

MURAT AYDEDE

COMPUTATION AND FUNCTIONALISM: SYNTACTIC THEORY OF MIND REVISITED

1. INTRODUCTION

There is a thesis often aired by some philosophers of psychology that syntax is all we need and there is no need to advert to intentional/semantic properties of symbols for purposes of psychological explanation. Indeed, the worry has been present since the first explicit articulation of so-called Computational Theory of Mind (CTM). Even Fodor, who has been the most ardent defender of the Language of Thought Hypothesis (LOTH) (which requires the CTM), has raised worries about its apparent consequences. The worry can be put in the form of a question, which Fodor called the “Eponymous Question” alluding to the title of a chapter in his (1994) book:

(EQ) If cognition is computational, how can psychological laws be intentional?

This question has been haunting people working in the field since the publication of a paper by Stich in 1978 in which he gave his celebrated “autonomy argument”. Then, as everybody knows, came Fodor’s notorious “Methodological Solipsism” in 1980, in which he argued for the *formality condition*: namely, thought processes are causal sequences of symbol tokenings in one’s language of thought (LOT), and the causal processes are sensitive only to the syntactic/formal properties of its symbols. Hence, he argued against what he called a “naturalistic psychology,” i.e. a psychology whose laws essentially advert to broad semantic properties of mental states they cover. The alternative, rationalist psychology, according to Fodor, was to advert only to formal characteristics of symbols, of which Fodor conceived as narrow computational roles of LOT symbols.

Stich’s 1983 book, *From Folk Psychology to Cognitive Science*, was the culmination of the worries. He turned these into a sustained argument against the possibility of a scientific *intentional* psychology (along with the common sense belief-desire psychology), and at the same time, for a syntactic way of doing psychology, i.e., for his Syntactic Theory of Mind (STM). He defended an eliminativist stance: STM involves the elimination of all intentional idioms proposed to be used in a scientific enterprise, hence it envisions a scientific psychology free of semantics. STM has been around for almost two decades now. It has generated a lot of discussion because it has usually been taken to articulate the paradox alleged to underlie the LOTH,

which was to vindicate intentional folk psychology through computationalism. For this reason, I will concentrate on Stich's book in what follows, and argue that the worries are altogether baseless, that a computational theory needs a semantic individuating scheme to get off the ground, and that the envisioned alternative, i.e. a pure CTM (or, STM) is a non-starter, and cannot do the required job. Although nowadays there are probably few adherents of STM, so far no one, to my mind, has left a lasting impression of having refuted the theory. Indeed, as recently as 1994, Fodor raised the worry and tried to answer it by showing the feasibility of its alternative, and not by directly attacking the syntactalist claim. In what follows, I intend to leave a lasting impression: I hope to refute the STM and show that the kind of semantic-free psychology it envisions is impossible, thus answering Fodor's Eponymous Question.

In his book, Stich has argued basically for two claims. First, the application conditions of such intentional common sense predicates as 'believes that *P*' and 'desires that *Q*' are essentially vague, context-sensitive, observer-relative, and thus are not suitable to be used as stable projectable predicates in the vocabulary of a scientific psychology. In particular, according to Stich, since observer relativity partly stems from the fact that content ascriptions are essentially based on similarity judgments along different dimensions (see below) between the ascriber and the ascribee, a consequence of ordinary content ascriptions is that a certain form of parochialism will profusely infect our psychological theories if we insist on having a content-based psychology which, according to Stich, essentially relies on ascribing such contents to agents covered by its generalizations. This means that content-based psychologies are bound to miss many important generalizations about the psychology of children, exotics, perceptually or cognitively handicapped people, higher animals, etc., since any content ascribed to these will necessarily reflect how cognitively similar the ascribee is to the ascriber. In short, Stich thinks that content-based psychologies won't make respectable science.

His second general claim is that we don't *need* to advert to the content of mental states in doing psychology, "syntax" will be enough. Furthermore, we had better advert only to the syntactic properties of mental states if we don't want to miss any psychological generalizations: STM offers a paradigm that has all the virtues of content-based psychologies and none of its vices.

These two claims are relatively independent of each other: in particular, the truth of the latter does not depend on the truth of the former. If Stich is right in his second claim, then the falsity of his former claim, i.e. his characterization of contentful mental states as scientifically problematic posits, would imply that content vocabulary is at best otiose in doing scientific psychology. It is therefore important to see whether Stich is right in his second claim. In what follows I will argue for three basic claims.

The first is that when we see more clearly the nature of STM, as presented by Stich, the claimed superiority of STM over content-based psychologies totally disappears. Put differently, I will be arguing for a conditional claim: *if* Stich is right in his claim that content-based psychologies have the disadvantages he enumerates, *then* STM-style theories have exactly the parallel problems; so it is false that the STM

framework is scientifically superior to content-based psychologies as conceived by Stich. Therefore, Stich will lose his primary motivation to promote STM.

Secondly, I will argue that STM can't do the required job: it lacks the necessary resources to type individuate particular psychological states *qua* mapped onto particular "syntactic objects" as Stich puts it.

Thirdly, I will show that the STM-theorist is, at any rate, committed to intentional vocabulary at some stage of theory construction. In other words, if the STM strategy is taken to claim, as Stich seems to intend, that it is possible and advisable to develop psychological theories without using any intentional scheme whatsoever no matter what the stage of theory construction is, then STM is false: psychological theory construction cannot get off the ground if the strictures of STM are firmly complied.

I will end by moralizing on Stich's failure, and point out that if STM (\approx Narrow Causal Account of typing symbols over which computation is defined, as I will show) and type-type identity theory are false, then a content-based psychology (= intentional psychology) is practically mandatory. Hence, if cognition is computational, psychological laws have got to be intentional!

Since all my arguments crucially depend on what exactly STM is, I will present it in a way that its purely functionalist structure becomes explicit. However, before embarking on my criticism, I need to say a little about how Stich views the problem space within which he criticizes content-based psychologies and thus motivates his own alternative, STM. In particular, it will turn out that the exact way in which Stich motivates his STM is very important, since my arguments against STM partly rely on his own strategy.

2. THE PROBLEM SPACE ACCORDING TO STICH

Stich takes what he calls the Mental Sentence Theories as his starting point, and assumes their basic framework throughout his discussion. After a lengthy presentation of a Fodor-style LOTH and computationalism in general he raises the following problem:

for a Fodor-style account of belief sentences to hang together, we must have some workable notion of what it is for two distinct people, speaking different languages, to have in their heads distinct tokens of the same sentence type. (Stich, 1983: 43-4)

On behalf of Fodor, he offers three possible solutions: One is the *Narrow Causal Account* (NCA), according to which two sentence tokens count as type identical if and only if (iff) they have the same narrow causal/functional role. Since this is going to be of some importance, let me elaborate on it a bit. According to Stich, "[t]o adopt this view of . . . psychology is to exclude any reference to noncausal relations . . . There can be no mention of a subject's social setting, natural environment, or personal history, nor of the psychological characteristics of other people" (1983: 22). This is what makes this kind of individuation *narrow causal*. It is narrow because the causal role in question is defined in terms of generalizations that detail nomological connections among *proximal* stimuli, behavior (like motor commands), and other central cognitive states. Moreover, the causal relations are given by a set of *counterfactual supporting* generalizations. Thus, for a mental state of an individual to count, say, as

the belief that *P*, it is not necessary that the state *actually* play a causal role in the individual's mental economy; all that is required to be true of the state is that it *would play* a certain causal role if some other conditions specified in the generalizations *were to obtain*. So the notion of functional/causal role of a mental state should be so understood as to include the *potential* causal interactions that the state would enter. Accordingly, two mental states of two distinct organisms count as of the same type if their potential causal interactions are the same, namely, if they are covered by more or less the same lawlike generalizations, despite the fact they may differ quite radically in their actual etiologies. Finally, the generalizations in question are hedged by *ceteris paribus* clauses.

The second account is what Stich calls the *Semantic Account* (SA) according to which two sentence tokens count as type identical iff they have the same semantic content. The third one is what might be called the *Physical Account* (PA) according to which the sentence tokens are of the same sentence type iff their physical properties, their shape, so to speak, are the same. After quickly dismissing the PA as hopeless, Stich makes the following remark:

[The] interesting question is how causal accounts and content accounts compare with each other. Do they categorize mental tokens differently, or do they inevitably come out with the same categorization? On this issue, opinions divide. According to Fodor the two sorts of classification schemes coincide, "plus or minus a bit." Indeed Fodor sees this as "*the* basic idea of modern cognitive science." Any thoroughgoing [i.e., content] functionalist in the philosophy of mind will also end up on this side of the divide. On the other side, denying that causal and content accounts converge, are Field, Lycan, Perry, McDowell, and the truth. (1983: 48-9)

Here Stich conceives of the SA as fixing the type identity of mental sentences according to their broad content. Twin-earth cases show that functionally identical twins' mental states may differ in their broad content. However, this is not going to be very important for what follows. Since, for many people in the field, narrow content is a construct out of broad content, Stich has the same line of argument against narrow content.¹

Here is Stich's argumentative strategy. Stich thinks that if folk psychology is to be scientifically vindicated through some version of a mental sentence theory, the SA of typing mental sentence tokens is indispensable. He then proceeds to show that the NCA and SA come up with radically different taxonomies. The way he does this is idiosyncratic. He constructs a series of thought experiments that are supposed to show that folk judgments about how to classify certain mental states do radically differ from the way the NCA would type them. Then relying on what these thought experiments seem to show, he proceeds to give an account, or rather a "descriptive analysis" of folk conception of belief as a paradigm case of a contentful mental state, i.e. as a paradigm case of mental state typed according to the SA.

