Oedipus satisfies (M) in virtue of believing the moral prescription, while in another situation, he is a putative counterexample to (M) when he thinks *try to marry Jocasta/Mother* (under the mode of presentation, #try to marry Jocasta#). So a broad theory can cover Oedipus' unsuccessful attempt to not marry Mother.[23] I conclude that Broad Psychology does not fail to explain events in laws like (M) that narrow psychology, on the other hand, captures (under narrow description), in the narrow version of those laws. So while (i) is true, (ii) is false. Hence, including Frege Cases in the clauses does not result in missed intentional-level explanation of Frege Cases.

At this point in the dialectic, let me turn to an important refinement to the solution to the Frege Cases.

3. Refinement—Frege Phenomenon Explanation

Recall that I introduced this issue as a challenge to (2):

(2). (a) Including Frege Cases in the *ceteris paribus* clauses does not result in Broad Psychology failing to predict Frege Cases. (b) Frege Cases can be

[23] One might also suspect that because Oedipus' *try to marry Jocasta/Mother* belief is included in the *ceteris paribus* clause of (M) Broad Psychology glosses over the fact that he tried to marry Jocasta. However, no missed intentional explanation arises from including the *try to marry Jocasta/Mother* thought in the *ceteris paribus* clause. Notice that the narrow thought, *try to marry Jocasta*, would obviously not have been subsumed by a narrow version of (M) either because the narrow version of (M) is about Mother, not Jocasta. Of course, the *try to marry Jocasta* thought would have satisfied the narrow prediction: '(N), If Oedipus wants to marry Jocasta and believes that he can try to marry Jocasta then he will try to marry Jocasta'. But because Oedipus has a *try to marry Mother/Jocasta* thought (which he has when he has #try to marry Jocasta#) he satisfies the broad version of (N).

predicted somewhere within the total theory (but not necessarily at the intentional level).

(I now divide (2) into parts (a) and (b) for reasons which will soon become apparent). My discussion thus far is only intended to serve as a defense of (2a). As we shall see, settling (2a) goes only part way towards a defense of (2b). For even bearing in mind the previous discussion, there seems to be the following gap in the Russellian's prediction of Frege Cases. Broad Psychology can predict the following:

(1) Oedipus will try not to marry Mother/Jocasta.
(2) Oedipus will try to marry Mother/Jocasta.

But (1) and (2) do not tell us that Oedipus will have a Frege Case; it just says that he is in deep trouble. Intuitively, a defining feature of a Frege Case is that the agent fails to be aware of a coreferentiality essential to the success of her behavior. A madman may know of the coreference and try to do both actions in any case. Intuitively, the madman is not having a Frege Case; what gets him into trouble isn't any ignorance about coreference but his failure to act rationally. Simply knowing that (1) and (2) characterize Oedipus is insufficient for us to tell whether he is a Frege Case or is simply being irrational. The proponent of narrow content will point out that in order to distinguish these scenarios we need a manner of belief individuation that is sensitive to narrow content or the mode of presentation (MOP) of the thought. For we need to know if Oedipus *narrowly* believes that Mother=Jocasta. Let us call this notion of belief 'belief*' to distinguish it from the Russellian notion of belief. According to the Russellian, anyone who believes that Jocasta=Jocasta or Mother=Mother believes that Mother=Jocasta. In contrast, whether one believes* that Mother=Jocasta, as opposed to simply believing* that Jocasta=Jocasta, depends upon belief individuation that is sensitive to modes of presentation (MOPs), or alternately, to narrow contents.[24] So at least at first blush, Broad Psychology cannot predict that Oedipus will have a Frege Case. It would seem that this is a fairly serious omission: knowledge that someone is having a Frege Case serves to explain his rationality despite the apparent irrationality of his actions. So it would seem that any psychological theory that fails to predict Frege Cases would be incomplete.

