

INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each original is also photographed in one exposure and is included in reduced form at the back of the book.

Photographs included in the original manuscript have been reproduced xerographically in this copy. Higher quality 6" x 9" black and white photographic prints are available for any photographs or illustrations appearing in this copy for an additional charge. Contact UMI directly to order.

U·M·I

University Microfilms International
A Bell & Howell Information Company
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
313/761-4700 800/521-0600

Order Number 9136195

**Mental state content and the ontological autonomy of
psychology**

Sullivan, Sonja Rebecca, Ph.D.

Yale University, 1991

Copyright ©1991 by Sullivan, Sonja Rebecca. All rights reserved.

U·M·I
300 N. Zeeb Rd.
Ann Arbor, MI 48106

Mental State Content and the Ontological Autonomy of Psychology

A Dissertation

Presented to the Faculty of the Graduate School

of

Yale University

in Candidacy for the Degree of

Doctor of Philosophy

by

Sonja Rebecca Sullivan

May 1991

© Copyright by Sonja Rebecca Sullivan 1991

ALL RIGHTS RESERVED

ABSTRACT

Mental State Content and the Ontological Autonomy of Psychology

Sonja Rebecca Sullivan

Yale University

May 1991

Much of the current work in philosophy of mind focuses on problems of intentionality. In particular, owing in part to recent work in artificial intelligence and cognitive science, one of the central topics of debate is whether or not it is possible to construct a rigorous scientific psychology that embodies a Realist account of intentional states, one that endorses the commonsense ontology of causally efficacious beliefs and desires. If a scientific psychology is to adopt such a Realist view of intentional states, an account of how the content of intentional states, which are routinely appealed to in commonsense explanations of behavior, is relevant to their causal powers is required. Consequently, the issue of how to determine and individuate the intentional content of mental states is central to the debate over whether such a Realist scientific psychology is possible.

The dissertation begins with a critical discussion of the views of Jerry Fodor, Tyler Burge, and Fred Dretske concerning the individuation of the intentional content of mental states. The positions they endorse vary widely, and I argue that each is unsatisfactory on independent grounds. Nonetheless, I argue that they share a fundamental difficulty: I argue that each of their theories presupposes some individuation or taxonomy of the world that is unwarranted given the purposes of

psychology. Since intentional content ultimately derives from the world, any presupposition concerning how the world should be individuated limits a priori how psychology can individuate mental state content. I claim that individuating mental state content and individuating the world are two sides of the same coin. One cannot determine how to individuate mental state content independent of determining how to individuate the world for the purposes of psychology. Thus, I maintain that psychology should be granted the same sort of autonomy in specifying the natural kinds it recognizes that we grant to physics, chemistry, or economics. I argue that once we grant psychology this ontological autonomy, it will be clear that the predictive and explanatory needs of a scientific psychology will require a radically idiosyncratic taxonomy of both mental state content -- and the world.

PREFACE

I would, first and foremost, like to record my debt to Ivan Fox, my advisor. This thesis is, in large part, the result of the many afternoons spent in discussion with him, not only of the various drafts I produced along the way, but of the literature in the field as well. To him I extend my gratitude, for his time, his interest, and his counsel. At this time, I would also like to thank Ruth Marcus for her participation in the early stages of this thesis.

CONTENTS

Preface.....	iii
List of Figures.....	vi
Chapter 1: Introduction.....	1
Chapter 2: Jerry Fodor's Notion of Narrow Content.....	14
I. The Commonsense "Folk" Psychology View of Beliefs.....	16
II. Content, Causation, and Fodor's Representational Theory of Mind.....	21
III. The Individuation of Mental State Content is Influenced by both Scientific and Folk Considerations.....	29
IV. A Scientific Psychology Needs a More Robust Notion of Content than Fodor's RTM Provides.....	44
Chapter 3: The "Communitarian" Notion of Content of Tyler Burge.....	66
I. Burge's Arguments Against an Individualistic Notion of Content.....	69
II. Burge's Arguments for a Social or "Communitarian" Notion of Mental State Content.....	81
III. The Substitution of Synonyms in Belief Contexts is Not Entailed by Burge's Position.....	94
IV. Burge's "Communitarian" Notion of Content is Plausible Neither as an Explication of Our Ordinary Mentalistic Notion Nor as a Scientific Construct.....	103
Chapter 4: Fred Dretske's Information Theoretic Account of Content..	123
I. Communication Theory and the Conditions for Information Transmission.....	130

II. Dretske's Account of Content is Based on Information Theory and a Notion of the Function of a State.....	153
III. Information Theory Leads to an Individualistic Notion of Content, Dretske's Comments Notwithstanding.....	169
Chapter 5: The Ontological Autonomy of Psychology.....	190
I. Burge Makes Linguistics the Arbiter of Ontology for Psychology.....	193
II. Fodor is Led Astray, and Comes to Reject the Ontological Autonomy of Psychology.....	206
III. Dretske's Concern for Knowledge Leads Him to Endorse the Ontological Subordination of Psychology.....	216
IV. Psychology Must Be Granted Autonomy in the Individuation of Both the World and Mental State Content.....	226
V. Psychology Will Employ a Radically Individualistic Notion of Content.....	231
Endnotes.....	251
Bibliography.....	261

LIST OF FIGURES

Figure 1: A Schematic View of the Relation Between a Psychological Taxonomy of the World for a Hypothetical Individual and the Taxonomies of the World of Various Other Fields.....236

CHAPTER 1
INTRODUCTION

In recent years, at least partly owing to the emergence of the fields of artificial intelligence and cognitive science, there has been a great deal of interest in the possibility of constructing an account of intentional states and their role in the etiology of behavior based on a computational, or more broadly, functionalist, theory of the mind. The suggestion is that a theory that views the mind as a computational system operating on representations will be able to formulate its generalizations at a level that seems natural from the point of psychology. Such a psychology will be able to construct generalizations that apply to all systems in which it seems that the behavior of the system is governed by internal states in accordance with what those states represent or indicate about the surrounding environment. Specifically, it is maintained that a psychology that adopts a computational stance will be able to vindicate the commonsense "folk psychology" practice of predicting and explaining behavior on the basis of beliefs and desires, or more generally, intentional states.

According to the commonsense folk psychology view, people do the things they do and say the things they say because of the particular beliefs and desires they have. Beliefs, desires, hopes, fears, etc. are genuine states of individuals that cause individuals to behave the way they do. But intentional states are not only causally efficacious states. Commonsense belief/desire psychology treats them as states that have content as well. Beliefs, for example, not only can cause

behavior, but are semantically evaluable as well. Beliefs are about this, or that, or the other thing, and the facts about this, that, or the other thing determine whether the belief is true or false. A belief that-P is true or false in virtue of whether or not P obtains. What one believes, for example, that George Bush was elected President in 1988, is what is often referred to as the "content" of the belief. The content of a belief, for our purposes, is what is believed, it is the state of affairs asserted (or denied) to obtain, or a function from possible worlds to states of affairs that make it true. Thus, a belief is made true or false in virtue of the relation between the content of the belief and the facts about the world in which it is held.

In the conceptual framework of commonsense belief/desire psychology, mental states are not states that merely have both causal powers and semantic properties. Rather, it is a central tenet of the folk psychology model that the pattern of causal interactions among beliefs and desires at least reflect the relations among their semantic contents. All of the generalizations that commonsense psychology makes about the connection between behavior and mental states are made on the basis of the semantic content of mental states. Commonsense psychology characterizes the causal powers of beliefs in terms of their content: Mental states with the same content have the same causal powers, and mental states with different contents have different causal powers.

The computational theory of the mind claims that the mind is a type of computational device. Mental processes are the manipulations and transformations of objects with representational content on the basis of certain of their physical properties, what Fodor has referred to as their "formal" properties. (A formal property of a state or object is

simply a physical property of the state or object that can influence the normal computational functioning of the system of which it is a part.) Mental states, the symbolic objects over which mental processes are defined, are claimed to have both semantic properties and causal powers, but according to the computational model, mental states have their causal powers in virtue of their formal properties. Although the semantic properties of mental states are, themselves, not involved in actual causal processes, the computational model claims that the semantic properties of mental states are tied to or linked, in some fashion, to the formal properties of mental states so that one can, by performing a series of formal manipulations, effect a predictable transformation of the semantic properties of the state. What makes the computational model of the mind initially plausible is that computers and electronic calculators are systems in which objects that have semantic properties interact causally with one another on the basis of their formal properties, but in ways that respect their semantic relations. The computational model of human cognition seeks to provide an account of how, within a physicalist framework, causal interactions can respect and effect semantic relations.

In order for a computational psychology, which accepts that the formal properties of mental states are responsible for their causal powers, to provide an adequate analysis of commonsense belief/desire explanations and predictions of behavior, which treat the semantic properties of mental states as preeminent in the determination of causal powers, the formal properties of mental states must, at minimum, co-vary with those semantic properties of mental states that are used for predicting and explaining behavior. A rigorous scientific

psychology that wants to be able to endorse the sort of content-based generalizations that commonsense psychology uses in predicting and explaining behavior must claim that thoughts that differ in their contents, and hence, according to commonsense psychology, in their causal powers, must also differ with respect to some feature that, according to the scientific theory, is involved in the actual causal mechanism. If a functionalist psychology is to be a plausible empirical model of how cognitive mental states are involved in the etiology of behavior, then there must be a plausible account of mental state content that is consistent with the realization that mental states have their causal powers, ultimately, in virtue of their formal properties. However, a psychological theory that posits a mere epiphenomenal co-variation of content with causal powers will not have succeeded in providing an account of mental causation. At best, it will provide an account of why predictions and explanations of behavior couched in terms of the semantic properties of mental states are, in practice, successful. An account of mental causation, that is, the causation of behavior by intentional states per se, must make content causally relevant, in some fashion, to behavior. Thus, the success of a functionalist psychology depends upon 1) how the semantic properties of those mental states that are, in fact, appealed to when behavior is explained in terms of beliefs and desires should be characterized or individuated, and whether or not such a characterization permits a co-variation of semantic properties and formal properties; and 2) whether or not the functionalist theory posits a sufficiently robust connection between content and causal powers to be an account of mental causation.

In large part, the debate over whether or not a functionalist

psychology can be constructed using the conceptual framework of commonsense belief/desire psychology has focused on the issue of how to characterize, identify, and individuate the contents of mental states. (In general, the issue of the strength of the content-causal power connection has been taken up only secondarily, in the case of those theories that claim to have successfully dealt with the issue of content identification, e.g., Dretske's information theoretic account.) It is claimed that unless a scientific psychology can employ a notion of content roughly compatible with the one used by commonsense psychology, it cannot be considered a viable account of what we ordinarily think of as intentionality and mental causation. According to a functionalist psychology, mental state types supervene on formal state types, so tokens of the same formal state type are tokens of the same mental state type. Environmental factors to which an individual is physically and perceptually insensitive cannot have an impact on the physical state of the individual, and consequently cannot have an impact on the formal state of the individual. Since, as Fodor once argued,¹ mental states are typically type individuated on the basis of their contents, a functionalist psychology requires a notion of content that ensures that only those environmental factors that affect formal state type can affect mental state content type. A functionalist psychology then requires that the notion of content used when predicting and explaining behavior be what has come to be referred to as a "narrow" notion of content, one that stipulates that only those environmental factors that affect an individual physically or perceptually are relevant to the determination of mental state content type. However, if predictions and explanations of behavior by

commonsense psychology utilize a notion of content in which the content of a mental state is determined, in part, by factors that do not affect the individual physically or perceptually, what has come to be referred to as a "wide" notion of content, then mental state content types will not supervene on formal state types. It is a consequence of this position, although a controversial one to be defended in the chapters that follow, that a functionalist psychology would not be possible were commonsense to use a wide notion of content.

Jerry Fodor argues that since the only even remotely plausible accounts of mental causation are those that adopt a machine functionalist position, we should presuppose that model when constructing a notion of content. Content, by hypothesis, will co-vary with the formal, computationally relevant physical properties of mental states. The issue for Fodor is not the empirical question of whether or not the notion of content that is (apparently) used by commonsense psychology does, in fact, co-vary with formal properties, but rather the theoretical question of how should we construe the notion of content (what notion of content should we construct) so that it both co-varies with formal properties and agrees, at least roughly, with commonsense intuitions about content. Any success that commonsense psychology has had in predicting and explaining behavior on the basis of beliefs and desires must be assumed to have resulted, according to Fodor, from its implicit use of a narrow notion of content.

In "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology"² Fodor claims that a narrow notion of content based on the opaque construal of content clauses will both meet the requirement that content co-vary with causal powers (because opaque

construal is, roughly, determined by how the individual conceives of the object in question), and accord reasonably well with common practice in attributing mental states. However, in Psychosemantics³ Fodor argues for a rather different notion of "narrow content," although he still maintains that "narrow content" is the only plausible candidate for a Realist construal of mental states. Apparently swayed by the examples of Tyler Burge, Fodor is no longer of the opinion that a scientific psychology will be able to explain and predict behavior on the basis of the opaque construal of content clauses. Consequently, Fodor suggests an alternate notion of narrow content, one that he feels will more accurately reflect common intuitions than simply opaque construal did, but that will still maintain the connection between content and formal properties. The new notion of narrow content Fodor suggests is a non-semantic function which, for general terms, maps contents to extensions, given a context, where the extension of a content in a context is understood to be identifiable and expressible using the kind predicates of the language of the context.

While there is no doubt that Fodor's Representational Theory of Mind (given either of his notions of narrow content) provides a plausible explanation of why content-based predictions and explanations are useful, since the hypothesized coordination of semantic and formal properties ensures the accuracy of content-based predictions of formal-property-caused behavior, I will argue that it cannot be said to provide an explanation of mental causation. Although Fodor claims that narrow content will play a central role in his functionalist psychology, the theorized correspondence between formal and semantic properties of mental states reduces content to playing what is, at

best, a peripheral role. The contents of propositional attitudes are irrelevant to the etiology of behavior. Fodor's assertions to the contrary notwithstanding, I will argue that his psychology will not need to couch its generalizations in terms of the semantic properties of mental states.

Tyler Burge claims that the notion of narrow content is "hopelessly oversimplified as a philosophical explication of ordinary mentalistic notions" (OB, p. 113), and that the notion of content that is operative when behavior is predicted or explained on the basis of propositional attitudes is a wide, socio-linguistic notion of content. Burge claims that the content of an individual's mental state cannot be determined without appealing to the specific environmental context in which the person exists. The notion of narrow or individualistic content cannot, according to Burge, adequately account for the ordinary attributions of content used in commonsense belief/desire psychology. Burge argues that because the particular environmental cum linguistic context in which an individual exists determines what mental state contents we are and are not willing to attribute to the individual, the content of mental states must be determined by wide socio-linguistic, or communitarian factors. Burge claims that there is no such thing as individualistic mental content. The mental contents of those who use language are determined by their local linguistic environments. The semantics of a language determine the mental state contents an individual who speaks that language can have. One cannot, according to Burge, determine sameness or difference of content from the formal properties of an individual's mental states. Mental state content-types do not supervene on mental states type-individuated on the basis

of their formal properties.

Burge's construal of content is, I maintain, untenable as an explication of either the ordinary notion of content or a rigorous psychological notion of content. I argue that there are a number of important ways in which Burge's theory conflicts with the sorts of beliefs and desires we, in fact, are and are not willing to attribute to individuals, both language users and non-users of language. I argue that Burge's theory, counterintuitively, makes the nature of mental causation different in users of language than in non-users of language, necessitating that psychology have two notions of content and two theories of mental content and two theories of mental causation. Furthermore, I claim that Burge's thought experiments are dependent upon an assumption that is incompatible with the evolution, no less the existence of, the shared semantics of language that his own theory presupposes.

Fred Dretske claims that the intentional character of mental states arises from their status as indicators, that is, as information carrying signals. Although factors internal to the individual do have a role to play in determining what the content of a mental state is, the content of mental states is fundamentally determined by the nomic regularities that ground the information relation. According to Dretske, the semantic information a signal carries is dependent upon the nature of that to which the signal is nomically related. By changing the nature of whatever the signal is nomically correlated with, one changes the information the signal carries, and consequently, according to Dretske, changes the content of the individual's mental state. Since the information a signal carries is grounded in nomic

dependences that are supposedly independent of the individual, information, and mental content, can vary without the variation being detectable within the individual. Like Burge's, Dretske's position is that because mental state content can vary independently of changes internal to the individual mental states do not, in fact, supervene on brain states. Mental content, according to Dretske, is wide.

Although I am, in general, quite sympathetic to the information theory based program of Dretske, I believe that Dretske has misconstrued the ramifications of an information theoretic account of intentional states. Dretske claims that in an information theory based account the contents of mental states will be determined on the basis of certain factors pertaining to the environment, regardless of whether or not the individual whose mental state is under consideration is sensitive to those factors. I argue that an information theoretic account of mental states cannot, in fact, condone the determination of content in the way Dretske suggests. Dretske's primary motivation for this position is, I believe, based on epistemological considerations, and not on considerations of what is necessary or desirable for a plausible psychological theory. I believe that psychology must determine for itself what individuation of content it is appropriate for it to use, based on an assessment of what is necessary for it to achieve its predictive and explanatory goals, without regard for the needs of other theories or fields. I claim that an information theoretic account of mental states must individuate content solely on the basis of those environmental factors that the individual is cognitively sensitive to.

While I argue there are problems peculiar to each of the theories of

Fodor, Burge, and Dretske that make those theories unsatisfactory, it is my contention that those various difficulties arise, in part, because of a common, more fundamental misconstrual of what is involved in determining the content of a mental state. In their discussions of what notion of content a scientific psychology must use if it is to vindicate commonsense belief/desire psychology, Fodor, Burge, and Dretske all take for granted that there are various different independently identifiable substances, properties, and/or kinds (natural and otherwise) in the world (which includes the individual). Although Fodor, Burge, and Dretske each argue for a position that is quite different from the others, underlying each of their positions regarding the identification and individuation of mental state content is the assumption that the individuation of the world for psychology is fixed independently of psychology. The point that they, in fact, debate is how those independently fixed and identified things do or do not affect the content of a mental state. In essence, they construe constructing a notion of content for psychology to be determining which of those independently taxonomized substances, properties, and kinds influence content and what sort of influence they have. But approaching the issue of mental state content with the presupposition that there is already a taxonomy of the world, an individuation of the world into kinds, is, I believe, fundamentally misguided. The question of what notion of content to use and how to individuate mental state contents in psychology is itself, I maintain, a question about what natural kinds there are in the world according to psychology. I believe, that integral to the question of what is the content of a mental state is the question of what there is in the world according to

psychology that gives content to the mental state. One cannot, I claim, separate the question of how to individuate mental state contents from the question of how to individuate the world. It is my contention that the type-individuation of mental state contents and the type-individuation of the world go hand in hand.

I want to argue that the type-individuation of mental state contents a psychology endorses must be based on the type-individuation of the world that it, itself, endorses, and that the criteria for both of these type-individuations must be determined by the predictive and explanatory needs of a presumptively autonomous psychological theory. The taxonomic types that other fields recognize are not, in my view, determinative of how psychology should type-individuate either the world or mental state content. The natural kinds which physics, chemistry, economics, or linguistics recognizes are not necessarily the natural kinds that psychology will recognize. The ontological demands of a scientific psychology may result in an individuation of this world that is different from the individuation that physics uses, chemistry uses, economics uses, linguistics uses, or that any other field uses.

Fodor, Burge, and Dretske are, I believe, each ontological absolutists. They do not recognize that the explanatory and predictive needs of psychology may mandate an individuation of the world that is different than any of the ones we employ elsewhere. Indeed, I argue that the difficulties that each of their theories has arise, in part, because of the assumption that psychology must taxonomize the world, and consequently mental states, in accordance with the taxonomy that physics, chemistry, etc. use. I argue that once we recognize that psychology must taxonomize not only mental state contents but the world

as well, and that the taxonomy it uses must be determined in accordance with its predictive and explanatory needs, it becomes apparent that psychology must use a narrow notion of content. Furthermore, I argue that content will be radically idiosyncratic, determined on a case by case basis, because the taxonomy of the world relevant to the determination of the content of mental states will vary with each individual, and that the variables of psychological generalizations will range over such individual-indexed contents. Although there will, in general, be a great deal of similarity of mental state contents among members of the same community, I believe that, in practice, there will be little if any content identity across individuals, as content is individuated for a scientific psychology. However, by endorsing a narrow, radically individualistic notion of content a computational psychology will not be jeopardizing its ability to vindicate commonsense folk psychology. A scientific psychology can vindicate common practice without endorsing the particular attributions of content that commonsense psychology makes. The fact that commonsense folk psychology claims that many different people often have the "same" belief does not require that a scientific psychology endorse the same equivalence relations between mental state contents. To vindicate commonsense psychology all that is required is that a scientific psychology endorse the explanation and prediction of behavior in terms of beliefs and desires, and that there be an account of when and why commonsense attributions of content differ from scientific ones. I contend that commonsense belief/desire psychology can be vindicated by a functionalist psychology only if it uses a radically individualist narrow notion of the content of mental states.

CHAPTER 2
JERRY FODOR'S NOTION OF CONTENT

In recent years, Jerry Fodor has been arguing that a rigorous scientific psychology based on a computational, i.e., machine functionalist, theory of the mind can vindicate commonsense belief/desire psychology by providing an account of intentionality and mental causation. In this chapter, I want to lay out the main tenets of Fodor's theory, and will suggest that Fodor's defense of mental causation and the need for content-based generalizations in a scientific psychology is inadequate. It will be my contention that Fodor's theory fails to provide an account of mental causation because the content of beliefs and desires are not given a sufficiently robust role to play in the causation of behavior. While it is my position that Fodor's theory cannot account for mental causation, nevertheless Fodor's theory does connect mental causation with individualism in psychology, a connection that I will advocate throughout this thesis.

Fodor claims that a scientific psychology based on machine functionalism will not only be able to explain how propositional attitudes are causally involved in the production of behavior, but will also be able to exonerate the commonsense practice of generalizing about those causal interactions on the basis of the semantic properties of propositional attitudes. The scientific psychology Fodor envisions is one that regards propositional attitudes as genuine states of individuals, having both semantic properties and causal powers. In

particular, propositional attitudes are to be identified as relations to representations, where a representation is understood to be a symbol, an object having both physical and semantic properties. Propositional attitudes are to be individuated and generalized over, for the purposes of explaining and predicting behavior, on the basis of the semantic properties of their representations. But, it is in virtue of (certain of) the physical features of representations that propositional attitudes have the causal powers they do.

Although Fodor believes that a psychology based on his Representational Theory of Mind (henceforth RTM) will have to couch its generalizations in terms of the semantic properties of propositional attitudes, I will argue that a careful examination of the role that semantic properties actually play in a psychology of the sort Fodor conceives of reveals that the semantic properties of propositional attitudes are of relatively little importance. RTM claims that when the semantic properties of propositional attitudes are individuated in the appropriate way there will be a "harmony" between the semantic properties of propositional attitudes and their causally relevant physical properties. Propositional attitudes that differ in their properly individuated semantic properties will also differ in their causally relevant physical properties, and those that differ in their causally relevant physical properties will also differ in their properly individuated semantic properties. If we assume, as Fodor does, that commonsense psychology properly identifies the semantic properties of propositional attitudes, then this hypothesized harmony between semantic and causal properties allows RTM to account for the fact that propositional attitudes have, in commonsense belief/desire psychology,

causal powers that respect the logical form of their contents. However, I will argue that because the harmony is only a parallelism between the causal powers of propositional attitudes and their appropriately individuated semantic properties, the explanatory significance of semantic properties is ultimately undermined in a psychology based on RTM. I will argue that RTM can, at best, provide only an account of an 'epiphenomenalistic' notion of mental causation and a tenuous justification for the retention of content-based generalizations in a scientific psychology.

I. THE COMMONSENSE "FOLK" PSYCHOLOGY VIEW OF BELIEFS

The commonsense belief/desire psychology that Fodor takes as the model for a scientific psychology has as its fundamental underlying assumption the idea that people act out of their beliefs and desires. People do the things they do and say the things they say because of the particular set of beliefs and desires they have. Beliefs, desires, hopes, fears, hunches, and the like are all things that people genuinely have, and they are the things that spur people to action. They are the things that cause people to behave in the ways they do, and people with similar beliefs and desires tend to behave in similar ways under similar circumstances. In ordinary situations when we explain why someone did something we almost invariably cite some belief or combination of beliefs and/or desires that the person had as the cause or causes of his behavior. Such explanations are essentially the only kind of explanations that commonsense psychology countenances. For example, if you ask me why Leonora hired a lawyer, I might tell you that she hired a

lawyer because she wants to get a divorce. You can ask for a more detailed explanation of why she hired a lawyer in order to get a divorce (perhaps you think she could have done it herself with the help of a \$19.95 How-to book), but you cannot ask me for an explanation that does not either explicitly or implicitly advert to what she believes, wants, hopes, or fears. Neither can you ask me why, in the general case, wanting something can be a cause of doing something. The only kind of explanation of Leonora's behavior that I can give is one couched in terms of her beliefs, desires, and fears. Commonsense belief/desire psychology countenances no way other than citing beliefs, desires, and the like for explaining behavior, because beliefs and desires are just the sorts of things that cause behavior.

It is, of course, the case that we sometimes explain behavior not by adverting to the beliefs and desires of people, but by adverting indirectly to such things as reflex arcs and sensations. For example, we are allowed to explain why Bob dropped the skillet simply by saying that it was hot, because we know that if you touch something that is sufficiently hot you will automatically recoil from it. Bob's behavior was caused by a reflex that occurred in response to touching something hot. It was not caused by anything he believed, wanted, or feared (though, in appropriate circumstances, believing or fearing that a skillet was hot might also be an explanation of skillet-dropping behavior). While reflex arcs, sensations and beliefs are all used in casual explanations of why someone did something, Fodor's interest is in the construction of a scientific psychology concerned with how beliefs, desires, and the like influence behavior. That the skillet's being hot caused Bob to drop it is not a fact about belief/desire psychology. It

is a fact about the neurological construction of the human body: certain sorts of stimuli result in involuntary responses. The scientific psychology Fodor is concerned with constructing is one that deals with how beliefs, desires, fears, etc. -- states that have semantic content -- can be causally efficacious with respect to behavior.

In addition to holding that beliefs can be the cause of behavior, commonsense psychology also treats beliefs and desires as having satisfaction conditions. For example, beliefs are held to be semantically evaluable, that is, they are the sort of things that are either true or false. Not only is it the case that Leonora's belief that her husband is unemployed can be the cause of her seeking a divorce, but it is also something that is either true or false. What makes her belief true or false, as the case may be, are various facts about the world. "[W]hat makes a belief true (/false) is something about its relation to the nonpsychological world. . . . Hence to say of a belief that it is true (/false) is to evaluate that belief in terms of its relation to the world" (Psycho, p. 11). In particular, what makes Leonora's belief, for example, true is the fact that her husband does not have a job. It is that fact about the world that makes her belief true. No other fact about the world determines the truth value of Leonora's belief that her husband's is unemployed. It is because her belief is about the joblessness of her husband that her belief is made true or false by the facts of her husband's employment. That her husband is unemployed is the content of her belief. Her belief is either true or false in virtue of the relation between the content of her belief -- that her husband is unemployed -- and the facts of her husband's employment. "If you know what the content of a belief is, then you know

what it is about the world that determines the semantic evaluation of the belief" (Psycho, p. 11). As Fodor sometimes says, her belief 'expresses the proposition' that her husband is unemployed, and the truth value of a belief is determined by the truth value of the proposition it expresses. (Because on this view beliefs, in particular, but desires, fears, etc. as well, "express" a proposition about the actual or some possible world and indicate a disposition that the subject has towards that situation, they are often called, following Russell, 'propositional attitudes', and I will frequently use that terminology.)

The content of a belief not only tells you what facts about the world determine its truth value, but it also determines the identity of the belief. Beliefs, desires, etc. are all identified on the basis of their contents. Beliefs that are about different things in the world are taken to be different beliefs. Your belief that Koch is incompetent, and my belief that D'Amato is incompetent are different beliefs because they are about different things in the world. The content of a belief is essential to its identity: "[P]ropositional attitudes have their contents essentially: the canonical way of picking out an attitude is to say (a) what sort of attitude it is (a belief, a desire, a hunch, or whatever); and (b) what the content of the attitude is (that Hamlet's uncle killed his father; that 2 is a prime number; that Hermia believes that Demetrius dislikes Lysander; or whatever)" (Psycho, p. 11).

What is perhaps most striking about commonsense psychology is not that propositional attitudes are imbued with both content and causal powers, but that the patterns of causal interactions among them mirror the

relations among their contents. The generalizations of commonsense psychology predict the pattern of causal interactions among beliefs and desires on the basis of their content relations. For example, if we know that Percy believes that all Irish have red hair, and he comes to believe that Margaret does not have red hair, other things being equal, we can reliably predict that he is likely to come to believe that Margaret is not Irish. We conclude that Percy is likely to come to believe that Margaret is not Irish on the basis of a commonsense rule of thumb that says, roughly:

If you believe that:	All A's are B
And you believe that:	S is not B
Then you are likely to come to believe:	S is not A

Beliefs and desires interact causally with one another in a way that respects the logical form of their contents, the logical form of the propositions they express. The generalizations of commonsense psychology, which achieve their predictive power by generalizing over all individuals -- all agents with the same beliefs and desires will behave in the same way -- and by abstracting over the contents of propositional attitudes -- "'If you want P and you believe not-P unless Q . . . you try to bring about Q,' whatever the P and Q may be" (Psycho, p. 13) -- are all predicated on there being this linkage between content relations and causal interactions.

II. CONTENT, CAUSATION, AND FODOR'S REPRESENTATIONAL THEORY OF MIND

All of the generalizations that commonsense psychology makes about the connection between behavior and propositional attitudes are made on the basis of the semantic content of propositional attitudes. Commonsense psychology ties causation to content: propositional attitudes that have different contents have different causal powers. In order for a rigorous scientific psychology to vindicate these sorts of content-dependent causal explanations of behavior, it "must permit the assignment of content to causally efficacious mental states and must recognize behavioral explanations in which covering generalizations refer to (or quantify over) the content of the mental states that they subsume (Psycho, pp. 14-15). To permit the assignment of content to mental states is, basically, to hold that mental states are representational states. Thus, a scientific psychology that vindicates commonsense psychology must be, fundamentally, a representational theory of propositional attitudes.

In general, representational theories of propositional attitudes differentiate between propositional attitudes in much the same way that commonsense psychology does: propositional attitudes can differ from one another in their contents, "so we can allow for the difference between thinking that Marvin is melancholy and thinking that Sam is (or that Albert isn't, or that it sometimes snows in Cincinnati)," and they can differ from one another in their attitudes towards the contents, "so we can allow for the difference between thinking, hoping, supposing, doubting, and pretending that Marvin is melancholy" (MS, p. 226). In

fact, "[s]o far as the representational theory of mind is concerned, it [content] is possibly the only thing that distinguishes Peter's thought that Sam is silly from his thought that Sally is depressed" (MS, p. 227). However, on anyone's account, the direct cause of a person's behavior, for example, of Leonora's behavior when she walks to her attorney's office, sits down and talks to him, must be some collection of physical processes and mechanisms. That which is the direct and immediate cause of her behavior is some physical process involving neurons in the brain firing, causing muscles to contract, etc., etc., none of which is at all sensitive to the semantic properties of propositional attitudes. Semantic properties, while borne by things that have physical properties, are not themselves physical properties. The fact that in the United States an octagonal road sign means 'stop' is not in any way determined by the physical properties of the sign. It could have just as easily been that a pentagonal sign meant stop. Content -- "a semantic notion par excellence" (MS, p. 227) -- cannot be involved in the actual causal mechanism because content per se is not a physical property.

A rigorous scientific psychology that wants to be able to endorse the sort of content-based generalizations that commonsense psychology uses in predicting behavior, must demonstrate that thoughts that differ in their contents, and hence in their causal powers, must also differ with respect to some feature that is involved in the actual causal mechanism. Peter's thought that Sam is silly must differ from his thought that Sally is depressed in two respects: in content and in some feature relevant to the causal (i.e., physical) mechanism by which behavior is produced, if commonsense psychology is to be vindicated. It must be

able to provide an account of how the semantic properties of propositional attitudes are related to the physical processes that are actually causally responsible for the production of behavior, such that we can predict and explain behavior on the basis of the one while it is actually the other that is causally responsible for the behavior in question. It must provide an explanation of the linkage between content and causal power.

The problem for Fodor, then, is to construct an account of what propositional attitudes are that is sufficiently physicalist to account for causation while still doing justice to the practices of commonsense psychology and our intuitions about what psychology is. In other words, RTM must provide an account of propositional attitudes that acknowledges that they are causally efficacious representational states whose causal interactions contrive to respect their content relations. Furthermore, it must tell us when two mental particulars are to count as the same propositional attitudes, and tell us what the physical characteristics of propositional attitudes are that gives them their causal powers. Although commonsense psychology does not provide any indication of what kind of underlying causal mechanism should be posited, Fodor does feel that there are a variety of considerations that dictate what the general outlines of a theory of propositional attitudes must be.

In the Introduction to Representations, Fodor argues that

. . . if we want a science of mental phenomena at all, we are required to so identify mental properties that the kinds they subsume are natural from the point of view of psychological theory construction. . . . Now, there is a level of abstraction at which the generalizations of psychology are most naturally pitched and, as things appear to be turning out, that level of abstraction collapses across the differences between physically quite different kinds of systems. Given the sorts of things we need to say

about . . . believing P's, it seems to be at best accidental, and at worst just false, that . . . beliefs are proprietary to creatures like us. . . . This is a state of affairs which cries out for a relational treatment of mental properties, one which identifies them in ways that abstract from the physiology of their bearer.

(SSA, pp. 8-9)

Whatever the specific doctrine about the nature of mental properties that a science adopts it must be one that identifies mental properties in a way that is independent of the particular physical features of any given bearer of mental properties. We need to identify mental properties in such a way that the properties expressed by typical mental predicates, e.g., believing-that-P, are projectible across systems of varying physical construction.

The account of propositional attitudes that RTM gives must be a relational account, at least in the sense that it must be possible to ascribe propositional attitudes to systems regardless of the particulars of their physical organization. Given that RTM claims that propositional attitudes are representational states, if we want propositional attitudes to apply across different physical systems, then it must be that the particular physical features of a representation are immaterial to its having the content it does. Furthermore, since commonsense psychology allows that different propositional attitudes can share the same content (representation), e.g., believing that Marvin is melancholy and doubting that Marvin is melancholy, it must also be that attitude type is projectible across physically different systems. According to Fodor, the way to achieve this projectibility is to identify propositional attitudes as relations to representations. Propositional attitudes are to be viewed as two-place relations between an organism and a representation internal to the organism that functions

as the object of the relation: to be in a propositional attitude is to stand in a relation to an internal representation. For example, to believe that Simon likes haggis is to stand in that relation identified as believing to an internal representation that represents Simon as liking haggis, or as Fodor sometimes says, to an internal representation the content of which is that Simon likes haggis. Both the type of relation an organism bears to the representation and the content of the representation that functions as the object of the relation are individuating features of propositional attitudes. A propositional attitude that involves, for example, the relation of believing is distinct from one that involves the relation of wanting, and one that has as its object a representation whose content is that Simon likes haggis is distinct from one that has a representation whose content is that Simon likes porridge. Propositional attitude tokens are identical only if they are identical relations to content identical representations. Propositional attitudes are individuated "by specifying a relation and a representation such that the subject bears the one to the other" (MS, p. 226).