According to Stich's analysis, the "content identity" of beliefs that is thought to be assumed by folk psychology is a myth. On the basis of the evidence he claims to have collected through his thought experiments, he claims that the notion of content according to folk psychology is such that it is only a similarity measure along three different dimensions that the folk implicitly assume. One dimension of similarity between contents is the *functional or causal-pattern* of contentful mental states: "A

pair of belief states count as similar along this dimension if they have similar patterns of potential causal interaction with (actual or possible) stimuli, with other (actual or possible) mental states, and with (actual or possible) behavior” (Stich, 1983: 88-9). The second dimension draws on the *ideological (doxastic) background* of the agents. Since these can greatly vary from person to person, the relation between two beliefs in two different people can only be a matter of similarity: “The ideological similarity of a pair of beliefs is a measure of the extent to which the beliefs are embedded in similar networks of belief” (89). The third dimension of similarity measure concerns the *reference or truth-conditions* of beliefs. According to Stich, since they are dependent on the speakers’ linguistic community, social embeddings, the causal history of the use of terms, the speakers/believers’ physical as well as cultural environments, etc., the reference will vary as these vary without necessarily affecting the functional role of a mental state. To the extent that these factors are similar, to that extent the contents of beliefs will be similar. Stich thinks that this is essentially what the SA of typing mental sentences comes down to.

It is now relatively easy to see how the two taxonomic schemes diverge. The NCA can capture only the *causal pattern similarity* dimension assumed in the SA. It can’t be sensitive to the other dimensions. Stich concludes that “the mental sentence theory of belief, if fleshed out with a narrow causal account of belief, just does not comport with our workaday folk psychological notion of belief—it is not an account of belief, as the term is ordinarily used” (1983: 49).

If the two taxonomic schemes differ, what scheme should a scientific psychology adopt? Stich argues that adopting the SA is ill advised, because mental states typed according to the SA will make bad science since a semantic taxonomy would only provide the psychologist with a theoretical vocabulary whose application is vague, unstable, context-sensitive, and observer relative. Who would want such a science, Stich argues, especially if there is a clear alternative that is free of such defects? According to Stich, the alternative is a psychology whose taxonomic scheme is based on the NCA. This is the STM paradigm. Hence Stich’s main conclusion: if a mature cognitive science is and ought to be committed to the NCA (\approx STM), then folk psychological notions like beliefs and desires are likely to be eliminated.

This is how Stich motivates and argues for his STM. It is therefore very important to see whether Stich is right in his claim that the STM paradigm is really superior in any of the respects in which he criticizes content-based psychologies. As I advertised, I will argue that Stich is wrong.

3. WHAT IS STM?

According to Stich, the core idea of STM can be captured in the following way:

the cognitive states whose interaction is (in part) responsible for behavior can be systematically mapped to abstract syntactic objects in such a way that causal interactions among cognitive states, as well as causal links with [proximal] stimuli and behavioral events, can be described in terms of the syntactic properties and relations of the abstract objects to which the cognitive states are mapped. More briefly, the idea is that the causal relations among cognitive states mirror formal relations among syntactic objects. (1983: 149)

Stich here considers two networks, one of which is the network consisting of the causal relations among brain state types, proximal stimuli and behavioral events. This network is supposed to be mirrored by another network expressed by a syntactic psychological theory **T**. This theory consists of at least three kinds of generalizations: (1) the ones that nomologically connect proximal stimuli to **B**-states (belief-like states) with *particular* syntactic objects mapped to them, (2) the ones that describe causal relations among **B**-states and **D**-states (desire-like states), and (3) the ones that nomologically connects **B**-and **D**-states to motor-gestures. Following Michael Devitt (1990), I will call these kinds of generalizations: I-T, T-T, and T-O generalizations respectively.²

If we want to put **T** into some canonical form, we may write out **T** as a single conjunctive sentence, replacing all the occurrences of the theoretical predicates such as “x has a **B**-state mapped to δ_1 ” and “x has a **D**-state mapped to δ_1 ” with expressions of the form:

$$\begin{aligned} x \text{ is in (some member of) } B(\delta_1), \\ x \text{ is in (some member of) } D(\delta_1). \end{aligned}$$

B and **D** are functions (in the set theoretic sense) that map a particular syntactic object, which the theorist had already specified for the job at hand, onto the set of x 's first order physical state types that have the functional role that **T** associates with that syntactic object. We may now express **T** in the following way:

$$\mathbf{T}[s_1, s_2, \dots, B(\delta_1), B(\delta_2), \dots, D(\delta_1), D(\delta_2), \dots, b_1, b_2, \dots] \quad (\text{i})$$

where s_i 's are proximal stimulus types and b_i 's behavioral event types (motor gestures), and δ_i 's are specific syntactic objects.

Roughly, this is the form an STM theory would take. Let us now see how STM is committed to the NCA of typing brain states hypothesized by the theorist, i.e., how we can get their explicit functional definitions.

From (i) it is easy to get the Ramsey sentence of **T** by quantifying over the functions **B** and **D**:

$$(\exists f_1)(\exists f_2) \mathbf{T}[s_1, s_2, \dots, f_1(\delta_1), f_1(\delta_2), \dots, f_2(\delta_1), f_2(\delta_2), \dots, b_1, b_2, \dots] \quad (\text{ii})$$

We can now get the explicit functional definition of **B**:

B =_{df.} The function f_1 , such that there is a function f_2 , such that the two uniquely satisfy

$$\mathbf{T}[s_1, s_2, \dots, x_1(\delta_1), x_1(\delta_2), \dots, x_2(\delta_1), \dots, b_1, b_2, \dots]'$$

Similarly for the definition of **D**.

Although this is the formal procedure to get the explicit functional definitions of **B** and **D**, what we really want is explicit functional definitions of ‘**B**(δ_i)’ and ‘**D**(δ_i)’ for

each i . The intuitive idea is this. Notice that in this formalism the existential quantification in getting the Ramsey sentence is over certain functions that map distinct syntactic objects to distinct sets of an organism's first-order physical states. Here, in a certain sense, syntactic objects are exploited as external indices that pick out certain states of an organism that have distinct functional roles as specified by theory **T**. Each specific syntactic object in virtue of its distinctive place in **T**'s generalizations specifies a unique functional role that the two functions **B** and **D** then map onto the underlying physical states of the organism. Intuitively, we may extract the functional definition of $B(\delta_i)$ for each i in the following way: since δ_i in the domain of **B**, in virtue of its place in **T**, is supposed to pick out a *unique* functional role that may be indexed by F_i ,

$B(\delta_i) =_{df}$ the set of first-order states that have F_i as determined by **T**.

Similarly for **D** and for each particular δ_i .

Now Stich does not present his STM in this way. Here I have used a procedure very similar to the one developed by Brian Loar (1982a) in his presentation of his own content functionalism. This is not accidental of course. In fact, this is the point. For, as should be obvious, Stich's STM, *structurally* at least, is nothing but a de-intentionalized version of Loar's content functionalism, except that Loar takes the causal role of "observational" beliefs to be fixed on the basis of *distal* stimuli. Where Stich uses abstract syntactic objects, Loar uses ("fine-grained") propositions, intentional objects *par excellence*. The type identity of specific abstract syntactic objects is given by their place in the theory. This is the way they are purely functionally defined according to their narrow causal profile.

In fact, the similarities between STM and Loar's content functionalism are, in one respect, stronger than that. Loar uses propositions in the initial stage of getting the functional theory first (and, for good reasons—see below). He then proposes a procedure by which all the propositions are replaced by purely formal expressions. The theory in this ultimate form structurally is almost an STM! Loar, of course, is no eliminativist. His aim is to naturalize intentionality by offering a sophisticated functionalist theory. So he thinks at some stage he should get rid of the intentional objects like propositions he had initially used. Once the theory is completed, it is supposed to provide sufficient (and, necessary?) conditions for a mental state to have a semantic content, which can ultimately be specified without using any intentional terminology. This is his strategy, and as far as it goes it is perfectly kosher. But if I am right in what I am going to say, it does not go very far, at least in its narrow version.

4. SOME CURIOUS ASPECTS OF STM

My presentation of Stich's STM may be taken to be tendentious. I presented it as a purely functionalist theory and said that the abstract syntactic objects, which the brain states are mapped onto, may be viewed as indices that are external to the underlying first-order brain states. But, STM is supposed to be a formal/syntactic theory very similar to the Computational Theory of Mind (CTM) Fodor has de-

veloped and defended. STM is supposed to be a de-intentionalized version of what Stich calls Mental Sentence Theories. Indeed STM has been taken in this way in the literature by its friends and foes. But if my presentation is right, STM is not in fact theoretically committed to there being syntactically complex “sentences” literally realized in the brain. If so, how could it be very similar to Fodor’s CTM? Stich writes:

It is not, strictly speaking, required for an STM theorist to view hypothesized neurological state tokens as mental sentence tokens, though talking of them in this way is often an all but unavoidable shorthand. (1983: 152)³

This is curious but actually understandable. Remember Stich’s question about how the tokens in different heads can be individuated as of the same sentence type. His solution is the NCA. But the NCA requires a theory first in which syntactic expressions figure as theoretical terms in the generalizations. However, once we have such a theory, it is easy to define the syntactic expressions functionally à la Loar. But once we do that, the question of whether the referents of such expressions do really have syntactic structure somehow realized in the brain becomes secondary and at best an open empirical question. For, if the functional theory is true, it seems that we can do everything we want that the Mental Sentence version of the theory can do.⁴

So STM as a purely functionalist theory is not committed to a semantic-free LOT. On the other hand, of course, whatever CTM is, it cannot be neutral with respect to the question of whether there are syntactically complex sentences realized in the brain. CTM should be so formulated that it essentially entails a positive answer to this question. The problem in fact stems from the widely shared conviction that the type identity of brain sentences can and should be given in terms of the NCA (for some, as well as in terms of other ways like the SA). Fodor, at least in his early writings, is explicit about how to formally type the LOT symbols: functionalism à la NCA is the answer.⁵ Below, I will argue that this can’t be done. So there is at least this dissimilarity between STM and CTM: whereas STM is non-committal about there being brain sentences, CTM, whatever it is, is essentially committed to it.