The problem with this objection, in broad strokes, is that the information about the agent's ignorance of the coreference is in fact available to a broad theory. This information may be unavailable intentionally, but it is available at the computational level. So the total theory can distinguish Oedipus from the madman. Indeed, the total theory can inform us about Frege Cases in an even richer way. To

[24] The problem is that the broad theory would consider anyone as instancing the thought that Mother=Jocasta who believes that Jocasta=Jocasta or Mother=Mother. This is because, according to Broad Psychology, Jocasta=Jocasta is intentionally identical to Mother=Jocasta.

illustrate this point I need to examine, in more detail, what it is to predict Frege Cases. Consider, again, the prediction:

> (M) *Ceteris paribus*, if Oedipus desires that he not marry Mother/Jocasta, and Oedipus believes that not marrying mother/Jocasta is the only way to bring this about, then he will try not to marry Mother/Jocasta.

Notice that this statement does not predict that Oedipus will have a Frege Case, it merely predicts that Oedipus will try not to marry Jocasta/Mother. It pays to keep in mind that *the intentional laws in which Frege Cases are included in the* ceteris paribus *clauses are not about Frege Cases per se; (M), for instance, is about people trying not to marry Mother*. After all, would (M) constitute an explanation for the phenomenon that when an agent fails to be aware of a coreference relevant to his behavioral success he will appear to act irrationally in his action(s)? Intuitively, only a generalization concerning this phenomenon could predict Frege Cases, and not, in contrast, a generalization concerning people trying not to marry Mother.

With this point in mind let us ask: where might a failure to predict Frege Cases arise from? Intuitively, it arises when a statement like the following is false according to the psychological theory:

> (FP) *Ceteris paribus*, if system S has distinct MOPs that represent entity *a*, but the MOPs are not linked in the system's data set as being coreferential, and S's behavioral success depends on the system's linking this data, then: S will appear to act irrationally in her *a* directed action.

Let us call explanation of the phenomenon behind Frege Cases 'Frege Phenomenon Explanation'. The reader may notice that (FP) is not a broad intentional generalization because it is sensitive to MOPs; more on this shortly. Before delving into this issue, I would like to clarify (FP) and then explain why it is significant. First, let me clarify that in speaking of MOPs that are 'linked in the systems data set' I mean that there is some level of mental processing, which need not be introspectively accessible, in which the way of conceiving of Cicero, and the way of conceiving of Tully, are encoded by the mind as being about the same object.

This being said, (FP) is introduced to illustrate that there are actually *two* ways that Frege Cases could figure as objects of explanation in psychological laws. In the case of laws like (M) the object of explanation is not Frege Cases qua Frege Cases, but some other phenomenon entirely—in this case the phenomenon that one tries not to marry Mother. In contrast, Frege Phenomenon Explanation has Frege Cases as the object of explanation. With this distinction in mind, we can now see that (2a) and (2b) are not equivalent: even if my previous argumentation succeeds in showing that including Frege Cases in the clauses does not lead to missed prediction, the broad theory may nonetheless fail to predict Frege Cases because it fails to include a generalization about Frege Cases along the lines of (FP).

At this point, it is natural to ask: can Broad Psychology incorporate a generalization along the lines of (FP)? After all, (FP) does not subsume agents by the Russellian propositions that their thoughts express, as it requires differentiating thoughts by their MOPs. Still, it can be regarded as a computational level generalization, as this is a theoretical level that is sensitive to MOPs. In this way, Broad Psychology can employ this generalization to predict that Oedipus will have a Frege Case. To recur to the case of the madman, (FP) can be summoned to distinguish irrational individuals who are having conflicting goals from rational agents who are experiencing Frege Cases. Those who are aware that the expressions co-refer will fail to satisfy the antecedent of (FP).

At this juncture, I expect the critic will object to these suggestions in the following way. It is inappropriate to take (FP) as a computational generalization, because (FP), and explanation of Frege Cases more generally, should be intentional. Notice that this charge is not really about the ability of Broad Psychology to predict Frege Cases; the present worry is that even if all the psychologically-relevant events are covered by Broad Psychology, certain phenomena are explained at the wrong level of analysis. This is really an objection to (3), a claim which, at the outset of the section, I argued needs to be proved by any answer to the Frege Cases.

4. Must Distinct Coreferring Names Differ in their Contents?

Recall that (3) was:

> There is no compelling philosophical argument that differences between coreferring concepts must have intentional, as opposed to computational, explanation.