By saying that propositional attitudes are relations to representations Fodor has provided only a partial account of propositional attitudes for a scientific psychology. RTM has to include an account of the physical features of mental representations that allow them to have the causal powers they do, and those physical features must be identified in a way that "abstracts from the physiology" of the bearers of those representations, since Fodor wants to allow that "there could, at least in logical principle, be nonbiological bearers of mental properties" (SSA, p. 11). Fodor claims that machine functionalism

provide just such an account. It provides an account of how we should identify the causal physical features of mental representation such that we do not restrict the possibility of having mental representations, and thus the possibility of having propositional attitudes, on the basis of physiology.

Machine functionalism claims that when we provide an account of what a propositional attitude is we must do so using only those techniques of functional state identification that can be used in specifying the program states of a computer. When we specify the program states of a computer, we "can advert to (a) inputs to the machine; (b) outputs of the machine; (c) any of a designated set of elementary machine operations; and (d) other program states of the machine" (SSA, p. 13). In particular, what one cannot do is make reference to any of the specific physical characteristics of the machine. For example, the functional specification of a Coke machine, one which, for the sake of simplicity, we will assume charges 75 cents for a Coke and only accepts quarters, might be:

	(a) <u>INPUTS</u>	(b) <u>OUTPUTS</u>	(c) <u>DO</u>	(d) <u>GO TO</u>	(r) <u>REGISTER</u>
State 0	Q	-	Add 25 to r	State 1	0
State 1	Q	-	Add 25 to r	State 2	25
State 2	Q	Coke	Set r to 0	State 0	50

(While in this simple example neither column (c) nor column (r) are necessary, one can easily see how they would be important if the machine accepted nickles, dimes, and quarters, and gave change.)

What is important about the machine analogy is that machines function the way they do on the basis of certain physical characteristics of the inputs, outputs, and internal states. More specifically, a machine

operates the way it does in virtue of the 'formal' properties of its inputs, outputs, and internal states, where formal properties of inputs, outputs, and internal states are understood as being those physical features of the inputs, outputs, etc., that the machine has been designed to be sensitive to. If we apply the machine analogy to the case of mental representations we get a model of representations as being symbols, having both semantic and formal properties, but where a representation has its causal powers in virtue of its formal properties.

While Fodor typically identifies the causally relevant properties of propositional attitudes in terms of the 'shape' of the symbols, he is not particularly illuminating on what precisely the formal properties of a symbol or propositional attitude are. His basic position seems to be to view formal properties in terms of being non-semantic ones, and since "we don't know what semantic properties there are," we can only have an "intuitive and metaphoric" understanding of what the formal properties of a symbol are (MS, p. 227). Nevertheless, it does seem that we can come up with a rather more precise characterization of the formal properties of a symbol than Fodor does. Consider the case of our Coke machine. It is designed to vend a soda when the right amount of money has been put in. But it determines when to vend a soda, not on the basis of the amount of money that it has taken in, but on the basis of the number of objects of a certain size and shape that it has taken in. It makes no difference to the functioning of the machine that the objects it generally takes in are coins, having monetary value, or that the thing it vends is a can of Coke rather than a can of Pepsi. These things are irrelevant to its "behavior". It will function in exactly the same way regardless of whether one puts in a genuine U.S. quarter

dollars, a counterfeit U.S. quarter dollar, or a Panamanian quarter balboa. To begin with, we know that formal properties must be physical properties, because only physical properties can have causal effects. Furthermore, they must be locally relevant physical features. In saying that the physical features must be locally relevant ones, I mean to exclude those physical properties that the machine operating on the symbol has no means of detecting.

The objects that our Coke machine accepts when determining when to vend a soda have a variety of physical features: they are round, of a certain thickness, have milled edges, weigh a certain amount, have surface contours, etc., etc. However, not all of these properties are relevant to the functioning of this particular machine. The Coke machine considers size, shape, and weight among those physical features that are relevant to its functioning, so it can distinguish between quarter dollars and wooden slugs. On the other hand, it does not consider the particular pattern of surface contours to be functionally relevant physical features, so it cannot distinguish between quarter dollars and quarter balboas. Although the Coke machine is not designed to be able to detect the difference in the surface contours of the objects it accepts, we can easily imagine a machine that would consider surface contours relevant to its functioning. However, we cannot construct a machine, no matter how sophisticated, that would be able to distinguish between a genuine U.S. quarter dollar and a (perfect) counterfeit U.S. quarter dollar. What makes a coin genuine or counterfeit is its ancestry, its history. As Fodor says in a different context, "Even God couldn't make a gen-u-ine United States ten cent piece; only the U.S. Treasury Department can do that" (Psycho, p. 45).

But being a genuine United States quarter is immaterial to the functioning of the Coke machine. What is important for the Coke machine is those physical features of the things it takes in that it can test for or examine.

In general then, the formal properties of a symbol are those physical properties of the symbol that are relevant to the functioning of the machine operating on them. Which physical properties of an object are the formal ones is relative to the system within which it operates. Those particular physical features of a representation that are its formal features -- the ones in virtue of which a propositional attitude has its causal powers -- will depend upon the exact mechanism that implements the pattern of causal interactions. It might be that the mechanism consists of wheels and cogs in one individual but of resistors and silicon chips in another, in which case the formal properties of the representations that they operate on will be different in the two cases. Which particular physical properties are the formal ones will depend upon the particulars of the mechanism, and -- at least to a first approximation -- such considerations are not of concern to psychology.

III. THE INDIVIDUATION OF MENTAL STATE CONTENT IS INFLUENCED BY BOTH SCIENTIFIC AND FOLK CONSIDERATIONS

If we accept the machine analogy for mental representations, it must be that propositional attitudes have their causal powers in virtue of the formal properties of the representations that are their objects. A rigorous scientific psychology that can endorse the commonsense psychology

practice of predicting and explaining behavior on the basis of the content of propositional attitudes must, at the very least, be able to show that there is a "harmony" between the semantic properties of propositional attitudes and the formal properties of propositional attitude representations. It must be able to show that by classifying propositional attitudes on the basis of their semantic properties, commonsense psychology is also, at least roughly, classifying them on the basis of their causally relevant properties, i.e., their formal properties. Propositional attitudes that are identified by commonsense psychology as being of the same type on the basis of their semantic properties, such as Tom's belief that Simon likes haggis and Sam's belief that Simon likes haggis, must also turn out to be of the same type on the basis of their formal properties, if a scientific psychology is to vindicate the use of semantic-property-based generalizations to predict formal-property-caused behavior.

In order to account for this hypothesized harmony between semantic properties and formal properties, there are two different issues that must be addressed. First, we need to determine what semantic properties of propositional attitudes display this harmony with the causally relevant features: What are the semantic properties of propositional attitudes that commonsense psychology uses when it "successfully" predicts/explains the behavior of individuals on the basis of those semantic properties? Second, we need to consider how and/or why there is this harmony between the formal properties and semantic properties of propositional attitudes. What is it that ensures that the semantic properties of propositional attitudes are paired with and co-vary with the formal ones?

Obviously, these two issues are not independent of one another when trying to formulate a scientific psychology. It must be the case, if a scientific psychology is to continue predicting behavior on the basis of semantic properties, that those semantic properties that are identified as the ones to be used for predicting and explaining behavior are also the ones for which we have an explanation of their being in harmony with formal properties. Any scientific psychology that wants to continue the commonsense psychological practice of generalizing on the basis of semantic content, must be able to provide an account of the semantic properties it will use as the basis of its generalizations that is both reasonable given what commonsense psychology does, and for which it can explain the harmony between formal and semantic properties.

Constructing a viable theory of which of the semantic properties of propositional attitudes can and should be used in the psychological generalizations that connect propositional attitudes and behavior is, perhaps, the most difficult task facing Fodor, or anyone trying to characterize a scientific psychology that vindicates commonsense psychology. If a scientific psychology is to make generalizations about how behavior is influenced by propositional attitudes on the basis of their semantic properties, it must be that propositional attitudes that have the same semantic properties have the same causal powers. On the other hand, given the machine analogy, having the same causal powers means having the same formal properties.¹ Therefore, the semantic properties of propositional attitudes that psychology uses in its generalizations must be ones that are reasonable given the common sense idea of the content of beliefs and desires and must classify or, as Fodor says, 'taxonomize' propositional attitudes in such a way that all

propositional attitudes that have the same formal properties/causal powers are in the same content category, and that all propositional attitudes that are in the same content category have the same formal properties/causal powers.

The 'formality condition' (MS, pp. 227-228) of the mechanistic model, that is, the requirement that only the formal properties of propositional attitudes can affect the working of mental processes, dictates that two individuals, or one individual at different times, can be considered to be in type distinct mental states, e.g., believing-that-P and believing-that-Q, only if the mental representations that are in their heads, that are the objects of their beliefs, differ in their formal properties, and two individuals whose mental representations differ in their formal properties must be considered to be in type distinct mental states. The individuation of mental states by content must be consistent with the individuation of mental state by their formal properties, if we are to be able to make predictions about behavior on the basis of content. This is not in any way meant to limit the possible range of semantic properties that a mental representation can have to those that are linked to the formal properties of mental representations. It simply means that there must be some subset of the semantic properties of representations that are linked to the formal properties of representations on the basis of which psychology can successfully individuate propositional attitudes for the purposes of predicting and explaining the mental causation of behavior.

In "Methodological Solipsism" Fodor sketches the outlines of what subset of the semantic properties of mental representations he feels a scientific psychology can and should use for type individuating

propositional attitudes. Roughly, the subset of semantic properties that Fodor maintains should be used are essentially those that commonsense psychology uses when predicting and explaining behavior, which, according to Fodor, are to be identified with the opaque construal of propositional attitude ascriptions.² The opaque construal of a propositional attitude ascription is one that is sensitive to the way an object referred to is picked out, but is not sensitive to the identity of that object. On the other hand, the transparent construal of a propositional attitude ascription is one that is sensitive to the identity of the object referred to, but is not sensitive to the way it is picked out. For example, the belief that the Morning Star rises in the east and the belief that the Evening Star rises in the east, given that 'the Morning Star' and 'the Evening Star' refer to the same thing, are different beliefs when construed opaquely, but the same belief when construed transparently. In common, everyday situations beliefs and desires are attributed to individuals using linguistic constructs that typically include an embedded sentence or sentential clause that names the belief or desire. For example, given the belief ascription "Leonora believes (that) the cougar is an endangered species," "the cougar is an endangered species" is understood to specify the content of her belief.³ Fodor claims that commonsense psychology typically predicts and explains behavior on the basis of the opaque construals of the embedded content-specifying sentences that occur in propositional attitude ascriptions. The content of a propositional attitude must be opaquely construed in order for it to be useful for predicting and explaining behavior.

[W]hen we articulate the generalizations in virtue of which behavior is contingent upon mental states, it is typically an opaque construal of the mental state

attributions that does the work; for example, it's a construal under which believing that a is F is logically independent of believing that b is F, even in the case where a = b.

(MS, p. 234)

In order to apply any of our commonsense generalizations to Leonora, we must construe "the cougar is an endangered species" opaquely. Even though "the puma is an endangered species" means the same thing as "the cougar is an endangered species," 'puma' being just another name for cougars, we cannot treat the belief that the cougar is an endangered species and the belief that the puma is an endangered species as the same belief when we go to predict Leonora's behavior.

When we explain or predict behavior in everyday situations using commonsense psychology, we do it by saying what the content of the person's beliefs or desires are, and we take the way the beliefs and desires are specified rather literally. The particular words or phrases that are used in attributing propositional attitudes are, according to Fodor, taken to express what the individual has in mind; they indicate the way the individual conceives of the object of the belief; they specify what features are to be found in the representation in the persons head, and we cannot assume that that representation has any specific features other than those stated in the belief ascription.

Suppose I know that John wants to meet the girl next door, and suppose I know that this is true when "wants to" is construed opaquely. Then, . . . I can make some reasonable predictions (guesses) about what John is likely to do: he's likely to say (viz., utter), "I want to meet the girl who lives next door." He's likely to call upon his neighbor. He's likely (at a minimum, and all things being equal) to exhibit next-door-directed behavior. . . .

On the other hand, suppose that all I know is that John wants to meet the girl next door where "wants to" is construed transparently; i.e., all I know is that it's true of the girl next door that John wants to meet her. Then

there is little or nothing that I can predict about how John is likely to proceed. . . . For example, I have no reason to predict that John will say such things as "I want to meet the girl who lives next door" since, let John be as cooperative and as truthful as you like, and let him be utterly a native speaker, still, he may believe that the girl he wants to meet languishes in Latvia. In which case, "I want to meet the girl who lives next door" is the last thin[g] it will occur to him to say.

(MS, p. 235)

Substituting a co-referring term for one of the terms in the content clause of a propositional ascription radically alters the predictions and explanations that commonsense psychology will make. The particular words used in the content clause of a propositional attitude ascription when opaquely construed tell us (roughly) how "the agent represents the objects of his wants (intentions, beliefs, etc.) to himself," and it is, by hypothesis, the particular representation that the agent has in mind that functions in the causation of his behavior (ibid.).

Although commonsense psychology does, in general, prohibit the free substitution of co-referring expressions in propositional attitude ascriptions when predicting behavior, it nevertheless does seem to accept that the referent of the content clause is important for the identification of propositional attitudes. Commonsense psychology takes into consideration what a belief, desire, etc. is about when determining its identity, that is, the object of a propositional attitude is one of the determinants of the kind of propositional attitude it is. The common sense view is that if you have a belief that is normally attributed by saying that you believe that the Mayor is a crook, then you have a belief about the Mayor. If your neighbor also has a belief that is commonly attributed by saying that he believes that the Mayor is a crook, then you and your neighbor believe the same thing,

i.e., your beliefs have both the same content and the same object. On the other hand, if your niece, who lives in a different city, has a belief that is commonly attributed by saying that she believes that the Mayor is a crook, her belief will generally not count as the same belief as the one you and your neighbor share because her belief is about a different person. She has a belief about the mayor of Detroit, while you (and your neighbor) have a belief about the mayor of Newark. Even if you and your niece each think of your respective mayors in exactly the same terms, and would use exactly the same words to express your do not count as the same belief because they involve representations of different people, i.e., they have different objects. According to the commonsense model, two individuals count as having the same belief if and only if their beliefs have both the same content and the same object. Thus, on the commonsense model, what the identity of a propositional attitude is is determined, in part, by factors external to the individual.

Commonsense psychology seems to have two conditions for the identity of propositional attitudes: first, the objects of the propositional attitude representations must be the same; and, second, the contents must be "opaquely" identical, that is, the contents must represent the object the belief is about in the same way. While it is not difficult to imagine that the formal properties of propositional attitudes could co-vary with the way the object is represented, it is rather more difficult to see how the formal properties of propositional attitudes could co-vary with the identity of the object represented in the propositional attitude, particularly in cases where the sameness or difference of the object(s) represented is unknown to anyone. The

machine model stipulates that information about the environment can influence the formal properties of mental representations only to the extent that that information is conveyed by "ambient environmental energies" that impinge on the senses in a fashion such that they cause the senses to produce signals whose formal properties have been determined by those energies (MS, p. 231). For example, the formal properties of a representation of a flower could be influenced by such environmental information as the color of it if one looks at it, the texture of it if one touches it, or the scent of it if one smells it, but not the color of it if one only touches it, or the texture of it if one only smells it. It is hard to see how the identity of the object a belief is about could influence the formal properties of a mental representation apart from the way it impinges on the senses, and the way it impinges on the senses determines the way it is represented.

Putnam's well-known Twin-Earth example, from "The Meaning of 'Meaning'",⁴ clearly illustrates the difficulty a scientific psychology will have if it tries to consider the identity of an object a propositional attitude is about, apart from the way it is represented, as one of the semantic properties of propositional attitudes that psychology should use when type individuating propositional attitudes. In Putnam's example Oscar₁ lives on Earth and has a doppelganger, Oscar₂, who lives on Twin-Earth. Oscar₁ and Oscar₂ are identical in all physical, functional, phenomenological, and experiential respects, and Twin-Earth is just like Earth except that the liquid that runs in rivers, flows from taps, falls from the sky, and quenches thirst is something other than H₂O, call it 'XYZ'. Oscar₁ and Oscar₂ live at a time prior to the discovery of the atomic structures of H₂O and XYZ on

their respective planets. Oscar1, living on Earth, has a belief about H₂O (though, of course he does not know the stuff he has a belief about as H₂O), a belief that is normally attributed by the English sentence "Oscar1 believes that water is wet"; Oscar2, living on Twin-Earth, has a belief about XYZ (though, again, he does not know the stuff his belief is about as XYZ), a belief that is normally attributed by the Twin-English sentence "Oscar2 believes that water is wet." Oscar1's belief is about H₂O, while Oscar2's belief is about XYZ. If we follow the commonsense psychology method of individuating beliefs, then Oscar1 and Oscar2 would have different beliefs, just as you and your niece had different beliefs.

In his discussion of the Twin-Earth example, Putnam distinguishes between what he calls "psychological states in the wide sense" and "psychological states in the narrow sense." A psychological state in the narrow sense is one that does not "presuppose the existence of any individual other than the subject to whom the state is ascribed" ("The Meaning of 'Meaning'," p. 10), while a psychological state in the wide sense does presuppose the existence of something other than the subject. Viewed as psychological states in the wide sense Oscar1's belief and Oscar2's belief clearly are different psychological states. Oscar1's belief, construed widely, presupposes the existence of water, i.e., the existence of the stuff called 'water' by Oscar1 and his fellow Earth dwellers, i.e., H₂O. On the other hand, Oscar2's belief, widely construed, presupposes the existence of the stuff called 'water' by Oscar2 and his fellow Twin-Earth dwellers, i.e., XYZ. Viewed as psychological states in the narrow sense, however, Oscar1's belief and Oscar2's belief are not different psychological states. The difference

between Oscar1's belief and Oscar2's belief is only apparent if we look beyond the individuals and into the surrounding environment. If the subject to whom the state is ascribed is the only one whose existence we can presuppose, then we cannot take into consideration any features outside of the individual when identifying the psychological state the subject is in. Thus, Oscar1 and Oscar2 are in the same state narrowly construed.

If we let the way we individuate objects influence the way we individuate propositional attitudes for the purposes of predicting and explaining behavior, then it must be that propositional attitudes that have distinct objects (given whatever scheme for object individuation that we have chosen to use) must also be formally distinct. Regardless of the way we choose to individuate objects, the individuation of objects must be coordinated with a difference of the formal properties of propositional attitudes. In Putnam's example, H_2O and XYZ are identified as distinct natural kinds using a scheme of individuation based on an idealized physics. If the distinction between H_2O and XYZ, as identified by physics, is to be important for the prediction and explanation of Oscar1's and Oscar2's behavior, then it must be that their beliefs are formally distinct. However, we stipulated at the beginning of the example that Oscar1 and Oscar2 are identical in all physical and functional respects. Oscar1 and Oscar2 are, if you will, identical physical machines that function in physically identical ways in all circumstances. Because formal properties are simply a subset of physical properties, the formal properties of Oscar1's belief about H_2O are not different from those of Oscar2's belief about XYZ. A scientific psychology can defend its use of a semantic criterion of propositional

attitude identity only if it can demonstrate that the taxonomy of propositional attitudes that results from the semantic criterion is isomorphic to a taxonomy that results from a formal criterion. A scientific psychology must be able to claim that propositional attitudes that have different contents also have different formal, i.e., physical, features. In the Twin-Earth case, however, we see that if the content of propositional attitudes is construed in the wide sense, we have two beliefs that have different contents but do not have different formal properties. From these considerations, Fodor concludes that a scientific psychology is not going to be able to use wide content as the notion of content it uses for individuating propositional attitudes.

On the other hand, Fodor claims that under a narrow construal of the content of psychological states there is a notion of content available to a scientific psychology because psychological states in the narrow sense are "those individuated in light of the formality condition; viz., without reference to such semantic properties as truth and reference" (MS, p. 247), and are therefore compatible with a mechanistic model of mental processes. Although Fodor is not explicit about this point, presumably what he has in mind is that when determining what psychological state a person is in, if we construe 'psychological state' narrowly, we cannot consider the meaning of any symbols in the person's head (or for that matter, even the fact that it is a symbol). We can only look at the physical features of the individual. As we noted earlier, among the physical features of a psychological state are its formal features. Thus, if we construe psychological states narrowly when determining what kind of states they are, we are in effect individuating them, roughly, on the basis of their formal properties.

Although individuating psychological states on the basis of a narrow construal seems to be compatible with the formality condition, that fact does not itself, in any obvious sense, vindicate the commonsense psychological practice of individuating by opaque content. In order for Fodor's scientific psychology to vindicate the use of opaque content for individuating propositional attitudes, it must be that individuating by opaque content is consistent with individuating by formal properties. Fodor claims that by having shown that an individuation of psychological states narrowly construed is compatible with formal individuation, he has also shown that the individuation of psychological states opaquely construed is compatible with formal individuation because his notion of opaque content corresponds to Putnam's notion of narrow content. "[T]he narrow sense [of belief] must be (what I've [Fodor has] been calling) fully opaque" (MS, p. 246).⁵ Fodor's argument for this "correspondence between narrowness and full opacity" is essentially that they must correspond "because only full opacity allows type identity of beliefs that have different truth conditions" (MS, p. 246). When construed narrowly, Oscar1's and Oscar2's beliefs are type identical, even though they have different truth conditions, but Oscar1's and Oscar2's beliefs can be considered content identical only if we construe the contents of their beliefs opaquely. If we want to be able to taxonomize on the basis of content and still be able to taxonomize beliefs that have different truth conditions as the same belief kind, we must taxonomize on the basis of opaque content. Regardless of what the full range of semantic properties of mental representations is, it is only those semantic properties that are relevant to an opaque construal of content clauses that we can be certain will adhere to the formality condition.

Thus Fodor concludes that the opaque construal of content clauses is the notion of content that a scientific psychology should use.

Although Fodor claims that he wants to "emphasize this correspondence between narrowness and full opacity, and not just in aid of terminological parsimony" (MS, p. 246), narrowness and opacity are distinct notions. Putnam's definition of a psychological state in the narrow sense as one that does not presuppose the existence of anything other than the individual to whom the state is attributed leads to a sort of methodological idealism. What is going on inside of the individual's head is all that we can assume that there is. Fodor, however, does not intend to be understood as endorsing idealism of any sort, including methodological idealism, but only as endorsing solipsism. Fodor is not advocating the claim that, as far as psychology is concerned, there is nothing besides the individual to whom a belief is attributed, but merely that psychology should be concerned only with the individual's conception of the objects of his beliefs. The claim Fodor is making is that what is important for psychology is what is important to the individual. Fodor endorses not methodological idealism, but rather the weaker methodological solipsism,⁶ the view that psychology should individuate mental state content on the basis of how the individual conceives of the objects of his propositional attitudes without taking into consideration what there is in the environment in which he is embedded.

Having noted that there is an important difference between the methodological idealism of Putnam and the methodological solipsism of Fodor, there is a central shared element in their views which is, I believe, the basis on which Fodor adopts Putnam's term 'narrow content.'

Both use the term to express the view that it is something about the individual in question that determines the identity of the propositional attitude. In determining propositional attitude type identity, the individual takes precedence over the environment. It is this emphasis that Putnam's characterization of narrow content places on the individual that, I believe, underlies Fodor's adoption of Putnam's terminology. This commonality notwithstanding, Fodor's notion of narrow content is not identical to Putnam's. Therefore, it would perhaps have been more appropriate for Fodor to refer to his position as "individualism" and to claim that he endorsed an "individualistic" notion of content rather than a narrow notion of content. (Indeed, this is the way in which Fodor has been standardly interpreted. For example, in the criticisms that Tyler Burge levies against Fodor, Fodor's position is referred to as "individualism" and Fodor is represented as advocating an "individualistic" notion of content.) Nevertheless, since Fodor explicitly and emphatically chooses to borrow Putnam's term 'narrow content' and put it to his own use in explicating his position on the notion of content that psychology should use, this is the term that will be used in discussing Fodor's position, and 'narrow content' should always be understood in Fodor's sense. Consequently, it should be kept in mind that "narrow content" in Fodor's sense implies only individualism and methodological solipsism, and not methodological idealism.

IV. A SCIENTIFIC PSYCHOLOGY NEEDS A MORE ROBUST NOTION OF CONTENT
THAN FODOR'S RTM PROVIDES

Even if we accept Fodor's argument for there being a notion of content that is available to a scientific psychology, and that the opaque reading of content clauses is that notion, given the strict parallel between opaque content and formal properties that RTM posits, one is lead to question whether or not psychology needs this notion of content at all. Why should psychology individuate propositional attitudes on the basis of their content for predicting and explaining behavior when it is the formal properties that are actually causally responsible for the production of behavior? Or, to put it in a slightly different way, why shouldn't psychology simply couch its generalizations about how behavior is dependent upon propositional attitudes in terms of the functionally circumscribed formal properties of internal states? What justification is there for insisting on using an indirect criterion of propositional attitude individuation -- one that requires a fairly elaborate theory explaining how it coordinates with the causal, i.e., physical, account of the mechanism underlying the mental causation of behavior -- when a more direct criterion is available? This question should not be understood as impugning the authenticity of propositional attitude ascriptions, for it is raised within the context of Fodor's theory that accepts that "ascriptions of beliefs, when true at all, are literal," not merely admonitions to adopt a heuristically useful stance (TCPA, p. 102). If we are to accept Fodor's claim that propositional attitudes are genuine states of

individuals that have their causal powers in virtue of their formal properties, Fodor must provide some justification for his insistence that psychology use generalizations that quantify over the contents of propositional attitudes rather than over their formal properties. Given the preordained harmony between the narrow content of propositional attitudes and their formal properties that is required by his theory, and his insistence that content does not figure in any psychological mechanism (cf. *Psycho*, p. 140), it would seem that a scientific psychology would have no need for content-based generalizations. While Fodor does offer a number of reasons for maintaining that psychology couch its generalizations in terms of the semantic properties of mental states, I believe that they are ultimately insufficient to show that there is any either predictive or explanatory justification for holding on to content-based generalizations.

In "Something on the State of the Art,"⁷ Fodor argues that psychology must continue to individuate propositional attitudes on the basis of content, rather than individuating on the basis of formal properties, because psychology cannot succeed in making the generalizations about mental causation that it wants to make without appealing to the semantic properties of mental states (cf. *SSA*, pp. 25-30).

[We were driven to functionalism (hence to the autonomy of psychology) by the suspicion that there are empirical generalizations about mental states that can't be formulated in the vocabulary of neurological or physical theories; neurology and physics don't, we supposed, provide projectable kind-predicates that subtend the domains of these generalizations. But now if we think about what these generalizations are like, what's striking is that all of the candidates -- literally all of them -- are generalizations

that apply to propositional attitudes in virtue of the content of propositional attitudes. We don't need the clever examples from linguistics or psychology to make this point; commonsense psychological etiologies will do. So consider: seeing that a is F is a normal cause of believing that a is F; the intention that it should be the case that so and so is a normal cause of actions whose goal is to bring it about that so and so; statements that P are normally caused by beliefs that P; observations that many of the xs are F often contribute to the etiology of the belief that all the xs are F; the belief that a thing is red is a normal cause of the inference that the thing is colored; and so on and on. The point of such examples is not, of course, that any of them are likely to figure in a serious cognitive psychology. It's rather that our attempts at a serious cognitive psychology are founded in the hope that this kind of generalization can be systematized and made rigorous; it's precisely this kind of generalization that we abandoned type physicalism in hopes of preserving. And, YOU CAN'T SAVE THESE GENERALIZATIONS WITHOUT APPEALING TO THE NOTION OF THE CONTENT OF A MENTAL STATE, since, as previously remarked, these generalizations are precisely such as apply to mental states in virtue of their contents.

(SSA, pp. 25-26)

Fodor's claim is that it is only if psychology adverts to the content of mental states that it can formulate the generalizations it needs. If psychology were to couch its generalizations in the vocabulary of neurological, physical, or functional theories, rather than in terms that advert to semantic properties, it would lose its ability to make predictive and explanatory generalizations. "[I]n order to specify the generalizations that mentalistic etiologies instantiate, we need to advert to the contents of mental states" (SSA, p. 30).

Unfortunately, the arguments for requiring that psychology couch its generalizations in terms of the contents of mental states given in the passages quoted above are unpersuasive. Arguably, they are not even consistent with Fodor's own account of RTM. In these arguments, Fodor not only gives as examples of the types of generalizations he wants preserved by a scientific psychology, ones that elsewhere he claims are

not part of psychology, but, more importantly, he fails to recognize the strength of the "harmony" hypothesis of his own account of machine functionalism. While the inappropriateness of the examples he gives, is not, itself, damaging to his claim, the incorrect identification of the kinds of generalizations psychology can make makes his case seem more persuasive than it is. Among the generalizations that Fodor gives as examples of the types of generalizations that a scientific psychology should preserve are "seeing that a is F is a normal cause of believing that a is F" and "observations that many of the xs are F often contribute to the etiology of the belief that all the xs are F." But neither of these generalizations are of a type that can be preserved by Fodor's RTM. Seeing and observing are success verbs, which by Fodor's own admission, can have no place in the generalizations of a psychology that embraces the formality condition.

[A]ll sorts of states which look, prima facie, to be mental states in good standing are going to turn out to be none of the psychologist's business if the formality condition is endorsed. . . . [S]trictly speaking, there can't be a psychology of perception if the formality condition is to be complied with. Seeing is an achievement; you can't see what's not there. From the point of view of the representational theory of the mind, this means that seeing involves relations between mental representations and their referents.

(MS, p. 227-228)

No psychology that adheres to the formality condition, the requirement that only the formal properties of propositional attitudes can affect the workings of mental processes, can have a generalization that connects seeing or observing to believing because identifying something as an instance of seeing or observing is beyond the capabilities of such a psychology.

More importantly, however, it is simply not true that psychology

must advert to the content of propositional attitudes in order to preserve the kinds of generalizations an RTM-based psychology will be able to make. While it may very well be true that the generalizations of psychology cannot be formulated in the vocabulary of neurology or physics, given the perfect correspondence between content and formal properties that RTM posits, it is just not the case that the generalizations of psychology cannot be formulated in the vocabulary of machine functionalism. Any generalization about mental states couched in terms of their contents can be replaced by a generalization couched in terms of their formal/functional properties without any loss in explanatory power. Propositional attitudes can be the same in content only if they can be identified with relations to formally identical representations, and they "can be distinct in content only if they can be identified with relations to formally distinct representations" (MS, p. 227). If the formal properties of mental representations covary with the content of mental representations, then any generalization made on the basis of the one can be recast in terms of the other. For example, imagine a game in which all and only the heart cards are red cards. We can equally well make generalizations about the role of those cards in the game on the basis of their color or on the basis of their shape. We can make a generalization about those cards, say, that they are wild cards and can match any other card, on the basis of either their color or their shape. Either way the game will proceed in exactly the same fashion. There is no situation in which you will have drawn a wild-card under one of the generalizations but not under the other. The same is true about the generalizations of psychology. The notion of content, according to Fodor, that psychology

can use is such that all and only mental representations that have content P have formal properties F. Thus, any generalization that can be made on the basis of being a mental representation with content P, can also be made on the basis of being a mental representation with formal properties F. While it is trivially true that we cannot make generalizations about propositional attitudes that apply to them "in virtue of their contents" without adverting to content, we can make generalizations about propositional attitudes that do not advert to contents that are predictively and explanatorily equivalent to those that do advert to contents. The success that psychology will experience in predicting and explaining behavior will be the same regardless of whether it couches its generalizations in terms of the content of mental representations or the formal properties of mental representations.

Fodor's failure to provide a persuasive defense of his insistence that psychology individuate propositional attitudes on the basis of their content is particularly damaging to his project because, in addition to holding that psychology must couch its generalizations in terms of content, he also wants to hold that "it's got to be possible to tell the whole story about mental causation (the whole story about the implementation of the generalizations that belief/desire psychologies articulate) without referring to the intentional properties of the mental states that such generalizations subsume" (Psycho, p. 139). It would seem, at least on the face of it, that if we can tell the whole story about mental causation without referring to the semantic properties of mental representations, then it should not be necessary for the generalizations of psychology to advert to

content. Or, to put it another way, since Fodor's account does not adequately defend the commonsense psychological practice of individuating mental states by semantic contents we may suspect (and, indeed, I shall later argue) that his conception or construal of mental causation is itself inadequate.

In From Folk Psychology to Cognitive Science, Stephen Stich criticizes Fodor on just this point, claiming that Fodor cannot consistently hold both that the generalizations of psychology apply to mental states in virtue of their content, and that mental states have their causal powers in virtue of their formal properties. I am going to quote both Stich's criticism of Fodor and Fodor's response to Stich at some length because I think Stich is right on the mark when he points out the conflict between the two claims Fodor makes, and Fodor's failure to respond to the criticism is confirmation that Fodor recognizes neither the strength of his harmony hypothesis nor the weakness of the role of content in an RTM-based psychology. I begin with Stich quoting Fodor:

It is very characteristic of current versions of RTM -- indeed it is one of the few respects in which they differ significantly from classical formulations of the doctrine -- to be explicit in requiring that only non-semantic properties of mental representations can figure in determining which mental operations apply to them. . . . I take [this] to be part and parcel of the idea that mental processes are computational, and I take the idea that mental processes are computational to be among the central tenets of cognitive science.

The upshot is that one can do quite a lot of cognitive science without ever raising the foundational -- or, indeed, any -- issue about the semanticity of mental representations. In a certain sense you can do the whole theory of

mental processes without raising such issues, what with mental operations being computational and mental processes being causal sequences of mental operations. . . .⁸

How is it possible for Fodor to have it both ways, for him to urge both that cognitive generalizations apply to mental states in virtue of their content and that "only non-semantic properties of mental representations can figure in determining which mental operations apply to them"? One way to take the bite out of this apparent contradiction would be to endorse the correlation thesis which holds that differences in content are mirrored by differences in syntax. If this were true, then generalizations couched in terms of content would, so to speak, be co-extensive, with generalizations couched in terms of syntax. And although strictly speaking it might be their syntactic properties which account for causal interactions among mental state tokens, there would be no harm in talking as though semantic properties were causally relevant, since if they were, the system would behave in exactly the same way. . . .