Having made the point, however, I want to talk of STM as if it were concerned with the functional individuation of syntactically complex brain sentences. Not only because, as Stich says, this is an all but unavoidable shorthand, but also because I want to see whether Mentalese expressions can be individuated on the basis of the NCA if the LOTH is true. Hence, my argument can equally be seen as an argument against Narrow Content Functionalism (NCF) in so far as it is pursued as part of a naturalistic semantic program run on a LOT story. So, in what follows, I will assume the framework of Mental Sentence Theories, and often treat STM in this form.

So far we have been talking about the functionalist nature of STM, and thus its commitment to the NCA of typing brain states. But what does this have to do with syntax? More particularly, how does Stich conceive of syntax when he talks about the syntactic type identity of brain sentences? Indeed, what makes his theory a “syntactic” theory? To this last question he answers in the following way:⁶

We would have no reason to view brain states as syntactically structured unless that structure can be exploited in capturing generalizations about the workings of mind/brain’s mechanisms. Attributing syntactic structure to brain state tokens—assigning them to syntactic types—is justified only if some

interesting set of causal interactions among those tokens is isomorphic to formal relations among abstract syntactic objects. (1991: 244)

Notice that if Stich is right about this, Fodor can't have any reason for postulating a separate computational level in which intentional laws of psychology are implemented. In particular, what is puzzling about Stich's answer is that he doesn't mention at all the Turing legacy which is the main driving force behind Fodor's insistence that the computational story, according to which thought processes are defined over the formal/syntactic properties of representations, is our only plausible story about how semantically coherent processes can be physically/mechanically possible.⁷ Stich's interest seems not to be in computationalism classically understood. This is understandable to a certain extent. For Stich doesn't think that there are any semantically coherent thought processes that need the attention of science because he doesn't think that there are any states with semantic content. Put this aside. He has a different line of answer.

When he talks about the syntactic type identity of brain sentences, he has a "rich" notion of syntax, according to which mere difference in lexical items (e.g. "Tully was bald" versus "Cicero was bald", or "*Fa*" versus "*Fb*") is enough to make the sentence tokens belong to different syntactic types.⁸ In particular, for Stich, the criterion according to which two sentence tokens in two different heads count as of the same type is a syntactic one. But since this criterion is captured by the NCA, the syntactic type identity of brain sentences is a matter of functional identity:

when mental states are viewed as tokens of syntactic types, the functional profile exhibited by a mental state can be equated what we have been calling its formal or syntactic properties. (Stich 1983: 190)

So it seems that, according to Stich, the very postulation of complex semantic-free sentences realized in the brain whose "syntactic" type identities are given purely functionally is what makes Stich's theory a syntactic theory. As I argued elsewhere, I don't think this notion of syntax is the one that is needed for a Fodorian Computational Theory of Mind: what is required for the LOTH is a combinatorial syntax that fixes the logical form of expressions.⁹ Nevertheless, "syntactic" typing of LOT symbols has always been understood on the basis of the NCA. But put this aside. The important question I will address below is whether the type identity of brain sentences can be given in terms of their narrow causal profile, *whatever it is called*. Now let us see whether the STM paradigm is any superior over content-based psychologies.

5. THE ALLEGED SUPERIORITY OF STM

From Stich's analysis of folk conception of belief individuation, it follows that predicates like 'believes that P' (1) are vague and unstable, (2) depend on a (observer-relative) similarity matrix along three different dimensions for their applicability, and (3) their application involves many unnecessary "fine-grained distinctions which contribute nothing in the explanation [and prediction] of behavior." From (2), it also follows that there are likely to be many important cognitive generalizations that will not be stateable in terms of such predicates. So a content-based psychology will

inherit all of these limitations. In contrast, the STM style theories, Stich claims, will have none of these defects.

In this section, I will argue that *if* Stich's criticism of content-based psychologies is right *then* exactly parallel problems equally plague STM. But for this, we first need to see, exactly, how Stich argues for the superiority of STM. In other words, since my claim is conditional, we need to see in some detail what makes its antecedent true according to Stich, and why he thinks that STM is free of similar problems.

5.1. How STM Is Supposed to Be Superior

In discussing how STM theories will succeed where the content theories fail, Stich again uses the thought experiments he had considered in showing how the content taxonomies radically differ from the ones based on the NCA. Much of the difference stems from the fact that whereas the individuation of content essentially depends on three different dimensions, the NCA is only committed to individuating mental states according to their narrow causal pattern. The other two dimensions, ideological and referential (or, truth-conditional) similarity are to be amputated. First, these last two are unnecessary and therefore contribute nothing to the explanation and prediction of behavior. Second, by getting rid of them, context-sensitivity is eliminated. That is because, as in every multi-dimensional similarity judgment, it is the context that decides which dimension is to be emphasized in deciding whether a given state in a particular situation counts as the belief that *P*. Sometimes referential similarity will count more, sometimes ideological similarity, or simply causal pattern similarity depending on the demands of the particular situation in which the question arises.

Stich puts the greatest emphasis on the problems created by the *ideological similarity dimension*. This involves what he calls the *holism problem* in the folk conception of belief. In order to bring out the problem vividly, let's focus on his most celebrated thought experiment: the case of Mrs. T. Mrs. T is an elderly woman who suffers from a progressive loss of memory. At the end, she does not "know" what an assassination is, what dying is, who McKinley was, etc. Nonetheless, she appears to remember/believe that McKinley was assassinated, because that is what she persistently says when asked "What happened to McKinley?" According to Stich, the folk psychology's clear verdict is that she does not believe that McKinley was assassinated. Stich's diagnosis is that when she ceased to have a certain set of relevant beliefs, she ceased to believe that McKinley was assassinated, despite the fact that she appears to respond correctly to the question. This, Stich says, shows that folk conception of belief attribution attends to the doxastic background of an agent. From this he infers that the type-identity of someone's belief is partly constituted by what other actual beliefs the individual happens to have. This is the notorious problem of content holism, according to Stich.

On the other hand, an individuating scheme based on the NCA, he claims, is and ought to be nonsensitive to the actual doxastic surrounding of a mental state it individuates. That is because the NCA taxonomizes her state underlying her utterance on the basis of its *potential* (narrow) causal interactions. Thus STM is able to account for her ability to infer, for instance, "McKinley was buried in Ohio" from

her “acknowledgment” of “McKinley was assassinated” and “if McKinley was assassinated then he is buried in Ohio.” So whereas content psychologies miss such important generalizations as those that cover Mrs. T, STM theories will be able to take such agents under their scope.

Stich also argues that with respect to the *reference* and *causal pattern similarity* dimensions STM is superior. But I do want to leave aside the discussion of Stich’s arguments about these latter dimensions here since I think they are pretty weak and don’t occupy any centrality in his discussion. Stich puts the greatest emphasis on the holism problem of content-based psychologies, and clearly thinks that STM theories are free of this serious problem. If it can be shown that the NCA is equally problematic in this respect then STM is undermined completely, and this is exactly what I intend to do next.

5.2. The Parallel Disadvantages of STM-Style Theories

5.2.1. Doxastic Similarity Dimension and the Holism Problem

Although most of Stich’s arguments for his case against belief turns on the “holistic” nature of commonsense individuation of beliefs, it is unfortunate that he does not elaborate on what, exactly, holistic nature of content comes down to. Much of what he says on the matter is provided through a handful of examples like the case of Mrs. T.

In what follows, I will first try to explain as clearly as possible the sense in which Stich thinks that commonsense individuation of beliefs is holistic. My discussion will show that he is obscure and vague about what he thinks the “holism problem” is. I will then indicate exactly why Stich thinks that the holistic nature of belief is a problem for content based-psychologies, and why he thinks that an STM-style theory is totally free of it. We need to be as precise as possible about this, because my argument against Stich depends on his own premises.

According to Stich, the identity of a particular belief, say, the belief that McKinley was assassinated, depends on what other actual beliefs a person happens to have. The doxastic surround of the belief that McKinley was assassinated is constitutive of its content. But Stich does not say what this doxastic surround is, how it is determined, nor how it is supposed to be constitutive of a given content. For instance, at any moment, a person who believes that McKinley was assassinated has also a very large stock of other beliefs. Do all of them contribute to the content of the belief in question, or only some of it? From the way Stich writes and uses the expression ‘doxastic surround of a belief’ and its cognates, it seems that only some portion of a person’s entire belief system is relevant to the determination of the content. Unfortunately, he gives almost no clue about how big the portion is supposed to be.

The case of Mrs. T is typical. She ceases to have many beliefs among which, for instance, is the belief that if someone is assassinated then she is dead. In fact, she no longer “knows” what dying is, what a presidency is, what presidents do, etc. In all of Stich’s similar examples, the beliefs that make up the doxastic surround of a particular belief have rather certain “direct relations” to the belief. They are not only semantically close, but also, in some loose sense (to be discussed later),

“conceptually” tied to the belief. In all such examples, the fact we are invited to observe is that when someone ceases to have *those* kinds of beliefs then someone ceases to have a particular belief.

Stich is crucially vague and not particularly careful here. If his claim is that the content of a particular belief is (partly) determined by the set of *all* beliefs one has, which I take it what holism at its extreme comes down to, then he has not provided any single reason, let alone a relatively elaborate argument, for his claim. On the other hand, if his claim is that only *some* beliefs determine content, as he seems to intend, then the identity conditions for belief are not holistic.