Let us ask: why would only intentional explanation do the job? Indeed, it is a common suspicion that only intentional explanation can rationalize, or make sense of, thoughts and behaviors. Dennis Aryo voices a variant of the common concern that corefering concepts must have intentional, as opposed to computational, explanation in the following passage:

> *Given Oedipus' beliefs and desires—i.e. given the facts about how he believes the world to be and the way he wishes it to be—he acts exactly how we would expect him to behave, despite the utter unlikelihood of the results* . . . And this, it seems, requires a way of individuating the content of mental states which is sensitive to different ways the same thing in the world might be represented (Aryo, p. 244). (Italics mine).

Aryo does not explain why narrow content, rather than MOPs, are needed to make sense of Oedipus' thoughts and behavior. But I do not see why this should be

the case—is there supposed to be some a priori reason why explanation of Oedipus' behavior must be intentional, as opposed to computational? While it is correct that rationalizations of one's thoughts and behaviors seem to be sensitive to ways of conceiving referents, MOPs, as well as narrow contents, can capture one's way of conceiving the world (Aryo, 1996, p. 244).

To be fair to Aryo, his point was just an aside—a pronouncement of skepticism in the conclusion of a critical paper on the problem of Frege cases. We should ask: what fuels such suspicions? It is likely that they are motivated by one or more of the following considerations:

(I). Coreferential expressions must differ in their contents, rather than merely in their LOT expression type/MOP, because a certain theory of narrow content or two-factor theory is correct (e.g. the mapping theory).

(II). Coreferential expressions must differ in their contents because it is counterintuitive to take differences between MOPs as sources of exceptions to intentional laws (for MOPs are precisely what intentional laws are supposed to be sensitive to).

(III). Psychological generalizations must be purely intentional (and narrow) because computational explanation faces problems with MOP individuation.

I will discuss each of these considerations in turn.

(I). It is arguable that the two most influential theories of narrow content are Ned Block's version of conceptual role semantics and the mapping theory of narrow content (developed by Fodor, Perry, and others). Given that an exhaustive survey of the various theories of narrow content is well beyond the scope of this paper, what I propose to do is comment on why these leading theories fail to give us reason to regard the differences between coreferring names as being, in fact, differences in their content.[25]

Let us first turn to the mapping, or character-based views of narrow content.[26] Examples like the following are frequently given to motivate such views: Steve thinks, 'I am about to be run over by a train' and Ellen thinks of Steve, 'He is about to be run over by a train'; although the same singular proposition is expressed, the agent's behavior differs. On the other hand, if they both think, 'I am about to be run over by a train' they will, *ceteris paribus*, engage in the same behavior. This

[25] I shall not discuss versions of narrow content based on descriptivism (including cluster versions), as they have been widely regarded as problematic given the arguments of Kripke, Putnam and others. For a nice overview of problems with different theories of narrow content see Segal, 2000. (Unfortunately, Segal's own disjunctive theory of narrow content faces twin cases.) For problems with Harman's long-arm conceptual role theory and other two factor views see Segal, 2000 and Block, 1994.

[26] For such views see, e.g. Fodor, 1987; Kaplan, 1990; Perry, 1977.

contrast is supposed to motivate the view that the narrow (cognitive) content of a thought is its character; a function from the context of thoughts to extensions.[27]

The crucial problem with such views of narrow content, in very broad strokes, is that states that function very differently within the subject's cognitive economy can satisfy the same mapping from contexts to extensions. To borrow an example from Howard Wettstein, consider Sally, who sees Jose from outside a window of a building; and suppose that Jose is outfitted so that she cannot tell from looking at his right profile and from looking at his left profile that it is the same person. Suppose that Sally sees one side of Jose from the window (outside of the building) and then walks to a doorway and sees Jose's other side, and assumes that he is an entirely different person. Sally, upon learning of the identity, may say, 'ah, *he* is the same man that *he* is'. Here, both tokens of 'he' have the same linguistic meaning, but, intuitively, the states have different conceptual roles.[28] Objections along roughly the same lines can also be provided in the cases of proper names and kind terms; in general, two systems may have computational states which satisfy the same mapping function but play very distinct computational roles. Indeed, the mapping theory does not even treat Oedipus' #Jocasta# and #Mother# tokens as differing in their narrow contents; thus, the theory will face Frege Cases.