We cannot get Fodor off the hook by assuming that he accepts the correlation thesis, however. For there are also places in which Fodor seems to reject the correlation thesis. . . . [We saw [in TSPG] Fodor endorsing the view that a pair of computers might be running the same machine language program, though one was simulating a chess game while the other was simulating the Six Day War. Thus a pair of mental sentence tokens might be syntactically or functionally identical though one is about Moshe Dayan and the other is about the king's bishop. Consider also the following quote:

Searle is certainly right that instantiating the same program that the brain does is not, in and of itself, a sufficient condition for having those propositional attitudes characteristic of the organism that has the brain. . . .^{9,10}

Stich's conclusion is not surprisingly, that "Fodor has fallen victim to an endemic ambiguity in discussions of content" and has, thus, become involved in endorsing an untenable position (FFPCS, p. 190).

In a footnote to a passage in Psychosemantics where Fodor has been discussing the claim that psychological laws should pick out the mental states they apply to by specifying the intentional contents of the states, even though the intentional properties of such states do not

figure in the causal mechanisms, (see Psycho, pp. 139-140), Fodor attempts to clarify his position in light of Stich's remarks.

In From Folk Psychology to Cognitive Science, Stich wrings his hands a lot about how I could hold that the counterfactual-supporting generalizations of psychology are uniformly intentional and also hold the 'solipsistic' principle that mental operations are computational (viz., formal/syntactic). "How is it possible for Fodor to have it both ways, for him to urge both that cognitive generalizations apply to mental states in virtue of their content and that 'only non-semantic properties of mental representations can figure in determining which mental operations apply to them'?" (FFPCS, 188).

But there's no contradiction. The vocabulary required to articulate the characteristic laws of a special science is -- almost invariably -- different from the vocabulary required to articulate the mechanisms by which these laws are sustained, the theory of the mechanisms being pitched -- to put it crudely -- one level down. So the typical laws of

psychology are intentional, and the typical operations of psychological mechanisms are computational, and everything's fine except that Stich has missed a distinction.

(Psycho, p. 166.)

In his reply, Fodor attempts to dismiss Stich's charge of inconsistency as simply being the result of Stich's failure to recognize when Fodor has been talking about the laws of psychology itself, and when he has been talking about the underlying mechanism that instantiates those laws. But Stich's criticism is not merely a matter of a distinction between the laws and the mechanisms that instantiate those laws. Fodor has claimed that it is possible to tell the whole story about mental causation without referring to the intentional properties of mental states, not that it must be possible to tell the whole story about the physical causation that underlies mental processes without referring to the intentional properties of mental states; and that only the non-semantic properties of mental representations can figure in determining

which mental operations apply to them, not that only the non-semantic properties of mental representations can figure in determining which computational operations apply to them. The difference between providing an account of the physical causation that must underly mental causation and providing an account of mental causation itself is not simply a matter of vocabulary. If Fodor maintains that we can tell the whole story about mental causation without adverting to content, then there is a real question as to how, without being inconsistent, he can also hold that the generalizations of psychology apply to mental states in virtue of their content.

Although Fodor, himself, fails to parry Stich's charges of inconsistency, and to defend adequately his insistence that psychology individuate propositional attitudes on the basis of their content, there is a sense in which the sort of RTM-based machine functionalism Fodor is interested in constructing does need (some) notion of content. It is not however, as Fodor claims, that the generalizations of a machine functionalist psychology apply to mental states in virtue of their content. Rather, it is that if what we want is a machine functionalist psychology its generalization will have to apply only to things that have content. The point that needs to be made about content is not that psychology cannot make the generalizations it needs and/or wants without adverting to content, but that psychology cannot determine which states ought to be candidates for subsumption by its generalizations without appealing to content. The issue is that the domain of things to which the generalizations of psychology ought to apply to cannot be properly circumscribed unless psychology appeals to the notion of content. Content is the only means by which psychology

can distinguishing those states that are propositional attitudes, and, therefore, within the domain of psychology, from those states that are not.

An RTM-based psychology will be dependent on content for the determination of which states fall within its domain in two closely related ways. First, although there is, by hypothesis, a one to one correspondence between the semantic properties of propositional attitudes and the formal properties of propositional attitudes, a correspondence which allows propositional attitudes to be identified on the basis of their formal properties, the possibility of identifying a propositional attitude on the basis of its formal properties is dependent upon there having been a prior "discovery" of what the formal specification of that propositional attitude is. In order to determine what the formal properties are that correspond to the semantic properties of a propositional attitude psychology first has to have picked out that propositional attitude. Psychology cannot do that initial identification of propositional attitudes on the basis of formal properties, because there are more states that have formal properties than there are states that have semantic properties.

The situation is not unlike identifying tigers on the basis of their stripes. Let us assume that there is a one to one correspondence between individual tigers and stripe patterns (it is my understanding that zoologists claim that this correspondence is empirically true: stripe patterns are the "finger prints" of tigers). If there is such a correspondence, then any given tiger can be reliably identified solely on the basis of its striping pattern. However, any given striping pattern will not reliably identify a tiger. Some striping patterns

will reliably identify individual zebras rather than individual tigers (assuming that zebras, like tigers, each have a unique stripe pattern "finger print," another claim that zoologists do make). There is nothing about striping patterns per se that pick out tigers rather than zebras. The possibility of identifying tigers on the basis of their striping patterns is dependent upon having previously determined what striping patterns correspond to individual tigers, and to have done that one has to have been able to pick out tigers on some basis other than their striping patterns. Just as not every striping pattern identifies a tiger, not every internal, formally specifiable mental state type identifies a propositional attitude. The possibility of psychology identifying propositional attitudes on the basis of their formal properties is dependent upon having previously determined what formal properties correspond to semantic properties, and to have done that psychology has to have been able to pick out propositional attitudes on some basis other than their formal properties. In the initial phase when the correspondence between formal properties and semantic properties of propositional attitudes is being identified, content is the only means available to psychology for picking out propositional attitudes.

Second, even if psychology has done the initial determination of which formal states it is interested in on the basis of content, it still cannot give up using content: psychology still needs content in order to determine which particular tokens of a (previously identified) formal state type actually are candidates for subsumption. The generalizations of psychology need to advert to content because if they do not use content to pick out those states they are to subsume, states

that we do not want to consider propositional attitudes, in spite of their having the right formal properties, will also be subsumed by those generalizations. Even if psychology has determined what formal properties are paired with and co-vary with the semantic properties of propositional attitudes, there is no guarantee that every state that has those formal properties is a propositional attitude. The correspondence between formal properties and semantic properties holds only for propositional attitudes, and not every state that has the specified set of formal properties is necessarily a propositional attitude. The same is true in the case of tigers. The correspondence between individual tigers and striping patterns holds only for (genuine) tigers, and not everything that has the correct striping pattern is a tiger. I can paint a pattern of stripes on a cougar that is identical to the pattern that identifies some particular tiger, but that does not mean that the cougar with stripes is a tiger. Neither does it mean that we cannot reliably identify tigers on the basis of their stripes. Striping patterns are a reliable way of identifying which tiger is in front of us.

If psychology were to identify mental states on the basis of their formal properties, rather than their semantic properties, it would be identifying things having a certain sort of formal organization as a natural kind for the purposes of psychology. It would be, in effect, claiming that domains consisting of creatures having certain sorts of formal states are natural from the point of view of psychology. However, just as type physicalism was pronounced untenable because it defined the domain of psychology in a way that ran counter to intuitions (as the domain of creatures having a certain neural

organization), so too would machine functionalism be pronounced untenable if it were to identify the domain of psychology as those creatures having a certain formal organization, because having the right formal organization is not sufficient for being included in the domain of a psychology of propositional attitudes.

A scientific psychology that identified mental states on the basis of their formal properties would, in effect, make the domain of psychology too large: it would permit the application of the generalizations of psychology to things that we feel should not be subsumed by a theory of the mental causation of behavior by beliefs and desires. For example (to borrow one of Fodor's from MS), imagine a machine that answers baseball questions, in the sense that if you ask it a question about baseball, it displays the answer on a VDT. If you ask it "Who had the most wins by a National League pitcher since Dizzy Dean?", it will respond "Robin Roberts, who won 28, had the most wins by a National League pitcher since Dizzy Dean." Even if the machine has been so constructed that the internal state it is in, when it replies "Robin Roberts . . ." has the same formal properties as the state you would be in were you to respond "Robin Roberts . . .", it would nevertheless be inappropriate to use the generalizations of psychology to explain why the machine responded as it did. The problem is that the machine isn't really answering questions about baseball. We can interpret the display on the VDT as being about Robin Roberts, but it is possible that we could equally as well interpret it as being about the Khmer Rouge. But, when you say "Robin Roberts . . ." you are referring to Robin Roberts, and you are talking about baseball. What makes it appropriate to apply the generalizations of psychology to you,

but not to the machine, is that your internal state has a fixed interpretation, one that is fixed by something about the causes of your being in a state that has the specified formal properties, "something about the way [Robin Roberts] and [you] are embedded in the world" (MS, p. 233). The internal state of the machine does not carry any information, because the interpretation of that internal state is not fixed. The VDT display is about whatever we decide to take it as being about.

If we identify propositional attitudes by their formal properties, then any system that operates on formal objects in the appropriate fashion, including the question answering machine, will be one to which we can legitimately apply the generalizations of psychology, regardless of whether or not they are representational systems. But, and I think Fodor is right in this regard, the natural domain of a psychology of propositional attitudes seems to be "something like the set of (real and possible) information processing systems" (SSA, p. 9). The generalizations of commonsense belief/desire psychology that involve propositional attitudes are clearly meant to apply only to states that have semantic content. If the natural domain for psychological theorizing is the set of information processing systems, then it must be that that which is being manipulated is information for the generalizations of psychology to apply. Contrary to Stich's claim, we can hold a correlation thesis, and "endorse the view that a pair of computers might be running the same machine language program, though one is simulating a chess game while the other is simulating the Six Day War" (FFPCS, p. 189). What makes one a simulation of a chess game while the other is a simulation of the Six Day War is our

interpretation of what they are doing. Neither machine is in a state that carries any information. Neither machine is in a state that falls within the domain of psychology, so the fact that we can choose to interpret one as simulating a chess game, while we interpret the other as simulating the Six Day War, has no bearing on the correlation between the formal properties and the semantic properties of those states that do fall within its domain, viz., propositional attitudes. If we individuate propositional attitudes by the semantic properties of mental representations, while maintaining that propositional attitudes are relational states that have both formal and semantic properties, we insure that it is only systems that manipulate semantically interpreted symbols that will fall within the domain of psychological theorizing. We need to hold on to individuating mental states by their content, rather than using their formal properties, in order to restrict the domain of psychology to information processing systems.

If we grant Fodor what I have argued we should not, and allow that it may be possible for a psychological theory to do away with content in an "after the fact" sort of way, Fodor might be able to defend his claim that it is possible to tell the whole story about mental causation without adverting to the intentional properties of mental representations. If psychology has already determined the precise correlation between formal and semantic properties in propositional attitudes, and it has ensured that the only states it is dealing with are those that actually are propositional attitudes, then one might be able to argue that propositional attitudes can be identified without adverting to content. However, it seems that there would be little point to making such a claim. It's not that psychology can do without

content, but only that it is harmless to talk about the causal relations of an uninterpreted calculus of functional states without referring to the intentional properties of mental representations, once content has done the necessary discriminations.

It should be noted that the notion of content that will remain in an RTM-based scientific psychology need not be identified with the sort of content that commonsense psychology uses in its generalizations. It is not that narrow content is used by both commonsense and scientific psychologies. Narrow content co-varies with the formal properties of mental states, and that fact explains why commonsense psychological generalizations which do apply to mental states in virtue of their (narrow) content are useful in predicting formal-property-caused behavior. But, since the only role that content has in an RTM-based scientific psychology is that of identifying which states fall within the domain of psychology, i.e., have semantic properties, what construal we give to that notion is immaterial. A computational psychology cares only that a state have content, not what the content of the state is. By couching its generalizations in terms of the narrow semantic properties of formally specified states, a computational psychology will be able both to limit the things its generalizations apply to to those states that have content, and to predict and explain behavior accurately in virtue of the fact that narrow content and formal properties co-vary. Nevertheless, we must not allow the fact that the generalizations are couched in terms of narrow content delude us into believing that narrow content has a predictive or explanatory role to play in the sort of computational psychology Fodor's RTM permits.

Although a scientific psychology based on Fodor's RTM will, I have

argued, need to couch its generalizations in terms of some notion of content (in order to ensure that its generalizations are applied only to those things having semantic properties), such a theory will not constitute a vindication of commonsense folk psychology and mental causation. In an RTM-based psychological theory, the role of content is simply too feeble for Fodor to claim that such a theory provides an explication of the causation of behavior by mental states. Mental causation, ordinarily understood, is not simply the causation of behavior by states or things that have semantic properties -- not even if those states' identity conditions are essentially semantic -- but by states whose semantic properties are explanatorily relevant to the causal effects on behavior that they have. Mental causation is the causation of behavior by states whose semantic properties are, in some way, responsible for the causal roles the states play in the life of the individual. Fodor's RTM simply does not allow for the semantic properties of mental states to be relevant in any such fashion. In Explaining Behavior, Fred Dretske criticizes Fodor's theory on precisely these grounds:

If a symbol's meaning is correlated with the symbol's physical properties -- if the semantics of symbols is faithfully reflected in their syntax, plus or minus a bit, as Fodor (1980) [MS] puts it -- then meanings may turn out to be predictively useful without being explanatorily relevant. If I know that the high note is the only passage in the aria that has a certain meaning, I can predict that the glass will shatter when a passage with a certain meaning is sung. The fact that the words have this meaning, however, will not explain why the glass shattered. Rather, a sound's having a certain meaning will co-occur with something else (that sound's having a sufficient pitch and amplitude) that does explain this physical effect. It may even turn out, if the semantic features co-occur often enough with the right syntactic features, that useful generalizations (useful for predictive purposes) can be formulated in semantic terms. It may even be useful, perhaps even essential for methodological purposes, to catalog or index

the causally relevant formal properties of our internal states in terms of their causally irrelevant meanings But this, even if it turns out to be a fact, will not transform meaning into a relevant explanatory notion. If beliefs and desires explain behavior in this way, then what we believe and desire (the content of our beliefs and desires), however useful it might be for predicting what we are going to do, will not be a part of the explanation of what we do. . . . On this account of the explanatory role of meaning, meaning would be as relevant -- i.e., wholly irrelevant -- to explanations of human and animal behavior as it now is to explanations in the science of acoustics.

(EB, p. 81)

Granting that the semantic properties of mental states cannot themselves be causally efficacious, what we need is an account of the semantic properties of mental states in which the fact that they have the semantic properties that they do is somehow relevant to why they have the causal roles that they do. Co-variation of content and causal role is a necessary part of a theory that can vindicate the commonsense notion of mental causation, but by itself co-variation is not sufficient for that vindication. An account of mental causation requires a theory of the semantic properties of mental states which not only satisfies the co-variation requirement, but in which the fact that they have the semantic properties they do is relevant to why they have the causal roles they do. For example, one might construct a theory in which the mechanism that determines or fixes the semantic properties a state has is also relevant to the determination or fixing of the causal role that the state will play in the life of the individual (this is, to a very rough approximation, the approach that Dretske takes). A notion of content that permits merely the covariation of content and causal role has only achieved half of what is necessary for the vindication of the commonsense practice of explaining what people do in terms of what they believe and desire.

While I have focused on Fodor's theory as he presents it in "Methodological Solipsism" and various other articles that date from approximately the same time, it should not be thought that the revised notion of content that Fodor presents in Psychosemantic (about which I will have more to say in Chapter 5) represents an improvement in terms of the explanatory relevance of content. Fodor alters the way he feels the notion of content should be construed in psychological contexts because of his desire to maintain, simultaneously, an extensional criterion of content identity, and the supervenience of mental states (which have their semantic properties essentially) on functional states (see Ch. 2 of Psychosemantics, in particular pp. 44-53). The revised notion of content that Fodor suggests should be adopted by psychology may be an improvement over opacity if we are simply concerned with devising a notion of content that agrees with our pretheoretic notions of content identity, and which is compatible with the formality condition of the machine functionalist model of the mind. But the revisions to the notion of content that Fodor suggests in Psychosemantics have no effect on the explanatory role, or lack thereof, that content plays in RTM. While Fodor's new notion of narrow content still co-varies with causal powers, it also still has no role to play in the determination of the causal powers of mental states.

It may be objected that Fodor's goal, both in "Methodological Solipsism" and in Psychosemantics, was simply to construct a theory of the semantic properties of propositional attitudes that was compatible with the supervenience of propositional attitudes on brain states. If Fodor's aim was simply to construct a theory of the semantic properties of propositional attitudes that permits the concomitance of content and

causal powers, then it would be unfair to criticize him for failing to do more than that. While it is certainly the case that Fodor's immediate project was to construct such a semantic theory, that was not the ultimate project Fodor had in mind. Fodor's contention has been that a semantic theory that permitted the co-variation of content and causal powers would vindicate, or at least partially vindicate, commonsense belief/desire psychology. But the problem is that concomitance alone cannot even partially vindicate commonsense belief/desire psychology. The vindication of commonsense psychology requires that content be relevant, in some fashion, to causal powers. A theory that succeeds in vindicating commonsense psychology must, as Fodor argues, be one in which content and causal powers do co-vary, but co-variation by itself cannot vindicate commonsense psychology. The vindication of commonsense psychology also requires that the concomitance of content and causal role be explained or accounted for, and that content be explanatorily relevant to why propositional attitudes have the causal powers that they do. Fodor's theory is unsatisfactory because it fails to fulfill these later requirements.

Fodor has argued that a scientific psychology can be built from the practices of commonsense belief/desire psychology. His position is that if we construe propositional attitudes as relations to representations (where representations are symbols having both formal and semantic features), and abstract across propositional attitudes on the basis of the opaque or narrow content of representations, we will have a science of psychology that vindicates the use of generalizations that apply to propositional attitudes in virtue of their content, by providing an account of how the causal interactions of propositional attitudes

contrive to respect their content relations. In many respects Fodor's account of propositional attitudes is intuitively appealing, and I believe he is correct in maintaining that propositional attitudes ascriptions must be construed narrowly or individualistically if we are to have any hope of accounting of mental causation. Nevertheless, his theory ultimately, is inadequate to the demands of laying the foundation of a scientific psychology of propositional attitudes. Given the peripheral role that content actually plays in such a psychology RTM cannot be seen as a vindication of either commonsense psychology or mental causation. A satisfactory account of mental causation must give content a more robust role than RTM does. Content and causal role must not merely co-vary, as they do in Fodor's RTM. Content must be relevant to the causal role that a state plays in the life of the individual if mental causation and commonsense belief/desire psychology are to be vindicated.

CHAPTER 3

THE "COMMUNITARIAN" NOTION OF CONTENT OF TYLER BURGE

Tyler Burge claims that the notion of narrow content such as Fodor endorses, what Burge calls 'individualistic' content, is "hopelessly oversimplified as a philosophical explication of ordinary mentalistic notions" and "cannot be seen as providing a means of individuating ordinary ('non-transparent') attributions of content" (OB, p. 113). In this chapter I want to defend the notion of narrow content and its relevance to a scientific psychology against Burge's claim that mental state content is essentially a function of the socio-linguistic environment. I will argue that neither Burge's claim that individualism need not be adopted by a science that endorses materialism, nor his claim that individualism is not consistent with our best contemporary theories of vision is borne out by the evidence at hand. Further, I will argue that Burge's own construal of content is implausible because it unnecessarily complicates psychological theorizing, conflicts in important ways with our intuitions about mental state attributions, and is incompatible with the existence and evolution of language.

Burge claims that narrow content cannot be used to individuate or taxonomize propositional attitude content because "identifying a person's mental states depends upon the nature of his physical environment -- or on the specification, by his fellows, of the nature

of that environment" (OB, p. 98). The content of a person's mental state is dependent upon and varies with both the nature of his physical environment and the activities of those of his socio-linguistic community. "[O]ne cannot explicate what propositional attitude contents a person has by taking into account only facts about him that are non-intentional and individualistic" (OB, p. 112). The content of an individual's intentional state is derived from the world, but what there is in the world that can be the source of that intentional content is determined by the linguistic practices of the community in which the individual exists. The content of a propositional attitude is a social construct. The nature of the socio-linguistic environment is the determinant of propositional attitude content. Burge's claim is that the predictions and explanations of behavior constructed by commonsense psychology advert to propositional attitudes whose contents are functions of the socio-linguistic environment in which the individual is embedded. A scientific psychology that embraces functionalism and the individuation of propositional attitudes on the basis of narrow content will not capture commonsense belief/desire psychology, as Fodor claims it will, because the notion of narrow content is alien to commonsense psychology. According to Burge, commonsense psychology is fundamentally non-individualistic: the attribution of propositional attitudes and the prediction and explanation of behavior by commonsense belief/desire psychology are inextricably bound to a socially constituted notion of content. Burge claims that there is no such thing as a 'psychological state in the narrow sense' in commonsense belief/desire psychology.

In arguing that an individualistic notion of content is both

inadequate as an explication of the commonsense notion of content and alien to commonsense psychology, Burge is not arguing that commonsense belief/desire psychology is inappropriate for use as the paradigm for a scientific theory of the mind. Instead, Burge claims that the wide, non-individualistic, or communitarian, notion of content that he argues is integral to the commonsense psychology model can be part of a rigorous scientific theory of the mind. Burge claims that the individuation of intentional content on the basis of a non-individualistic criterion of content can play a systematic role in scientific theorizing (IP, pp. 6-7). It should be emphasized that the scientific psychology that Burge envisions being built on commonsense psychology and a communitarian notion of content, is meant to be the same sort of methodologically rigorous scientific theory that Fodor envisions.

In taking psychology as it is, I am assuming that it seeks to refine, deepen, generalize and systematize some of the statements of informed common sense about people's mental activity. . . . Psychology accepts that people remember events and truths, that they categorize objects, that they draw inferences, that they act on beliefs and preferences. And it attempts to find deep regularities in these activities, to specify mechanisms that underly them, and to provide systematic accounts of how these activities relate to one another.

(IP, p. 8)

Given Burge's view, there is no justification for adopting a narrow notion of content for use in psychological explanations: narrow content, according to Burge, is neither part of the ontology of commonsense psychology nor required on methodological grounds by a scientific psychology. Narrow content is revisionistic with respect to the commonsense notion of content and superfluous to a scientific psychology.¹

While most of Burge's writings on the issue of intentional content have been devoted to defending his claim that the commonsense psychology notion of content is a wide, communitarian notion of content, he does at least briefly defend his claim that such a notion of content can be incorporated into a rigorous, systematic scientific theory, as well as present a more general argument against individualism (though not one that supports, in particular, the communitarian notion of content that Burge advocates) that centers on visual perception. I will first briefly look at these latter arguments, indicating why he has, in one case, failed to demonstrate that a communitarian notion of content is compatible with a scientific theory, and, in the other case, has failed to provide a general argument against individualism, before turning to a more detailed examination of his argument that the commonsense notion of content is a wide, communitarian notion.

I. BURGE'S ARGUMENTS AGAINST AN INDIVIDUALISTIC NOTION OF CONTENT

In "Individualism and Psychology"² Burge presents two arguments against individualism that are substantially different from his more typical arguments that are based on language. First, he criticizes a number of arguments that have been given in support of the claim that, contrary to Burge's claim, psychology must individuate the states that it appeals to in its explanations on the basis of an individualistic criterion of content (type-)identity. Second, he presents what he

claims is a general argument against individualism based on visual perception, and which he claims supports and is supported by David Marr's theory of visual perception. I will begin by looking at his two "metaphysical" points against the arguments for individualism, and then consider his general argument against individualism based on visual perception.

The first "metaphysical" point that Burge makes concerns the extent to which materialism supports, or does not support, individualism. The argument is sometimes made that if materialism about mental states and events is true, then mental states and events must supervene on the underlying physical, i.e., brain, states and events: the determinants of behavior must supervene on brain states, so if intentional states are to be determinants of behavior, then they must supervene on brain states as well. Since individuation of intentional states on the basis of a wide, communitarian criterion of content results in an individuation in which intentional states do not supervene on brain states, such a criterion of content identity cannot be true if materialism is true. An individualistic criterion of content identity, it is claimed, is the only criterion that results in intentional states supervening on brain states. Thus, individualism must be true if materialism is true.

Burge argues that this argument from the truth of materialism to the truth of individualism is faulty because "what supervenes on what has at least as much to do with how the relevant entities are individuated as with what they are made of" (IP, p. 13). Burge claims that materialism supports individualism only if we antecedently stipulate that mental states must supervene on brain states, but whether or not

mental states supervene on brain states has to do with how we individuate mental states. Materialism and supervenience are separate doctrines, and, according to Burge, it is only if we accept supervenience in addition to materialism that materialism supports individualism. Burge claims that we can accept materialism, without accepting supervenience, and gives two examples of situations in which we individuate physical events in such a way that they do not supervene on their "constituent" physical events.

[C]onsider the Battle of Hastings. Suppose that we preserve every human body, every piece of turf, every weapon, every physical structure and all the physical interactions among them, from the first confrontation to the last death or withdrawal on the day of the battle. Suppose that, counterfactually, we imagine all these physical events and props placed in California (perhaps at the same time in 1066). Suppose that the physical activity is artificially induced by brilliant scientists transported to earth by Martian film producers. The distal causes of the battle have nothing to do with the causes of the Battle of Hastings. I think it plausible (and certainly coherent) to say that in such circumstances, not the Battle of Hastings, but only a physical facsimile would have taken place. I think that even if the location in Hastings were maintained, sufficiently different counterfactual causal antedecedents would suffice to vary the identity of the battle. The battle is individuated partly in terms of its causes. Though the battle does not supervene on its physical constituents, we have little hesitation about counting it a physical event.

(IP, p. 14)

Burge grants that the way we individuate battles is probably tied up with the intentional states of the participants (the soldier behaved the way he did because he believed he was defending his home against the Norman invaders), but claims that the same point can be made in situations that have nothing to do with intentional considerations:

Consider the emergence of North America from the ocean. Suppose that we delimit what count as constituent

(say, micro-) physical events of this larger event. It seems that if the surrounding physical conditions and laws are artfully enough contrived, we can counterfactually conceive these same constituent events (or the constituent physical objects' undergoing physically identical changes in the same places) in such a way that they are embedded in a much larger land mass, so that the physical constituents of North America do not make up any salient part of this larger mass. The emergence of North America would not have occurred in such a case, even though its "constituent" physical events were, in isolation, physically identical with the actual events. We individuate the emergence of continents or other land masses in such a way that they are not supervenient on their physical constituents. but such events are nonetheless physical.

(IP, p. 14)

Certainly, there is little question that we often distinguish between physical objects or events on the basis of factors that have nothing to do with their physical properties. We routinely distinguish between a genuine United States twenty dollar bill and a counterfeit twenty, and between a Cadillac that belonged to Elvis Presley and an identical one that did not belong to him. Burge is correct in pointing out that materialism per se does not warrant individualism -- it does so only on the assumption of supervenience. However, and this brings us to Burge's second "metaphysical" point, materialism does not warrant individualism in a scientific psychology only if it is possible to have a scientific field that embraces materialism and meets the methodological constraints on scientific theorizing without embracing supervenience. If scientific fields that embrace materialism must, on methodological grounds, also embrace supervenience, then materialism does warrant individualism in a scientific psychology.

This second "metaphysical" point that Burge makes is that the identity of states and events and their causal powers can be affected by changes that do not affect their constituent physical properties.

He wants to argue that various fields, including scientific ones, that routinely accept that the states and events they are concerned with are physical states and events that interact with one another on the basis of their physical properties, nevertheless do not individuate those states and events in a fashion that makes them supervene on their constituent physical features: According to Burge, scientific fields can and do embrace materialism without embracing supervenience. Burge claims that his example from geology is an example of a scientific field that recognizes that the events and states it is concerned with are physical and interact on the basis of physical properties, but, nevertheless individuates those states and events on the basis of relational (wide) properties.

The example from geology provides a useful countermodel. It shows that one can accept the causal principles and thereby experience no bewilderment whatsoever in rejecting individualism. A continent moves and is moved by local impacts from rocks, waves, molecules. Yet we can conceive of holding constant the continent's peripheral impacts and chemically constituent events and objects, without holding identical the continent or certain of its macro-changes -- because the continent's spatial relations to other land masses affect the way we individuate it. . . .

The intended analogy to mental events should be evident. We may agree that a person's mental events and behavior are causally affected by the person's environment only through local causal effects on the person's body. Without the slightest conceptual discomfort we may individuate mental events so as to allow distinct events (types and tokens) with indistinguishable chemistries, or even physiologies, for the subject's body. Information from and about the environment is transmitted only through proximal simulations, but the information is individuated partly by reference to the nature of normal distal stimuli. Causation is local. Individuation may presuppose facts about the specific nature of a subject's environment.

(IP, pp. 16-17)

The problem with Burge's analogy is that the emergence of the North

American continent is not an event type over which geology generalizes. Nothing is the emergence of the North American continent but the emergence of the North American continent. In order for something to be the emergence of the North American continent it must involve the North American continent, but there is only one thing that is the North American continent. So there is only one event that is a token of the type "emergence of the North American continent." Geology generalizes over such things as rock types, plates, and currents in the mantle. It explains and predicts geological events in terms of, for example, plate drift. Geology does not explain or predict anything on the basis of being the North American continent. Being the North American continent or the emergence of the North American continent is no more a state or event that geology appeals to and generalizes over than is being my grandmother's diamond engagement ring. Geology just does not individuate those states and events over which it generalizes in the wide fashion that Burge attributes to it.

Leaving aside the fact that Burge has failed to demonstrate that there are scientific fields that individuate widely while still maintaining that the states and events they are concerned with are physical and interact solely on the basis of their physical properties, the general point that Burge is arguing for is that a field can endorse materialism but still individuate widely if it is also willing to make it generalizations sensitive to those same wide considerations.

Where intentional psychological explanation is itself causal, it may well presuppose that the causal transactions to which its generalizations apply bear some necessary relation to some underlying physical transactions (or other). Without a set of physical transactions, none of the intentional transactions would transpire. But it does not follow that the kinds invoked in explaining causal interactions among intentional

states (or between physical states and intentional states -- for example, in vision or in action) supervene on the underlying physiological transactions. The same physical transactions in a given person may in principle mediate, or underly, transactions involving different intentional states -- if the environmental features that enter into the individuation of the intentional states and that are critical in the explanatory generalizations that invoke those states vary in appropriate ways.

(IP, p. 17)

In other words, if we are willing to make our generalizations sensitive to "wide" considerations, then we can individuate states and events widely without running into problems with accepted principles of causation. Thus, for example, geology could individuate in such a way that the emergence of the North American continent was an event type, if the generalizations that geology used were ones that distinguished between being the North American continent and being some physically identical but numerically distinct continent. Geology would have one generalization for events involving North America, one for events involving North America₁, one for events involving North America₂, . . ., one for events involving North America_n. Such a system would be possible, but, without further argument on Burge's part, it seems untenable to claim that such a system could be part of a scientific field concerned with explicating underlying causal mechanisms. The difference between being an event that involves North America and being an event that involves North America_n, is not a difference that is independently identifiable, since the difference between being North America and being North America_n is not, itself, independently identifiable. Geology certainly does not use such a system of state and event individuation for its generalizations. Continent identity is irrelevant to the generalizations that geology

formulates and utilizes.

Contrary to Burge's claim, he has failed to demonstrate that there is a scientific field that individuates widely yet maintains that the states and events it is concerned with are physical, and interact solely on the basis of their physical properties. Obviously, his failure to show that there is such a field does not mean that there can not be such a scientific field. Nevertheless, given that there seems not to be an existent scientific field that accepts materialism while denying supervenience, to assume without a convincing argument that there can be such a scientific field is unwarranted.

The general argument against individualism that Burge presents in "Individualism and Psychology" is based on the claim that psychological theories of visual perception, such as Marr's, are non-individualistic theories, because they make essential reference to distal stimuli and "essential assumptions about contingent facts regarding the subject's physical environment" (IP, p. 29). "Since perceptual processes provide the input for many higher cognitive processes, it is reasonable to think that if the theory of vision treats intentional states non-individualistically, other parts of cognitive psychology will do likewise" (IP, pp. 25-26). Thus, according to Burge, we have good reason to believe that cognitive psychology in general will be non-individualistic because vision is non-individualistic. While it is certainly true of Marr's theory of vision, and of any non-idealist theory of perception, that essential reference is made to the external environment, that fact alone does not constitute an argument against individualism. In order for the fact that a theory of vision makes essential reference to the external environment to constitute an

implicit denial of individualism, it must be that the sort of reference to the external environment that the theory makes is a sort of reference that is prohibited by individualism. However, if we consider the sort of reference to the external world that Burge claims a theory of vision will make, it becomes clear that, in fact, individualism is not only compatible with such reference, but makes similar reference to the external environment.

Burge asks us to imagine the following situation. P is a person who lives in a world that has both O's (small shadows of a certain shape on a gently countoured surface) and C's (small cracks of a size and shape similar to O's), although C's occur only very rarely relative to O's in P's world. We are told to assume that O* is the visual state P is in when P visually perceives an instance of O, and that states of type O* represent O's as such, as O's. "P's visual representations are properly explained and specified as shadow representations" (IP, p. 42). Under ideal circumstances P could visually distinguish between O's and C's, but under less than ideal circumstances, the only circumstances under which P has, in fact, been confronted with C's, P has no means of discriminating instances of C from instances of O. In less than ideal circumstances instances of C give rise to visual states of type O*. "On such occasions, P mistakenly sees an instance of C as an O" (IP, p. 41).

Burge then asks us to imagine a counterfactual situation in which P is physically and functionally identical in all respects, non-intentionally described, to P in the actual situation, but in which the environment in which P lives has been changed. In the counterfactual environment

[t]here are no instances of the relevant shadows visible to P; and the laws of optics differ in such a way that P's physical visual stimulations (and the rest of P's physical makeup) are unaffected. Suppose that the physical visual stimulations that in the actual case are derived from instances of O -- at the relevant sort of shadows -- are counterfactually caused by and explained in terms of C's, relevantly sized cracks. Counterfactually, the cracks take the place of the shadows. On the few occasions where, in the actual case, P misperceives [cracks as shadows]³ P is counterfactually confronted with cracks; and the optical circumstances that lead to the visual impressions on those occasions are, we may suppose, normal for the counterfactual environment. On such counterfactual occasions, P would be visually representing small cracks as small cracks. P would never have visual representations of the relevant sort of shadows. One can suppose that even if there were the relevant sort of shadows in the counterfactual environment, the different laws of optics in that environment would not enable P ever to see them. But since P's visual states would be the normal products of normal processes and would provide as good empirical basis for learning about the counterfactual environment as P has for learning about the actual environment, it would be absurd to hold that (counterfactually) P misperceives the prevalent cracks as shadows on gently contoured surfaces. Counterfactually, P correctly sees the cracks as cracks. So P's intentional perceptual states differ between actual and counterfactual situations.