I think that part of the reason why Stich is so vague and careless is that he does not care about this distinction, some or all. According to Stich, it seems, the very fact that the content of a belief depends for its identity on at least some other actual beliefs the agent has is enough to make serious trouble for any psychology that hopes to essentially advert to content in its generalizations. For one thing, given Quine’s influence on him, Stich clearly thinks that the distinction between those beliefs that determine content and those that don’t can be anything but sharp and principled. If so, content can at best be a matter of degree. And this is enough to make trouble: a content psychology is possible, at best, for those who are doxastically similar. But even for such a psychology, vagueness will still continue to be a serious problem, since it is almost certain that doxastic similarity never actually achieves doxastic identity among people. Whatever the case is with Stich’s analysis of belief, however, he certainly thinks that his alternative paradigm of doing psychology, STM, does not have any such problem.

Why does Stich think that the framework provided by STM has no such “holistic” problem. Here is a typical remark by Stich:¹⁰

In chapter 7, section 3 our focus was on ideological similarity, and the persistent problem was that as subjects became increasingly ideologically distant from ourselves, we lost our folk psychological grip on how to characterize their beliefs. For a syntactic theory, however, ideological similarity poses no problem, since the characterization of a B-state does not depend on the other B-states that the subject happens to have. A B-state will count as a token of a wff [well-formed formula] if its potential causal links fit the pattern detailed in the theorist’s generalizations, regardless of the further B-states the subject may have or lack. (1983: 158)

Stich, then, goes on to clarify how this can be so by working on the example provided by Mrs. T:¹¹

If we assume that before the onset of her disease the B-state which commonly caused her to say “McKinley was assassinated” obeyed generalizations like (4)-(6), then if the illness simply destroys B-states . . . without affecting the causal potential of the tokens which remain, the very same generalizations will be true of her after the illness has become quite severe. In chapter 7 we imagined a little experiment in which, shortly before her death, we tell Mrs. T, “If McKinley was assassinated, then he is buried in Ohio,” and she replies, “Well, then, he is buried in Ohio.” This is readily explainable by (5) [the syntactic version of psychologized Modus Ponens] . . . So if the generalization is there, it can be captured by a syntactic theory. But as we saw, there is no comfortable way to capture this generalization in the language of folk psychology . . . Thus a cognitive science that adopts the STM paradigm can aspire to broadly applicable developmental, clinical, and comparative theories, all of which are problematic for a content-based theory because of the constraints of ideological similarity. (1983: 158-9)

Is it true that ideological similarity poses no parallel problems for STM-style theories? I think not. It is time to see why.

5.2.2. *Holism and STM*

Here is the structure of the argument for my claim that STM, contrary to Stich's advertisement, has exactly the parallel problem.

- (1) The STM framework is committed to the NCA of type individuating B-states *qua* mapped to particular syntactic objects like, say, 'Fa', through the generalizations that cover them.
- (2) The NCA is capable of individuating such states *only if* it has enough generalizations of a certain sort, which I will call, S-generalizations.
- (3) If STM has S-generalizations among its stock of generalizations then it has all the parallel problems that Stich complains about content-based psychologies regarding the dimension of ideological similarity.

In the remainder of this section, I will make this argument stick. I take it that (1) is common ground (see above). Let me first argue for premise (2).

All the T-T generalizations Stich ever considers, by way of giving examples or otherwise, have rather a certain character: they all quantify over *particular* syntactic objects, i.e., they all use meta-variables to refer to classes of actually specified sentences that have a certain common "logical form". Even in the quote above, it is apparent that when he talks about the causal interactions of the token that underlies Mrs. T's utterance of 'McKinley was assassinated', the generalizations Stich has in mind are of this kind. Let me call the generalizations that quantify over *particular* brain sentences "L-generalizations", since these apply to any sentences that have a certain "logical" form. L-generalizations are all blind to the primitive non-logical vocabulary that the STM-theorist specifies.

It should be obvious that if all the T-T generalizations that go into the specification of the causal role F in the individuation of $B(\delta_i)$ for any i (see above) are of this type, i.e., if they are all L-generalizations, then there cannot be a *unique* causal role for each particular $B(\delta_i)$, which means that there can be no type individuation of B-states with particular syntactic objects mapped onto them. Here is why: with only L-generalizations in force, *any* sentence token has *potential* inferential (causal) connections to *any* other one. Put differently, since, on Stich's own admission, the generalizations in the theory detail not only the actual but also the potential causal interactions of any particular B-state, and since any sentence token can potentially be "inferred" from any other (i.e., causally connected through L-generalizations to any other), L-generalizations all by themselves cannot type individuate particular B-states.¹² All they can specify is at most the "logical" form or syntactic type of sentence tokens. As we will see in the next section, this situation does not change even when we add the I-T and T-O generalizations to L-generalizations: together they are still incapable of providing unique causal roles for particular B-states.¹³ For one thing, as I will point out later, there can be no such (narrow) I-T/T-O generalizations. But, for our purposes here, more importantly, even if there are such generalizations they can help at best to identify a very small subset of particular B-states whose character is rather "observational". However, Stich himself is pessimistic about there being any such subset (see below). My point is that S-generalizations are necessary (not sufficient) for type individuating at least some B-states, and this will do for premise (2).

What is needed, of course, is a different kind of T-T generalizations in addition to L-generalizations, T-T generalizations that are not blind, so to speak, to the primitive non-logical vocabulary of the STM-theorist, generalizations that detail (part of) the causal role that is unique to, say, the B-state mapped to 'Fa'. It is obvious that such "low-level" generalizations will typically be the syntactic parallels of such "content generalizations" (*C-generalizations*) as¹⁴

- (i) For all subjects S and for all x, if S comes to believe that x is a cow, then S will typically come to believe that x is an animal,
- (ii) For all subjects S and for all x, if S comes to believe that x is a bachelor, then S will typically come to believe that x is unmarried,

and so on.¹⁵ Let me call the *syntactic parallel* of this kind of C-generalizations "*S-generalizations*". Stich is committed to such generalizations, otherwise there is no individuation of particular B-states. Hence, premise (2).

Let's now take up premise (3): If Stich is committed to S-generalizations, then his STM framework has exactly the same "holism" problem which he claims plague content-based psychologies. There are different ways of showing this, but at the end they all come to the same thing. Let me begin with the obvious version.

S-generalizations are *low-level* generalizations. What makes them low level is the following fact. Subjects that are covered by such generalizations are also covered by L-generalizations if the subjects have certain *actual* B-states. For instance,¹⁶

IF S has the belief* that #for all x, if x is a cow, then x is an animal#,
and
S comes to believe* that #Samantha is a cow#,
THEN
S will typically come to have the belief* that #Samantha is an animal#.

What might license this inference* is, of course, the existence of high-level L-generalizations that Stich (*mutatis mutandis*) specifies among his examples:

- (5) For all subjects S, and all wffs A and B, if S has a B-state mapped to $A \rightarrow B$ and if S comes to have a B-state mapped to A, then S will come to have a B-state mapped to B. (1983: 155)

As we may recall, according to Stich the "holism" problem that plagues the content-based theories consists in the fact that the type identity of a particular belief (partly) depends on what other actual beliefs the subject has. Stich thinks that this fact is the source of the problem. In contrast, he claims, the NCA of typing particular B-states has no such commitment to there being any actual B-states surrounding a particular B-state in terms of which its type-identity is determined.

But, *if* every subject who is covered by S-generalizations is also covered by the relevant L-generalizations, in the way I've just indicated, then the STM-theorist *is* committed to their being actual B-states for determining the type-identity of particular B-states, and thus committed to construct syntactic theories only for those who more or less share their doxastic* background. In other words, in the STM paradigm, the "syntactic" type identity of sentence tokens, contrary to Stich's advertisement, *is* acutely sensitive to the *actual* particular B-states that surround them. This is a problem that is exactly parallel to what Stich calls the "holism" problem of belief individuation. And so, STM must incur all the parallel problems which Stich claims

seriously bother content psychologists: Sharing a particular B-state can only be a matter of degree, therefore, those that are doxastically* dissimilar to us cannot be covered by STM-theories. What are we to do with the children, exotics, cognitively handicapped, higher animals, etc.? Furthermore, unless Stich can come up with a principled distinction between those B-states that contribute to the syntactic type identity of a sentence token and those that don't, the vagueness that already exists in the conditions that type identify sentence tokens will be greatly aggravated. Again, we have exactly the parallel problem here. If Stich is right in his criticism of content-based theories regarding "holism" problem, it is false that STM theories are any superior in just that respect.

However, one might object: It is not necessary that for any subject who is covered by S-generalizations is also covered by the relevant L-generalizations in the way I have just indicated, and therefore, it is not necessary for an STM theory to be committed to there being *actual* B-states, which a subject must have, for the type individuation of sentence tokens. It may be that the S-generalization (i) above may be true of a subject even though she may not have any *actual* belief* that #all cows are animals#. In such a case, the syntactic type identity of a sentence token may be given in terms of such *dispositions* as the likes of (i) and (ii) specify without any recourse to high-level L-generalizations. How does this evade the problem? Well, let me show that it doesn't.

Although I don't have to make my point in the way I will do, I think it is important to cast the issue in terms of that perspective. But nothing important will hang on it. STM has usually been brought up as a de-intentionalized version of a language of thought story, or CTM. We have seen that STM is not in fact committed to there being (semantic-free) sentences literally realized in the brain. But it may be taken in this way, and this is the assumption we are now operating under.

Anyone who is sympathetic to the computational paradigm must keep in mind that CTM is a "rules and representations" framework: any relatively higher level mental processes consist in transformation of syntactically structured representations according to rules that are causally sensitive only to the formal properties of representations over which they are defined. In other words, the typical computational treatment of such inferences as expressed by (i) or (ii) will take the form of applying some relatively high level rule like Modus Ponens to *actually tokened* complex sentences.