In contrast to the mapping theory of narrow content, Ned Block's version of conceptual role semantics provides a notion of narrow content that is designed to track sameness and difference of computationally relevant causal powers. Block, as noted, accepts referential generalizations and, in addition to this, aims to provide another sort of content as well. He explains: 'The internal factor, conceptual role, is a matter of the causal role of the expression in reasoning and deliberation, and in general, in the way the expression combines and interacts with other expressions so as to mediate between sensory inputs and behavioral outputs'.[29] Indeed, Block intends the specification of the causal roles to be in terms of internal computations in Mentalese (Block, 1994, pp. 97–99). As it happens, conceptual roles seem to be individuated by the very same features that type the LOT states that serve as MOPs on my own view. So I obviously think Block is on the right track. But for Block's theory of narrow content to genuinely challenge (3) he must provide reason to believe that these entities individuated by their conceptual roles are really narrow contents, rather than merely being computational states (MOPs) that are items in LOT syntax.

So let us ask: what reason is there to regard these narrow computational states as being narrow contents rather than merely LOT syntax? We might think that

---

[27] Such theories of narrow content take inspiration from the case of indexicals, but they are intended to extend to other expressions as well. To keep things brief, I will not delve into Kaplan's theory of indexicals and the related controversy surrounding whether, in fact, such a theory can serve as a basis for narrow content of non-indexical expressions more generally. This would take us too far afield; even assuming an extension is plausible, (as I suspect) the problems that I mention below apply.

[28] Wettstein, 1991, p. 191.

[29] Ned Block, 1994, p. 93.

because, according to conceptual role theories, entities that individuate narrow contents are called 'conceptual' or 'inferential', since such entities are standardly thought to be semantic, the narrow states get to be contents. But, as Block himself underscores, one cannot maintain that that which individuates the conceptual roles is itself semantic, for if one is concerned with reducing (or more aptly, naturalizing) intentionality in terms of computational states, one cannot explain semantic properties in terms of properties that are themselves semantic. He writes:

> Calling the causal roles CRS [Conceptual Role Semantics] appeals to 'conceptual' or 'inferential' shouldn't mislead anyone into supposing that the theory's description of them can appeal to their meanings—that would defeat the point of reductionist theories. The project of giving a nonsemantic (and non-intentional) description of these roles is certainly daunting, but the reader would do well to note that it is no more daunting than the programs of various popular philosophical theories (Block, 1994, p. 97).

In sum, naturalistic or reductive CRS programs must regard the features that individuate the narrow contents as being non-semantic. So we have no rationale for rejecting (3) here.

Block's reason for regarding the narrow states as being contents is that narrow meaning, 'determines the nature of the referential factor . . . '. More specifically, ' . . . what theory of reference is true is a fact about how referring terms function in our thought process. This is an aspect of conceptual role. So it is the conceptual role of referring expressions that determines what theory of reference is true. Conclusion, the conceptual role factor determines the nature of the referential factor' (Block, 1994, p. 109).[30] Block's remarks are unsatisfying; the fact that we are agents who have referring terms functioning in a particular way in our thought process does not suggest that the narrow states that we have, are, in fact, contents. For there is an alternate story that could be told: such conceptual roles are merely computational states that, only when supplemented with a reference relation linking the internal states to the world, have contents. And these contents are broad. On this alternate account, narrow contents drop out of the picture.[31]

---

[30] It should be noted that Block seems to hedge his bets on whether narrow contents are really contents: 'Nothing in my position requires me to regard narrow meaning and narrow content as (respectively) kinds of meaning and content. As mentioned earlier, I regard them as aspects of or as determinants of meaning and content' (Block, 1994, p. 92).

[31] Perhaps the proponent of narrow content would say that the states are contents because they are ways of representing the world. However, computationally individuated states, without a reference relation that links the states to referents, are not themselves ways of representing the world. They are uninterpreted symbols.

I suppose that one could simply stipulate that the computational states are contents, but this would not answer the charge that these entities seem to be just syntax and it would surely not provide the need argument that differences between distinct coreferring names *must* be differences in content.