(IP, pp. 42-43)

Burge's contention is that a theory of vision will claim that P has visual representations of different things in the actual and the counterfactual situation, while individualism will claim that P's visual states were representations of the same thing in both the actual and the counterfactual situations. Consequently, according to Burge, individualism conflicts with our best theories of visual perception, and we should prefer holding on to our theories of vision to holding on to individualism.

Even granting Burge that in the actual situation P visually represents shadows as shadows, and in the counterfactual situation P visually represents small cracks as small cracks, a claim which is, at

best, questionable, the example still fails as an argument against individualism. Individualism does not deny that the external environment plays an essential role in the determination of intentional state content. Fundamentally, individualism is a claim about what the relationship is between the content of an internal state and the external environment that is ultimately responsible for the internal state having the content it does. Individualism claims that only those differences in the external environment that can make a (causal) difference to the individual can affect the content of the individual's mental state. According to individualism, the content of an internal state is to be determined by looking at what in the environment can and cannot influence whether or not the individual is in that particular internal state. If some aspect of the external environment cannot causally affect the individual, then that aspect of the environment will not play a part in determining the content of the individual's internal state.

Individualism, by appealing to what in the environment can and cannot causally affect the internal state of the individual, is itself making essential reference to the distal causes of the individual's internal states, and is making essential assumptions about contingent facts concerning the subject's physical environment, more specifically, it is making essential assumptions about the natural laws that specify what things can and cannot causally affect the individual. Burge's example involves the alteration of the causal laws that determine what things in the environment can and cannot causally affect the internal state of the individual. It is, precisely, by changing what can and cannot causally affect the internal state of the individual that,

according to individualism, one can change the content of that internal state. While it is more typical to imagine altering the way the environment causally affects the individual by imagining the sensory organs of the individual to be different than they actually are, one can also change the way the environment causally affects the individual by changing natural laws, such as the laws of optics. In his example, Burge has to change not only what there is in the environment (C's verses O's), but the laws of optics as well in order to make it even plausible that a theory of vision would conclude that the visual state is a representation of something different in the counterfactual situation than in the actual situation. But, such a change in the laws of optics -- in the way the individual causally interacts with the environment -- is precisely the sort of change that individualism claims alters the content of an internal state. Individualism, like theories of vision, will under such circumstances claim that the visual state of the individual in the counterfactual situation represents something different than it does in the actual situation.

The similarity between the way a theory of vision such as Marr's determines the representational content of a visual state and the way individualism determines content of a mental state bears emphasizing. When Marr is discussing how it was determined that a set of zero-crossing segments constitutes a representation of an edge, Marr says "the theory starts with the observation that physical edges produce roughly coincident zero-crossings in channels of neighboring sizes" (Marr, p. 70). In other words, the theory concludes that a set of zero-crossing segments is a representation of an edge by determining what things in the world can cause those sorts of zero-crossing

segments. Individualism employs the same technique. Individualism claims that what an internal state is a representation of is determined, in part, by what things in the world can cause that state in the individual. By changing the environment and how the individual can be causally affected by the environment, one will change not only what a theory of vision has to say about the representational content of a visual state, but what an individualistic theory of mental state content has to say about the content of the mental state.

II. BURGE'S ARGUMENTS FOR A SOCIAL OR "COMMUNITARIAN" NOTION OF CONTENT

Burge bases his position on the social character of propositional attitude content on a number of related thought experiments which he claims demonstrate that we can change the content of a subject's propositional attitudes, while holding the subject, non-intentionally described, constant, by changing the linguistic practices of the community in which the individual exists. By changing what it is in the world that a word applies to, we can change the content of an individual's mental state even if that change in the relation between the word and the world does not have any overt effect on the individual. Burge's claim is not simply that the attributed propositional attitude content varies (or that the object in the world that the propositional attitude is about varies), but that the content of the propositional attitude itself varies. "It is to be re-emphasized that the variations in propositional attitudes envisaged are

not exhausted by variations in the entities to which the individuals' mental contents are related. The contents themselves vary" (OB, pp. 99-100). Burge is arguing that the content of an intentional states is determined by the nature of the world as that nature is individuated or identified by the linguistic practices of the community. By changing the ontological commitments of a language, you can change the contents of an individual's mental states, even if neither the world, nor the individual, nor the relation between the individual and the world has been altered. If by varying the environment while holding fixed the individual's physical, behavioral, phenomenistic, and functional histories, we vary the content of the individual's propositional attitude, then any theory that individuates propositional attitudes using a criterion that ignores the surrounding context will not provide an adequate account of the commonsense psychology practice of predicting and explaining behavior on the basis of mental state content because it will have failed to capture the essential social nature of propositional attitude contents.

Each of the thought experiments involves the attribution of a propositional attitude to an individual using a that-clause containing a term that the individual either misunderstands or incompletely understands. Burge maintains that if an individual meets minimal socially accepted standards for using a term -- natural kind or otherwise -- there is no basis in intuition for the denial that those terms, understood as they ordinarily are, express the way the individual thinks about the relevant stuffs/things, notwithstanding the fact that he does not fully know the socially determined meanings of the terms.

It does not follow from the assumption that the subject thought that a word means something that it does not (or misapplies the word, or is disposed to misexplain its meaning) that the word cannot be used in literally describing his mental contents. It does not follow from the assumption that a person has in mind something that a word does not denote or express that the word cannot occur obliquely (and be interpreted literally) in that clauses that provide some of his mental contents.

(IM, P. 101)

The terms and expressions that a linguistically competent individual uses, when literally interpreted, can provide the contents of his propositional attitudes even though he only partially understands, or even misunderstands, some of them. The content of an individual's belief is not limited to what he thinks. Although I choose the word I will use to express my belief on the basis of what I think the various candidate words mean, the content of the belief I have (and am trying to express) is not expressed by what I think the word I have chosen means, but by what the word I have chosen actually means. The content of the belief I have is expressed by the socially determined meaning of the word I have chosen to use to express my belief, even though my choice of that word to express my belief was made on the basis of my belief that that word expressed something other than what it actually does.

In "Individualism and the Mental,"⁴ Burge asks us to imagine an individual, call him 'Alfred', competent in English, rational and intelligent, who, through causal contact with both doctors and laymen, has acquired a large number of beliefs commonly attributed using belief ascriptions with that-clauses containing the word 'arthritis'. For example, he believes that he has arthritis in his wrists and ankles, that one of the symptoms of arthritis is a stiffening of the joints,

that certain aches and pains are characteristic of arthritis, etc. In addition to a whole range of unsurprising beliefs that are specified using 'arthritis', he also has a belief that he would express by saying that he has developed arthritis in his thigh. Burge then asks us to consider a counterfactual situation in which Alfred has proceeded through life just as he has done in the actual situation, right up to and including the moment when he first comes to believe that he has arthritis in his thigh. Everything that has happened to him, that he has said, that has been said to him, that he has felt, wanted, done, etc., is exactly as it is in the actual situation, non-intentionally described. The only difference between the counterfactual situation and the actual situation has to do with the linguistic environment in which he finds himself. In the actual situation, 'arthritis' is used as it normally is in English, and does not apply to ailments that occur outside of the joints. In the counterfactual situation, 'arthritis' is used for both arthritis and various other rheumatoid ailments, including the one that is afflicting his thigh. Furthermore, we are to assume that in the counterfactual situation there is no other word in the linguistic community that is used to pick out just arthritis. Burge claims that it is 'reasonable to suppose' that

[i]n the counterfactual situation, the patient lacks some -- probably all -- of the attitudes commonly attributed with content clauses containing 'arthritis' in oblique occurrence. He lacks the occurrent thoughts or beliefs that he has arthritis in his thigh, that he has had arthritis for years, that stiffening joints and various sorts of aches are symptoms of arthritis, that his father had arthritis, and so on.

(IM, p. 78)

Although he may sincerely utter the sentence form "I believe that I have arthritis in my thigh," he does not believe that he has arthritis

in his thigh, or, for that matter, any place else in his body.

There is no question that in the two situations what has been said is different. In the actual situation Alfred has said that he believes that he has arthritis in his thigh. In the counterfactual situation, Alfred has not said that he believes that he has arthritis in his thigh, because in the counterfactual situation the word 'arthritis' does not mean arthritis. If we want to express what Alfred has said he believes in the counterfactual situation in the language of the actual situation, we need a new term, say 'tharthritis'. So, if we were to express in the language of the actual situation the belief that Alfred has said he has in the counterfactual situation we would have to say that Alfred has attributed to himself the belief that he has tharthritis in his thigh, not arthritis.

In both the actual and the counterfactual situations, belief ascriptions of the form 'Alfred believes that he has arthritis in his thigh' are taken to be the correct or appropriate means by which to attribute Alfred's belief. However, the content clauses of these belief ascriptions mean different things, because 'arthritis' means something different in the actual situation than it does in the counterfactual situation. In the actual situation 'arthritis' means arthritis, while in the counterfactual situation 'arthritis' means something else, say tharthritis. In the counterfactual situation the linguistic environment does not even have a term whose meaning or extension is equivalent to the meaning or extension of 'arthritis' in the actual situation. Burge claims that our general intuitions are not only that we in the actual situation cannot express the belief that Alfred attributes to himself in the counterfactual situation unless we

introduce a new term into our language, but that we cannot use our word 'arthritis' to attribute the belief Alfred has in the counterfactual situation though we can and do use it in the actual situation, nor can we use the counterfactual word 'arthritis' to attribute the belief Alfred has in the actual situation though we do use it in the counterfactual situation. Burge claims that "in the counterfactual case we cannot correctly ascribe any content clause containing an oblique occurrence of 'arthritis' " because such "ascriptions of content clauses to the patient . . . would not constitute attributions of the same contents we actually attribute" (IM, p. 79). Although Burge does not elaborate much on this point, saying only that it is hard to see how, in the counterfactual situation, the patient could have acquired the notion of arthritis, implicit in his remark is the claim that differences in the meaning of the ascribed content clauses entail differences in the attributed content. It is the literal interpretation of the ascribed content clause that specifies the attributed content. It should be noted that Burge's point cannot be supported by the claim that different content clauses generally attribute different propositional attitude contents. It cannot be that there is ever a situation in which two different content clauses can equally well be used to attribute a single content. Different content clauses must always attribute different contents. If by ascribing different content clauses in the actual and the counterfactual situations we thereby attribute different propositional attitude contents, then we cannot correctly ascribe a content clause that includes the actual-English term 'arthritis' in the counterfactual situation, given that we can correctly ascribe a content clause that

includes the counterfactual-English term 'arthritis'. If we were to use the English term 'arthritis' in attributing a belief to Alfred in the counterfactual situation we would be attributing to him a belief that differs in its content from the one Alfred attributes to himself using the counterfactual-English word 'arthritis.' Our use of 'arthritis' in a belief ascription would fail to capture the belief that Alfred and his counterfactual peers attribute to him using the same word forms.

However, Burge's claim is not simply that there is a mismatch between the beliefs attributed in the two situations using the word form 'arthritis,' but that the actual content of Alfred's belief differs in the counterfactual situation from in the actual situation. Our hypothetical unwillingness to use the term 'arthritis' in attributing a belief to Alfred in the counterfactual situation, while we would naturally use it in attributing a belief to Alfred in the actual situation, is taken by Burge to be indicative of Alfred's belief having a different content in the actual situation than in the counterfactual situation: if we are not willing to use the same term, it must be that the contents differ. It is not simply that we do not want to use the word 'arthritis' in attributing a belief to Alfred in the counterfactual situation (regardless of whether or not it "correctly" expresses the content of his belief) because we would be attributing a different belief than he would, or could, attribute to himself. Burge's claim is that the literal interpretation of 'arthritis' cannot correctly express the content of Alfred's belief in the counterfactual situation, even though it does correctly express the content of his belief in the actual situation.

However we describe the patient's attitudes in the counterfactual situation, it will not be with a term or phrase extensionally equivalent with 'arthritis'. So the patient's counterfactual attitude contents differ from his actual ones.

The upshot of these reflections is that the patient's mental contents differ while his entire physical and non-intentional mental histories, considered in isolation from their social context, remain the same. . . . The differences seem to stem from differences "outside" the patient considered as an isolated physical organism, causal mechanism, or seat of consciousness. The difference in his mental contents is attributable to differences in his social environment.

(IM, p. 79)

Burge's claim is that in the counterfactual situation Alfred does not and cannot have the same notion of arthritis as he does in the actual situation, and therefore does not have the same mental state content. By changing the way the language of a community carves up the world we can change the contents of individuals' propositional attitudes.

While this example does involve what is arguably a de re attitude, Burge claims that whether or not the propositional attitude is de re is immaterial to the thrust of the example because it can be run in any situation "where it is intuitively possible to attribute a mental state or event whose content involves a notion that the subject incompletely understands" (IM, p. 79). Furthermore, Burge wants to claim that since the thought experiment can be run in the other direction, that is, from complete understanding in the actual situation to incomplete understanding in the counterfactual situation, "even those propositional attitudes not infected by incomplete understanding depend for their content on social factors that are independent of the individual, asocially and non-intentionally described. For if the social environment had been appropriately different, the content of

those attitudes would have been different" (IM, pp. 84-5).

In "Other Bodies"⁵ Burge presents a second version of his thought experiment, based on Putnam's Twin-Earth example, which he claims demonstrates that one can vary the contents of an individual's propositional attitudes by varying the physical environment while holding constant the individual's physical, behavioral, phenomenistic, and functional histories. By changing the chemical composition of what a word refers to, without changing its phenomenal properties, we can change the content of an individual's intentional state. The imagined situation in Burge's Twin-Earth example is essentially the same as in Putnam's original example, except that on each planet the chemical structure of the stuff called 'water' has been discovered and is widely known throughout the lay community. On Earth the scientific community has discovered that the stuff commonly called 'water' has the chemical structure H_2O , while on Twin-Earth the scientific community has discovered that the stuff commonly called 'water' has the chemical structure XYZ. (In order to distinguish the two waters without having to use their chemical formulae, Earth water will be called 'water', while Twin-Earth water will be called 'twater'.)

Although the chemical structures of water and twater are widely known in the lay communities of Earth and Twin-Earth, respectively, there are scattered individuals on each planet who are somewhat removed from the mainstream of society and have not learned of these scientific discoveries. Adam is such an individual on Earth, and Adam_{te} is his doppelganger on Twin-Earth. Neither one know the chemical structure of the stuff he calls 'water'. Burge claims that the propositional

attitude that Adam is expressing when he utters the words "I hope there is water within twenty miles" has a different content than the propositional attitude that Adam_{te} is expressing when he utters "I hope there is water within twenty miles."

When Adam says or consciously thinks the words, 'There is some water within twenty miles, I hope', Adam_{te} says or consciously thinks the same word forms. But there are differences. . . . Adam's occurrences of 'water' apply to water and mean water, whereas Adam_{te}'s apply to twater and mean twater. And, . . . the differences affect oblique occurrences in 'that'-clauses that provide the contents of their mental states and events. Adam hopes that there is some water (oblique occurrence) within twenty miles. Adam_{te} hopes that there is some twater within twenty miles. That is, even as we suppose that 'water' and 'twater' are not logically exchangeable with co-extensive expressions salva veritate, we have a difference between their thoughts (thought contents).

(OB, p. 101)

The contents of their propositional attitudes are different because 'water' as it functions in Adam's utterance means water, while 'water' as it functions in Adam_{te}'s utterance means twater.

Burge argues that there are two general considerations that support his claim that Adam_{te} cannot have the propositional attitudes that Adam does. First, he claims that Adam_{te} cannot have the same propositional attitude as Adam because Adam_{te} does not have the same natural kind concept as Adam does.

[I]t is hard to see how Adam_{te} could have acquired thoughts involving the concept of water . . . There is no water on Twin-Earth, so he has never had any contact with water. Nor has he had contact with anyone else who has had contact with water. Further, no one on Twin-Earth so much as uses a word which means water. It is not just that water does not fall in the extension of any of the Twin-Earthians' terms. The point is that none of their terms even translates into our (non-indexical) word 'water'. . . . It would thus be a mystery how a Twin-Earthian could share any of Adam's attitudes that involve the notion of water. They have not had any of the normal means of acquiring the

concept. The correct view is that they have acquired, by entirely normal means, a concept expressed in their language that bears some striking, superficial similarities to ours. But it is different.

(OB, pp. 109-10)

According to Burge, Adam's thought involves the natural kind concept of water, and since Adam_{te} does not (cannot) have the natural kind concept of water, Adam_{te}'s propositional attitude cannot have the same content as Adam's does. Secondly, if Adam_{te} had propositional attitudes that involved the concept of water, rather than twater, many of those attitudes would be false because their truth conditions would involve water, rather than twater. If when Adam_{te} utters "I think there is water in the well," he is expressing the belief that there is water, i.e., H₂O, in the well, his belief will be false because there is no water in the well, only twater.

As in the example Burge gives in "Individualism and the Mental," the difference in the content of their thoughts is not supposed to be due to their thoughts being de re. Burge claims that it is not in virtue of their thoughts being de re that their contents differ, but, rather, it is because the meaning of the that-clauses that express the contents of their beliefs is different. The differences in the mental states of Adam and Adam_{te} are due to differences in their physical surroundings that are transmitted by their respective social and linguistic communities -- the mental states of their peers, and the conventional meanings of the words they use -- to the contents of their mental states.

Finally, in "Intellectual Norms and the Foundations of Mind"⁶ Burge argues that even in situations where socio-linguistic practice is the primary determinant of the meaning of a term one must nevertheless

appeal to factors outside of the individual to determine the content of his mental states. Even in situations that do not involve natural kind terms, we still must have recourse to the surrounding context in order to determine what the content of a given individual's mental state is. Burge asks us to imagine a person, call him 'Albert', who possesses a normal competency in English. Albert's initial instruction in the use of the word 'sofa' was ostensive, but in the normal course of life he has picked up and accepted various common truisms about sofas, such as, that sofas are pieces of furniture meant or made for sitting. However, at some point in his life he comes to doubt that sofas are pieces of furniture meant for sitting. He comes to believe that sofas are really religious artifacts or works of art, and that many of the normal remarks people make about sofas conceal, or represent delusions concerning, the real function of sofas. He admits that many sofas have been sat on, but maintains that most sofas would break if sat on, and, in any event, it is not the function of sofas to be sat on.

When Albert first expresses his opinion concerning the function of sofas to his peers he meets with a great deal of resistance. Being a firm believer in empirical testing, Albert suggests that his hypothesis be subject to testing, and so constructs various sophisticated sociological experiments that all agree are adequate proving grounds for his hypothesis. Upon running the proposed experiments, and seeing the results, Albert admits that his theory about sofas was wrong.

Next we are asked to imagine a person, Bertrand, who is physically identical to Albert. Everything that Albert has heard, said, done, or had happen to him Bertrand has heard, said, done, or had happen to him. Bertrand, like Albert, learned to use the word 'sofa' by ostension, and

has heard and come to believe such statements as "sofas are pieces of furniture made for sitting." However, in the world in which Bertrand finds himself, most of these statements have been jokes or figures of speech, not meant to be taken seriously, which he has simply misconstrued. The things that Bertrand has learned to call 'sofas' are, in fact, religious objects or works of art (Burge calls them 'safos'), something widely known by members of his community. In Bertrand's world, we are to assume that there are no sofas, and 'sofa' means safo. At some point in his life Bertrand, like Albert, comes to doubt the truth of the utterance "sofas are pieces of furniture made for sitting."

According to Burge, Albert has come to believe, mistakenly, that sofas do not function primarily as pieces of furniture to be sat upon. Bertrand, on the other hand does not believe that sofas do not function primarily as pieces of furniture. Bertrand's thoughts "do not involve the notion of sofa and could not correctly be ascribed with 'sofa' in oblique position" (IN, p. 708). Essentially, Burge is claiming that 'sofa' correctly expresses the way in which Albert conceives of the object of his doubt, while 'safo' correctly expresses the way in which Bertrand conceives of the object of his doubt. Albert has thoughts literally attributable to him using a content clause containing 'sofa'. Bertrand, on the other hand, has no thoughts literally attributable to him using the term 'sofa' in oblique position. Albert and Bertrand are physically identical but have different mental contents.

III. THE SUBSTITUTION OF SYNONYMS IN BELIEFS CONTEXTS IN NOT ENTAILED
BY BURGE'S POSITION

The position that Burge is arguing for in these three articles, as well as elsewhere, is quite radical. He is, in essence, advocating a sort of linguistic hegemony: language is not merely the means by which we attempt to express the content of a mental state, it determines what the content of a mental state can be. The content of a propositional attitude of a linguistically competent individual is determined by what Burge calls the "literal interpretation" of the that-clause used, or that would be used, in attributing the propositional attitude to the individual. Burge claims that the particular words used in the content clauses of propositional attitude attributions, when interpreted literally, "characterize a subject's mental states and events in such a way as to take into account the way he views or thinks about objects in his environment" (OB, p. 109), regardless of whether or not the subject incompletely understands the meaning of the terms, is ignorant of expert knowledge about the structure of the stuffs and things to which the terms apply, or has a non-standard theory about those stuffs and things. His position is that if we are willing to use a given sentence to attribute a mental state, say a belief, to someone who is competent in the use of language, then the particular words of the sentence we have chosen to use accurately express the content of that individual's belief, regardless of whether or not the individual in

question has any mistaken impressions about the words he, or we, use. If an individual is linguistically competent, then those concepts that -- in the eyes of the linguistic community -- are expressed by the words used in reporting his belief are the actual concepts involved in the belief, regardless of what the individual does or does not know about the things in question. All that is required is that the subject be competent in the use of language. So long as the individual is a normally competent member of the linguistic community, "the expressions the subject uses sometimes provide the content of his mental states or events even though he only partially understands, or even misunderstands, some of them" (IM, p. 114).

Burge's claim that propositional attitudes should be individuated on the basis of a literal interpretation of their content clauses has been the source of not inconsiderable confusion concerning precisely what is being asserted. The most natural and obvious construal of "literal interpretation", one borne out by many of Burge's own remarks, is as conventional meaning. For example, in "Belief and Synonymy"⁷ he writes that the "[c]ommunal conventions about the meaning of a speaker's words tend to override what a speaker mistakenly associates with his words in determining what he says and even, sometimes, believes" (BS, pp. 134-5), while in "Other Bodies" he writes "[t]he difference in their mental states and events seems to be a product primarily of differences in their physical environments, mediated by differences in their social environments -- in the mental states of their fellows and conventional meanings of words they and their fellows employ" (OB, p. 102). However, as various philosophers have pointed out, if the literal interpretation of content clauses is simply their

conventional meaning, Burge's position entails -- or at least endorses -- the substitutivity of synonyms and co-referring expressions in intentional contexts. For example, in "Cognitive Science and the Twin-Earth Problem,"⁸ Jerry Fodor discusses Burge's example from "Individualism and the Mental" involving someone who thinks that all contracts must be written, and criticizes him on the grounds that interpreting content clauses in terms of their conventional meaning requires that we attribute explicitly contradictory beliefs to people who are rational, intelligent, and mentally competent:

Still, it seems to me, we cannot grant Burge his intuitions what belief Jones uses "Smith just signed a contract" to express. For, surely, Jones expresses the same concept by "contract" when he says that as when he says, for example, "I deny that verbal contracts bind". But if the concept of contract expressed in this latter case is our concept of contract (and if, by assumption, being binding when verbal is constitutive of our concept of contract) then the belief that Jones is expressing when he denies that verbal contracts bind is explicitly self-contradictory. Specifically, the belief expressed is that what is binding when verbal is not binding when verbal. Notice, moreover, that we have to read this belief de dicto; it is not just that Jones believes of something which is as a matter of fact so and so that it is not so and so . . . If it means anything to say that Jones has our concept of contract, it must mean that we should construe his utterances of "contract" in the same way we would construe our own. If, however, we do translate that way, we get self-contradictions whenever Jones says of verbal contracts what, by Burge's own assumption, Jones believes to be true of them: viz., that there aren't any.

(CSTP, p. 107)

Fodor's point is that if what Jones has in mind whenever he uses the term 'contract' is what we have in mind whenever we use that term, then Jones must have an explicitly self-contradictory belief in mind when he denies that verbal contracts bind,⁹ since we have in mind something that is, among other things, binding when verbal whenever we use the term

'contract'. If the content of Jones' belief is given by the conventional meaning of the content clause used to express his belief, then his belief has a content that is explicitly self-contradictory.

In his reply to Fodor, Burge rejects this presentation and discussion of the contract example by Fodor, claiming that it represents a misreading of his contract example (see "Two Thought Experiments Reviewed"¹⁰). Burge asserts, "I am not assuming, absurdly, that whenever someone sincerely utters words that mean that p, he believes that p. The thought experiments depend on no such inference" (TTER, p. 288). Certainly, there is no question of Burge asserting that whenever someone sincerely utters words that mean that p, the individual believes that p. Burge makes it quite clear that if the individual in question is not a minimally competent user of one or more of the terms that occur in a sentence he utters, then we may not be able to attribute to him the belief expressed by the sentence he uttered. For example, if Jones has overheard Justice Blackmun say that Smith has a contract with Ace Wrenching, but Jones has no idea of what a contract is, he still may sincerely utter "Smith has a contract with Ace Wrenching," believing, simply, that it expresses something true. Though Jones sincerely utters "Smith has a contract with Ace Wrenching," Burge would agree that we cannot attribute to him the belief that Smith just signed a contract because Jones does not have adequate mastery of the terms he has used. If an individual fails to meet minimum linguistic competency requirements we cannot attribute to that individual attitudes expressed by the sentences he sincerely utters.

Allowing for this slight caveat, Fodor's point is still valid: if

'literal interpretation' means conventional social meaning, Burge cannot consistently deny substitutivity of synonyms in intentional contexts. However, Burge does explicitly deny substitutivity of synonyms in belief contexts in several places, devoting most of an article to the subject (viz., "Belief and Synonymy"). Unfortunately, Burge leaves the reconciliation of his seemingly conflicting claims -- that content clauses should be interpreted literally, and that substitutivity of synonyms and co-referring expressions is invalid in content clauses -- as, shall we say, an exercise for the reader. Consequently, I make no representations that the following discussion is an explication of Burge's actual views on the matter. Based, as it is, on an extrapolative reading of "Intellectual Norms and the Foundations of Mind," "Belief and Synonymy," "Self-Reference and Translation,"¹¹ and, to a lesser extent, "Belief De Re,"¹² it is meant simply as one possible way of construing his position coherently and consistently.

While there is no question that Burge does identify literal interpretation with conventional meaning in a number of places, I want to suggest that, in light of the position he argues for in "Intellectual Norms and the Foundations of Mind", that identification needs to be reconstructed as a "rough and ready" approximation of how we are to understand 'literal interpretation' rather than an exact explication of what the literal interpretation of a content clause is. In "Intellectual Norms" Burge introduces the notion of the 'cognitive value' of a term or expression, and claims that

[i]ntentional mental states and events are individuated in terms of cognitive value. We have no other systematic, cognitively informative way of individuating them. Since communally accepted characterizations as well as expert

opinion can be doubted, the ultimate authority regarding the application, explication, and individuation of a subject's intentional mental events does not derive solely from the actual motions, behavior, actions, usage, practices, understanding, or even (except trivially) thoughts of any person or social group.

(IN, p. 720)

According to Burge, expressions have both a cognitive value and a conventional linguistic meaning. Burge claims that we need the notion of the cognitive value of expressions in addition to that of the conventional linguistic meaning of expression in order to account for the fact that certain sorts of necessary truths can be genuinely doubted. Burge argues that since doubts about "meaning-giving normative characterizations", e.g., 'sofas are pieces of furniture made or meant for sitting', can be informative whereas doubts about the corresponding identity judgements, e.g., 'pieces of furniture made or meant for sitting are pieces of furniture made or meant for sitting', cannot, we must assign those two doubts different contents, different cognitive values.

In short, such characterizations have different cognitive values from those of the corresponding identity judgments. So in interpreting a specification of a belief that sofas are pieces of furniture . . . meant for sitting, one must assign different cognitive values or units of potential information -- to the conventionally synonymous phrases ('sofa' and 'piece of furniture . . . meant for sitting') as they occur in such specifications. Thus cognitive value and conventional meaning should be distinguished.

(IN, p. 715)

Furthermore, Burge claims that cognitive value and conventional linguistic meaning are not only distinct, but that the linguistic meaning of a term can change while its cognitive value remains unchanged:

[I]t is possible to doubt that sofas are all and only pieces of furniture of a certain construction meant or made for sitting. If the doubt were to prove well founded, the conventional meaning of 'sofa' would be forced to change. But despite the change, it might remain appropriate, before and after the change, to attribute propositional attitudes involving the notion of sofa. Both before and after, A and his opponents would agree that these are sofas. Before and after, they would be characterized as having disagreed over whether all and only sofas are furnishings of a certain structure made or meant for sitting.

This situation bears some analogy to cases of theoretical change in science. Dalton and his predecessors defined 'atom' (and its translations) in terms of indivisibility. Major theoretical changes intervened. The definition was discarded. Despite the change, we want to say, Dalton wrongly thought that atoms were indivisible: despite his erroneous definition, he had the "concept" of atom (not merely the referent of 'atom').

(IN, p. 715-716)

The arguments Burge gives for making this distinction between cognitive value and conventional linguistic meaning are variants of Frege's arguments for distinguishing between the sense of an expression and its referent, and I want suggest that the cognitive value of an expression (or what Burge sometimes refers to as the concept expressed by an expression) is roughly analogous to Frege's notion of the sense of an expression. According to Burge, the conventional linguistic meaning of an expression is fixed by some sort of consensus among those whom Burge calls the 'most competent' speakers of the language, whereas (I am maintaining) the cognitive value is fixed by some sort of Fregean "deep rationale underlying the expression's use and understanding -- a rationale that might not have been understood by anyone" (IN, footnote 15, p. 715).

The identification of 'literal interpretation' with cognitive value is further supported by Burge's discussion of non-substitutivity of synonyms in belief contexts in "Belief and Synonymy." In "Belief and

Synonymy" Burge claims that propositional attitude ascriptions are self-referential: expressions that occur within the content clauses of belief ascriptions "function autonomously", that is, they denote themselves. However, the content clauses of propositional attitude ascriptions "are to be interpreted as we (the reporters) use them outside of belief contexts. . . . Our presupposed interpretation is of the denoted expression" (BS, p. 127). To say that content clauses are to be interpreted literally is to say that we should take content clauses to denote themselves, to denote the very words of the content clauses. But the propositional attitude contents we take them to specify are given by a normal interpretation of the denoted expression. Since, according to Burge, the self-referential character of autonomous sentences is essential to their "cognitive content" or "information value" one cannot substitute a synonym or co-referring term for a term in the sentence without altering the information conveyed by the sentence, because such substitutions do not preserve self-reference (see "Self-Reference and Translation" for Burge's discussion of this issue). If we are to view the content clauses of propositional attitude ascriptions as being self-referential, then one cannot substitute co-referring expressions in such clauses without altering their cognitive content. Although 'water' and 'H₂O' refer to the same natural kind, "Adam hopes that there is water within twenty miles" cannot be replaced by "Adam hopes that there is H₂O within twenty miles", or vice versa, because 'water' and 'H₂O' both occur in self-referential contexts where cognitive content is dependent upon self-reference, not reference. Substitution of synonyms preserves reference in ascriptions but not self-reference or cognitive content. Thus,

Fodor's claim that Burge's position requires us to attribute an explicitly contradictory belief to a rational individual is unfounded. The belief specified by the content clause 'a fortnight is ten days' is content distinct from one specified by the content clause 'a period of fourteen days is ten days', with only the latter being explicitly self-contradictory.

If we assume that Burge has, in fact, meant cognitive value when he has used the phrase 'literal interpretation,' rather than conventional meaning, we can not only account for why he might have identified literal interpretation with conventional linguistic meaning, but we can also account for his insistence that linguistic competence is the criterion by which to determine whether or not an individual has the notion expressed by a particular word. The identification of literal interpretation with conventional meaning can be accounted for because Burge claims that

Usually, the best understanding one can achieve of a cognitive value is that offered by accepted normative characterizations and whatever background information accompanies them. Thus full understanding of cognitive value is normally not distinct from ideal understanding of ordinary usage and meaning.

(IN, p. 718)

If under ideal conditions of understanding linguistic meaning and cognitive value are not distinct (though Burge does insist that they are nevertheless still individuated differently), then the identification of literal interpretation with conventional linguistic meaning, rather than cognitive value, is misleading but not wholly inaccurate. If we assume that literal interpretation is to be equated with cognitive value, where cognitive value is fixed by the rationale

underlying the use of an expression, Burge's claim that an individual need only be a competent user of an expression to think with the concept (cognitive value) makes at least some kind of sense. Burge might want to claim that if an individual has mastered the use of a term, then he has implicitly mastered the rationale governing its use.

IV. BURGE'S "COMMUNITARIAN" NOTION OF CONTENT IS PLAUSIBLE NEITHER AS
AN EXPLICATION OF OUR ORDINARY MENTALISTIC NOTION NOR AS A
SCIENTIFIC CONSTRUCT

While Burge's position may not, in fact, commit him either to endorsing the substitutivity of synonyms or co-referring terms in belief contexts or to attributing explicitly self-contradictory beliefs to rational individuals, it does require that we accept a fundamental division in psychology on the issue of the relation between mental state content and behavior, a division that seems counterintuitive and unwarranted. Furthermore, if we examine the interpretation Burge gives to his thought experiments, it becomes apparent that it is dependent upon assumptions about language, thought, and the connection between them that are, at best, problematic. Taken together these considerations make Burge's view of content as socially constituted plausible neither for use in a scientific psychology nor as an account of the notion of content as it is used in commonsense belief/desire psychology.

Burge's claim that the content of a mental state is determined by

the words we use to attribute that mental state to an individual applies only to individuals who are competent members of a linguistic community. If an individual is a competent user of a term and generally competent in the language, then the contents of his beliefs are determined by the peculiarities of his language. On the other hand, if the individual in question is not linguistically competent, then the words we use in attributing a mental state to that individual cannot be used to determine the content of that mental state. Since, according to Burge, we cannot determine mental state contents on the basis of the words uttered by or about a person (or animal) who is not linguistically competent -- because "mastery of the language and responsibility to its precepts have not been developed" (IM, p. 90) -- psychology must have some other means of determining what mental state contents to attribute in such cases, and it must recognize that the content of those mental states is determined without reference to any socio-linguistic community. While some may want to deny that (non-human) animals have mental states with content (though Burge is certainly not among their number -- See IN, footnote 14, p. 713, and IM, pp. 96 and 114), the need for two notions of content arises just in case we are not prepared to deny that children of 2 or 3 (or 6 or 18 depending upon when we decide that they are 'linguistically competent') have mental states with content. However, this division of the world into language users and non-language users for determining the content of intentional states not only runs counter to common psychological practice, but entails that psychology accept that there are two distinct means by which internal states acquire content.