Of course, this is only one possible implementation story that can be given for such generalizations like (i) or (ii) at the computational level. Another possibility is that the rules that govern the inference from #x is a cow# to #x is an animal# is rather more specific and low-level, rather like the syntactic analogues of Carnap's "meaning postulates" implemented as rules.¹⁷ But, either way, according to CTM, you need rules to manage inferential processes defined over data structures.¹⁸

My point is simply this. On a computational paradigm S-generalizations can be cashed out *either* by postulating high-level laws and actual beliefs* upon which they operate *or* by postulating the syntactic analogues of meaning postulates in the form of low-level rules. So, once it is obvious that an STM-theorist is committed to such S-generalizations, the STM-theorist is no longer in the position that Stich claims is free of the problems confronting the content psychologist. Let me illustrate.

As we may remember, Stich claims that unlike content psychologies, the STM framework is capable of covering Mrs. T's mental states in its generalizations. That is because, he says, according to STM, the type identity of the state underlying her utterance "McKinley was assassinated" does not depend on further actual doxastic* states she has. In fact, as the example is constructed, she has almost none. The type individuation of Mrs. T's state proceeds according to its inferential**potential*, not according to what it *actually* inferentially* interacts with. So far so good. For instance, all the L-generalizations that cover her state detail just this potential. But, of course, with only L-generalizations, the STM-theorist cannot individuate the state. It is obvious that in addition to L-generalizations, the STM-theorist needs some such S-generalizations as

- (iii) For all subjects S, if S comes to believe* that #someone is assassinated#, then S will typically come to believe* that #someone is dead#

In order to type individuate the state underlying Mrs. T's utterance of "McKinley was assassinated". But, of course, we see that it is precisely this kind of generalizations that become inapplicable to Mrs. T when we come to see that she ceases to have many relevant beliefs. This can easily be explained on the version of the computational story that derives the S-generalizations from actual beliefs* and high-level L-generalizations. But Stich would probably insist that this is the wrong version. Well, then, let us look at the other version where S-generalizations are implemented as specific "dedicated" rules rather like the syntactic analogues of "meaning postulates".

The question now is whether there are any such rules intact in Mrs. T's case. As we may remember, it becomes apparent *under questioning* that she does not "know" whether an assassinated person is dead, what dying is, who McKinley was, etc. When she is asked whether McKinley was dead, she answers "I don't know". What better evidence can there be for the fact that the above S-generalization is broken? In the case of S-generalizations, appeal to potential causal profile doesn't even begin to help since it is precisely this potential that is lost in her case. But then, if such generalizations do no longer cover the mental states of people like Mrs. T, of course, we can't tell the computational story along the lines we have been assuming given that the other version is out. But either way, the important point is that the S-generalizations do simply not hold in Mrs. T's case. If so, however, the STM theorist is in exactly the same boat as the content psychologist: there is simply no saying what "syntactic" state Mrs. T is in, since the STM-theorist is no longer able to type individuate her state.

The same is true, similarly, for people who are doxastically* dissimilar to us like children, exotics, cognitively handicapped, higher animals, etc. In so far as the S-generalizations that are true of them are not available or non-existent, there is no type individuation of their syntactic mental states, hence they are beyond the reach of STM-theories.

So here is the score. Contrary to Stich's claim and advertisement, because of their commitment to the NCA, STM theories are committed to the type individuation of particular B-states (depending on the computational story preferred) either according to what other actual B-states the subjects have, or according to what

S-generalizations are true of them. The first option makes STM equally sensitive to the *actual* doxastic background of subjects. The second option restricts the scope of STM-theories to those for whom S-generalizations exist, or are specifiable, thus, again, to those who are doxastically/dispositionally similar. But the consequences of both options are just the same for the prospects of STM if the prospects of content psychologies are as Stich claims them to be.

6. THE NCA AND THE TYPE INDIVIDUATION OF BRAIN SENTENCES

The introduction of S-generalizations generates a fatal dilemma for the STM-theorist: Either (1) the type-individuation of syntactic objects is possible but only for individual systems (or clusters of systems) given separately for each, or (2) the individuation cannot be carried out interpersonally, i.e., for sufficiently large populations. Both horns are equally destructive for the prospects of STM. I argue for this conclusion elsewhere extensively (Aydede, 2000). But the gist of the argument can be conveyed intuitively.

Remember that the theory **T** consists of generalizations of three sorts (I-T, T-T, and T-O). The T-T generalizations are divided into two: L- and S-generalizations. We have also seen that the heaviest burden in the individuation of particular syntactic objects is carried by the S-generalizations. Now if STM is to provide solid foundations for a semantic-free scientific psychology, then, minimally, **T** must consist of only those generalizations that satisfy the following constraints simultaneously.

- a) they together must secure a unique causal role for each distinct syntactic object (i.e. there must be sufficiently enough of them for securing the uniqueness),
- b) they must be interpersonally applicable (i.e., they must be true of a sufficiently large intentional population, if not all intentional organisms),
- c) they must be lawlike (minimally, they must go beyond being statistical summaries of what causes what).

Is there such a theory? If one reflects on the question for a moment, one can see that the answer is negative. Let's suppose that we want the theory to be true of what we might otherwise characterize as the common folk. (So **T** can be thought of as more or less the de-intentionalized and cleaned-up version of Folk Psychology.) This entails that there are S-generalizations in **T** that satisfy (a)–(c). But what could they be?

The most plausible candidates that come to mind are the syntactic parallels of content generalizations that detail “analytical connections” among concepts like:

- (iii) For all *S* and *x*, if *S* comes to believe* that #*x* is assassinated#, then, *ceteris paribus*, *S* will come to believe* that #*x* is dead#.
- (iv) For all *S* and *x*, if *S* comes to believe* that #*x* is a bachelor#, then, *ceteris paribus*, *S* will come to believe* that #*x* is unmarried#. Etc.

One immediate problem with this suggestion is that if the STM-theorist is to pick out these S-generalizations by appealing to analyticity, then he is being unfaithful to his

own program and tenets: the semantic notions are being used in the construction of **T** in crucial ways. Secondly, if Quine is right about there being no principled distinction between analytic and synthetic statements (as Stich himself thinks), then this strategy is unavailable to the STM-theorist.

But perhaps Quine is wrong and the first difficulty can be circumvented by giving a syntactic (or, at any rate, non-semantic) criterion to pick out such S-generalizations. This is another way of asking: what are the criteria for choosing the S-generalizations that are to go into **T**? Perhaps, these are given all by (a)–(c), or at least (b)–(c). Indeed, it may seem plausible to think that (iii) and (iv) are lawlike and interpersonally applicable, whereas the following are not:

(v) For all *S* and *x*, if *S* comes to believe* that #*x* is a tiger#, then, *ceteris paribus*, *S* will come to believe* that #*x* is dangerous#.

(vi) For all *S* and *x*, if *S* comes to believe* that #*x* is a bachelor#, then, *ceteris paribus*, *S* will come to believe* that #*x* is a neurotic#.

Although I am not quite sure what exactly being lawlike comes to, nevertheless it seems intuitively plausible to regard (v) as violating (c), even if it satisfies (b), which I doubt. (vi) may be true of a few eccentrics, but clearly violates both (b) and (c). Perhaps this kind of approach can be made to work. That is a lot ‘ifs’ and ‘perhaps’, I admit, but let’s be charitable and carry on.

But still, it is clear that on this strategy there won’t be *enough* generalizations to secure *unique* causal roles for each possible syntactic object that an STM-theorist would posit. We can see this if for a moment we drop the ban on talking about analytic generalizations and see (iii) and (iv) as such, i.e. if we take (b)–(c) as reconstructing analyticity in non-semantic terms. So supposing that S-generalizations that are to be put in **T** on this proposal will intuitively detail the analytic conceptual connections, we may ask: are there enough of these? Let’s first consider whether S- and L-generalizations all by themselves can secure uniqueness. It seems that they can’t. For supposing that they can is tantamount to assuming that each concept can be defined completely. But given the failure of philosophy to define any concept of any significance in the last two millennia, as Fodor once pointed out in his characteristic way, this is simply not true. Second, it seems intuitively clear that even if there are indeed analytic connections, they are very scarce. This is admitted even by the proponents of content functionalism like Block.¹⁹ The obvious remedy to this is to drop (c) and appeal to S-generalizations that detail “empirical” or “contingent” connections among concepts* while satisfying (b) nonetheless. So, with this move, generalizations like (v) are now in, but not (vi) and the like. This is already dangerous holistic terrain, but there is *still* no guarantee that **T** will secure a unique role for all syntactic objects. Indeed, what needs to be done is to poll all those beliefs of the form ‘(x)(Fx → Gx)’ that are more or less accepted by all those who share a “common psychology”. Indeed with this move comes the admission that STM-theories will only be possible for those who are sufficiently doxastically similar. Putting aside the serious worries about the arbitrariness and vagueness this would create and the impossibility of developing STM-theories for those who are doxastically dissimilar, the main question still lingers: are there sufficiently enough of these generalizations to secure uniqueness? The answer is by no means obvious.

But perhaps with the help of I-T and T-O generalizations this problem can be overcome. But this is surely an illusion. Take, for instance, the causal generalizations that are supposed to connect lawfully a set of proximal stimuli to, say, #Clinton is not faring well#, or any similarly specific sentence. Whatever the laws of psychophysics may tell us with respect to a very restricted range of psychophysically available properties, they will certainly be silent for the vast majority of symbol types figuring in full-blown propositional attitudes*. The problem partly stems from stimuli being proximal. There are certainly no scientifically well-delineated sets of proximal stimuli nomically correlated with the objects of beliefs*. This is to say that no such set could constitute a natural kind which would lawfully correlate with the objects of beliefs*. The other part of the problem is the holism involved in belief* fixation. Which proximal stimuli will cause which symbol(s) to be tokened in the belief*-box is determined by what other symbols actually happen to be there and by the overall internal organization of the belief*-box (simplicity, conservatism, etc.).