Now let us turn to the second attempt to justify the view that differences between coreferring names must be differences in content. Rather than appealing to a particular theory of narrow content, in (II), the critic simply claims that my view is counterintuitive. For as an anonymous reviewer has noted, on the standard view of intentional laws, intentional laws are supposed to be sensitive to MOPs, and it seems absurd, in light of this, to suggest that differences between MOPs are supposed to be treated as sources of *exceptions* to intentional laws. For intentional laws, by their very nature, are supposed to be sensitive to MOPs.

I have a number of replies to this objection. First, the objector is assuming that intentional laws track MOPs, and this is precisely what is up for debate, for the broad psychologist denies this. Now, I am happy to grant that the standard intuition about intentional laws supports the view that the laws are sensitive to MOPs. But here I would ask: how far should these intuitions go? That is, how extensively should they influence the dialectical status of the broad/narrow content debate? I assume that both sides would agree that it is not built into the very concept of mental content that content *must be* narrow. Theories of broad content may strike a proponent of narrow content as being flawed, but certainly, they do not seem conceptually incoherent, like talk of round squares. But if it is not a conceptual truth that content is narrow, then it seems fair to say that a selection of a theory of content is an all-things-considered judgment, covering a variety of issues.

Now admittedly, philosophical intuitions should weigh into this judgment. However, while I'm happy to admit that there's an intuitive pull behind the neo-Fregean picture, I venture that there is also reason to regard as intuitive a theory that glosses over intentional differences between coreferring thoughts. For, to return to Block's observation, (pp. 430–431), there is a clear sense in which different individuals, no matter how they represent things, satisfy the same intentional generalization. (For example, no matter how people conceive of fire, they satisfy the generalization that people want to leave a crowded theater when they believe there is fire.) A purely narrow theory loses this sense, and thus, is susceptible to a counterintuitiveness charge as well.

Some might respond that it is open to the proponent of narrow content to claim that there are a number of broad intentional laws as well. (This would amount to occupying a two-factor theory.) However, if this move is made, it is very difficult for the critic to maintain the original objection. For the original objection was that intentional laws just are the sort of entities which should be sensitive to ways of conceiving of things; intentional laws should not take differences between coreferring thoughts as the source of exceptions. But broad laws will require such exceptions, and this new response grants that some intentional laws are broad.

(III) Now let us turn to a third, entirely different, motivation for rejecting (3). One may claim that explanation of differences between coreferring names must be intentional (and narrow) because computational explanation faces problems with MOP individuation. It is the main thesis of a recent paper by Murat Aydede, that an appeal to MOPs in a ' ... solution to the Frege Cases succumbs to the problem of providing interpersonally applicable functional roles for MOPs' (Aydede, 1998

p. 1). Here is the problem, in a nutshell: in order to include Frege Cases in the *ceteris paribus* clauses there must be some way to take the coreferring concepts as being different tokens of distinct MOP types. Any plausible theory of MOP state typing must cut states at the level of grain of sameness and difference of computationally-relevant causal powers. But this rather innocent demand cuts computational states extremely thin. For example, merely adding a new belief to one's existing data set can generate differences in the computational role of a thought, even if the thought is, in fact, seemingly unrelated to the added belief (for example, my having the thought, cows are ugly, may impact the computational role of my thoughts about coffee). Insofar as the addition of a single, seemingly unrelated belief impacts the computational role of a thought, it seems that we cannot rule out one's entire data set as contributing to the individuation of any given thought. As a consequence, only systems having all and only the same beliefs will have any single thought in common. Because systems sharing all and only the same beliefs simply don't appear in nature, LOT expressions will not be shared.