In predicting and explaining the behavior of young children, dogs,

dolphins, and the like, we use exactly the same sorts of mental state ascriptions as we do when we predict and explain the behavior of adults who are competent in the use of language. Not only do we use the same sorts of ascriptions, but we go through the same chains of reasoning and arrive at the same conclusions and predictions. Furthermore, it seems that we are at least as successful in our predictions and explanations of the behavior of non-language users on the basis of these ascriptions as we are in our prediction and explanations of the behavior of language users. There is nothing in our use of commonsense psychology that distinguishes between those who speak a language and those who do not. But it is not that we treat non-language users as language users gratuitously. The behavior of dogs, dolphins, and certainly children long before they can talk cannot be explained or predicted without appealing to intentional states. Their behavior is simply too complex. We attribute to them beliefs and desires that have content for exactly the same reason that we attribute beliefs and desires to adults: it is only on the assumption that they have intentional states that their behavior is comprehensible.¹³

Nevertheless, given Burge's theory of content, we cannot be attributing the same mental content to a non-language user as we are to a language user when we say of each, for example, that he believes that there is a squirrel in the yard. The content of the language user's belief is determined by the literal interpretation of 'there is a squirrel in the yard', but that cannot be what determines the content of the non-language user's belief, because the non-language user is not a responsible member of the linguistic community. What the content is of the non-language user's belief that there is a squirrel in the yard

must be determined by something other than the literal interpretation of the content clause of the ascription. In fact, it seems that regardless of what does determine the content of a non-language user's belief, any given content clause must be attributing a different content when it is used in specifying the belief of a non-language user than when it is used in specifying the belief of a language user.

For example, imagine that Alfred, from Burge's arthritis example, is also under the mistaken impression that beaches are stretches of nearly level, pebbly shore. We can assume that all of the beaches he has ever seen have been pebbly, and that he has simply undergeneralized. Now, given a dog who is familiar with all of the same beaches as Alfred is, we can imagine attributing to both Alfred and the dog, the same belief, say the belief that when Leonora puts the top down on her car she is going to the beach. According to Burge, Alfred believes that Leonora is going to the beach, regardless of the fact that Alfred has an inaccurate conception of beaches. If, however, Alfred were in a linguistic community in which the word 'beach' meant, or referred to, pebbly stretches of nearly level shore, then Alfred's belief content would be that Simon was going to the, say, pebbleach. But what about the dog's belief? What makes it the case that Alfred's belief contents differ in the two contexts is the fact that 'beach' has different meanings in the two contexts. However, that difference cannot influence the content of the dog's belief. His belief contents are, presumably, individualistic. Furthermore, since the meaning of 'beach' is the only difference between the two environments the dog's belief content must be the same in both contexts. There simply is no other difference -- the world is the very same world, it is simply that

the linguistic communities have decided to carve up that world in slightly different ways. But, if the dog has the same belief content in both contexts, then in neither context can the dog have the same belief content as Alfred. If the dog has the same belief regardless of which linguistic environment he is in, and if the dog's belief in the actual situation can be specified by saying that he believes that Leonora is going to the beach, then it must be equally correct to say of the dog in the counterfactual situation that he believes that Leonora is going to the beach. Similarly, with specifying the dog's belief in the counterfactual situation by saying that he believes that Leonora is going to the pebbbleach. However, since in the actual situation Alfred's belief can be specified only by saying that he believes that Leonora is going to the beach, and in the counterfactual situation his belief can only be specified by saying that he believes that Leonora is going to the pebbleach, then in neither the actual nor the counterfactual situation can Alfred and the dog have the same belief. The content clause 'Leonora is going to the beach' specifies a different content when it is used to attribute a belief to a language user, like Alfred, than when used to attribute a belief to a non-language user, like the dog. 'Leonora is going to the beach' and 'Leonora is going to the pebbleach' specify the same content when used to attribute the belief of a non-language user, but different contents when used to attribute the belief of a language user, contents that are both distinct from the non-language user's content. The same content clause will be used to attribute two different belief contents, one for language users and one for non-language users, while different content clauses can be used to specify the same belief content in the case of

non-language users. ¹⁴

Burge's theory of socio-linguistic content requires that a single content clause within a single context be used to attribute two different mental states, mental states that have different contents. However, all of our predictions and explanations of behavior assume that the two different mental state contents will have the same effect on behavior. Psychology will acknowledge two types of mental state content, but will use a single content clause to pick out both types, and will make the same predictions and explanations of non-verbal behavior regardless of the fact that the contents of the two mental states differ. The world will be the same, the words we use to attribute the mental states will be the same, and the behavior will be the same, but the mental contents will be different. If psychology has to have two separate notions of content, one for language users and one for non-language users, then it is also going to have to have some sort of theory about why the nature of how a mental state content is determined changes when an individual goes from being a non-language user to a language user. In particular, it is going to have to explain why content goes from being determined by the world and the way the individual relates to and conceives of or taxonomizes that world -- as it is in the case of non-language users like the dog, to being determined by the world as it is conceived of or taxonomized by the socio-linguistic community, regardless of how the individual conceives of the world -- as it is in the case of language users. Burge's theory implicitly acknowledges that the mental contents of non-language users are determined by the world as they taxonomize it, but it denies language users that same right of taxonomic self-determination.¹⁵

The problem with Burge's notion of content is precisely that it views verbal behavior as that which psychology should explain when dealing with linguistically competent individuals. By adopting a view of psychology and mental state content that emphasizes verbal behavior over non-verbal behavior in the case of linguists, Burge's theory forces the construction of another account of content that can be used when dealing with non-linguists. However, the need for two notions of content does not arise if, rather than adopting a socio-linguistic individuation for the mental state contents of linguists, we adopt the individualistic individuation, that under Burge's theory we will need for non-linguists, for all the mental state contents of both linguists and non-linguists. If we attribute mental state contents on the basis of literal interpretations of what the individual says, then we must have an alternate notion of content for non-linguists, since the literal interpretation method is unavailable in those cases. On the other hand, there seems no reason not to use whatever method of content specification we would use for non-linguists, for linguists, Using the individualistic notion of content, it seems possible to account for not only the non-verbal behavior of linguists and non-linguists, but the verbal behavior of linguists as well.

Burge objects to this sort of an individualistic notion of content on two grounds. He claims that such readings of content clauses are not borne out by the actual behavior of linguistically competent individuals, and that the individualistic reading of content clauses forces the metalinguistic reinterpretation of such belief attributions as "Simon believes that a fortnight is ten days" if, while maintaining that his belief is, in fact, false, we do not want to attribute to

Simon a belief that is analytically false. In "Belief and Synonymy" Burge argues that although it is true that Simon took 'fortnight' to mean "period of ten days" when he used it to express his belief that a fortnight is ten days, that fact should not be taken as persuasive when we are determining how to interpret his assertion. Burge claims that we must also take into consideration "the willingness of the speaker to submit his statement to the arbitration of a dictionary" (BS, p. 130).

Which of these considerations take precedence in determining how to interpret the speaker's statement and attitude in the present instance? The speaker's behavior, as we are imagining it, provides the key. His willingness to defer to a dictionary or the intuitions of other speakers and his and our unmixed feelings, after consulting the relevant authority, that he has made a mistake suggest that the latter consideration is crucial.

(BS, p. 131)

Burge again makes much the same point in "Individualism and the Mental":

The subject's willingness to submit his statement and belief to the arbitration of an authority suggests a willingness to have his words taken in the normal way -- regardless of mistaken associations with the word. Typically, the subject will regard recourse to a dictionary and to the rest of us, as at once a check on his usage and his belief. When the verdict goes against him, he will not usually plead that we have simply misunderstood his views.

(IM, p. 101)

Burge's claim is that an individual's willingness to have his belief statement judged by the linguistic standards of his community indicates that the content of his belief is to be understood as being given by the literal interpretation of the content clause of the belief attribution. Burge's other objection is that the only way we can claim that Simon's belief is false when he says that he believes that a fortnight is ten days, if we adopt an individualistic method of content

determination, is to reinterpret his assertion metalinguistically. However, and I think Burge is right on this point, reinterpreting Simon's belief statement as about the word 'fortnight' does not do justice to the belief Simon has. It is undoubtedly true that Simon does believe that 'fortnight' means "period of ten days", but that certainly is not the belief that he is expressing when he says he thinks that a fortnight is ten days. Since, according to Burge, an individualistic notion of content forces one into untenable metalinguistic reinterpretation, he claims that such a notion of content cannot be acceptable.

Granting Burge both that individuals will (often) consider a dictionary a check on the truth or falsity of their beliefs¹⁶, and that metalinguistic reinterpretation fails to capture the content of such beliefs as Simon's that a fortnight is ten days, nevertheless, neither point provides a definitive reason for denying the plausibility of an individualistic notion of content. Rather than viewing Simon's belief that a fortnight is ten days either as being analytically false (as Burge does) or as being metalinguistic and false, we can simply view it as being (roughly) empirical and false, analogous to his beliefs that shrews are rodents or that this page is red. When Simon says that he believes that a fortnight is ten days, he is attributing to himself a belief about a publicly accessible, abstract object, namely, a fortnight. His belief is that the object that is referred to by the word 'fortnight' -- and he intends his use of the word to refer to what it is customarily used to refer to -- has a certain property, viz., being ten days. His belief about fortnights is analogous to his belief about shrews: his belief that shrews are rodents is a belief about

those things commonly referred to as shrews, and claims that they have the property of being rodents. Simon conceives of fortnights as having the property of being ten days, just as he conceives of shrews as having the property of being rodents. His belief that fortnights have the property of being ten days is subject to revision in light of the "facts" about fortnights, as is his belief that shrews have the property of being rodents. In both cases, Simon believes that the object in question has a property that it does not. Simon has simply made a mistake about a property of fortnights. Granting that a fortnight only has one property, nevertheless, Simon has simply gotten that property wrong.

If we view Simon's belief that a fortnight is ten days as, essentially, an empirical belief, like his belief that shrews are rodents, we neither have to attribute to him a belief that is analytically false, nor do we have to reinterpret his belief metalinguistically in order to admit that it is false. We can attribute to him a belief based on an individualistic notion of content without having to say either that his belief is true or that it is really a metalinguistic belief. Even though Simon's conception of a shrew, that is, what he has in mind, when he says that he believes that shrews are rodents is that of something that is a rodent, we have no hesitation saying that his belief is false, not metalinguistic and false, but false in virtue of the facts about shrews. The same can be said of Simon's belief about fortnights. It is a belief about fortnights that is false in virtue of the facts about fortnights. His belief is neither metalinguistic nor analytically false.

Taken as a belief about an empirical matter, the fact that Simon is

willing to accept the verdict of a dictionary on the truth of his belief is unremarkable. When he learns that his belief is false, that a fortnight is fourteen days not ten, he will not plead that we have simply misunderstood his belief, any more than he would when he learns that shrews are not rodents. Dictionaries, as well as the rest of us, are not merely conveyors of the meaning of words. One can learn facts about the world from them as well. Many good dictionaries could have told Simon not only that a fortnight is not ten days but that a shrew is not rodents as well. What Simon defers on is the truth of his belief, and the facts about the world, not on how he conceived of the object of his belief.

Neither of Burge's objections to an individualistic notion of content are persuasive, and given that his view of content as socially constructed divides the domain of psychological theorizing into two distinct categories, there seems to be no reason for adopting it. However, prior to leaving the discussion of Burge's view of content, two other issues needs to be addressed.

It is certainly a common occurrence to wonder whether what you mean by something is what everyone else means by it. For example, I can plausibly wonder whether the belief that I have and express by saying that I believe that Simon has influenza is the same belief that everyone else has when they say that they believe that Simon has influenza. I may begin to wonder if perhaps the belief that everyone else has when they say that Simon has influenza is the belief that I would express by saying that I believe that Simon is possessed by the devil. But if Burge is correct, and the content of a mental state is determined by the literal interpretation of what one says, then such

doubt is rendered incoherent. I cannot be skeptical about whether the belief that I have and that is attributed by saying, for example, that I believe that a spatula is useful in the kitchen is the same belief that Mary has that is attributed by saying that she believes that a spatula is useful in the kitchen. The content of my belief that spatulas are useful in the kitchen is necessarily the same as the content of Mary's belief that spatulas are useful in the kitchen, the Pope's belief that spatulas are useful in the kitchen, everyone's belief that spatulas are useful in the kitchen. Burge's theory makes a perfectly reasonable type of doubt, the doubt that what I believe is the same as what you believe when we both say the same things, impossible to entertain.

At the same time that Burge's theory renders it impossible to wonder if what you believe is the same as what everyone else believes, it introduces the possibility of a new kind of doubt, a doubt that seems unreasonable from a commonsense point of view. If the content of what I believe is determined by my socio-linguistic community, I can wonder if what I actually believe is what I think I believe. For example, imagine that Leonora, a native speaker of English, is also a competent speaker of French, and would certainly count as a competent user of the phrase "une crise de foie". Leonora and a number of other people (each a competent speaker of French) all observe Marchale over the period of an evening, and each comes to have a belief that he or she would express by saying that Marchale is having une crise de foie. Now, it happens that Leonora is under the impression (correct, we might add) that une crise de foie is a general state of irritability or being out of sorts, typically brought on by long-term overindulgence in rich

food, that tends to result in social gaffs, contrariness, etc. However, since what determines what Leonora believes is not, in any direct sense, something about Leonora, we can imagine Leonora coming to wonder if she has any idea of what she, in fact, believes. For instance, Leonora might begin to wonder if, perhaps, her belief that Marchale is having une crise de foie, a belief that she took to be entirely innocent, isn't really a rather naughty thought. More generally, neither Leonora nor we have any way of knowing what the contents of any of our beliefs are. So long as there is the possibility that we might misunderstand a word we use, we cannot but doubt that we know what we think. Thus, we are forced to a new type of skepticism, not about the world, but about our own minds.

In arguing for his position Burge would have us believe that the interpretations he gives to his various thought experiments are intuitively obvious. However, his interpretations are dependent upon, among other things, the assumption that, for the most part, the only notions a linguistically competent individual can acquire are those that his socio-linguistic community has terms for. The notions that a linguistically competent individual can have are, by in large, limited to those that are expressed by the terms in use in his socio-linguistic community. The linguistic individual cannot acquire notions that are not recognized by his linguistic community, nor can he construct from his own imagination such notions, with the possible exception of a notion of a natural kind that the individual has had direct experience of.

Burge has to limit the notions a linguistically competent individual can acquire if he is to be able to maintain that the various

individuals in his thought experiments have, or are even likely to have, different thoughts. If an individual were able to acquire or construct notions that were outside the scope of his linguistic community, then one could not assume to know what specific notion was involved in a propositional attitude from the words the individual uttered when attempting to express his thought. The words an individual has available for use in the content clause of a belief attribution are determined by the linguistic community in which he finds himself and his own competence in the use of the language. If, however, an individual can have propositional attitudes that involve notions that are not expressed by the words he has available to express those thoughts in language, he will be forced to use content clauses that do not accurately express the content of his propositional attitudes. It is only if the notions that can be involved in the propositional attitudes of linguistically competent individuals are restricted to those expressed by the terms of the language, can there be any plausibility to Burge's claim that the propositional attitude contents of linguistically competent individuals are given by a literal interpretation of the words they use when expressing their attitudes.

In most of his thought experiments, Burge relies very heavily on the notion of a linguistic community and the idea of the social meaning of words. It is that notion which, in the eyes of the linguistic community, is expressed by a word that determines what an individual can and does think. Community meaning is the authority in matters of both language and thought. However, the claim that an individual cannot pick up, acquire, or construct any notion that is not expressed by a word or words already in use in his linguistic community calls

into question how it is possible for the language of such a community to expand or grow, if, in fact, such a linguistic community could arise in the first place. If Burge could have limited this claim so that it only applied to natural kind concepts, the claim, itself, would be more plausible. One might want to claim that one can only have a concept of water if one has experience, either direct or indirect, of water. However, Burge specifically makes this claims about what notions an individual can have with respect to the notions of arthritis and sofa, neither of which are a natural kind, and to the extent that he wants his thought experiments to be applicable in a wide range of terms -- natural kind words, color adjectives, social role terms, terms for historical styles, etc. (cp. IM, p. 79) -- he cannot, in fact, limit the claim.

Social meaning clearly derives from individual practice. A word comes to have the meaning it does because of what individuals have used it to express. But, if an individual cannot come to have a notion that is not already part of his linguistic community, it is hard to see how a language could acquire new terms, ones that express new notions or concepts. Taken together, the claim that individuals can only have notions already in their linguistic community and the view that social meaning is derived from the practices of individuals entail either that language springs full-blown into a community, rather like Athena from the head of Zeus, or that it arises spontaneously from the meaningless utterances in individuals. Needless to say, the former is absurd and the latter is implausible. If language is to develop gradually in a community from the practices of individuals without individuals being able to have notions not already expressed in their language, it would

seem that in order for a new term to enter a language individual members of the linguistic community would have to start uttering a new sound -- without it expressing anything or meaning anything. Only at some point later would or could this sound be taken to express a notion or concept new to the language. But whatever that new notion turned out to be it could not in any sense be derived from what individuals have used it to express, since they can only have notions already expressed in the language. It seems, at best, somewhat farfetched to claim that a sound that does not express any notion or concept not only would be picked up and gain acceptance in the linguistic community, but would experience a convergence of, or at least display some sort of consistency in, the circumstances in which it was uttered so as to allow it to have a (single) meaning.

In "Intellectual Norms" Burge discusses at some length his position on how a linguistic community ultimately arrives at a formulation of the conventional meaning of a term in its language. I should be understood that Burge's discussion there has no bearing on the problem his theory has in accounting for the existence and growth of language given the inability of individuals to have propositional attitudes involving notions that are not expressed by terms in their language. In the section in "Intellectual Norms" Burge is involved in presenting his view on how a community arrives at a characterization of the meaning of a term already in use in the linguistic community. Burge is not concerned with how it acquired its meaning, but only with how the community arrives at a characterization of its (already existent) meaning. "The conventional linguistic meaning of a term has been correctly specified when . . . the most competent speakers have reached

equilibrium on a characterization" (IN, p. 704, emphasis added).

The contention that the only notions that can be part of the contents of the thoughts of a linguistically competent individual are those expressed by the terms in his language plays a central role not only in Burge's interpretation of his thought experiments, but in the underlying considerations that prompts Burge to advocate the socio-linguistic individuation of mental state contents as well. The underlying motivation for Burge's claim that we should interpret propositional attitude attributions literally is his claim that neither metalinguistic reinterpretation of belief attributions nor the substitutivity of synonyms in belief contexts is tenable. In "Belief and Synonymy," Burge claims that the need for and motivation behind the arguments for metalinguistic reinterpretation of such belief attributions as

(1) Alfred used to believe that a fortnight was ten days.

is the assumption that an individual fully and infallibly knows what it is that he believes, what the content of his belief is. Burge's point seems to be that if one fully and infallibly knows what the content of one's belief is, then it should make no difference what particular terms are used to express that content. If the content of Alfred's belief is given by 'a fortnight is ten days' and Alfred knows -- fully -- the content of his belief then it must be equally correct to say of Alfred that he believes that fourteen days is ten days, since 'fortnight' means fourteen days. However, if we do not want to attribute such explicitly false beliefs to generally rational individuals, and it seems perfectly plausible for Alfred to say that he did believe that a fortnight was ten days but that he never believed

that fourteen days was ten days, Burge claims that we must reinterpret such belief attributions as (1) metalinguistically so as to prohibit the substitution of 'fourteen days' for 'fortnight'. Burge argues that the various justifications for metalinguistic reinterpretation are unpersuasive, and claims (implicitly) that one must either hold on to substitutivity in belief contexts and attribute explicit analytically false beliefs to rational individuals, e.g., the belief that fourteen days is ten days, or abandon substitutivity and endorse a literal interpretation of propositional attitude attributions. Burge claims that since the former is untenable -- we do not want to attribute such explicit analytically false beliefs to individuals -- so we must embrace the latter option. However, according to Burge, abandonment of substitutivity undermines the possibility of maintaining that individuals know what they believe.

Burge's argument that we must abandon both substitutivity and the idea that individuals know what they think is dependent upon two assumptions: that there are only two possible ways of interpreting belief ascriptions such as (1), metalinguistic and literal; and, that if individuals fully and infallibly know the contents of their beliefs, the substitutivity of synonyms in belief contexts must hold. As I have already indicated, the former assumption is unwarranted, and arguably untenable. Neither of the possibilities that Burge allows seems to correctly characterize what it is that someone believes when she believes, for example, that a fortnight is ten days, or that a sofa is a religious artifact. As for the latter assumption, if we grant that individuals can have propositional attitude contents that involve notions not expressed by terms of their languages, or even

simply not expressed by the terms they use when indicating their attitudes, then it too must be unwarranted. If the notions involved in the content of the belief of an individual can differ from those that are expressed by the terms of which she is a competent user, and she chooses the terms she will use in expressing her belief on the basis of the content of her belief, then if she has a mistaken (or missing) association about the meaning of a term, believing it to accurately express the notion she has in mind, but no corresponding mistaken (or missing) association concerning the meaning of a synonym of that term, then substitutivity will fail even though the individual fully understands the content of her belief. Substitutivity fails because neither synonym correctly expresses what the individual has in mind, though she thinks one of them does. For example, imagine that Leonora believes that a fortnight is ten days, that is, she thinks a fortnight has the property of being ten days, and that she thinks that the term 'fortnight' not only refers to that which is the object of her belief, but that, taken in its usual sense, it accurately expresses the way she conceives of a fortnight. In other words, she takes 'fortnight' to refer to fortnights but to express the notion of a period of ten days. However, she knows that the expression 'a period of fourteen days', taken in its usual sense, does not express the way she conceives of the object of her belief. Substitutivity fails not because she does not know the content of her belief, but because neither 'fortnight' nor 'a period of fourteen days' accurately expresses the way she conceives of a fortnight, though she thinks, mistakenly, that 'fortnight' does. We can acknowledge the failure of substitutivity without being forced into metalinguistic

reinterpretation or the denial of the view that people, in an ordinary sense, know the contents of their propositional attitudes.

It seems that there is little doubt that Burge's theory of social content cannot be taken as a serious candidate for adoption by a scientific psychology, but neither does it seem a plausible account of the ordinary mentalistic notion of content. Certainly we use language to express our thoughts, and expect to be taken at our word. But that does not necessarily mean that the linguistic community has authority over the contents of our beliefs. We expect to be taken at our word because that is one of the underlying assumptions of linguistic communication, regardless of whether we are discussing the weather, the structure of water, or what someone believes. Burge makes the content of a linguistically competent individual's propositional attitude a property of the community. There need be nothing different about Adam, himself, for the content of his propositional attitude -- that which is causally responsible for his behavior to be different than it is. Psychology, even commonsense belief/desire psychology, is concerned with the mental causation of behavior. By individuating the contents of the propositional attitudes of linguistically competent individuals on the basis of a literal interpretation of propositional attitude ascriptions, Burge's theory divides the domain of psychology into two parts, in essence claiming that the nature of the mental causation of behavior in non-linguists is fundamentally different from the mental causation of behavior in linguists.

CHAPTER 4

FRED DRETSKE'S INFORMATION THEORETIC ACCOUNT OF CONTENT

The viability of the computational theory of mind requires an account of how the content of intentional states is relevant to their causal powers. Furthermore, as a model of the cognitive mental states and processes routinely appealed to in commonsense explanations of behavior, the notion of content used by the computational theory of mind in its account of mental causation must accord with both common usage and the formality condition. Since the formality condition claims that two individuals can be in type identical mental states only if they are in type identical functional states, the requirement that the notion of content used be compatible with the formality condition amounts to requiring a narrow notion of content. In other words, the computational theory of the mind is a plausible model only if an account of how content influences causal powers can be constructed that incorporates a notion of content that is both narrow and accords with common usage. Burge's position is that the notion of content that is, in fact, used in commonsense explanations of behavior is a wide, communitarian notion of content. Had Burge's arguments been persuasive, we would have had good reason to doubt the appropriateness of the computational model. As we saw in the last chapter, however, Burge's position not only does not accord with

common usage in a variety of situations, but has a number of implausible ramifications as well. Consequently, he fails to cast doubt on the computational model. Fodor, on the other hand, claims that since the computational theory is the only even remotely plausible account of mental causation, we must presuppose an individuation of semantic content that follows computational properties. He maintains that the success of commonsense belief/desire psychology in predicting and explaining behavior is due to its implicit use of a narrow, individualistic notion of content. While the notion of content that Fodor endorses is, by explicit design, compatible with the formality condition, his theory denies that the content of intentional states is relevant to their causal effects on behavior. Although the narrow semantic properties of mental states invariably co-vary with their formal properties, there is no causal or nomic dependence between the semantic properties of mental states and their formal properties. Consequently, the semantic properties of mental states cannot enter into a scientific explanation of the causation of behavior. Content is predictively useful but causally inert.

Recently, Fred Dretske has suggested that the various attempts at reconciling commonsense belief/desire psychology with a scientific cognitive psychology have been, in essence, attempting to reconcile the wrong things. The reconciliation of commonsense psychology with a scientific theory has typically been seen to involve accounting for the co-variation of the content of a state and the causal powers of that state. Dretske claims that this view of what is required for vindicating commonsense psychology, a view which Dretske attributes

to, among others, Fodor, Loar, Stich, and even himself in earlier works, is fundamentally misguided (EB, p. 82). There is no direct connection between content and causal powers, so it is not surprising that theories that attempt to connect them cannot do so in a fashion that is sufficiently robust to be plausible as part of an account of mental causation. This lack of a direct connection between content and causal powers, however, does not, Dretske claims, mean that we must abandon commonsense psychology. It is Dretske's contention that in commonsense psychology beliefs and desires are not appealed to as causal explanations of how we move; the commonsense psychology model does not imply that there is a direct connection between content and causal powers.

[Beliefs (and desires) are not (typically) invoked to explain physical movements. They are brought into play to explain what we do: why we moved our finger, pressed the key, turned off the lights, or insulted our host. And though we can't do these things without some associated movement . . ., what we do must be carefully distinguished from whatever movements are required for the doing.

(ERC, p. 34)

Content, Dretske claims, is supposed to play a causal role in what we do -- how we behave -- not in how we move, and behavior, according to Dretske, "is to be identified with a complex causal process, a structure wherein certain internal conditions or events (C) produce certain external movements or changes (M)" (EB, p. 21). Content, Dretske claims, is supposed to explain causally not M, but why C causes M rather than some other movement or change (EB, p. 38). Once we "get clear about what is supposed to be causally explained by contentful inner states" (ERC, p. 33), we will be able to construct a theory in which content does have the causal power to influence what

we do.

Although Dretske advocates a somewhat different approach to reconciling ordinary folk psychology with a scientific psychology than Fodor does, nevertheless, Fodor and Dretske are engaged in fundamentally the same project: the vindication of commonsense belief/desire psychology. Fodor's interest is in constructing a notion of content that is plausible as an explication of our ordinary notion of content and which can meet the general methodological and theoretical constraints placed on any notion that is to be incorporated into a scientific theory. Dretske's interest is somewhat broader. He is interested in not only constructing a theory of content that is plausible as an explication of our ordinary notion, but in constructing a general theory of how content plays a causal role in behavior that incorporates his theory of content.

If our ordinary way of explaining behavior is to have any chance of being absorbed into a scientifically more respectable framework of explanation, some sense will have to be made of the way the content of our inner states figures in the explanation of behavior. If Clyde's trip to the refrigerator is to be explained by his belief that there is a beer left (and, of course, his desire to have a beer), then this belief, having this content, must be assigned causal powers. It must, together with relevant desires, intentions and collateral beliefs, have the power to move Clyde off the couch and into the kitchen. If it lacks this power, then this familiar way of understanding why people do what they do will someday take its place, along with witchcraft and demonic possession, on mankind's conceptual curio shelf.

(ERC, p. 31)

If, then, my body and I are not to march off in different directions, we must suppose that my reason for going to the kitchen -- to get a drink -- is, or is intimately related to, those events in my central nervous system that cause my limbs to move so as to bring me to the kitchen. My reasons, my beliefs, desires, purposes, and intentions, are -- indeed they must be -- the cause of my body's movements. What appeared to be two drummers must really be a single drummer.

(EB, p. ix)

Dretske was led to reexamine the role of content in commonsense psychology precisely because of the failure of various theories of content, including his own earlier view, to provide a causal role for content to play in behavior (see EB, pp. 79-83 and ERC, pp. 32-33). "The project is to see how reasons -- our beliefs, desires, purposes, and plans -- operate in a world of causes, and to exhibit the role of reasons in causal explanations of behavior" (EB, p. x).

Given his motivation, needless to say, Dretske explicitly designs his theory to avoid the problem of making content irrelevant to the causation of behavior that Fodor's theory suffers from. Dretske claims that beliefs are a type of representation. Following Armstrong¹ (and Ramsey²), Dretske says that beliefs are the internal "maps by means of which we steer" (EB, p. 79). In order for something to be a genuine intentional state, such as a belief, it must be, according to Dretske, that its representational content, that is, what it says about the world, must be relevant to "the direction in which one steers" (EB, p. 79). Intentional states are semantic structures that have an executive function, structures whose semantic content is a causal determinant of behavior. Intentional states are not simply, as Fodor's theory would have, those internal states or structures of a system that both play a causal role in determining the output of the system, and have semantic content. Beliefs, on Dretske's view, are representations "whose role in the determination of output, and hence in the explanation of behavior, is shaped by the relations underlying its representational content or meaning" (EB, p. 79).

In order to construct a naturalized account of intentional states, one in which intentional states are representations whose contents are

causally relevant to or responsible for their causal roles, Dretske must show that there is a principled way of identifying the content of an intentional state that does not itself rely upon intentional notions, and he must show how it is possible for this content or meaning of a state to determine its causal role in a system. Dretske argues that one can construct a naturalized account of how the states of a purely physical system can acquire representational content, and how those states, in virtue of their content, can be part of a causal explanation of behavior using only materialistic notions of information, function, and learning: A semantic theory of information derived from communication theory will provide an account of when -- that is, under what conditions -- one thing (event, state, condition) carries information about another, as well as what information is carried; a notion of function will determine which piece of information one state carries about another is that state's representational content; and a notion of discrimination or associative learning will be used to explain how it is possible for the content of a state to determine its causal role in a system.

While there is much in Dretske theory that warrants examination, it is my intention to focus on the notion of content that Dretske has developed and incorporated into that theory. In Dretske's own opinion, unless the notion of content that is constructed is plausible, the vindication of commonsense psychology is not possible.

[W]hat I propose to do here is to look for some recognizable notion of content, something that will (perhaps with a little dressing up) pass muster as the meaning or content of a physical state, something that also plays a role in the causal determination of behavior. If nothing can be found that plays this dual role, or what is found is too anemic to be plausibly identified with the content of a belief or desire, then so much the worse for the explanatory framework,

ordinary folk psychology, that relies on such elements for its understanding of why we do what we do.

(ERC, p.33)

It will be my contention that the wide notion of content that Dretske constructs is not plausible even within the confines of his own theory. The components of Dretske's theory neither individually nor collectively require a wide notion of content. Although I believe that Dretske is wrong to advocate a wide notion of content, I believe that the basic structure of Dretske's theory, and the connection it posits between content and causation is fundamentally correct. The weakness in Dretske's position lies not in the general framework, but in the notion of content that he uses. Ultimately, my goal is to retain Dretske's basic idea that in intentional systems mental states acquire their causal roles in virtue of what they indicate about the world. But the achievement of that goal is possible only if the notion of content that is, in fact, required by such a theory is plausible both as a scientific construct and as an explication of our ordinary notion of the content of a belief or desire. In this chapter, I will argue that the sort of information-cum-function account of intentionality and mental causation that Dretske proposes requires, in fact, a narrow notion of content. In the next chapter, I will develop a particular notion of narrow, individualistic content that I believe is plausible from the standpoints of both science and ordinary intuitions, and can be coupled with the Dretske's basic framework. Since the information theoretic foundation on which Dretske builds his account of how content can be relevant to causation is assumed as the basis for my argument for a narrow notion of content in Chapter 5, I will to give a detailed account of the semantic theory

of information that Dretske constructs.

I. COMMUNICATION THEORY AND THE CONDITIONS FOR INFORMATION TRANSMISSION

Mathematical information theory, or communication theory, is concerned with measuring the amount of information generated by or associated with the various events or states of affairs that can occur (or exist) at some source, as well as measuring the amount of information from that source that is transmitted or carried by some signal. Communication theory identifies the amount of information associated with or generated by the occurrence of some specific event at the source with the reduction in uncertainty, or the elimination of prior possibilities, at the source. The greater the degree of reduction in uncertainty that the occurrence of an event represents, the larger the amount of information associated with that event. For example, consider two sources of information, a (fair) die, and a (fair) coin. There are six possible, equally probable, outcomes of a role of the die, but only two possible, equally probable, outcomes of a flip of the coin. The degree of uncertainty concerning the outcome of any given role of the die is much greater than the degree of uncertainty concerning the outcome of a flip of the coin. There is only a (roughly) 17 percent chance of any particular number coming up on a given role of the die, while there is a 50 percent chance of either of the two sides coming up on a given flip of the coin. Thus, according to communication theory, a role of the die generates more information than a flip of the coin, because the role of the die

eliminates five out of six (prior) possibilities, while the flip of the coin eliminates only one out of two (prior) possibility. However, communication theory is, itself, not concerned with the amount of information generated by the occurrence of any particular state or event at the source. Rather, it is concerned with the range of possible amounts of information generated by a source, as well as the average amount of information generated by a source or transmitted by a signal. Nevertheless, Dretske claims that information theory can be used as a foundation for a semantic theory of content. According to Dretske, information theory can be adapted not only for determining the amount of information associated with particular events or states of affairs, but for determining what information (i.e., what content) is associated with those states or events.