The history of behaviorism also provides an overwhelming inductive evidence that there are no such laws to be stated. No one has ever succeeded actually stating a single such law! Similarly for the supposed generalizations that would lawfully connect basic motor-gestures to particular symbol types in the belief*- and desire*-boxes. To be sure, behaviorists were after lawful stimuli/behavior connections, which is different. But the moral must be the same, since their failure primarily stemmed from an inability to find projectable predicates to apply to all and only those proximal stimuli under physical descriptions that lawfully govern a given piece of behavior. They assumed that such stimuli directly and lawfully control the relevant piece of behavior: they wanted to bypass mediating internal states. They failed primarily because of the holism problem again. Nothing changes, however, if you assume that it is particular beliefs*, rather than behavior, that are directly under the lawful control of proximal stimuli: the routes from stimuli are equally holistic in each case.

Perhaps I am laboring this point needlessly. It should be clear that there are no laws to be stated with respect to proximal inputs/outputs for the full range of particular symbol types deployed in central cognitive processing as direct objects of propositional attitudes*.²⁰ And even if there may happen to be some, they will be so few and fragile that they will be of very little help in type individuating all the symbol types we may need in psychological explanations.

But if there is no unique role for each distinct syntactic object, there is no individuation of them suitable for the purposes of a common psychology. This is the second horn of the dilemma I stated above.

Given all this, an STM-theorist, in order to secure uniqueness and to meet the charge of arbitrariness involved in selecting the S-generalizations, may be forced to advert increasingly more to the kind of S-generalizations whose scopes are increasingly narrower, and at the limit, true of only single persons. T could then be stated for each individual with a certain distinctive psychology. But who would be so eager for the prospects of scientific STM-theories, if they won't be interpersonally applicable? This was the first horn.

Either way the NCA-cum-STM is in serious trouble. I am very doubtful whether it can ever be saved, and I am pretty sure it's not worth the try.

It is important to note that the requirements (b) and (c) pull in opposite directions. To the extent to which an STM-theorist can furnish interpersonally applicable generalizations and secure a unique role for each symbol type, to that extent he goes against the requirement that the generalizations be lawlike. And to the extent to which he can give lawlike generalizations and secure unique roles, to that extent he violates the condition that they be interpersonally applicable. I don't think that there is an optimum point in the continuum between these extremes such that you can meet both of these requirements and secure a unique functional role for each possible symbol type.²¹

Let me indicate two more problems for the kind of type individuation of mental sentences Stich envisages for STM. As remarked, Stich proposes a cluster view of identifying neural states as particular sentence tokens: "to count as a token of a sentence type, a neurological state must satisfy some substantial number of the cluster of generalizations included in a theory, without specifying any particular generalizations that must be satisfied, nor exactly how many must be satisfied" (1983: 152). He admits that this introduces vagueness into the identity conditions of mental sentences. However, the problem this may cause is more than just introducing vagueness. It risks downright misclassification. Consider again S-generalizations, for, in a certain sense, they are expected to do the heaviest work in the individuation of sentences on the STM framework. The problem is that there may be two sentence tokens satisfying almost the same generalizations but nevertheless differing in type because they satisfy a few different "essential" generalizations. Consider the token belief* that #...gay...# and the token belief* that #...lesbian...#, it is likely that they have very similar causal roles. What may be distinguishing them are just a few (counterfactual supporting) S-generalizations such as 'B*(#gay#) → B*(#male homosexual#)' and 'B*(#lesbian#) → B*(#female homosexual#)'. What reasons could Stich give us that such cases are not seriously troublesome or do not really arise? I can see none.

Another problem is one that Stich himself raised against content functionalism. On Stich's own admission, given two subjects with the same B- and D-states, the potential as well as actual causal patterns (concerning especially the ones captured by L-generalizations) that their B- and D-states will exhibit are very likely to differ. This is the problem parallel to the one that the content functionalism faces: the kinds and the degree of complexity of inferences that people can draw (i.e., their logical/analytical acumen) vary greatly from person to person. If any attempt to incorporate these different causal patterns into a functionalist theory in a principled way will be, as Stich says, "ad hoc and implausible", how could Stich think that an STM-theorist's parallel attempts will not similarly be *ad hoc* and implausible? Notice that insisting here that B-states are not beliefs cannot even begin to help: the explanation of a certain kind of mental activity on the basis of purely syntactic transformations of some complex abstract objects mapped onto B- and D-states is exactly what STM theories are supposed to be good at.

This completes my second argument against STM.

7. WHY A PURELY SYNTACTIC PSYCHOLOGY CANNOT GET OFF THE GROUND

Throughout Stich's 1983 book, there are various passages in which Stich seems to argue that an STM-theorist had better refrain altogether from using intentional notions even in the theory construction stage. Here is a typical one: "cognitive psychologists can and do develop the theory of mental processes without attending to the semanticity of formulae in the mental code" (1983: 193). In fact, Stich's discussion of what he calls the Weak Representational Theory of Mind (RTM) is an attempt to show that the assumption that the formulae have semantic content is frivolous at any stage of theory development.

Many people seem to think that functionalism in scientific psychology can be carried out without ever raising any semantic worries. In this section, I will argue that this in fact can't be done. In particular, I will show that the construction of an STM-style theory cannot be carried out without using intentional notions. This problem is one that seems to belong to "the context of discovery", but nevertheless it will be instructive to see why an STM-theorist is committed to using intentional notions in at least theory construction stage. I already detailed the reasons why STM is seriously problematic otherwise.

In fact, it is for a very simple reason that within a strictly STM paradigm theory construction cannot get off the ground without using any intentional idioms. STM is a purely functional theory. As such, all the theoretical predicates that denote functionally defined particular brain state types depend for their reference on the entire theory being *in situ*. In other words, within the STM paradigm, the only legitimate way to refer to the nodes of the causal network of brain states is by way of theoretical terms whose applicability entirely depends on the theorist's having almost the whole functional theory first. That is his point when Stich insists that the type identity of a sentence token (a brain state token) entirely depends which and how many generalizations cover it:

It is only against the background of a systematic mapping of state types to sentence types that any given state token counts as a token of a particular sentence type... No one neurological state can count as a token of a sentence type unless many neurological states count as tokens of many different sentence types. But this holism... is quite distinct from the holism imposed on the folk psychological notion of belief by the embedded appeal to ideological similarity. For the status of a state as a token of a sentence does not depend on what other cognitive states a subject currently happens to be in. It depends only on the causal interactions that the state would exhibit with stimuli, with behavior, and with other states. (1983: 153)

But there is no way to start theory construction without having an initial and independent way of referring to the nodes of the causal network of the brain states about which nothing is known in the initial stages. In other words, when there is no theory yet, the prospective "theoretical terms" can't refer. This presents a dilemma. On the one hand, the STM-theorist wants to theorize about the functional organization of particular brain state types. For this she must have an independent way of referring to them, independent of a more or less completed theory. On the other hand, as far as she refrains from using an intentional scheme, she can't even guess what she is talking about when she uses terms like 'the B-state mapped to *Fa*'. That is

because the theorist has no independent way of identifying the nodes of the network of brain state types. This network is completely unknown.

The problem stems from the STM paradigm itself. Notice that if there were an independent way of picking out the nodes (brain state types) in the causal network that does not presuppose a more or less complete specification of which nodes are connected to which others and how (i.e. their potential functional roles determined by the generalizations of the theory), then we would use this scheme in our way of saying what generalizations there are, i.e. in our theory construction. This is exactly what Brian Loar does (1982a) in presenting his semi-broad content functionalism: he uses propositions to pick out those brain states and state whatever generalizations there are that need to be stated. Once he gets the generalizations into place, he gets rid of propositions in favor of syntactic objects. Then, of course, he is in a position to specify, theoretically at least, all the functional roles there are without using any semantic terms. Once he does that the result is almost an STM theory very much like what Stich envisions.

So it should be obvious that the way out of this dilemma can only be semantic, not syntactic. The upshot is that pure “syntactic” (psycho)-functionalism in scientific theory construction à la Stich cannot be carried out without assuming the truth of content (semantic) functionalism (à la Loar). They stand or fall together, which is not to say that narrow content functionalism has got to be true (see above).

If what I have said so far is right, the lesson to be drawn is that syntactic functionalism is not an option in psychological *theory construction* somehow at the discretion of the psychologist. When we reflect upon the historical rise of functionalism in the philosophy of mind, that this is so should be obvious. Functionalism was developed as a response to the inherent difficulties in behaviorism and (type-type) identity theories. It was conceived as a metaphysical theory saying *what* mental states *are*. Functionalism identified mental states with functional states. But that was not enough. Functionalism had to be able to provide the identity conditions of mental state types. This required providing identity conditions for functional roles. Functionalists had to be able to say what functional roles uniquely define what types of mental states. But this required having a theory first. Some versions of functionalism took this theory to be folk psychology made explicit with all the intentional/mental terms employed as theoretical terms. Then, Ramsefying this theory was the major step in explicitly getting the identity conditions for functional roles. Similarly with psychofunctionalism: the theory to be Ramsefied was conceived to be a theory to be developed by scientific psychology. The underlying idea was the same. Once such a theory was at our disposal with all its intentional/mental terms employed as theoretical terms, we could explicitly get the identity conditions for the functional roles by Ramsefying it. In all this, the construction of the theory to be Ramsefied was conceived along with using all the intentional vocabulary available to the theorists. And that was OK, because functionalism was competing against dualism, eliminativism and reductionism (type identity theory). That is the reason why functionalism at its core is essentially an intentional realist theory. But Stich’s STM tries to reverse the situation, it wants to develop functional theories without ever using intentional terms; in this, however, Stich is putting the cart before the horse. As we have seen,

this turns out to be impossible, because the remaining vocabulary to be used in theory construction cannot do the required job. In a sense, in fact, Stich cuts the branch he is sitting on.