Murat Aydede claims that this lack of shared LOT expressions ruins the prospects for including Frege Cases in the clauses. But this claim is off the mark; all that is needed to include Frege Cases in the clauses is *intrapersonal* typing. For instance, to include Oedipus in the *ceteris paribus* clause of (M) all that is needed is to distinguish Oedipus' #Jocasta# and #Mother# thoughts at a given time; but it is uncontroversial that a functionalist (and holistic) individuation condition can distinguish thoughts synchronically within the same system.[32]

At this point, the critic may retort that 'interpersonal Frege Cases' can be concocted. For instance, Aydede gives a hypothetical case of a psychological generalization about people running in the direction of Superman when danger is near. He asks us to consider the hypothetical generalization:

---

[32] The critic might respond that even if Aydede's worry about shared LOT expressions doesn't impact the problem of Frege cases, LOT nonetheless faces a serious problem with 'publicity' because different individuals cannot have LOT expressions that are of the same type. Because this issue doesn't directly involve the Frege cases, I must leave it aside. But in a different work, I have argued that computational psychology can be public, despite any failure to arrive at a theory of shared LOT types (Schneider, 2003). To summarize, I argue that computational psychology is public because it appeals to the method of functional decomposition, and systems can have a similar decomposition despite having type distinct symbols. (For example, witness theories of memory, which abstract away from the particular memories encoded and focus on the functional decomposition of the memory system.) Second, computational psychology is public because computational theories in cognitive science have some laws that abstract away from MOPs, subsuming different individuals by their broad contents. (See footnote 18, for an example involving the moon illusion.) Here, their purpose is to single out phenomena that are supposed to be explained by lower-level computational accounts that appeal to functional decomposition. I argue that it is only in explaining the detailed workings of particular systems that computational explanation needs to appeal to MOP types. And here, a publicity requirement is not appropriate, for it is natural to expect any explanation to be highly system specific.

(P) When people feel threatened by perceived danger, and they think Superman will help them, they will run in the direction of Superman.

Now consider Lois Lane, who does not realize that the man she calls 'Kent' is also called 'Superman', and, because Superman/Kent is not in costume, fails to run to Superman/Kent when danger is near. As discussed, Lois is an exception to the generalization and will need to be included in the ceteris paribus clause. Now suppose that Sam is standing next to Lois when danger is near. Sam, on the other hand, is aware of the coreferentiality, and, while Lois stands still, darts in Superman/Kent's direction. This 'interpersonal Frege Case' is supposed to raise the following problem for Broad Psychology: lacking a manner of typing LOT tokens interpersonally, how is broad psychology to explain the differences in behavior between Sam and Lois?

Like the previous problem, this problem is not genuine: *intrapersonal* typing is all this is needed to solve *interpersonal* Frege Cases. We can say that as Sam stands with Lois, sensing danger, he has a MOP that corresponds to the English name, 'Superman', and one that corresponds to 'Kent' and that he knows that the same individual has these two English names. And Lois has a MOP that corresponds to the English name 'Superman' and a distinct one that corresponds to 'Kent' and she does not believe that the two English names pick out the same individual. Now, lacking a manner of interpersonal typing we do not know whether Lois and Sam have thought tokens which are of the same LOT type, (after all, many modes of presentation may correspond to one English expression), although we do know that the thought tokens refer to the same entity. However, despite this lack of knowledge the 'interpersonal' Frege Case can easily be solved. We have already established that we can handle a Frege Case for Lois via inclusion in the clauses. All that is needed to do so is intrapersonal typing. And intrapersonal typing is all that is needed to subsume Sam's thoughts in the generalization. He is subsumed in the same generalization that Lois is an exception to because their tokens share the same broad contents.[33] Hence, such objections can be set aside.

5. Taking Stock

Now let us take stock of things. This project set out to defend the plausibility of giving computational explanations of Frege Cases. I believe that we now have reason to question the bleak view that is generally present in the literature concerning the prospects for Broad Psychology to manage the Frege Cases. On my view including the cases in the *ceteris paribus* clauses is justified by a larger

[33] *Mutatis mutandis*, different systems may satisfy (FP), despite the holism of computational state individuation, insofar as each system has two (intrapersonally distinct) MOPs that represent the same entity.

theoretical decision for intentional laws with a canonical form that is broad. Any decision to include them in the clauses is the result of an overall assessment of the debate about which canonical form intentional laws should take, broad or narrow. While such a decision is a global affair, I have focused on the part of this theoretical decision that is internal to Frege Cases—the part which involves the issue of whether computational explanation of coreferring thoughts will suffice.