According to communication theory, the amount of information generated by a state or event is determined by the extent to which that state or event represents a reduction of possibilities at the source. However, the amount of information generated by any particular state or event, and by extension, by any source, is not determined solely by the reduction of possibilities. How likely each possibility is to occur plays a role in determining the amount of information that would be generated by each possible outcome. The more likely an outcome, the less information associated with its occurrence. For example, if one is flipping an unfair coin, say, one that is strongly biased towards tails, the amount of information generated by a flip of the coin coming up tails is less than the amount of information generated by a fair coin coming up tails, and much less than the amount of information generated by the unfair coin

coming up heads. Intuitively this seems reasonable: the more likely an event, the less surprising its occurrence is, and the less information its occurrence generates. The more probable a possible outcome is, the less information generated by its happening. The amount of information generated by the occurrence of a particular event depends upon the prior probability of that event occurring relative to the other possible outcomes. In general, the amount of information generated by the occurrence of a particular event or outcome is inversely related to the probability of that event or outcome happening. Specifically, communication theory claims that the amount of information, I , generated by an event, e , is:

$$I = \log 1/\text{pr}(e)$$

where $\text{pr}(e)$ is the prior probability of e occurring relative to those things that could have happened.³

While the amount of information generated by the occurrence of some event is determined by its probability, the amount of information generated by that event and transmitted to some receiving station will depend on not only the amount of information generated at the source, but also on the amount of information lost in the transmission process. All of the information generated at the source may not make it to the receiving station. For example, imagine the case of a student taking an exam. The exam is graded and given a 'C'. However, the student signed up to take the course pass/fail, and, consequently, is told only that she has passed the exam. She is told that she has passed on the basis of having gotten a 'C' on the exam, but in passing the exam more information was generated than is transmitted to her. She would have received the same information, that is, that she had

passed, if she had received any grade other than 'F'. The information she receives only eliminates 1 of the possible events at the source, whereas the information at the source eliminated four out of the five (prior) possibilities. Thus, she receives less information about the exam, than is associated with the exam.

The information generated at a source but not transmitted to the receiver (or carried by a particular signal) is called 'equivocation'. The amount of information generated at a source, \underline{s} , that is transmitted to a receiver is:

$$I_r(e) = I(e) - E$$

where $I_r(e)$ is the amount of information generated at \underline{e} and received at \underline{r} , and E is the equivocation between source and receiver.

Equivocation is essentially the degree to which events at the source can vary independently of events at the receiver. The less closely variations in events occurring at the source are mirrored by variations in events occurring at the receiver, the greater the equivocation.⁴ The equivocation between a receiver and a source is the degree of uncertainty about what actually happened at the source given what has happened at the receiver. For example, in the case of the student, there are only two possible states or events at the receiver (the student), she can be told that she has passed or she can be told that she has failed. On the other hand, there are five possible events or states ('A', 'B', 'C', 'D', and 'F') at the source. When she is told that she has passed, she has not been told what actually happened at the source (the exam). She has been told only that one of four events happened ('A', 'B', 'C', or 'D'). What

happened at the receiver is somewhat independent of what happened at the source. Thus, the information she receives is equivocal with respect to the information generated at the source. The equivocation, E , between a particular event at a receiver and a source is determined by looking at the conditional probabilities of each of the possible events at the source given what happened at the receiver.

$$E_r = \sum P(s_i/r) \cdot \log P(s_i/r)$$

where $P(s_i/r)$ is the conditional probability of s_i having occurred at the source given that r occurred at the receiver. If one can be certain that s_i occurred at the source given that r occurred at the receiver (the conditional probability of s_i is 1) then r is unequivocal with respect to the source. However, if one cannot be certain that any one s_i occurred at the source given r (a conditional probability of s_i of less than 1), then r is equivocal with respect to the source.

In order to use the above formula to calculate an actual numerical value for the amount of information generated by some state of affairs or carried by a signal, r , about that state of affairs, one has to know (1) what the alternative possibilities at the source were; (2) what the (prior) probability of each of those alternatives was; and (3) what the conditional probability of each of those alternatives is given r . While in the carefully circumscribed situations with which communication theory is concerned identifying the various possible events that could happen at the source, as well as the probability of each of those events occurring, is at least imaginable, in real life situations no such identification seems even remotely

possible. Even if we already know that Edith is the source of the information a signal carries, how are we to determine what the possible states of the source (i.e., Edith) are? Is playing tennis a possibility at the source, or should that possibility be identified as playing tennis poorly, or losing at tennis. Similarly, how should the alternative possibilities be identified? Is playing golf to be identified as an alternative, or should each or the innumerable ways in which she could be playing golf be counted as an alternative? Furthermore, what about such things as turning into a cockroach, is that a possibility as well? The amount of information generated by some real-life event cannot be calculated unless there is some a principled way of determining what the possibilities at the source are, and how the possibilities at the source should be individuated.

Nevertheless, Dretske claims that the formula of communication theory can be used as the basis for a semantic theory of information, because such a theory requires only that we be able to compare the relative amounts of information generated by a source and carried by a signal about that source. For such comparative purposes, one need not know what the possibilities at the source are, or what their relative probabilities are because those factors will contribute equally to the amount of information generated by the source, and the amount carried by the signal. All one needs to be able to do is calculate the equivocation between the source and the signal, a calculation that Dretske claims can, in fact, be performed.

Dretske claims that on the basis of these comparative judgments of the amount of information generated by a source and transmitted by a signal about that source, a theory of the information content of

signals can be constructed. From communication theory, we know that a signal that indicates a reduction in the number of possibilities at a source from \underline{n} possible states to $\underline{n} - \underline{m}$ possible states cannot indicate a reduction of the number of possibilities at the source from \underline{n} to $\underline{n} - (\underline{m} + 1)$. A signal can only carry information about a source that represents a reduction in the possibilities at the source no greater than what the signal can carry about that source. A signal, \underline{r} , cannot carry the information \underline{p} about a source, \underline{s} , if the amount of information generated by or associated with \underline{s} 's being \underline{p} (i.e., the reduction in the number of possibilities at the source that \underline{s} 's being \underline{p} entails) is greater than the amount of information about that source that the signal can carry (i.e., the reduction in the number of possibilities at the source that the signal can indicate). The signal 'Pass' that a student receives about an exam cannot carry the information 'C' about the exam, because the signal 'Pass' reduces the number of possible states at the source from five to four, while the source being 'C' reduces the number of possible states at the source from five to one. "[T]he amount of information a signal carries about a source sets an upper bound on what information a signal can carry about that source" (KFI, p. 62). If a certain piece of information about a particular source indicates a greater reduction in the number of possibilities at the source than can be carried by a given signal, then that signal does not carry that piece of information.

Although information theory puts this minimal constraint on what information a signal carries, it is clearly not enough to determine what information a signal actually carries. Nevertheless, if we are interested in constructing a semantic theory based on information

theory, information theory does specify two of the conditions that a definition of the information content of a signal must embody. In order for a signal to carry the information F about s,

- (A) The signal must carry as much information about s as would be generated by s's being F; and
- (B) s is F.

While (A) and (B) are clearly both necessary in order for a signal to carry the information F about s, they are not jointly sufficient. For example, suppose that there are 9 different colored pieces in a box (three different shapes, each of which comes in three different colors) one of which is selected, say the blue tetrahedron. A signal that carried the information that the piece selected was a tetrahedron would carry exactly the same amount of information as a signal that carried the information that the piece was blue, since both represent a reduction of 9 possibilities to three. So, although the signal carries enough information to carry the information that the piece is a tetrahedron, the information that it actually carries is that the piece is blue, not that it is a tetrahedron. Dretske suggests that the way to correct for this situation is to require something like,

- (C) The quantity of information the signal carries about s is (or includes) that quantity generated by s's being F (and not, say, by s's being G).

Dretske openly admits that (C) is wholly inadequate as part of a rigorous definition of what it is for a signal to carry a particular piece of information. "It is not clear, for example, what it could mean to say that one quantity (amount of information the signal carries) is (or includes) another quantity (amount of information generated) when this is meant to imply something more than a numerical

comparison" (KFI, p. 64). However, Dretske claims that since (A), (B), and (C) are merely meant to be conditions that an adequate definition will have to meet, so long as the definition can meet the underlying intentions of the conditions, difficulties with the particular formulation of any of the three conditions should not be a problem. In the case of (C), Dretske claims that the condition is merely meant to capture the intuition that a signal has to carry the information that was generated by s's being F, not just the right amount of information, in order for it to carry the information F about s.

Dretske claims that the only definition that can meet all three of the conditions mentioned is the following one:

A signal (state, event, condition), r, carries the information F, about a source s, if and only if the conditional probability of s's being F, given r (and k), is 1 (but, given k alone, less than 1).⁵

(KFI, p. 65)

where k stands for what the individual already knows about s. The idea behind including what the individual already knows about the source in the definition of the information content of the signal is to capture the intuition that what one can learn about a source "from" a given signal depends upon what you already know about the source. For example, if I already know that you live in New York City, and you tell me that you live at 44 Horatio Street, I apparently can learn exactly where you live just from being told your street address. If, however, I do not already know what town or city you live in, I cannot learn where you live by being told that you live at 44 Horatio Street.⁶ Dretske claims the above definition of what it is for a signal to carry a particular piece of information satisfies each of

the three conditions that he identified. It satisfies (A) because if the conditional probability of s's being F (given r is 1), then the equivocation of r must be 0 (r occurs when and only when s is F) so the signal carries as much information about s as is generated by its being F. It satisfies (B) because the conditional probability of s being F can be 1 only if s is F. It satisfies (C) because "whatever other quantities of information the signal may carry about s, our definition assures us that the signal includes the right quantity (the quantity associated with s's being F) in virtue of excluding just those situations that motivated the imposition of this requirement" (KFI, p. 65-6). It is only by carrying the "right" quantity of information that the signal will make the conditional probability of s being F 1.

By identifying the information a signal, state, or event, r, carries with that property whose conditional probability is 1 given r, but less than 1 otherwise, it might seem that Dretske is suggesting a correlation theory of information. If the statistical correlation between events at point A and events at point B is 1, then B carries information about A, and A about B. However, Dretske explicitly denies that correlation alone, even pervasive correlation, is sufficient, either in his use of information theory or in communication theory itself, for there to be an "information link" between two points.

The transmission of information requires, not simply a set of de facto correlations, but a network of nomic dependencies between the condition at that source and the properties of the signal. The conditional probabilities used to compute noise, equivocation, and amount of information (and therefore the conditional probabilities defining the information content of the signal) are all determined by the lawful relations

that exist between source and signal. Correlations are irrelevant unless these correlations are a symptom of lawful connections.

(KFI, p. 76-77)

In technological applications of communication theory statistical correlations are used to determine the relevant possibilities and probabilities, but what those possibilities and probabilities are is determined by the lawful regularities between source and signal. Statistical correlations are relevant to the flow of information "only insofar as these correlations are manifestations of underlying lawful regularities" (KFI, p. 247). When Dretske states that a signal carries the information F about s if and only if the conditional probability of s being F is 1, he does not intend for the connection between the signal and the state of s to be simply a statistical one. According to Dretske, in order for the conditional probability of s being F, given r, to be 1, there must be "a nomic (lawful) regularity between these event types, a regularity which nominally precludes r's occurrence when s is not F" (KFI, p. 245). To say that there is a conditional probability between s and r is to say that there is "a lawful (exceptionless) dependence between events of this sort" (KFI, p. 245). Correlations, even exceptionless correlations, are not sufficient for the conditional probability (regardless of what information is being considered) to be 1, and hence for the transmission of information. There must be a nomic dependence between source and signal in order for the signal to carry any information about the source.

This requirement that there be a nomic dependence of the signal on the source should not, however, be construed as requiring that there be a strict causal connection between signal and source, where a

strict causal connection requires that a cause be an essential part of some nomically sufficient condition for the effect. It is nomological dependence that is responsible for an information relation existing between some signal, r, and a source, s, not causal dependence. In general, r will be nomologically dependent on s, and therefore an information relation will exist between them, in virtue of there being an underlying causal process that makes what happens at r causally dependent upon what happens at s. Nevertheless, a causal dependence is neither necessary nor sufficient for an information relation to exist between r and s. "One can have full information without causality, and one can have no information with causality" (KFI, p. 33). For example, if you are looking at the back of a playing card, say the 5 of hearts, although it is the 5 of hearts that is the cause of your visual image, the visual image carries no information about which playing card you are looking at. There are 52 possible cards that you could be looking at, and the one that is causing you visual image gives you no information about which one you are looking at. Since information is determined by a reduction in possibilities, if all of the possibilities at a source produce a signal or effect of the same type, in this case type identical visual images, then the signal carries no information about the source, because it does not reduce the number of possibilities at the source. On the other hand, consider the case of two television sets tuned to the same channel. What is happening at set A carries information about set B, and vice versa, because A and B are both dependent upon what is happening at the broadcast station in the same way. Although there is no causal or physical link between A and B, there is an information relation

between them (KFI, pp. 38-9).

It may be objected that by requiring that the conditional probability of s being F given r (a token of type R) be 1 in order for r to carry that information about s, Dretske precludes the possibility of any signal carrying the information F about s. The information relation, and the nomic regularities that ground that relation, exist between event types. Signals of type R, carry the information F in virtue of the lawful regularities that exist between signals of type R and events having the property F. But, one can always imagine circumstances under which a token of type R will occur, even though there is no s that has the property F, thus ensuring that the conditional probability of s being F given r will be less than 1. One need not turn to science fiction for cases of this sort. Imagine a voltmeter attached to a circuit containing a resistor, one that creates a drop of 7 volts. The leads of the voltmeter are connected to points at opposite ends of the resistor. A voltmeter functions to measure the voltage drop between the two points in a circuit. With the leads of the voltmeter attached to opposite ends of a resistor, the voltmeter will function to measure the voltage drop across the resistor. When there is a voltage difference between the two points to which the voltmeter leads are connected, a current flow is generated within the voltmeter, setting up a magnetic field, which produces a torque on an armature (against the force of a restraining spring) on which a pointer is mounted. While it is obvious that the position of the pointer will depend on the voltage difference between the two points to which the leads are connected, it also depends upon such things as the number of windings around the voltmeters

electromagnets, and the elasticity of the restraining spring. If either of those things were to vary, then the pointer could come to rest at the same position on the scale even though the voltage drop across the the resistor was not 7. The conditional probability of the voltage drop being 7 given that the pointer comes to rest at '7' on the voltmeter scale would seem to be less than 1, since the pointer might come to rest at '7' on the voltmeter scale even though the voltage drop was only 5, if the elasticity of the restraining spring had changed. Dretske's reply to this objection is that the conditional probability of s being F given r (and k) must be 1 only with respect to what he calls the relevant alternatives of the source. Not everything that could possibly have an effect on a signal is a source of information for that signal. Some of the things that a signal is causally or nomologically dependent upon are not to be considered possible information sources. "The view that the information requirement on knowledge is too severe, that it cannot (in most practical situations) be satisfied, ultimately rests on the confusion between: (1) the information (about a source) a signal carries, and (2) the channel on which the delivery of that information depends" (KFI, p. 111). Some of the things on which the delivery of a signal depends are not relevant alternative sources of information. They are part of the channel on which the signal depends, about which the signal carries no information. The elasticity of the spring in a voltmeter is a part of the channel over which information about the voltage drop across two leads is transmitted. A change in the elasticity of the restraining spring is not a relevant alternative source of information for the signal. In determining the conditional

probability of the voltage drop being 7 given that the pointer is at rest at '7', one does not have to take into consideration the possibility that the elasticity of the restraining spring has changed. Even though the conditional probability of s being F given r is less than 1 relative to all (logically) possible alternatives, so long as the probability is 1 relative to the relevant alternatives, r will carry the information F about s. A signal can carry information about a source just in case it eliminates all relevant possibilities.

Dretske claims that anything that is a part of what he calls the "channel of communication" -- which is defined as "that set of existing conditions (on which the signal depends) that either (1) generate no (relevant) information, or (2) generate only redundant information (from the point of view of the receiver" (KFI, p. 115) -- is not something about which a signal carries information. According to information theory, a signal transmits information to the extent that it reduces the number of possible states, event, conditions, etc. that might obtain at the source. In order to determine the amount of information carried by a signal one has to look at the possibilities at the source, their prior possibilities, and their conditional probabilities given the signal received. If any of the "possibilities" at the source involve or depend upon a condition, that reliably does not to change (the force of "reliably" in this context will become clear in a moment), being different from what it actually is, then the prior probability of that "possibility" can be considered to be 0 -- making it not really a possibility at all, in Dretske's words, not a relevant possibility. Even though it is logically, even physically, possible for that condition to be different from what it

is, the fact that it reliably does not change makes it part of the channel conditions on which the signal depends. The signal carries no information about the condition because no reduction in the number of possibilities is possible: there is only one possible conditions. "It is important to emphasize that what qualifies a condition as a channel condition is not that it is known to be stable, not that it is known to lack relevant alternative states, but that it is stable, lacks relevant alternative states, in fact generates no (new) information" (KFI, p. 119). Failing to eliminate an irrelevant possibility can have no effect on the degree to which the signal indicates a reduction in the possibilities at the source. Although a signal can carry information about only those things on which it is nomologically dependent, not everything on which a signal is nomologically dependent will it carry information about. A signal carries information only about those things that it is nomologically dependent upon and which have genuine possible alternative states.

Needless to say, Dretske is aware that his invocation of a notion of a relevant or genuine possibility is likely to raise suspicion. What determines whether or not something is to count as a genuine possibility? In particular, given that information is supposed to be both an absolute and an objective commodity, how is one to distinguish between relevant and irrelevant possibilities? Dretske claims that, ultimately, what is to count as a channel condition, and, consequently, what is to count as a relevant alternative is a matter of opinion. The evaluation of a system of communication is dependent upon "social" and "pragmatic" considerations.

Whether or not a signal carries a piece of information depends on what the channel is between source and receiver, and the question of whether an existing condition is stable or permanent enough to qualify as part of the channel, as a condition which itself generates no (new) information, is a question of degree, a question about which people (given their differing interests and purposes) can reasonably disagree, a question that may not have an objectively correct answer. When a possibility becomes a relevant possibility is an issue that is, in part at least, responsive to the interests, purposes, and, yes, values of those with a stake in the communication process.

(KFI, pp. 132-133)

Dretske claims that this "social or pragmatic dimension" does not impugn the absolute nature of information because the pragmatism has to do only with our use of the absolute concept, not with the concept itself.

To be empty is to have nothing in it, and in this respect, the emptiness of something is not a matter of degree. Nonetheless, whether something counts as a thing for purposes of determining emptiness is a matter of degree, a question about which people (given their differing interests and purposes) might easily disagree, a question that may not have an objectively correct answer. To be empty is to be devoid of all relevant things. The concept, though absolute, has a built-in plasticity (in the idea of a "relevant" thing) that is responsive to the interests and purposes of people applying it. Knowledge and information are no different. To know, or to have received information, is to have eliminated all relevant alternative possibilities. These concepts are absolute. What is not absolute is the way we apply them to concrete situations -- the way we determine what will qualify as a relevant alternative.

(KFI, p. 133)

Information is not a matter of degree, neither is whether or not something counts as a relevant alternative. A signal either does or does not carry a particular piece of information, and an alternative either does or does not count as a genuine alternative. Neither is a matter of degree. The "flexibility" or "plasticity" of information that Dretske sees is, he claims, due to the fact that the criteria we use to determine whether or not something counts as a genuine alternative is interest relative.

Although in his discussion of channel conditions Dretske focuses almost exclusively on the condition of the physical system causally responsible for the transmission of the signal, in fact, Dretske wants a rather wider range of things to fall under the heading of channel conditions. Not only are channel conditions, in his view, to include those static conditions of the physical system, such as the number of windings around an electromagnet, but they are also to include those things whose affect on the signal is "well below the level of precision" of the system (KFI, p. 252), and those things that exist only outside of the system's "natural habitat." A couple of examples should make it obvious why Dretske wants to include this broader range of things under the heading of channel conditions rather than including just the static physical condition of the means of transmission. In the case of the voltmeter, the position of the pointer is dependent upon the resistance of the leads that connect the voltmeter to the resistor. Unlike the number of windings around the internal electromagnet of the voltmeter, however, the resistance of the leads is temperature dependent, and so is constantly changing. It is not the case that it generates no (new) information. Strictly speaking, variations in the resistance of the leads due to temperature changes should make the voltmeter signal equivocal. Nevertheless, although the resistance of the leads is constantly changing, the effect of temperature on the resistance of the leads is negligible given the precision of the voltmeter. For example, if the voltmeter is designed to measure voltage differences greater than or equal to .1 volts, then variations in the resistance of the leads due to temperature that result in a .001 volt error in the measurement are

not, according to Dretske, to be considered to make the signal equivocal. Presumably, Dretske holds that such variations do not affect whether or not the signal is equivocal because the signal does not carry such precise information. When the pointer on the voltmeter registers 7 volts, it does not carry the information that the voltage drop is precisely 7 volts. When the pointer registers 7 volts, the information it carries is that the voltage drop is 7 volts, to the nearest tenth of a volt.

He also wants to eliminate as relevant possibilities those things that exist only outside the systems "natural habitat." A signal is not to be made equivocal on the basis of the fact that there are things outside of the system's natural environment that would cause the same signal type were the system in some other environment. A signal's reliability must be assessed within its natural, i.e., a restricted, context.

[An assembly-line worker's] ability to recognize resistors inside the factory is to be explained by the fact that in the factory, on the assembly line, nothing is allowed to appear that he might confuse with a resistor (inside the factory there is no equivocation). Nevertheless, since he does not know the difference between a resistor and a capacitor, and since some capacitors look very much like resistors, he cannot (outside the factory) recognize a resistor (even those he correctly calls resistors). If, then, we include his being in the factory as one of the channel conditions (something we might explicitly do by saying "in the factory he can tell"), as something we tacitly hold fixed in reckoning equivocation, there is nothing to prevent us from saying that (in the factory) he is getting the information that s is a resistor. We often (explicitly or implicitly) make similar provisos about an organism's natural habitat in characterizing its cognitive capacities.

(KFI, pp. 253-254)

The fact that a signal will be equivocal in an larger context does not make the signal equivocal in a more limited context. A system's natural habitat is, according to Dretske, a channel condition.

Laboratory experiments and parallel universes are not to be permitted to count as relevant alternatives when determining whether or not a signal is equivocal.

While there is some basis in intuition for Dretske's claim that the first two types of conditions, which are concerned with the physical condition of the signal transmission apparatus, should be ruled out as possible sources of equivocation, the inclusion of context as a channel condition is unwarranted without additional argument. The differences between context and the other sorts of channel conditions are sufficient that the justification Dretske gives for eliminating the other sorts of channel conditions as relevant alternatives when determining equivocation is inapplicable in the case of context. (To avoid confusion, in the present discussion I will use 'channel condition(s)' to refer to only the first two types of things that Dretske identifies as channel conditions, and will use 'context' or 'context conditions' to refer to the generally stable environmental context in which the system is functioning.) A channel condition is ostensibly something that is a part of the channel of communication, which Dretske defined as the set of conditions on which the signal depends that either generates no new information or only redundant information. While the physical condition of the transmission apparatus is something on which the signal nomically depends, the context in which the system is embedded is not something on which the signal depends. Specifically, it is the information the signal carries in virtue of there being a reliable, but not nomic, correlation between the signal and the source. Information that the signal carries in virtue of a nomic correlation is left unchanged by a

change in the context conditions. It is the information carried by a signal, not the signal itself, that is dependent upon the context. Changing a channel condition, e.g., the elasticity of a spring in a voltmeter, will change the signal, but not the information transmitted; changing the context, e.g., moving the assembly-line worker from inside the factory to outside the factory, leaves the signal unchanged, but will alter (some of) the information carried by the signal.

In the case of the voltmeter, the signal is the angle of deflection of the armature, while the information carried by the signal is the voltage difference between the two points to which the leads are connected. Attaching the leads of the voltmeter to the opposite ends of a resistor in a circuit results in an angle of deflection of the armature of, say, 87 degrees, which, assuming the voltmeter is working properly, that is, it is functioning according to the way it was designed to function, carries the information "5 volts." If, while holding the context constant,⁷ we change a channel condition, say replace the spring in the voltmeter with a weaker one, then attaching the voltmeter to the same resistor will result in a different signal, say, an angle of deflection of 105 degrees, but if we make the same assumption about the voltmeter's condition, i.e., it is functioning the way it was designed to, then this new signal -- an angle of deflection of 105 degrees -- still carries the old information, "5 volts." The signal type is still nomenclally correlated with exactly the same set of event, condition, and state types as it was prior to the alternation in the channel condition. The angle of deflection is still nomenclally correlated with the elasticity of the spring. Note

that by changing this channel condition every signal that is transmitted over the channel will be altered, and they will all be altered in a consistent and predictable way. Changing the channel condition changed the code or translation table for the signals of this system, not the information transmitted by the system.

In the case of the assembly-line worker the context is the factory, and the channel conditions are presumably (Dretske does not say) the perceptual and cognitive apparatus of the worker. Inside the factory we present the worker with a resistor. The signal in the assembly-line worker is an internal brain state, call it 'b57,' and, assuming that the worker's perceptual and cognitive apparatus are working properly, the information carried by b57 (a token of type B57) is, on Dretske's view, "resistor" because B57 is nomically correlated with something having a certain size and appearance, and inside the factory there is a reliable correlation between things of that size and appearance and resistors. If, while holding the channel conditions constant we change the context in which the worker finds himself by placing him, let us say, in an electronic parts store, and present him with the same resistor, the same signal, b57, will occur in the head of the worker, but some of the information carried by b57 will be different. In the electronic parts store signals of type B57 carry the information "resistor or capacitor" because in the electronic parts store having a certain size and appearance is reliably correlated with both resistors and capacitors. The signal is the same -- the signal is not dependent on the context, only the information has changed. Of course it is not the case that all of the information b57 carries has been changed by changing the context conditions. In

addition to carrying (according to Dretske) the information "resistor" inside the factory, b57 also carries the information "object of such and such size and appearance." This latter piece of information, unlike the former, is carried by b57 in virtue of their being a nomic correlation between B57 and object of such and such size and appearance. Changing the context conditions does nothing to alter this piece of information that b57 carries. Changing the context conditions changes the information b57 carries because changing the context changes what objects of such and such a size and appearance of reliably correlated with in the environment. Inside the factory objects of such and such size and appearance are correlated with resistors, while in the electronic parts store they are correlated with resistors and capacitors. Note also that unlike changing a channel condition, changing the context affects only a single information carrying signal type. Only those signal types that were reliably correlated with the event, state, or condition types that were altered by the change of context will have the information they carry altered. Moving the assembly-line worker from the factory to the electronic parts store affects only the information carried by tokens of type B57. It has no effect on any other information carrying signal or on the information tokens of type B57 carry in virtue of a nomic correlations. If inside the factory B6 carries the information "astroturf" or B98 the information "Big Mac," putting him in the electronic parts store will have no effect on the information carried by those signals. Changing context conditions produces highly specific, localized changes in the information that can be carried by certain signals by changing what things in the local environment

signal types can be reliably correlated with. Changing channel conditions, on the other hand, does not change what things signal types can be correlated with, but only changes the mapping between signal types and things in the world.

II. DRETSKE'S ACCOUNT OF CONTENT IS BASED ON INFORMATION THEORY AND A NOTION OF THE FUNCTION OF A STATE

Dretske claims that in a very basic or rudimentary sense, all signals that carry information are intentional states: "[A]ll information-processing systems occupy intentional states of a certain low order. To describe a physical state as carrying information about a source is to describe it as occupying a certain intentional state relative to that source" (KFI, p. 172). Signals can carry the information "F" about a source, s, without carrying the information "G" about s even though "F" and "G" are coextensive. Whether or not a signal carries a particular piece of information is dependent upon there being a nomic correlation between it and the signal. If "F" and "G" are coextensive, but only "F" is nomically correlated with a signal, then the signal will carry the information "F" without carrying the information "G." "So, for example, even though all Elmer's children have measles, and S carries the information (has the content) that t is one of Elmer's children, S may not carry the information (have the content) that t has the measles" (KFI, pp. 172-173). Ultimately, it is from this intentional character of signals

that mental phenomena derive their intentionality, their representational powers.⁸

While both signals and intentional states,⁹ such as beliefs, have the property of distinguishing between extensionally equivalent properties, beliefs and the like have the further property of "expressing" (indicating) a single piece of information. Beliefs have a single informational content. Granting for the moment that there is a principled way of restricting the range of possible sources of a signal to a set of relevant alternatives, nevertheless, by defining what information a signal carries in terms of conditional probability, Dretske ensures that every signal will simultaneously carry many pieces of information about a variety of different sources. Even after having eliminated as possible sources of a signal all those things that involve altered channel conditions (including context conditions), it will still be the case that any number of things will stand in an information relation to a given signal. For example, given the way an electrically operated fuel gauge works a number of different things will have a conditional probability of 1 given any position of the pointer.

Electrically operated fuel gauges indicate not only the amount of fuel left in the tank but also the amount of electrical current flowing in the wires connecting the gauge to the tank, the amount of torque on the armature to which the pointer is affixed, and the magnitude of the magnetic field surrounding this armature. Given the way these gauges operate, they cannot indicate (i.e., have their behavior depend on) the amount of fuel in the tank without indicating (exhibiting at least the same degree of dependency on) these related conditions.

(EB, p. 59)

The conditional probability of the tank being half full cannot be 1 without the conditional probability of, for example, the magnetic

field surrounding the armature being some set value thereby being 1 as well. The signal, the position of the pointer, carries information about the amount of fuel in the tank, the amount of current flowing through the wires, the torque on the armature, the magnitude of the magnetic field surrounding the armature, and any number of other things. All of these things are sources about which the signal carries information. To the extent that the signal cannot carry information about one of these sources without carrying information about the others as well, they are all sources of the signal. In order for Dretske to be able to construct an account of intentional states from the notion of what information a signal carries that he has suggested, he has to show that there is a principled way of selecting from the many pieces of information a signal carries that piece of information that will be the content of the intentional state.¹⁰ "To occupy a belief state a system must somehow discriminate among the various pieces of information embodied in a physical structure and select one of these pieces for special treatment -- as the content of that higher-order intentional state that is to be identified as the belief" (KFI 174).

Dretske claims that intentional states have a single piece of information as their contents in virtue of being representations. According to Dretske, a representation is the "expressive" element in a system (a representational system, henceforth 'RS') whose function it is to indicate how things stand with respect to something else. The ability of an RS to indicate how things stand with respect to something else is derived from the ability of representations to indicate -- carry information about -- how things stand with respect

to something else. Representations are, in essence, a type of signal. However, not all signals are representations. A signal is a representation only if (1) it is part of a representational system, that is, a system whose function it is to indicate how things stand with respect to something else; and, (2) it is in virtue of carrying the information it does that the signal is part of the RS (EB, p. 52). That piece of information that the signal carries that is responsible, in some sense, for the signal's being a part of the RS is the representational content of the signal-as-representation, and what the piece of information is that is responsible for the signal's inclusion in the RS is determined by what it is the function of the RS to indicate. The content of a representation is that piece of information that the RS of which it is a part has the function of indicating. Thus, intentional states, as representations, have a single content determined by what it is the function of the system of which they are a part to indicate.

Considered as simply an indicator, the position of the pointer of the "fuel guage" is equally an indicator of electric current, torque, and magnitude of a magnetic field, as well as the amount of fuel in the tank. "Nevertheless, we take one of these indicated conditions to be what the gauge represents, one of these correlated conditions to define what kind of gauge it is" (EB, p. 59). Since what we are interested in is the amount of fuel in the tank, we designed a system, the fuel gauge, whose function it is to provide this piece of information. What it is the function of the system to indicate determines which of the many pieces of information a signal may indicate it represents. Thus, since the gauge is supposed to indicate

the amount of fuel in the tank, that is, the function of this system is to indicate the amount of fuel, then that is the piece of information that the signal carries that is its representational content.

Dretske claims that there are three basic types of representational systems distinguished from one another on the basis of (1) what makes it the case that the representational elements in the system indicate what they do, that is, in virtue of what is it the case that the representational elements carry information about the source; and (2) what makes it the case that the the system, or its representational element, has the function of indicating what it does. Type I RS's are wholly conventional or derivative. They derive their ability to represent solely from those who create and use them. They have no intrinsic ability to carry the information they are used to convey. Type II systems are conventional as well but their ability to represent what it is their function to represent is dependent, in part, on what information they carry. They are systems assigned, by those who created and use them, the function of indicating one of the things they carry information about, one of the things they are a natural sign of.¹¹ Type III RS's are wholly intrinsic. They succeed in indicating what they do because they are natural signs of those things, and they have the function of indicating what they do in virtue of being assigned that function by the system of which they are a part. It is systems of Type III that are to be the focus of the remaining discussion, but a brief overview of the other types will be useful in understanding the particular features of RS's of Type III. Given that something is a representation only in virtue of being part

of an RS, I will talk about representations of each of these three types, as well as representational systems of each type. A representation of a particular type is simply a (the) representational element of an RS of the corresponding type.

Type I representations are part of systems that are created and used by individuals outside of the system to indicate how things stand with respect to some other thing. An RS of Type I is constructed by assigning to various object (what Dretske calls the representational elements of the system) the function of indicating something that they do not carry information about. Dretske gives the example of constructing a system whose function is to represent the relative positions and movements of the players in a basketball game. By stipulating that this dime is Oscar Robertson, this nickle is Kareem Abdul-Jabbar, this penny the opposing center, and these pieces of popcorn the other players, he can, by moving them about, represent the positions and movements of those players in a game played last year.

The coins and the popcorn have been assigned a temporary function, the function of indicating (by their positions and movements) the relative positions and movements of certain players during a particular game. But these elements, the coins and the popcorn, obviously enjoy no intrinsic power to do what they have been assigned the function of doing -- indicating the positions and the movements of various players in a game long since over. Whatever success they enjoy in the performance of their job obviously derives from me, from my memory of the game being represented and my skill in translating that knowledge into the chosen idiom. The popcorn and the coins indicate, and in this sense perform their assigned function, only insofar as I am a reliable conduit for information about the situation being represented and a reliable and well-intentioned manipulator of the expressive medium.

(EB, p. 53)

In and of themselves, the elements of this system have no intrinsic ability to indicate what they have been assigned the the function of

indicating. They carry no information about some long over game; there is no nomic correlation between the position of Oscar Robertson and the position of this dime. Independent of me, or someone else, this dime has no ability to indicate what Oscar Robertson did. It is in virtue of me that the dime has the ability to indicate Oscar Robertson, and the function of indicating Oscar Robertson.

Representations of Type II are, like representations of Type I, part of systems that are created and used by individuals outside of the system to indicate how things stand with respect to some other thing. Like Type I representations, Type II representations are assigned the function of indicating something about some other condition by those who created and use the system of which they are a part. However, unlike Type I representations, Type II representations have the ability to indicate what it is their assigned function to indicate in virtue of "the way they are objectively related to the conditions they signify" (EB, p. 54). Type II representations are information carrying signals, they are natural signs. The curvature of a bimetallic strip is an indicator of temperature. The degree of curvature carries information about the ambient temperature. If I want to have a device whose operation will be dependent on temperature, I can construct such a device by making its operation dependent upon an element that can indicate the temperature. If I want the temperature sensitive device to function automatically, that is, without my intervention, the element that is to indicate the temperature must have the ability to indicate the temperature independent of me. The bimetallic strip is just such an element, and by putting it into a system whose operation is supposed to be

dependent on the information about the ambient temperature that the bimetallic strip carries, I give the bimetallic strip the function of indicating one of the things it is an indicator of. Through me and the role I assign it in a system I design, the bimetallic strip becomes a representation of the ambient temperature. Although the bimetallic strip owes its function (in the device) to me, the possibility of it having that function is intrinsic to it.