Admittedly, my point in this section is one that belongs to the context of discovery. It might be claimed that as such it is not that important: what matters is whether the ultimate STM-style theory, when completed, is committed to any intentional scheme. The STM theorist might use any tools (intentional or otherwise) that would help in getting the theory, i.e. in the context of discovery. But once the theory is completed and successful, it should not matter how it was gotten in the first place. For instance, as long as it belongs to a discovery stage, an STM-theorist might use a procedure like Loar's. It is the form of the ultimate completed theory that counts.

Well, I have two points to make against this. First, given Stich's criticism of content-based psychologies, it should be obvious that the brain states initially typed according to an intentional scheme will exhibit all the vagueness, context-sensitivity, and parochialism that Stich claims will pertain to a semantic taxonomy. So he can't avail himself to the SA of typing even in the context of discovery. Second, it is simply absurd to assume that a taxonomic scheme will be semantic-free if at the end it is essentially obtained by a SA and then gotten rid of à la Loar. The ultimate theory, if really successful, would be nothing but a (partial) scheme for a naturalized psychosemantics (e.g., in the tradition of two-factor semantic theories).²²

8. IF COGNITION IS COMPUTATIONAL, HOW CAN PSYCHOLOGICAL LAWS BE INTENTIONAL?

This is what Fodor called the "Eponymous Question" in his (1994). As I said, this question has in fact been around, constantly popping up here and there, and haunting people working in the field, for more almost two decades now, mostly thanks to Stich and Fodor.²³

This question is also related to certain puzzles computationalism has created *vis-à-vis* mental causation. According to the computational picture of mind (CRT, LOTH), mental processes are defined over mental symbols physically realized in the brain. But computationalism says that for these mental processes to qualify as computational, it is the non-semantic, in particular syntactic, properties of symbols that the processes must be causally sensitive to. In fact, given a physicalist framework, it is not even clear what it would be like for mental processes to be causally sensitive to the semantic properties of symbols, which are relational, i.e., hold between the symbol (or the organism) and environment. Given the locality of causation, thought processes can be causally sensitive to only syntactic (at any rate, non-semantic) properties of symbols that are implemented neurally. If so, even though mental symbols are causally efficacious in reasoning and causation of behavior, it is in virtue of having certain syntactic properties, but not in virtue of having semantic properties, that they are so. Thus, as far as the science of psychology is in the business of causal explanation, the relevant properties of mental states in virtue of which they are covered by causal psychological laws are all non-semantic, or so it seems on the face of it. This is another way of seeing Stich's motivation in arguing

against content-based psychologies and promoting his STM over them. As we have seen, Stich calls the Narrow Functional Account of typing symbol tokens “syntactic” typing, presumably meaning just non-semantic and non-physical. And this sort of typing, on his view, is what the STM (or CTM for that matter) is committed to. He then claims that STM/CTM is all a scientific psychology needs; hence, *pace* Fodor, there is no need to appeal to semantic/intentional properties of syntactically structured brain symbols in stating the laws of psychology. He accuses Fodor of having it both ways.²⁴ We are now in a position to see how it is possible to have it both ways, i.e., to see what the answer is to the Eponymous Question.

Let us suppose that computational psychology is correct. Any scientific computational psychology needs to postulate states in terms of which it can explain (and predict) behavior (construed broadly—bodily, verbal, mental behavior). This seems to call for covering laws or generalizations that subsume those states under an appropriate description. This is at least the assumption shared by all parties in the debate, and I will not challenge it. This means that these states, under the relevant description, are projectable, i.e. natural kinds from the perspective of the theory. As such they must have identity conditions. Computational psychology characterizes these states as symbol tokens realized in the heads of cognitive organisms. *Qua* symbols they have both syntactic and semantic properties. OK then, how are we to type them to suit the psychological laws covering their tokens? We have seen that they cannot be typed, in the required sense, by their narrow functional properties: NFA is hopeless. The Physical Account (PA) of typing them is hopeless too. Stich and almost everybody in the field agrees. The PA seems to commit one to a very strong version of type-type identity theory for propositional attitudes with specific content (like the belief *that snow is white*) cast across people. In this form, the PA has no defenders as far as I can tell. Our only other option, then, the Semantic Account, is in fact mandatory if psychological processes are to be computational. In other words, *if Stich’s original question, i.e. the question of what it is for two symbol tokens of Mentalese in different heads to be of the same type, has an answer, it must be some version of the SA.*²⁵ It must be on the basis of their semantic properties we type symbol tokens across systems.

Therefore, I conclude that computational psychology (CTM, for that matter) itself is essentially committed to semantic type individuation of symbol tokens across systems. And it is across systems that a scientific psychology casts its laws over. Hence, the necessity for intentional psychology whose laws advert to the semantic properties of representations. If mental representations can be typed interpersonally *only* on the basis of their semantic properties, CTM cannot be an alternative to replace intentional psychology. Hence the answer to Fodor’s Eponymous Question.²⁶

*University of Florida, Department of Philosophy,
330 Griffin-Floyd Hall, P.O. Box 118545
Gainesville, FL 32611-8545, maydede@ufl.edu*

9. NOTES

¹ In his (1991) Stich argues against Fodor that narrow content taxonomies will differ from the narrow causal taxonomies, which he calls “fat syntax” taxonomies. The problem, according to Stich, stems directly from the SA, narrow or wide.

² Generalizations detailing the causal relations between proximal input events and T-states (thought-like states), among T-states, and between T-states and proximal output events. See Devitt (1990).

³ See also (1983: 78-9), where Stich writes: “mental sentence theorists typically leave the notion of an internalized sentence token as little more than a metaphor. And it may well turn out that when the metaphor has been unpacked, it claims no more than that beliefs are relations to complex internal states whose components can occur as parts of other beliefs.” Here, it is not clear what the contrast Stich is trying to convey is supposed to be. Of course, this is what is literally intended by a LOT theorist and more: under a suitable mapping function internal states are literally interpretable as constituting a symbol *system* with a combinatorial syntax and semantics so that the processes defined over these are sensitive only to their formal/syntactic properties. That is the essence of LOTH and computationalism, see my (1997a, and 1998).

⁴ Indeed, this was the very point of Brian Loar in his polemical article (1982b) written against Fodor’s LOTH. He says that from a philosophical point of view his non-committal content functionalism is weaker than the LOT version of it and thereby should be preferred. He does not reject the LOTH, but he claims that its motivation cannot be due to its having more explanatory and predictive power. For, with respect to these, his pure functionalism is equally good. Loar views the LOTH as a scientific hypothesis, and as such he leaves its truth as an open question.

⁵ Here is a passage (among many others) from an early piece by Fodor: “For purposes of (narrow) ascriptions of content, the essential properties of a mental state are its functional properties (the ones it has in virtue of its causal role vis-à-vis behavior and other mental states). Since, as a matter of fact, there is a certain set of functional properties that mental states normally have when they are brought about by, for example, visual encounters with banks, we can specify a set of functionally equivalent experiences by reference to some such contingent fact as that they are like seeing a bank. Something could have these functional properties (hence this content) without being caused by a bank (cf. bank hallucinations). To this extent, the description ‘prescinds from semantic relations’ . . . But that’s alright since [methodological solipsism] isn’t the claim that we must pretend, in describing mental states, that the world does not exist; what it claims is that the properties of mental states which are essential for determining how mental operations apply to them are their functional (hence formally specifiable, hence nonsemantic) properties. A mental operation may apply to a mental state in virtue of its being like seeing a bank but not in virtue of its being seeing a bank” (1980b: 102). Here Fodor is taking formal/syntactic typing, narrow functional/causal typing, and narrow semantic typing to be all virtually equivalent.

⁶ He makes the same point in his (1983): “The core idea of the STM—the idea that makes it syntactic—is that generalizations detailing causal relations among the hypothesized neurological states are to be specified indirectly via the formal relations among the syntactic objects to which the neurological state types are mapped” (151).

⁷ For an extensive elaboration of how computation is to be understood in the role it plays in LOTH and modern cognitive psychology, see my (1997a) and (1998). In the former work, I also give an analysis of the notion of syntax as it is deployed by LOTH.

⁸ Note that Stich’s claim is stronger than merely saying that lexically different sentence tokens have the syntactic property of just *being different*. His claim is that they belong to different *specific* syntactic categories.

⁹ See my (1997a). Stich, in his (1991) article, calls the type identity of sentences that gets fixed on the basis of their narrow causal profile their “fat syntactic” identity. This is supposed to be contrasted by their “skinny syntax”. The latter is to be fixed by the T-T generalizations alone: no causal relations to proximal stimuli and behavioral events can be used in the individuation of sentences. Stich insists that it is the fat syntactic type identity that would do the work for STM-style theories. As I said, I will argue that the NCA cannot fix the type identity of mental sentence tokens whether or not what gets fixed is their (fat) “syntactic” type. Devitt (1990) has argued that even if their type identity can be so fixed, what gets fixed

would be their narrow semantics not their syntax. Devitt's discussion also contains a very helpful criticism of Stich's notion of syntax.

¹⁰ For some others, see Stich (1983: 53-60 and 137-44).

¹¹ For a similar and more striking discussion of the commitments of the NCA of typing where Stich goes through a similar example, see (1983: 53-4). The generalizations (4)-(6) Stich mentions here are all what I will call below, L-generalizations. They advert to the logical form of the sentences, hence are blind to the non-logical primitives the theorist postulates.