As noted, the literature on Frege Cases has been quite negative on this score. Against these critics, I've argued the following: first, Broad Psychology can treat Frege Cases as tolerable exceptions, rather than counterexamples, to broad intentional laws. Second, there is no missed prediction of Frege Cases: Broad Psychology does not fail to explain events in its laws that narrow psychology, on the other hand, captures (under narrow description). Further, although Broad Psychology must include Frege Cases in the clauses it can still offer generalizations along the lines of (FP) to predict Frege Cases. And finally, I've responded to various arguments that explanation of Frege Cases must be intentional that are based upon suspicions that a certain leading account of narrow content is correct, intuitions about the nature of intentional laws, and lastly, upon a skepticism about interpersonal typing of MOPs.

Department of Philosophy
Moravian College

References

Aryo, D. 1996: Sticking up for Oedipus: Fodor on intentional generalizations and broad content, *Mind & Language*, 11, no.3.

Aydede, M. 1997: Has Fodor really changed his mind on narrow content? *Mind & Language*, 12, 422–58.

Aydede, M. 1998: Fodor on concepts and Frege puzzles. *Pacific Philosophical Quarterly*, 79, 289–294.

Aydede, M. and Robbins, P. 2001: Are Frege Cases exceptions to intentional generalizations? *The Canadian Journal of Philosophy*, 31, no. 1.

Block, N. 1994: Advertisement for a semantics for psychology. In S. Stich and E. Warfield (eds), *Mental Representations; a Reader*, Cambridge: Blackwell.

Braun, D. 2001: Russellianism and prediction, *Philosophical Studies* 105, 59–105.

Churchland, P. 1995: *The Engine of Reason; the Seat of the Soul*. Cambridge: MIT Press.

Cummins, R. 2000: 'How does it work?' versus 'what are the laws?' Two conceptions of psychological explanation. In F. Keil and R. Wilson (eds), *Explanation and Cognition*. Cambridge: MIT Press, 114–144.

Crimmons, M. 1992. *Talk About Beliefs*. Cambridge: MIT Press.

Egan, F. 1995: Computation and content. *Philosophical Review* 104, 181–203.

Fodor, J. 1987: *Psychosemantics*. Cambridge: MIT Press.

Fodor, J. 1991: You can fool some of the people all of the time, everything else being equal: hedged laws and psychological explanations. *Mind*, 100, 19–34.

Fodor, J. 1994: *The Elm and the Expert: Mentalese and its Semantics*. Boston: MIT Press.

Kaplan, D. 1977: Thoughts on demonstratives. In P. Yourgrau (ed.), *Demonstratives*. Oxford: Oxford University Press.

Loar, B. 1996: Social content and psychological content. In A. Pessin and S. Goldberg, (eds), *The Twin Earth Chronicles*. London: M. E. Sharpe.

Perry, J. 1977: Frege on demonstratives. *Philosophical Review*, 86, 474–497.

Pessin, A. and Goldberg, S. (eds) 1996: *The Twin Earth Chronicles*. London: M. E. Sharpe.

Pietroski, P. and Rey, G. 1995: When other things aren't equal: saving ceteris paribus laws from vacuity. *The British Journal for the Philosophy of Science*, 46, 81–111.

Pylyshyn, Z. 1986: *Computation and Cognition*, London: MIT Press.

Ramsey, W., Stich, S. and Garon, J. 1990: Connectionism, eliminitivism, and the future of folk psychology. *Philosophical Perspectives*, 4, 499–533.

Richard, M. 1990: *Propositional Attitudes: An Essay on Thoughts and How We Ascribe Them*, Cambridge: Cambridge University Press.

Salmon, N. 1986: *Frege's Puzzle*. Cambridge: MIT Press.

Schiffer, S. 1991: Ceteris Paribus laws. *Mind*, 100, 1–17.

Schneider, S. 2003: *Meaning and the Computational Mind: the Relation Between Intentional and Computational States in a Scientific Psychology*. Ph.D. diss, Rutgers University.

Segal, G. 2001: *A Slim Book About Narrow Content*. Cambridge: MIT Press.

Wettstein, H. 1991: Has semantics rested on a mistake? *The Journal of Philosophy*, LXXXIII, 4, 191.