Representations of Type III are ones that not only have an intrinsic power to indicate what it is they do indicate, but what it is their function to represent is determined by the system itself.

Natural systems of representation, systems of Type III, are ones which have their own intrinsic indicator functions, functions that derive from the way the indicators are developed and used by the system of which they are a part. In contrast with systems of Type I and II, these functions are not assigned. They do not depend on the way others may use or regard the indicator elements.

(EB, p. 62)

Unlike representational systems of Types I and II, which derive their ability to represent (at least partially) from the purposes or intentions of the designers and users of the system, according to Dretske, representational systems of Type III arise without the assistance of intentional agents. RS's of Type III are self-designing, assigning functions to their own indicator states, creating representations for themselves. The intentional characteristics of such RS's do not derive from pre-existing intentional agents. It is in these self-designing representational systems that one finds, Dretske claims, a "source, not merely a reflection, of intentionality" (EB, p. 67). RS's of Type III are the foundation from which thoughts and beliefs, the "higher-order" intentional states, arise.

The plausibility of the existence of RS's of Type III depends upon there being a satisfactory account of the possibility of "intrinsic indicator functions." In the case of RS's of Types I and II, what it was their function to indicate could be determined simply by consulting the (intentional) agent who designed the system. In the case of RS's of Type III, this obviously is not an option. There must be a principled way of determining what it is the function of an internal indicator to indicate that does not rely on an external agent. Dretske claims that what it is the function of an RS of Type III to indicate is that condition upon which the successful completion of the behavior that the state causes depends. "Only by using an indicator in the production of movements whose successful outcome depends on what is being indicated" can what it is the function of an internal state to indicate be identified (EB, p. 70).

Dretske claims that there are two different means by which a system can acquire the function of indicating something that do not presuppose or involve intentionality: evolutionary development and learning. In the case of evolution or natural selection, systems in which an indicator of F, call it 'C,' causes behavior M are selected for because M is beneficial behavior when F is the case, but generally not beneficial when F is not the case. C has been selected as a cause of M because C indicates F (or at least is a better indicator of F than its competitors). Thus, C has the function of indicating F, and has acquired that function without the mediation of intentional agents. In the case of learning, within a system, an internal indicator, C, of F is "recruited" as a cause of behavior, M, which is beneficial when F is the case but not beneficial when F is not the

case, because C indicates F. The system makes C a cause of M because C is a good sign of when M is likely to be "successful" (in virtue of the fact that C indicated F, which is a condition on which the success of M depends).

While systems that have acquired indicator functions for their internal states in either of these ways are all RS's of Type III, according to Dretske, it is only those that have been shaped by learning that qualify as systems having thoughts and beliefs. Dretske argues that in the case of a system shaped by evolution, although C, an indicator of F, causes M, and has the function of indicating F (because F is the condition that C indicates on which M's success depends), the fact that C indicates F is irrelevant to the fact that, in this system, C causes M. "Meaning, though it is there, is not relevantly engaged in the production of output" (EB, p. 94). In a system shaped by evolution, the fact that C indicates F has nothing to do with why this C in this system causes this M. The reason that this system has an internal state C that causes M has to do with its genetic makeup, not with what C is an indicator of (EB, p. 92). Drawing on a distinction made by Richard Lewontin¹² and elaborated on by Elliot Sober,¹³ Dretske claims that in the case of individual systems that cannot learn, the fact that C is an indicator of F is only part of a selectional explanation of why there are systems that have C causing M. The fact that C is an indicator of F has nothing to do with a developmental explanation of why in this particular system a token of type C causes a token of type M. In order for something to qualify as a belief, it must be, according to Dretske, that the fact that C is an indicator of F is a part of a developmental explanation

of why in this particular system this C is a cause of this M. What an internal state, C, indicates must be an essential part of the explanation of why, in a particular individual, C causes M rather than, say, N, in order for C to be a belief, not merely a representation. An internal indicator is a belief only if its representational content (what it indicates) is part of a developmental explanation of why the system behaves the way it does, and it is only in systems that learn that, according to Dretske, one will find content influencing causal role in this fashion.

While Dretske does admit that there is no clear or sharp distinction between selection and development, between behavior determined by genetic factors that were selected for and behavior determined by environmental influences (EB, p. 92), when we examine Sober's explication of the distinction between a selectional explanation and a developmental one, it becomes quite apparent that Sober's distinction does not correlate with the distinction that Dretske wants to draw between systems that do not learn and those that do. According to Sober, a developmental explanation is one which explains why a particular population has, here and now, a particular feature by recounting, for each individual member of the population, the sequence of (physical) stages that led up to each individual member of the population having the particular feature in question. A selectional explanation, on the other hand, explains why a particular population has, here and now, a particular feature by citing those environmental factors that resulted in individuals with that feature being here now, and individuals without that feature not being here now. Sober gives the following example, which Dretske cites

(see EB, pp. 92-93), to help clarify the distinction.

You observe that all the children in a room read at the third grade level. What could be the explanation? Two strategies of analysis are possible. A developmental account will take the children one at a time and describe the earlier experiences and psychological conditions that caused each to attain that particular level of reading proficiency. These individual stories may then be aggregated. You may explain why all the children in the room read at the third grade level by showing why Sam, Aaron, Marisa, and Alexander each do.

A selectional explanation would proceed very differently. Suppose it were true that individuals would not be admitted to the room unless they could read at the third grade level. This would explain why all the individuals in the room read at that level. But, unlike the developmental story, the selectional account would not explain the population-level fact by aggregating individual explanations. The selectional theory explains why all the people in the room read at the third grade level, but not by showing why Sam, Aaron, Marisa, and Alexander do.

(Sober, p. 149)

While the distinction that Sober seems to be illustrating in this example is that between an explanation at the level of the individual (a developmental account) and one at the level of the group (a selectional account), Dretske appears to identify Sober's notion of a developmental account with a developmental stage account. In the paragraph following his presentation of Sober's example, Dretske discusses the developmental explanation one would give of why tokens of type C cause movements of type M in a moth, and it seems clear that Dretske has a developmental stage theory in mind.

The moth has the kind of nervous system it has, the kind in which an internal representation of an approaching bat causes evasive movements, because it developed from a fertilized egg which contained genetic instructions for this kind of neural circuitry, circuitry in which the occurrence of C will cause M. This is a developmental explanation, a causal explanation of why, in today's moths, tokens of type C produce movements of type M. These genetically coded instructions regulated the way in which development occurred, channeling the proliferation and specialization of cells along pathways that produced a nervous system with these special features.

(EB, p. 93)

If Dretske does, in fact, want to identify developmental accounts with developmental stage accounts, then it is quite clear that Dretske cannot use the role of content in a developmental explanation of why C causes M as the criterion by which to distinguish genuine beliefs from mere representations. Such a criterion will be ineffective not because content is appealed to in developmental stage accounts of moth behavior, but because content is not appealed to in developmental stage accounts of the behavior of systems that learn. The indicator status of a state is not relevant to a developmental stage explanation of why C causes M in any system, regardless of whether its behavior is shaped by evolution, by learning, or by both.

Developmental stage explanations are essentially antithetical to the fundamental underlying premise of learning, which is that the environment plays the pivotal role in shaping the individual.

Developmental stage theories are predicated on the assumption that the sequence of states an organism occupies is not the fortuitous result of the experiences that happen to impinge. Regardless of wide possible variation in the character and order of experience, the organism will change in a certain way. The idea of a developmental pathway is precisely the idea of regularly occurring changes that are insulated from environmental influences. . . . The state of the environment need not be entirely irrelevant, of course. But the more irrelevant it is, the more attractive this kind of developmental theory will be.

(Sober, p. 153)

A developmental stage explanation of why some particular individual has some particular feature will consist fundamentally of citing those physiological and psychological changes that take place in all individuals (of the relevant species). The particular experiences of the individual in question will be important only to the extent that they influences such things as the rate at which those changes

take place. The environment plays a subsidiary role in developmental stage explanations. It is only important to the extent that it can influence the sequence of stages that an individual is "preprogrammed" to go through.

Dretske claims that only systems that learn have genuine beliefs, because it is only in systems that learn that the indicator properties of internal states are relevant to a developmental explanation of why those states have the causal properties they do (EB, p. 94). However, the basic premise of learning is that the environment is the primary shaper of the individual, while the basic premise of developmental stage theories is that the environment plays, at best, a secondary role in shaping the individual. The more important the environment in shaping the individual, the less plausible a developmental stage explanation of why the individual has the particular features he does becomes. Thus it is clear that no distinction can be made between individuals that have genuine beliefs and those that do not on the basis of whether or not content is relevant to a developmental stage account of why they behave the way they do, because content is not relevant to any such account.

Alternatively, it may be suggested that we should assume that Dretske intends to be relying on the distinction that Sober makes between an explanation at the level of the individual and one at the level of the group. However, even if we make that assumption, we still find that an appeal to the role of content in shaping behavior does not provide a principled way of distinguishing between beliefs and mere representations. The process through which content influences the causal role of an indicator state in a system that learns is

identical to the process through which content influences the causal role of an indicator state in a system shaped by evolution alone. Content is no more relevantly engaged in shaping the output of a system that learns than it is in shaping the output of a system that does not.

An examination of Dretske's own rather tentative explanation of how, in individuals that learn, the recruitment of C as a cause of M is possible, reveals that, contrary to Dretske's stated position, the indicator status of an internal state is relevant to learning in exactly the same way that the indicator status of an internal state is relevant to evolution. Beneficial C-M connections are selected for in the learning process in the same way that beneficial C-M connections are selected for in evolution.

The parallel distributed processing (PDP) networks, networks of interconnected nodes in which the strength of connections between nodes is continually reweighted (during "learning") so that, eventually, given inputs will yield desired outputs, provide intriguing and suggestive models for this recruitment process. . . . In these models, the internal indicators would be patterns of activation of the network's input nodes, and recruitment would proceed by selection (by appropriate reweighting between nodes) of the desired input (i.e., an F indicator) for an appropriate activation of effector mechanism (M).

(EB, p. 98)

Regardless of whether C was already a cause of M, perhaps along with various other internal states that are "pruned" away during the process of selecting of C, or was simply a state that was correlated with successful executions of M, C is selected as a cause of M in virtue of what it has indicated. C was recruited, i.e., selected, as a cause of M because M is, in general, productive behavior when and only when F is the case, and C is a good indicator of F (or at least a better indicator of F than anything else that is available). The ability of C to cause M

does not develop in response to what C indicates. Rather, it is C's ability to cause M that is responsible for C being selected, on the basis of what it indicates, as a cause of M.

In systems that learn, content is engaged in the determination of output through the process of selecting beneficial C-M connections. In systems that do not learn at the individual level, content is engaged in the determination of output through the process of selecting beneficial C-M connections as well. The only difference is that in systems that do not learn each individual has only one possible C-M connection, so that the selection of beneficial C-M connections can be achieved only through the selection of individuals. Learning and evolution are essentially the same process: learning is simply evolution on the fast track. Compressed within me is the equivalent of generation upon generation of moths. Just as those moths that were wired in a beneficial way were selected for, while those that were wired in a less beneficial way were not selected for, so too, internal processes within me that prove to be beneficial (going to the door when and only when the 'someone at the door' indicator is on) have been selected for while those that are not beneficial (going to the door when the 'telephone is ringing' indicator is active) have not been selected for. While Dretske (appealing to work done by Robert Cummins¹⁴) notes that natural selection does not explain why individuals have the properties for which they were selected (EB, p. 92), he seems not to recognize that the same is true of learning. Learning does not explain why an individual has the properties that permit it to learn to have C cause M. Rather, it is the fact that an individual has certain properties that permits it to learn to have C

cause M.

We clearly want content to be relevant, in some way or another, to the causal powers of beliefs -- one of the main reasons that Fodor's theory was found unacceptable (see Chapter 2) was precisely because it made content, though predictively useful, causally irrelevant in the determination of behavior -- but requiring that internal states acquire their causal roles through learning does not increase the role that content plays in determining behavior over the role it plays under natural selection. In both cases the system behaves the way it does now because of the particular C-M connections that were selected for. This is not to deny that there is a difference between natural selection and learning, but the difference is not one of the degree to which content is engaged in the determination of output. The indicator properties of C are engaged in the same way in the same process.

III. INFORMATION THEORY LEADS TO AN INDIVIDUALISTIC NOTION OF CONTENT, DRETSKE'S COMMENTS NOTWITHSTANDING

Leaving aside, for the present, the question of how (if one felt it necessary to do so) one might distinguish between a moth's internal representations and my internal representations so that my representations are beliefs but the moth's are not,¹⁵ I want to turn back to Dretske's theory of how the content of a representation is

determined. According to Dretske, a state represents (has as its content) what it has acquired the function, the job, of indicating for the system. In order for a state to be a belief, an intentional state, that state must have acquired a control function in the production of output in virtue of what it indicates. A belief is a state that not only has a functional, i.e., causal, role in the production of some type of behavior, but has acquired the function, i.e., job, of representing one of the things it indicates because that thing is relevant to the "success" of the behavior. The content of a belief is what it has become the function of that state to indicate in that system. Thus, according to Dretske, there are two factors that influence the content of a belief. The meaning or representational content of internal states is determined by "a combination of (1) their relations (usually causal or informational relations) to the external situations they represent and (2) their functional (or conceptual) role in the production of output (including their internal relations of each other)" (EB, pp. 150-151). On the basis of (2) Dretske is lead to endorse a holistic view of mental state content, while on the basis of (1) and (2) he is lead to endorse a wide notion of mental state content. While meaning holism is a genuine consequence of Dretske's use of functional role and function or purpose in the determination of content, the wide view of content is not even warranted by, no less an inevitable consequence of, his theory.

In explaining how the content of an internal indicator, C, was determined it was tacitly assumed that C had been recruited as a cause of a single piece of behavior M. While such an assumption seems not

unwarranted in the case of moths -- the only causal role that C, a bat indicator, has is to produce evasive flight maneuvers -- as one moves up the phylogenic scale this assumption becomes less and less tenable. A single internal structure, C, may have a causal role in the production of many different pieces of behavior, and a single piece of behavior may have many different indicators, either jointly or separately, as its causes. The content of a structure is what it has the function of indicating, and what it has the function of indicating is determined by which of the things it indicates is relevant to the "successful" performance of the behavior it is causally involved in producing. Thus, anytime an indicator changes its causal or functional role there is the possibility that its content (what it has the function of indicating) will also change. But given that the causal role of an indicator includes, according to Dretske, its interaction with other indicators, anytime a system integrates new indicators the contents of the old ones will change in so far as what they have the function (job) of indicating has changed. It is not that what these states indicate has changed -- that is fixed by the nomic dependences that exist between signal and world. Rather, what will be altered is which of the things they indicate they have the function (job) of indicating, which of the things they indicate they have as their content. "A belief having the putative fact that F as its content, an internal state whose function it is to indicate that condition F exists, will inevitably change this content as it becomes more tightly integrated with other states having corresponding indicator functions" (EB, p. 150). As more and more indicators are employed, each indicator can become more specialized in its function.

The function of the indicators can become more narrowly defined. A state that carries the information "square," "rectangle," and "quadralateral," but which has the function of indicating "quadralateral" can change what its function is to indicate, coming to have the function of indicating "square," if the system acquires a new indicator that can take over the function of indicating "quadralateral."

The present view of belief, a view that identifies what we believe with what it is the function of certain elements to indicate, not only implies that beliefs have this holistic character, it reveals why they have it. As beliefs become integrated into more tightly structured cognitive systems, their indicator functions become more interdependent.

(EB, p. 150)

To the extent that what it is the function of a structure to indicate is dependent upon the causal connections of other indicator structures, its content will be dependent upon those other structures, their causal connections, and what they have the function of indicating. Mental state content is holistic because the function of any given internal indicator is dependent upon not only its causal role in the system, but on the functions and causal roles of the other internal indicators that are part of the system.

The use of the notion of the function, job, or purpose of a state in determining its content not only leads to content holism, but leads to content indeterminacy as well. According to Dretske, the content of a structure is that condition which the structure indicates in virtue of which the structure was selected as a cause of M. The fact that C causes M is relevant to the determination of C's content only if C was selected as a cause of M, and selected in virtue of what C

indicates. C can cause M, i.e., have as one of its causal/functional roles in the system causing M, without C's content being, in any way, dependent upon or determined by M. In order to be able to determine what the content of C is, one has to know, not how it functions in the system (its causal role in the system) -- a matter which is, at least in theory, empirically determinable, but rather which of its causal connections in the system it acquired in virtue of what it indicates -- a matter which, Dretske admits, may be difficult, or even impossible, to determine.

There is no clear dividing line between the way something functions and its function. After a thing functions in that way long enough so that it is clear that it is being selected or being used in a way that depends on its continued performance of that task, then we can say that it has acquired the function of performing that task. In the case of indicators and their functions, these questions will not always (or perhaps ever) have precise answers. But this result, far from being an objection to this account of meaning, is, I submit, one of its virtues. For this is precisely the sort of thing one should expect to find in the case of beliefs and the concepts on which they depend.

(EB, p. 155)

Thus, to the extent that what it is the function of a structure to do -- as opposed to how it functions -- is indeterminate, the content of that structure will also be indeterminate.

In addition to the indeterminacy that arises from the difficulty of determining when a structure has acquired the function (job) of causing M, rather than merely incidentally causing M, Dretske's theory involves another sort of content indeterminacy or relativity, one that arises from the role of channel conditions in determining content. According to Dretske, the information that a structure carries, and hence the range of possible contents that a belief built on that

structure can have, is determined by conditional probability. A signal (structure) carries the information F if and only if the conditional probability of the source of the signal being F is 1 given the signal (and k, whatever else the individual knows), but less than 1 otherwise. In order for the conditional probability of the source being F to be 1 on the basis of the signal, it must be that the signal can only occur when the source is F. In other words, the signal must be unequivocal, at least with respect to F. But whether or not a signal is equivocal is a pragmatic question, a question that can have different answers for different people in different circumstances. Since equivocation determines what information a signal can carry, what information a signal carries will also be, to some degree, a pragmatic question.

To the extent that whether or not a signal is equivocal is a question that has no objectively correct answer, then there will be no objectively correct answer to the question of what information a signal carries, or to the question of what content a mental state has. For example, what information the signal that the assembly-line worker receives in the factory is unequivocal with respect to is not a question that can be objectively determined. You may feel that the signal the worker receives is unequivocal with respect to the information "resistor" (because you do not view the possibility of a capacitor appearing in the factory as a relevant alternative), while I may feel that the very same signal is unequivocal only with respect to the information 'object of such and such size and appearance' (because I am unwilling to rule out as a genuine alternative a capacitor appearing in the factory). If, as Dretske claims, there is no correct

answer to the question of whether or not something is stable enough not to count as a genuine alternative -- it being a matter of one's interests and standards, then there is no correct answer to the question of what the assembly-line worker believes. You can legitimately and correctly attribute to him the belief that something is a resistor, while I can legitimately and correctly refuse to attribute to him the belief that anything is a resistor, attributing to him the belief only that something is an object having a certain size and appearance. There is no objective fact about what the assembly-line worker believes.

Although I am willing to grant Dretske that a certain (limited) degree of content indeterminacy or relativity may not be avoidable in a theory (the facts about content may, themselves, be indeterminate and/or relative), the question arises in the case of Dretske's theory, if perhaps intentional characterizations are not involved in the specification of either the selection/success criterion for determining the function of a structure, or the channel conditions on which the transmission of some piece of information depends. Recall that Dretske's goal is to construct a naturalized theory of intentionality, one that does not involve any prior intentional characterizations. In order to achieve this goal Dretske must provide principled criteria by which to determine what information a signal type carries, and what it is the function of something to indicate, that do not depend upon prior characterizations involving intentional notions. While in the case of determining what the function (job) of a structure is, Dretske is silent, he makes a number of remarks concerning channel conditions that call into question whether Dretske

will be able to specify the channel conditions for the transmission of some piece of information without invoking intentional or semantic notions.

Dretske claims that

the question of whether an existing condition is stable or permanent enough to qualify as part of the channel, as a condition which itself generates no (new) information, is a question of degree, a question about which people (given their differing interests and purposes) can reasonably disagree, a question that may not have an objectively correct answer. When a possibility becomes a relevant possibility is an issue that is, in part at least, responsive to the interests, purposes, and, yes, values of those with a stake in the communication process. The flow of information, just like the cognitive exploits it makes possible, is a process that exhibits some sensitivity to the variable purposes of those who send and receive this information.

(KFI, p. 132-133)

It is not the fact that Dretske's theory makes the content of mental states relative to a variable standard of equivocation that is worrisome. The concern over Dretske's theory arises not from the fact that mental state content is relative per se, but from the fact that it is relative to the interests, purposes, and values of those who have a stake in the communication process. As Dretske points out, one can make information relative without thereby making it subjective, or impugning its objectivity or naturalism.

[T]he meanings we ascribe to signs is [sic] relative. It is relative to what the speaker already knows about possible alternatives. This, however, doesn't mean that natural meaning is subjective. A person's weight isn't subjective just because it is relative, just because people weigh less on the moon than they do on earth.

(EB, p. 58)

However, the reason that a person's weight can be relative, but not subjective, is the fact that one's weight is a function of something that is objective, i.e., the value of gravitational acceleration

acting on the person. The question about Dretske's theory is whether or not the interests, purposes, and values which determine what things are part of the channel and what are relevant or genuine alternatives can be specified without appealing to the beliefs, desires, fears, etc. of those with a stake in the communication process. If in defining which things are to be considered part of the channel of communication and which are not in any given instance of communication, one must make reference to what else the individual knows, believes, etc., then Dretske will not have succeeded in baking "a mental cake using only physical yeast and flour" (KFI, p. xi). The fact that Dretske has not provided an account of how function and channel conditions are to be specified so as not to involve intentional notions, does not, of course, undermine his claim to have provided a naturalized account of intentionality. Nevertheless, it does take on the appearance of a substantial lacuna in his theory because of the central roles of channel conditions and function in the theory, and his invocation of interests, purposes, and values, notions which, at least on the face of it, would seem themselves to involve intentional notions.

Dretske is motivated to claim that a signal need only eliminate all other relevant possibilities in order to carry a given piece of information because without that claim, it will turn out that very little, if anything can actually be known. If one knows something, it must be that the conditional probability of that thing being the case is one. One can not know something that is not the case. But, if knowledge is to be derived from information, and Dretske defines what it is to know something as "K knows that s is F = K's belief that s is

F is caused (or causally sustained) by the information that s is F" (KFI, p. 86), then a signal must be capable of carrying information with a conditional probability of one. By introducing the notion of channel conditions, and claiming that channel conditions do not affect the equivocation of a signal, Dretske ensures that knowledge is possible. However, he does so at the cost of calling into question whether or not he has succeeded in constructing a naturalized account of intentionality using a notion of information understood as "an objective commodity, something whose generation, transmission, and reception do not require or in any presuppose interpretive processes" (KFI, p. vii).

Although it is the fact that Dretske's theory makes content ultimately dependent upon the interests, purposes, and values of those involved in communication that is most problematic, given his stated goal of constructing a naturalized account of intentionality, it is not this aspect of content that makes Dretske's theory a potential threat to the viability of the computational model of the mind. Rather, the computational model is jeopardized by Dretske's claim that the information theory foundation of his account leads to a wide notion of content, a notion of content inimical to the formality condition. From the point of view of a computational theory of mind, the sort of functional information-theoretic model of intentionality that Dretske has developed is the most promising account available, notwithstanding any shortcomings that Dretske's own theory might have. Consequently, if functional information-theoretic accounts of intentionality invariable lead to a wide notion of content, the plausibility of the computational model of the mind would be

substantially reduced. Contrary to what Dretske implicitly claims, however, such accounts do not entail wide content: they require, in fact, a narrow or individualistic notion of content.

According to Dretske, the content of a cognitive structure is a product of both the information the structure carries and the function of the structure in the system. Consequently, in order to assess whether or not Dretske's theory of content actually does entail a wide notion of content we need to look at both the information and the function "components" of content. If either of these components requires or entails a wide notion of content, then Dretske's claim that psychology needs such a notion of content would be legitimized, and so much the worse for computational theories of the mind. However, close scrutiny reveals that information theory leads to a narrow, individualistic notion of content, and that the notion of function as developed by Dretske need not and cannot convert that individualistic content into a wide notion of content. I am going to postpone the discussion of function until the next chapter (specifically, Chapter 5, Section III) because I want to treat it in the context of a more general point that applies to not only Dretske, but Fodor and Burge as well. For the moment, however, I want to consider whether or not the information "component" of Dretske's theory leads to or requires a wide notion of content.

Dretske claims that differences in the nature of the environment can alter the content of an individual's intentional state, even in those situations in which the difference in the environment has no physical, functional, or phenomenological effect on the individual. What the content of a belief can be is limited by what that structure

is an indicator of, what information that structure carries, and what information a structure carries is determined by the "nomic dependencies that define the information-response characteristics" (KFI, p. 266) of that structure in the local environment. Two individuals, identical in all respects will have different mental state contents if their local environments are different in such a way that different things are at the "other end" of the nomic dependencies. Dretske's discussion of the case of the assembly-line worker hints that this is the sort of view of content he has in mind. Inside the factory, Dretske wants to say that signals carry the information "resistor" because a capacitor is not a relevant alternative inside the factory. Outside the factory, however, capacitors are relevant alternatives, so that outside the factory the same signal type cannot carry the information "resistor." It can carry only some less specific piece of information, say, "resistor or capacitor" or "object of such and such size and appearance." Those whose "natural habitat" is inside the factory can have beliefs that things are resistors. Those whose "natural habitat" is outside the factory, however, cannot have beliefs that things are resistors, because outside the factory the signals they receive have an information relation, not to resistors, but to the union of resistors and capacitors. I do not mean to be saying that the content of some particular individual's belief will change if he is suddenly transported to another environment. What the content of a belief is is determined by what it has the function of indicating, and that is fixed at the time of its recruitment in accordance with what it was an indicator of then.

While the case of the assembly-line worker may not definitively demonstrate that Dretske endorses a wide view of mental state content, his discussion of his version of Putnam's Twin Earth example does. Not only does Dretske's interpretation of the Twin Earth case make it clear that he has a wide view of content, he explicitly endorses Putnam's claim that two individuals, in different contexts, can be identical in all relevant respects, yet have different concepts, beliefs, and meanings (EoB, p. 15).

Suppose there is a place (call it Twin Earth) in which there are two substances XYZ and H_2O , chemically quite different but both having the superficial properties of water. By "superficial" properties I mean the properties we ordinarily rely on (outside the laboratory) to identify something as water. Some of the lakes and rivers on Twin Earth are filled with H_2O ; others are filled with XYZ. Some houses have H_2O running out of their taps; others have XYZ. It rains H_2O in some parts of the country, XYZ in other parts. In some places there is a mixture. Both substances are called "water" by Twin Earthlings since they are (apart from elaborate chemical analysis) indistinguishable. Both substances quench thirst, taste the same, boil and freeze at (almost) the same temperature, and so on.

Consider, now, some Twin Earthling (call him Tommy) being taught what water is on a part of Twin Earth in which both H_2O and XYZ are available. As it turns out (quite by accident), he is taught to identify water (or what the Twin Earthlings call "water") by being exposed to only H_2O . After learning what water is (to the complete satisfaction of his teachers), he emigrates to a part of Twin Earth where there is to be found only H_2O . Or (to make the point in even clearer terms) we may suppose that Tommy is miraculously transported to Earth, where there is to be found only H_2O . Since there are no other significant differences between Twin Earth and Earth, Tommy blends in without any trouble. Everything Tommy says about water (using the word "water") will correspond with what his new-found friends say and believe about water (also using the word "water").

The question, of course, is not what Tommy says, but what Tommy believes. Tommy does not have the same concept as his Earthling associates. What Tommy believes when he says, "This is water" is not what his Earthling friends believe when they say, "This is water." What Tommy means by "water" is either H_2O or XYZ. This, of course, is how we (knowing all the facts of the case) would describe it, not Tommy. If

asked, Tommy will say that he means water by "water," and he surely does mean this. But the point is that more things qualify as water for Tommy than for his Earthling friends. If we should imagine that some XYZ was also suddenly transported to Earth, Tommy's belief of this substance that it was water would be true while his Earthling friend's belief that it was water would be false.

The information-theoretic explanation of this difference is to be found in the difference in the kind of information to which Tommy and his Earthling friends were made responsive during their respective learning periods. Even though it turns out (quite by accident) that Tommy and his Earthling friends were exposed to the same substance throughout the training period (viz., H_2O), the information that it was H_2O was made available to the Earthlings but not to Tommy. On Twin Earth this information was not available because on Twin Earth (but not on Earth) signals carried the information, not that s was H_2O , but that s was either H_2O or XYZ. It was this latter, essentially disjunctive, piece of information to which Tommy became selectively responsive during training. Since XYZ is not to be found on Earth (and, we are supposing, cannot be brought to Earth by anything short of a miracle), Earthlings acquire a different concept because their discriminatory responses were shaped by a different piece of information -- the information, namely, that this was H_2O . Since the regularities prevailing in these two worlds are different, the kind of information to be found in physically indistinguishable signals is different. Hence the semantic content of structures developed in response to these signals is also different. That is why Tommy's concept, though developed in response to the same sort of physical stimuli (the sort associated with seeing, tasting, and feeling water), though (in fact) developed in association with the same substance (H_2O), is quite different from the Earthling concept. They both use the same word to express what they mean, but they mean something different. At least they have concepts with different extensions.¹⁶ There is no way of discovering this difference by looking "inside their heads" -- by examining the physical properties of their internal states. For different extensions (hence different concepts) are a result of the different sort of information to which they were exposed during learning, and this difference is a difference, not of what is in their head, but of the informationally related regularities that dominated the environment in which they learned.

(KFI, pp. 225 - 227)

Because, according to Dretske, the signals that Earthlings receive carry the information " H_2O " while the signals the Twin Earthlings receive -- even those who have been exposed only to H_2O -- carry the

information "H₂O or XYZ," Earthlings and Twin-Earthlings do not have beliefs that have the same content when they think to themselves "I hope there is water nearby."

Dretske's claim that the information that Earthlings receive is different from that which Twin Earthlings receive is based on the idea that the signals the Earthlings receive are unequivocal with respect to one piece of information -- which Dretske claims is the information "H₂O," while the signals the Twin Earthlings receive are unequivocal only with respect to some different piece of information -- which Dretske claims is the "essentially disjunctive" piece of information "H₂O or XYZ" (a claim about which I will have more to say in a moment). Even though Twin Earthling Tommy and his Earthling friends were both exposed to the same substance throughout the training period (or even throughout their entire lives) the signals that the Earthlings receive carry the information "H₂O" rather than the information "H₂O or XYZ," while the signals Tommy receives carry only the less specific information "H₂O or XYZ," because XYZ is not considered to be a relevant alternative on Earth. Even though it is the case that were there XYZ on Earth, the signals the Earthlings receive would be equivocal with respect to the information "H₂O," the signals the Earthlings do receive are not equivocal because XYZ does not, in fact, occur around here. In order to carry the information "H₂O" the signal need only rule out other "relevant alternatives." Since on Twin Earth XYZ is a relevant alternative -- it exists there, but Tommy has simply not run into any yet -- in order to carry the information "H₂O" the signal would have to rule out XYZ which, by hypothesis, it does not. The signal on Twin Earth merely succeeds in

ruling out all relevant alternatives to H₂O or XYZ.

Since the lawful regularities that prevail in these two worlds are different, the kind of information to be found in physically indistinguishable signals is also different. Hence the concepts developed in response to these physically indistinguishable signals are also different. This is why Tommy and his Earthling friends, although they say the same thing, although they were exposed to exactly the same liquid during learning (viz., H₂O), and although they developed their ideas in exactly the same way, have quite different beliefs about the liquid they see and describe.

(EoB, p. 16)

Tommy and his Earthling friends receive different pieces of information, are made responsive to different pieces of information, and so develop different concepts.

As I have already remarked, the elimination of something as a relevant alternative on the basis of its not being part of the local environment is questionable. Nevertheless, even if we are willing to grant that context determines relevant alternatives, and that the signals Tommy received on Twin Earth carried different information than the corresponding physically indistinguishable signals Tommy's Earthling friends received on Earth, the Twin Earth example does not demonstrate that mental state content is wide. Dretske's interpretation of the example is not even consistent with his own semantic theory of information. Dretske misidentifies the information carried by signals on Twin Earth, and, even leaving the misidentification aside, the example fails to show that Tommy and his Earthling friends have different belief contents.

Dretske claims that Tommy receives signals that carry the "essentially disjunctive" piece of information "H₂O," while his Earthling friends receive signals that carry the (simple) information "H₂O." But Dretske's claim that the signals that Tommy receives carry

an essentially disjunctive piece of information presupposes the "correctness" of wide individuation. The signals Tommy receives carry information that is disjunctive only relative to some presupposed wide type individuation, corresponding, in this case, to the type individuation of chemistry. Relative to a different presupposed wide type individuation, say that of nuclear physics, not only are the signals Tommy receives disjunctive, but the signals his Earthling friends receive are disjunctive as well. The type individuation of chemistry only distinguishes between H_2O and XYZ, but the type individuation of nuclear physics not only distinguished between H_2O and XYZ, it also distinguishes between H_2O , H_2^2O , and H_2^3O . Relative to the type individuation of nuclear physics signals on Twin Earth carry the information " H_2O , H_2^2O , H_2^3O , or XYZ," while those on Earth carry the information " H_2O , H_2^2O , or H_2^3O ."

The distinction between H_2O and XYZ, or between H_2O , H_2^2O , and H_2^3O is not a distinction that has any relevance for the identification of the information the signals carry. By hypothesis, on Twin Earth the type of signal received is the same regardless of whether the source is a sample of H_2O , a sample of XYZ, or, for that matter, a sample of H_2^2O , or H_2^3O . The difference identified by chemistry (or physics) between H_2O and XYZ is a difference that has no effect on the signal or the information it carries. The nomic dependence in virtue of which the signal carries the information it does is between the signal and some property, call it 'bwater' -- a property that, as it happens, is one shared by substances that chemistry and physics think of as being of different types. The distinction between H_2O and XYZ does not show up in the set of nomic dependences that determine the

information content of the signal. H_2O and XYZ are type identical as far as the signal is concerned. The signal on Twin Earth carries the information "bwater," not the disjunctive " H_2O or XYZ."