¹² It is ironical, and in fact a bit puzzling, that Stich himself makes the parallel point in criticizing content functionalism: "There are literally infinitely many inferential paths leading both to and from every belief" (1983: 24). His point is that since every particular belief is potentially connected to every other, the generalizations detailing this potential will not be able to define beliefs with particular content.

¹³ In fact, the situation is even more complicated given that there is already a build-in vagueness in the "syntactic" individuation of particular B-states: for Stich a sentence token to count as of a particular type, it must satisfy a substantial number of generalizations. Stich seems to propose a cluster theory of type individuating sentence tokens, and this, as Stich himself admits, brings with itself a certain amount of vagueness. See below.

¹⁴ A parallel distinction is drawn by Loar (1982a) between "L-constraints" and "M-constraints".

¹⁵ These are supposed to be "ceteris paribus" generalizations. I'll generally ignore this in what follows.

¹⁶ In what follows, in order to avoid long and cumbersome ways of expressing the same thing, I will simply adopt the following convention: I will mark an intentional expression with a '*' to express whatever its syntactic parallel may be. Also, I will hedge a content sentence with '#s in order to indicate that I intend its syntactic parallel, i.e., whatever syntactic object or sentence might go in its stead.

¹⁷ There are many versions of this approach in AI. Frames, scripts, etc. are all versions of the same underlying idea. The tradition of "semantic representation" in linguistics again relies on the idea that lexical items can be semantically decomposed.

¹⁸ Rules may or may not be explicitly represented. CTM is neutral on this. However, given that the rules that implement S-generalizations reflect important pieces of "semantic knowledge" they are unlikely to be hard-wired.

¹⁹ Block (1993: 3-4) writes: "Fodor and Lepore seem to assume . . . that . . . the inferential role theorist has the option of appealing to analyticity as a way of discriminating the inferential liaisons that are in inferential roles from those that are out. But if we stick to traditional ideas about the extension of 'analytic', there aren't *enough* analyticities. Consider the putative analytic truths involving 'cat'—'Cats are animals', 'Cats are living beings', 'Cats are grown up kittens', etc. The problem is that abstracting from the words 'cat', 'kitten', etc., appearing in these sentences, there is nothing here to distinguish 'cat' from 'dog'. Corresponding to 'Cats are grown up kittens', we have 'Dogs are grown up puppies'. Sure, 'nothing is both a cat and a dog' can be used, but so can 'nothing is both a dog and a cat'. Even if 'Cats are feline', and 'Dogs are canine' are analytic, this is of no help without other analytic truths that distinguish 'feline' and 'canine' . . .". See also his (1986: 628-9). Cf. Loar (1982a: 81ff).

²⁰ It is very curious that more or less the same criticism is given by Stich himself for the claim made by content functionalists that there are such generalizations: "[t]here is generally no characteristic environmental stimulus which typically causes a belief. There is no bit of sensory stimulation which typically causes, say, the belief that the economy is in bad shape, or the belief that Mozart was a freemason . . . Nor do most beliefs have typical behavioral effects. My belief that Ouagadougou is the capital of Upper Volta does not cause me to do much of anything" (1983: 24). Later on, he argues (1983: 180-1), on familiar grounds, that there can be no principled distinction between beliefs whose content is "observational" and those whose content is "theoretical". So, according to Stich, even for allegedly "observational" beliefs there seems not to be any particular set of stimuli nomologically connected to them.

²¹ This is in fact more or less acknowledged by leading functional role semanticists like Block. Hence the destructive holism to which they are said to be committed.

²² On this last point, see also Higginbotham (1988) and Crane (1990).

²³ See, among many others, Stich (1983, 1991), Field (1978), Schiffer (1987), Fodor (1980a, 1989, 1994), Devitt (1991), Jacob (1997) who take issue with the EQ one way or other.

²⁴ Stich (1983). See also his (1991). Devitt (1991) joins Stich in accusing Fodor of trying to have it both ways but only with respect to processes governing thoughts without I/Os.

²⁵ It is of course possible that Stich's question doesn't have an answer. I surely haven't argued here independently for the truth of the SA. In other words, if Stich is right about the fate of the SA, and if I am right about the fate of the PA and NFA, then scientific cognitive psychology as we know it today is impossible. I can't take this option seriously, in particular I can't take seriously a priori arguments against the cogency of the foundations of what appears to be an enormously successful and fruitful scientific approach to cognition. Cognitive psychology seems to be into intentional talk up to its neck. I take it that there is an enormous *prima facie* evidence for the truth of the intentional assumptions of present day cognitive psychology. I take this to be the best argument for the SA albeit a non-demonstrative one. I left Stich's positive arguments against the SA aside in the beginning of the paper. What needs to be done, of course, is to address Stich's criticisms in order to begin to give an independent argument for the SA.

²⁶ There are, to be sure, problems with any version of the SA, as is well known. Suppose that the SA is broad as in (late) Fodor. Then we have problems with Frege cases as well as Twin-Earth cases. A narrow SA would be equally problematic, as we have seen, if it relies on narrow functional roles of vehicles as their narrow semantic content. On the other hand, a Fodor-style notion of narrow content as mapping from context to broad content can perhaps handle Twin-Earth cases at best, but not Frege cases (see my 1997b, 2000). But being problematic is one thing, being wrong is another: I think that a SA that works can after all be salvaged in the face of apparent difficulties.

10. REFERENCES

- Aydede, Murat (1997a), "Language of Thought: The Connectionist Contribution," *Minds and Machines*, Vol. 7, No. 1, pp. 1-45.
- Aydede, Murat (1997b). "Has Fodor Really Changed His Mind on Narrow Content?," *Mind and Language*, Vol. 12, No. 3/4, pp. 422-58.
- Aydede, Murat (1998). "Language of Thought Hypothesis" in *Stanford Encyclopedia of Philosophy*, edited by Edward Zalta, Stanford, CA: CSLI E-Publications.
- Aydede, Murat (2000). "On the Type/Token Relation of Mental Representations," *Facta Philosophica: International Journal of Contemporary Philosophy*, Vol. 2, No. 1, pp. 23-49.
- Block, Ned (1986). "Advertisement for a Semantics for Psychology" in *Studies in the Philosophy of Mind: Midwest Studies in Philosophy*, Vol. 10, French, P., T. Euhling and H. Wettstein (Eds.), Minneapolis: University of Minnesota Press, 1986.
- Block, Ned (1993). "Holism, Hyper-analyticity and Hyper-compositionality," *Mind and Language*, Vol. 8, No. 1, pp. 1-26.
- Crane, Tim (1990). "The Language of Thought: No Syntax Without Semantics," *Mind and Language*, Vol. 5, No. 3, pp. 187-212.
- Devitt, Michael (1990). "A Narrow Representational Theory of the Mind," in *Mind and Cognition*, W. G. Lycan (Ed.), Oxford, UK: Basil Blackwell, 1990.
- Devitt, Michael (1991). "Why Fodor Can't Have It Both Ways" in *Meaning in Mind: Fodor and his Critics*, B. Loewer and G. Rey (Eds.), Oxford, UK: Basil Blackwell, 1991.
- Devitt, Michael (1996). *Coming to Our Senses: A Program for Semantic Localism*, Cambridge, UK: Cambridge University Press.
- Field, H. (1978). "Mental Representation," *Erkenntnis* 13, 1: 9-61.
- Fodor, Jerry A. (1980a). "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology" in *RePresentations: Philosophical Essays on the Foundations of Cognitive Science*, J. Fodor, Cambridge, Massachusetts: MIT Press, 1981. (Originally appeared in *Behavioral and Brain Sciences* 3, 1, 1980.)
- Fodor, Jerry A. (1980b). "Methodological Solipsism: Replies to Commentators", *Behavioral and Brain Sciences* 3, pp. 99-109.

- Fodor, Jerry A. (1989). "Substitution Arguments and the Individuation of Belief" in *A Theory of Content and Other Essays*, J. Fodor, Cambridge, Massachusetts: MIT Press, 1990. (Originally appeared in *Method, Reason and Language*, G. Boolos (Ed.), Cambridge, UK: Cambridge University Press, 1989.)
- Fodor, Jerry A. (1994). *The Elm and the Expert*, Cambridge, Massachusetts: MIT Press.
- Higginbotham, J. (1988). "Is Semantics Necessary?", *Proceedings of the Aristotelian Society* 88, pp. 129-41.
- Jacob, Pierre (1997). *What Minds Can Do: Intentionality in a Non-Intentional World*, Cambridge, UK: Cambridge University Press.
- Loar, Brian F. (1982a) *Mind and Meaning*, Cambridge, UK: Cambridge University Press, 1982.
- Loar, Brian F. (1982b). "Must Beliefs Be Sentences?" in *Proceedings of the Philosophy of Science Association for 1982*, Asquith, P. and T. Nickles (Eds.), East Lansing, Michigan, 1983.
- Schiffer, Stephen (1987). *Remnants of Meaning*, Cambridge, Massachusetts: MIT Press.
- Stich, Stephen P. (1978). "Autonomous Psychology and the Belief-Desire Thesis" in *Mind and Cognition*, W. G. Lycan (Ed.), Oxford, UK: Basil Blackwell, 1990. (Originally appeared in *The Monist* 61, pp. 573-591, 1978.)
- Stich, Stephen P. (1983). *From Folk Psychology to Cognitive Science: The Case Against Belief*, Cambridge, Massachusetts: MIT Press.
- Stich, Stephen P. (1991). "Narrow Content Meets Fat Syntax" in *Meaning in Mind: Fodor and his Critics*, B. Loewer and G. Rey (Eds.), Oxford, UK: Basil Blackwell, 1991.