Even if we revise Dretske's Twin Earth example, expunging it of the claim that Tommy receives an essentially disjunctive piece of information, the example still does not show that Tommy and his Earthling friends have different beliefs. The fact that Tommy and his Earthling friends receive signals that carry different information ("bwater" and " H_2O ," respectively), in and of itself, does not necessarily mean that Tommy and his Earthling friends have different concepts. The signals that each of them receives carries a great deal of information in addition to the stated " H_2O " and "bwater." In particular, they both carry the information "colorless, tasteless, odorless liquid that falls from the sky, quenches thirst, etc., etc." This information is nomically nested in the signals that they all receive. In virtue of what is it legitimate to say that the Earthlings have the concept " H_2O " while the Twin Earthlings have only the concept "bwater"? The question of what concept(s) Earthlings and Twin Earthlings have acquired is not merely a question of what information was carried by the signals they received. The concepts they have are dependent upon what information they succeeded in extracting from the signals they received.

The fact that a signal carries the information F does not mean that that is the concept an individual will develop out of it. What concept a person has depends not only on what information they have received, but on what information that they have extracted from the signal. In Knowledge and the Flow of Information, as well as in a

number of articles, Dretske goes to great lengths discussing the difference between the information carried by a signal and the information extracted from that signal.¹⁷

To learn what a dingbat is, and hence to acquire the conceptual resources necessary for believing that something is a dingbat, one must not only be exposed to dingbats, but to the information that they are dingbats. Not only must this information be made available, it must be picked up and used by the learner to guide his discriminatory and identificatory responses if he is to be credited with the relevant concept.
(EoB, p. 11)

Any given piece of information a signal carries may be one that a system, although it receives that information, is unable to pick up, or to attend to (see KFI, p. 144). While there is no doubt that only Earthlings can acquire the information "H₂O" from the incoming signals, it may not be that that is the concept they have acquired.

Consider, for example, a child being taught to recognize and identify birds. The child is shown a number of robins at close range and in such a way that their distinctive markings and silhouette are clearly visible. A few bluejays are thrown in for contrast. . . . After a satisfactory training period the child spots a sparrow in a nearby tree, points at it excitedly, and says "robin." What the child says is false, of course. The bird is not a robin. But we are not now interested in assessing the truth or falsity of what the child says, but rather the truth or falsity of what the child believes. To determine this we have to know what the child believes, and it is not at all clear that the child is accurately expressing what she believes with the word "robin."

Does the child believe the bird (the sparrow) to be a robin? Or does she, perhaps, simply believe it to be a brown bird of some sort (a nonblue bird)? Given the rather limited range of contrasts to which the child was exposed during training (only bluejays), it is not clear what information she was responding to when she succeeded in identifying all the robins in the sample class. Hence it is unclear what concept the child is expressing with the word "robin," what belief the child has when it points to the sparrow and says "robin."

(KFI, p. 195-196)

The point is that although what a signal indicates determines the range of concepts that an individual can acquire from that signal, one cannot determine from the signal what concept has been developed. In order to determine what concepts an individual has acquired one has to know what information the individual has succeeded in extracting from the signal, what information that the signal carries that they are responsive to.

There is certainly no doubt that Dretske is aware of the distinction between information received and information extracted. Nevertheless, he fails to recognize, not only that it is relevant to the case of Tommy and his Twin Earth friends, but that it, in fact, forces a narrow interpretation of mental content rather than the wide one he explicitly favors. Dretske claims that Tommy was made responsive during training to the information "bwater (H_2O/XYZ)" while his Earthling friends were made responsive during training to the information " H_2O ," but there seems to be no basis for the claim that they are responsive to different pieces of information. While, if we grant that the signals they received carried different information, there is no question that they have responded to different information carrying signals. But responding to an information carrying signal and being responsive to the information carried by the signal are two different things. The claim that they are responsive to different pieces of information cannot be made solely on the basis of the fact that the signals they received carried some different pieces of information. Responsiveness is a functional property of the individual. If they were, in fact, made responsive to different pieces of information, then one would assume that there would be at

least some (conceivable) situations in which their responses differed from one another. If, as is the case, there are no circumstances under which Tommy and his Earthling friends will respond differently, there are no grounds for the claim that they have been made responsive to different pieces of information. If, in fact, Tommy's Earthling friends did have the concept "H₂O," then Tommy could not have the same concept as his friends, but Tommy's Earthling friends do not have the concept H₂O. Nor do I think that Tommy has acquired the concept "bwater." What both Tommy and his Earthling friends have been made responsive to, and consequently what concept they have acquired, is something along the lines of "colorless, odorless, liquid that quenches thirst, condenses on the outside of cold lemonade pitchers, etc., etc., etc." The indicator relation, itself, does not determine whether or not the content of a belief is to be wide or narrow. It only places a limit on what the content of a belief can be. Consequently, even if information were wide, and, as I have argued, it seems in fact not to be, content, contrary to what Dretske claims, would still have to be determined individualistically from the range of possible contents.

CHAPTER 5
THE ONTOLOGICAL AUTONOMY OF PSYCHOLOGY

I want to suggest that the conclusion we should draw from the various difficulties I have thusfar claimed the theories of Fodor, Burge, and Dretske encounter is that in a very fundamental sense the whole debate over mental state contents has been conducted at the wrong level. The question, as we have seen, that is generally debated is "How should we individuate mental state content for the purposes of psychology?" The argument is taken fundamentally to be that of how to identify and individuate mental state content. However, in casting the issue in terms of how to individuate content, attention is, I believe, focused in the wrong direction. The question that we should be addressing, if we are interested in identifying and individuating mental content is "How should we individuate things in the world for the purposes of determining the content of mental states?" Among not only those views that I have discussed at length in earlier chapters, but such theories as Ruth Millikan's 'consumer'-driven theory¹ as well, there is certainly no conflict over whether or not content is dependent on the world. There is a consensus, with which I whole heartedly concur, that the content of mental states ultimately is derived from and is dependent upon the world. Intentionality is derived from the connections between internal states and the external

world. The disagreement among the various theories is over which characterization of the world should be used in determine the content of mental states. But given that content is determined, in one way or another, by the world, the question that we need to answer first is not which characterization to use, but what is there in the world that gives content to these internal states. If intentional states derive their content from the world, then in order to determine content we first need to determine what there is in the world.² We need to determine how to categorize or taxonomize the world in order to be able to individuate mental content. My claim is that we should be asking what the taxonomy of the world is according to psychology. One cannot begin to type-individuate mental state contents for the purposes of psychology until one has determined what there is in the world that can give meaning to mental states.

My point is that what the other sciences say there is in the world for their purposes need not determine what there is in the world for the purposes of psychology. The psychological theory we are interested in is a scientific field whose goal is the explanation and prediction of cognitively mediated behavior (the explanation of behavior in terms of the beliefs and desires that caused it). As a scientific field, psychology must be allowed to determine for itself what there is in the world. The taxonomic categories into which psychology divides the world must be appropriate to the needs of its theories. A priori, there is no reason to believe that the way it is appropriate for psychology to individuate the environment in order to specify law-instantiating relations between the environment and individuals will be the same as the way chemistry individuates the

environment for specifying its lawful generalizations.

The taxonomic autonomy which I am claiming that psychology should be permitted is in no way unique to psychology. Taxonomic autonomy is, in fact, the norm. It is routinely assumed in most, if not all, other fields. Botany is not expected to employ the taxonomic categories of geometry. Biology is not expected to explain the phenomena in its domain in terms of those states quantum mechanics recognizes as appropriate to quantum mechanical questions. Economics is not required to couch its generalizations about money, banking, and the monetary behavior of individuals or populations in categories commensurate with those of chemistry. Similarly, psychology, a priori, need not use the taxonomy of the world according to the other sciences, including linguistics, for its explanations, unless that taxonomy is useful for psychology. The taxonomy of the world that psychology posits must be one motivated by the needs of psychology, by the needs of its explanatory theories. The taxonomy of the world from the point of view of psychology should be one that facilitates the achievement of psychology's explanatory and predictive goals.

I believe that Fodor, Burge, and Dretske view the question of how to individuate mental state content from the same (limited) perspective: they do not allow that the ontological demands of a scientific psychology may lead to an individuation of the world that is different from that of physics, chemistry, or linguistics. Each accepts that the ontology available to psychology is that provided by some other field, although they differ on what that other field should be. They are absolutists about ontology. But the assumption that the taxonomic categories which psychology uses when determining the

content of an individual's mental state will be the same as the taxonomic categories which, say, physics, chemistry, or linguistics uses is unwarranted. We cannot simply assume that the way we divide up the external world for the purposes of physics, economics, or linguistics will be identical to the way we divide up the world for the purposes of psychology. In fact, we do not even require that the taxonomy of one field recognize or be reducible to that of another. Taxonomic categories and their memberships are determinate only relative to some theory. It cannot be assumed that the taxonomy of the world that psychology should use for the purposes of determining the content of mental states is identical or reducible to the taxonomy of the world as defined by the semantics of some language, or another field, scientific or otherwise. Consequently, without a prior demonstration that the taxonomy of the world according to physics, chemistry, economics, linguistics, or any other field, is appropriate for psychology, arguments for or against various psychological theories that presuppose one of these taxonomies are premature. Nevertheless, a brief review of the arguments of Burge, Fodor, and Dretske will show that each does, in fact, presuppose an individuation of the world according to other fields in their discussions of how to individuate mental state content for psychology.

I. BURGE MAKES LINGUISTICS THE ARBITER OF ONTOLOGY FOR PSYCHOLOGY

Burge claims that when commonsense psychology predicts and explains the behavior of individuals on the basis of their beliefs and desires,

that is, when it explains or predicts what people have done or will do on the basis of the particular beliefs and desires they have, it bases those predictions and explanations on a socio-linguistic notion of the content of intentional states. His position is that if we are willing to use a given sentence to attribute a mental state, say a belief, to someone who is competent in the use of language, the particular words of the sentence we have chosen to use accurately express the content of that individual's belief, regardless of whether or not the individual in question has any mistaken impressions about the words he, or we, use. Burge maintains that if an individual is linguistically competent, then his beliefs, etc. involve those concepts that are expressed by the words used in reporting his beliefs, regardless of what the individual does or does not know about the things in question.

It should be clearly understood that Burge is not simply claiming that the mental state attributed to the individual changes if we change the semantics of the language in which the attribution is made. Obviously, if we change the semantics of, for example, 'sofa', the mental state state attributed by "Bob wants a white leather sofa" will have changed, because what state has been attributed is determined by the semantics of the language used to make the attribution. But Burge is not merely claiming that the semantics of a language determine what mental states can and cannot be attributed in that language. His position is that the semantics of a language determine what mental state contents an individual who is competent in that language can have. Burge is arguing that the content of an intentional state is determined by the nature of the world as that nature is individuated

or taxonomized for the purposes of linguistics, for the purposes of constructing a theory of syntax and semantics. Burge's position is, essentially, that mental state content is fixed by semantic fiat. By changing the semantics of a language, you can change the contents of the mental states of an individual who speaks that language, even if neither the world, nor the individual, nor the relation between the individual and the world has been altered.

Burge's arguments for a linguistic individuation of mental state content are based on thought experiments that rely on intuitions about what we are or are not willing to say. He claims, for example, that because we are unwilling to use our word 'arthritis' to attribute a belief to Alfred in the counterfactual situation (a situation, remember, in which neither 'arthritis' nor any other word refers only to arthritis), although we would, according to Burge, "naturally" use it in attributing a belief to Alfred in the actual situation, Alfred's belief contents must vary between the actual and the counterfactual situations. Burge is asserting that the contents of an individual's beliefs are determined by the socio-linguistic practices of the community in which he lives, not per se by factors concerning the individual himself. Burge maintains that in the counterfactual situation Alfred cannot have the same belief content as he does in the actual situation because the counterfactual linguistic community in which he lives does not recognize arthritis as a taxonomic category, i.e., does not have a word for arthritis. The way the world is carved up into types or taxonomic categories by the semantics of the language a person speaks determines what mental state contents he can have.

In Chapter 3, I argued that Burge's claim that our best and most

well-developed theories of vision are non-individualistic is based on a misunderstanding of individualism, and then discussed a number of different ways in which Burge's own approach conflicts with the way we commonly treat beliefs and belief ascriptions. These conflicts, I have argued, arise directly from his insistence that the way the world is individuated by linguistics is the only taxonomy applicable when evaluating the content of a speaker's mental state. But this sort of parochialism about mental state content is untenable. In effect, Burge's approach makes linguistics the arbiter of the ontology for psychology. Burge's position, in this respect, not only conflicts with our actual practice of attributing mental states (we often use the same content-clause for attributing beliefs to individuals whom we know have somewhat different ideas about the extension of the objects referred to by the content-clause), but depends upon an argument that is internally flawed as well.

We routinely attribute beliefs and desires to animals using exactly the same content-clauses as we do when attributing beliefs and desires to users of language. Furthermore, we reason about animal behavior in the same fashion, give the same explanations, make the same predictions, and experience the same degree of success in those predictions and explanations as we do when we are dealing with a language user. There is nothing about the way we treat intelligent animals in commonsense psychology that distinguishes them from users of language.³ But as Burge grants, in the case of attributions of intentional states to children and animals the contents of their intentional states are not determined or limited by the semantics of a shared language. Nevertheless, we have no hesitation to say, for

example, that Tiny, a dog who has seen only pebbly stretches of nearly level shore, and Blacky, a dog who has seen only sandy stretches of nearly level shore, can both have beliefs correctly attributable with the word 'beach.' We are willing to use the word 'beach' in attributing their beliefs, not because we take them to have beliefs that involve the concept expressed by the English word 'beach,' but because 'beach' is the best we can do in our language, and we believe that it is close enough for practical purposes. The dogs just do not happen to categorize things in the world in exactly the same way that English does. Their beliefs reflect a taxonomy of the world that do not perfectly match the taxonomy of the world derived from the semantics of English. However, in practice we simply do not require every individual to whom we attribute a specific content-clause to individuate the world in the way implicitly specified by the semantics of the language of the attribution. At least in the case of children and animals, we do, as a matter of practice, allow a degree of variation in taxonomic categories given the same content-clauses.

It should be noted that admitting some variation in the way individuals taxonomize the world does not hinder our ability to use commonsense psychology to make adequately accurate predictions and explanations of the behavior of those individuals. All that commonsense psychology requires is that there not be too great a divergence between the taxonomies used by the various individuals to whom a given content is attributed. The predictive and explanatory goals of commonsense psychology are not adversely affected by modest taxonomic variations. When we attribute beliefs and desires to animals we do so on the anthropomorphic assumption that they are

basically like us. If pressed on an attribution of an intentional state content to an animal ("Do we really want to attribute a belief about going to the vet to a dog who has no concept of health care?" "Should we really attribute the belief that there is a mouse in the cellar to a cat that does not distinguish between mice and shrews?") we may become somewhat uncomfortable with our attributions. But even when we have been forced to acknowledge the (strict) inaccuracy or inappropriateness of an attribution, we will nevertheless often continue to use that very attribution in our predictions and explanations of the animal's behavior. The cat may not distinguish between mice and shrews, but that fact makes no difference in terms of the sorts of predictions and explanations of behavior that commonsense psychology typically makes. The degree of success that commonsense psychology aspires to -- and achieves -- does not require that every individual to whom we attribute a mental state content with a given clause taxonomize the world in exactly the way specified by the words we use in the belief attribution. Although taxonomic variation will disrupt commonsense psychology if it is great enough, the sort of taxonomic variations that typically occur are not generally sufficient to interfere with the functioning of commonsense psychology. The purposes of commonsense psychology require, not identity of taxonomic scheme, but mere similarity. Thus, the fact that in ordinary practice we make attributions in the only language we have should not be taken as a theoretical position inherent in or to explanation of behavior in terms of beliefs and desires.

It is not only when there is no common language, however, that common practices indicate that individuals can, and do, categorize the

the world in ways that do not accord precisely with the taxonomy implicit in the semantics of a shared language. On the one hand, we find it plausible that someone might wonder whether the belief he expresses by saying that Simon has influenza, is the same belief that everyone else has when they say that they believe that Simon has influenza. On the other hand, we find it implausible that the mere possibility that someone might not fully understand a word he uses forces him to doubt whether he ever believes what he thinks he believes (See Chapter 3). These two facts are a clear indication that we are willing to allow taxonomic self-determination in matters of mental state content to users of language as well as non-users of language. The very fact that we consider my doubt about whether my expressed belief is the same as what everyone else believes a reasonable, legitimate doubt for me to have, indicates that in actual practice we acknowledge that people who are competent speakers of a language may nevertheless have different belief contents which they express with the same sentence. What we believe may not be what our (sincere) linguistic utterances mean as types.

I do not intend to be denying that the intuitions Burge appeals to about what we are or are not inclined to say about what someone believes are genuine, but I do want to deny that those intuitions are relevant to the debate over intentional content. There is no doubt that some speakers of English may be somewhat hesitant to say of Alfred in the counterfactual situation that he believes that he has arthritis in his thigh, to say of Adam_{te} that he believes that there is water nearby, or to say of Bertrand that he doubts that sofas are pieces of furniture made or meant for sitting. But, by the same token, if

pressed, we may also be somewhat hesitant to say of Alfred in the actual situation that he believes that he has arthritis in his thigh, to say of Adam that he believes that there is water nearby, or to say of Albert that he doubts that sofas are pieces of furniture. Just as we may become uncomfortable attributing to a cat a belief about mice when we are confronted with the fact that the cat does not distinguish between mice and shrews, so too we may become uncomfortable with attributing a belief about arthritis to someone who does not distinguish between arthritis and various other rheumatoid ailments, a belief about water to someone who has no concept of natural kinds, a belief about sofas to someone whose beliefs depart radically from that of the community. To draw conclusions about the content of intentional states on the basis of our discomfiture over certain everyday attributions of intentional states is misguided. If forced to examine our attributions of intentional state contents carefully enough, I think we could come to question the accuracy of almost every attribution. The purpose of the attributions of intentional states that commonsense psychology sanctions is not per se to accurately and precisely specify the content of the individual's mental state.

The purpose of commonsense attributions of beliefs and desires is, fundamentally, practical communication. Attributions of beliefs and desires in ordinary discourse are made to facilitate the audience's understanding of how someone has or might behave. I tell you that A believes that P to explain to you why A did something, or to help you make predictions about what S might do. More generally, to help you avoid confusion in your dealings with A. What I say A believes is dependent upon not only facts about A, but on what I know or believe

about the use to which you will put the information I give you. If I know that you are interested in understanding why, for example, Albert is so upset, and I also know that you know nothing about Albert's strange beliefs about sofas, telling you that Albert thinks that Bob sat on his sofa is not going to help you very much. What will help explain Albert's state to you is if I tell you that Albert thinks Bob sat on a sacred object, or that he thinks Bob desecrated a religious artifact. On the other hand if you want to know why Albert has been collecting furniture catalogs recently, telling you that he wants a new sofa is going to satisfy your curiosity, whereas telling you that he wants a new religious artifact will not. Obviously, in neither of the cases have I given you a full picture of the causes of Albert's behavior, but if all you are interested in is making some sort of sense out of what you witness him doing, I do not have to give you the whole picture. All I need to do, and often all I will do is tell you the cause of his behavior in terms most likely to satisfy your interest. The decision to use "religious artifact" in one attribution, but "sofa" in another is not based on, or indicative of, Albert's intentional states involving different concepts. How we express Albert's beliefs and desires is determined, in part, by what we feel is expedient given the audience and the audience's goals. It should be noted that this also explains or can account for why we say what we do about animals and children.

Our ordinary attributions of intentional states are, fundamentally for the purpose of communication, and communication is served only if we adopt the categorial scheme and goals of the audience. What expressions we use in common attributions of beliefs and desires is

determined by both what we know about our subject (the individual to whom we are attributing the belief or desire), and what we know about our audience. A scientific psychology, in contrast, is concerned with explaining and predicting how an individual's behavior is determined by what that individual thinks -- believes, hopes, fears, wants, etc. Its objective is not to convey information about one individual to others so that they can understand or interact with the individual successfully. The goals and purposes of common attributions of beliefs and desires are different from those of a scientific psychology. Consequently, the significance for psychology that Burge wants to attach to common attributions is unwarranted.

The relation between folk psychology and a scientific psychology is roughly the same as that between folk astronomy and scientific astronomy. Folk astronomy describes the motion of the sun, stars, and planets relative to us. It describes the sun as rising in the east, setting in the west, and moving north and south with the seasons. But the science of astronomy is not expected to endorse such descriptions. The descriptions a scientific astronomy is expected to use are ones that facilitate constructing the generalizations of a scientific field. The same is true of a scientific psychology. Folk psychology describes beliefs and desires relative to us, in a way that we will find useful in our daily activities. But, a scientific psychology should adopt a way of describing the phenomena it is concerned that facilitates constructing scientific generalizations, regardless of whether or not doing so will break with folk practices. A scientific psychology needs to adopt a perspective that will facilitate explaining how the behavior of an individual is dependent upon the

beliefs of the individual. If a scientific psychology takes into consideration the perspective of the audience when specifying content, it will not succeed in its attempts to explain the effects of intentional states on behavior. What we say in an ordinary situation about what someone believes is influenced by factors that, given the goals of a scientific psychology, are irrelevant. Consequently, what, in an ordinary situation, we say someone believes cannot be used as a reliable indicator of intentional state content. Contrary to what Burge claims, what we say in common situations of intentional state attributions have little significance for a scientific psychology.

Burge argues from certain intuitions about what we are willing to say about someone's intentional states to the conclusion that mental state content is determined by the semantic rules of language. In Burge's arthritis example, the individual is held constant between the actual and the counterfactual situation. By hypothesis, there are no physical, functional, or phenomenological differences in the individual between the actual and the counterfactual situation. Nevertheless, Burge asserts, we are not willing to use the same sentence to attribute a belief in the counterfactual situation as we use in the actual situation. Since the only difference between the actual and the counterfactual situations is the semantics of the language spoken, our unwillingness to attribute the same belief to the individual must, according to Burge, be caused by the difference in the languages. In the actual situation, we attribute to Alfred the belief that he has arthritis in his thigh, because that is the meaning of the sentence he utters when expressing his belief. In the counterfactual situation, the belief that (counterfactually) is

attributed to Alfred is the belief that he has tharthritis in his thigh because that is the meaning of the sentence he utters when expressing his belief. The fact that we in the actual situation are disinclined to attribute a belief using the same content clause to Alfred in both the actual and the counterfactual situations Burge takes as proof of a difference in belief contents. Since the only difference between the situations is the semantics of the languages, the content of an individual's belief must, Burge maintains, be fixed by the semantics of his language. Note that Burge's analysis makes use of the fact that the meaning of 'arthritis' in English does not vary from individual to individual. But from whence did this constant meaning come? The fact that linguistic/sentence types have a single, meaning is stipulative, not empirical. It is a convention of language that terms of a single language have a constant meaning. The constant meaning that is attached to term types is a function of the divergent linguistic behavior of individuals. The constant meaning of a term type is derived, using some function weighted to account for the preeminence, etc. of the various speakers, from the idiosyncratic linguistic behavior of individuals.

Burge, himself, must recognize the contribution of individual linguistic behavior to linguistic meaning. But the very fact that individual linguistic behavior contributes to linguistic meaning depends upon that individual linguistic behavior being meaningful. Constant meaning of linguistic types cannot be derived from meaningless sounds. Linguistic behavior is, itself, a type of behavior influenced by mental states. If individual linguistic behavior can be meaningful prior to term types being assigned a

constant meaning, then mental state content cannot be inherently linguistic. Psychology is concerned with how mental state contents influence behavior. But since content is not inherently linguistic, psychology need not be concerned with the constant meaning of linguistic types. Psychology is concerned with the individual mental state contents that produce the individually divergent, meaningful linguistic behavior in light of which linguistics determines a constant meaning. One cannot, as Burge does, argue from the fact that term types have constant meanings to the conclusion that mental state contents are constant, because constant meaning of term types presupposes individual linguistic behavior caused by mental states having content relevant to the linguistic behavior.

Burge argues, on the basis of what we may or may not be inclined to say when reporting someone's belief, that the notion of content used in folk psychology is a socio-linguistic one. He claims that no psychological theory, scientific or otherwise, that fails to acknowledge "the role of the social environment" (IM, p. 73) in determining mental state content can be a plausible account of intentional states. The position he has adopted is, in essence, that the taxonomy of the world according to linguistics is the only appropriate taxonomy to use when assessing mental state content. Against this particular view, I have argued that in our folk psychology attributions of intentional states do not use, with any degree of consistency, the notion of socio-linguistic content Burge advocates. Since actual practice conflicts with Burge's theory of content, we, at the very least, cannot view his theory as an explication of the commonsense notion of content. Furthermore, I have argued that his

theory is not even internally consistent. His argument for a socio-linguistic notion of content depends upon word types having a single, shared meaning. On anyone's account, including Burge's, the fact that we all mean the same thing is stipulated, not discovered. Word type meaning is determined on the basis of the divergent (although not widely divergent) linguistic behavior of individuals. But linguistic behavior is caused by beliefs and desires, and if the behavior is divergent, so it would seem that the beliefs and desires are different. One cannot argue from the stipulated fact that a word has a single meaning to the claim that there is no variation in individual mental state content when the single meaning is, itself, a product of divergent behavior, and therefore, divergent mental states.

II. FODOR IS LED ASTRAY, AND COMES TO DENY THE ONTOLOGICAL AUTONOMY OF PSYCHOLOGY

In "Methodological Solipsism" Fodor claims that any scientific psychology that can account for mental causation, that is, the causation of behavior by intentional states, must pair the causal properties of intentional states with their semantic contents. Fodor argues that since the only accounts of how beliefs cause behavior that are even remotely plausible are those that view the mind as a physical system that functions much like a computational device, the computational model of the mind must be presupposed in determining how to individuate mental contents. According to Fodor, the actual

functioning of a computational system depends upon the physical properties of its states, not on any semantic properties those states may have. However, since mental states are to be taxonomized on the basis of their semantic properties, even though it is their formal properties that are causally efficacious, Fodor claims that meaning or intentional content can be appealed to in psychological explanations only if content is tailored to the causally individuated formal states within the individual. Content must co-vary with the formal, computationally relevant physical properties of mental states. Since, according to Fodor, it is only under a narrow or individualistic notion of content that the necessary co-variation is plausible, mental content -- to the extent that it is related to behavior -- must, therefore, be construed narrowly or individualistically by a scientific psychology. Any success that commonsense psychology has had in predicting and explaining behavior is, according to Fodor, due to the fact that it individuates mental state contents in a more or less narrow fashion.

The natural development of the view Fodor expresses in "Methodological Solipsism" would lead to a custom-made ontology of the world for the purposes of psychology. Mental state content ultimately derives from the world, so if the type-individuation of mental states on the basis of content follows the type-individuation of mental states on the basis of causally relevant physical properties, then the type-individuation of the world should also follow the causally relevant properties of mental states. However, apparently as a result of confronting the arguments and examples of Burge, Fodor abandons the taxonomic autonomy implicit in "Methodological Solipsism." By the

time of the writing of Psychosemantics, Fodor, like Burge maintains that, for the purposes of determining mental state content, the world should be taxonomized, not according to the causally/computationally relevant physical properties of the states of the intentional system in question, but, instead, according to some pre-existing scheme of what there is in the world.

In Psychosemantics, Fodor accepts the assertion that Adam and Twin-Adam have beliefs that have different extensions. Although the difference between water (H_2O) and twater (XYZ) is immaterial to Adam and Twin-Adam, Adam has a belief about water, and only water, while Twin-Adam has a belief about twater, and only twater (See Psycho, pp. 44-53). Fodor argues that in spite of the fact that their thoughts have different extensions, their thoughts are nevertheless "narrow" content identical. Furthermore, Fodor claims, we can acknowledge that their narrow content identical thoughts have different extensions without thereby denying that content 'determines' extension. Narrow content is, according to Fodor, a function from contexts and thoughts onto truth conditions. Two thoughts are narrow content identical if and only if "they effect the same mapping of thoughts and contexts onto truth conditions" (Psycho, p. 48). Adam and Twin-Adam are in narrow content identical mental states because the extensions of their thoughts would be the same were they in the same context. "Given the neurological identity between us, in a world where I am in my Twin's context my 'water'-thoughts are about XYZ iff his are. (And, of course, vice versa: In a world in which my Twin is in my context, given the neurological identity between us, it must be that his water-thoughts are about H_2O iff mine are.)" (Psycho, p. 48).⁴

Narrow content, according to Fodor, is still to co-vary with the formal properties of mental states, but what the extension of a narrow content is in a context is defined in accordance with the language or physics of the local environment. Fodor maintains that on Earth water is a psychologically relevant taxonomic category. On Twin-Earth, twater is the psychologically relevant taxonomic category. But Fodor does not allow for there to be a single non-disjunctive property that includes both water and twater; for example, he does not allow for there to be a taxonomic category based solely on the phenomenal properties of these chemically distinct substances. Content does vary with context, but for each environment or world there is a single taxonomy dictated by the meaning of the term-types of the language which they use, and everyone's mental contents follow that taxonomy of the world.

[T]he English expression 'the thought that water is wet' can be used to specify the narrow content of a mental state that my Twin and I share (even though, qua anchored to H_2O , it doesn't, of course, express that content). In particular, it can be used to pick out the content of my Twin's 'water'-thought via the truth conditions that it would have had if my Twin had been plugged into my world. Roughly speaking, this tactic works because the narrow thought that water is wet is the unique narrow thought that yields the truth condition H_2O is wet when anchored to my context and the truth condition XYZ is wet when anchored to his.

(Psycho, p. 51)

In other words, Fodor is claiming that the English expression 'the thought that water is wet' can be used to specify the thought type that Adam and Twin-Adam both have because the English expression 'the thought that water is wet' identifies a unique narrow thought function, one that has the value H_2O is wet on Earth, but XYZ is wet on Twin-Earth. Since narrow contents and narrow thoughts are constructs of Fodor's scientific psychology, he is in essence claiming

that there is a direct mapping from the taxonomic categories of language to the taxonomic categories of psychology. If there is a direct mapping from the semantics of language to the taxonomic categories of psychology, then there must, according to his own view, also be a mapping from the semantic of language to the causally relevant physical properties of mental states, since the taxonomic categories of psychology follow the causally relevant physical properties of mental states. However, this latter mapping requires that there be a nomic correlation between causally relevant physical properties of mental states and the taxonomic categories into which the semantics of language divide the world. But whether or not there is such a mapping is an empirical matter that cannot simply be stipulated or assumed to exist.

For Fodor to be able to claim that Adam and Twin-Adam have thoughts with the same causal powers, even though they have thoughts with different extensions, he needs a notion of content that maps from same causal powers to different extensions (different wide contents). Fodor suggests that the way to achieve this mapping is to construe narrow content as a function that specifies an extension when evaluated in a context. The problem with this suggestion, however, is that there are no independent grounds on which to determine sameness or difference of context. In explicating what he means when he says that content determines extension relative to a context, Fodor declares

[I]t's presumably common ground that there's something about the relation between Twin-Earth and Twin-Me in virtue of which his 'water'-thoughts are about XYZ even though my water-thoughts are not. Call this condition that's satisfied by {Twin-me, Twin-Earth} condition C (because it determines the Context of his 'water'-thoughts). Similarly, there must be something about the relation between me and Earth in virtue of which my water-thoughts are about H₂O even though

my Twin's 'water'-thoughts are not. Call this condition satisfied by {me, Earth} condition C'. I don't want to worry, just now, about the problem of how to articulate conditions C and C'. Some story about constraints on the causal relations between H₂O tokenings and water-thought tokenings (and between XYZ tokenings and 'water'-thought tokenings) would be the obvious proposal; but it doesn't matter much for the purposes now at hand.

(Psycho, p. 48)

According to Fodor, context is determined by some condition, C. Note, however, that condition C itself presupposes the extension of the thought in question. Condition C, is according to Fodor, whatever it is about the relation between Adam and Earth that makes his thought be about H₂O, rather than XYZ or some property that H₂O and XYZ share, and condition C will (probably) be specified in terms of the causal relation between H₂O tokenings and water-thought tokenings. Put more generally, condition C is whatever it is about the relation between an individual and the environment that determines the extension of the thought, and that condition is to be specified in terms of the causal relation between tokenings of the extension and tokenings of the thought. But in order to specify such a causal relation, one has to know the extension. Thus, contexts are themselves individuated by extension. Therefore, Fodor's claim that in the same context two thoughts with the same content will have the same extension is vacuous.

If sameness of context just is sameness of extension, we are back to the question of what determines extension. The original intuition that led Fodor to endorse a notion of narrow content as a function, one that determines "an equivalence class of mechanisms" (Psycho, p. 52), is that Adam and Twin-Adam have thoughts with different extensions. But the claim that their thoughts have different extensions presupposes an individuation of the world into types. While it is certainly true

that H₂O and XYZ are type distinct substances according to chemistry, whether or not they are type distinct substances according to psychology has, per se, nothing to do with what chemistry has to say about the matter. What the extension of their thoughts is is a matter to be determined by psychology, not chemistry. The fact that a chemist, if presented with a drop of every liquid sample that had or could bring about a water/twater-thought, would identify two distinct substances is irrelevant to the issue of whether or not water/twater thoughts are extensionally equivalent. It is only if H₂O and XYZ are type distinct from the point of view of psychology that we should claim that Adam and Twin-Adam have thoughts with different extensions.

Fodor's abandonment of the position of "Methodological Solipsism," that psychology need not be concerned with the characterizations that linguistics or physics would give to content clauses, for the position is Psychosemantics, that psychology needs a non-semantic notion of "content" that will map narrow content-types onto sets of kind-types of physics, chemistry, etc. is, I believe, ill-advised. To begin with, if psychology is concerned with narrow content and narrow content is a non-semantic property, as Fodor claims it is (Psycho, pp. 50-53), then the semantic properties of mental states are immaterial for a psychology concerned with how behavior is influenced by intentional states. In Chapter 2 I argued that, because of the unexplicated and non-causal harmony between content and causal powers, the semantic properties of mental states did not play a sufficiently robust role in Fodor's original theory for it to be an account of mental causation. Given this new notion of non-semantic content, it is evident that the semantic properties of mental state have no

influence on behavior. Secondly, by claiming that narrow content is a function that maps to different extensions (in different contexts), Fodor is, in essence, claiming that psychological kind predicates are reducible to the kind types of physics, chemistry, the semantics of language, etc. If, as Fodor maintains, narrow content is a function from contexts to truth conditions, then Adam has the narrow thought that water is wet iff Adam either is on Earth and his thought is about H₂O, or is on Twin-Earth and his thought is about XYZ, or is a brain in a vat and his thought is about, say, electrode 28, or But this is precisely the sort of reductionist program that Fodor argued against so vehemently in The Language of Thought. He discusses at some length a case from economics, and draws a parallel to psychology:

Gresham's law says something about what will happen in monetary exchanges under certain conditions. I am willing to believe that physics is general in the sense that it implies that any event which consists of a monetary exchange (hence any event which falls under Gresham's law) has a true description in the vocabulary of physics and in virtue of which it falls under the laws of physics. But banal considerations suggest that a physical description which covers all such events must be wildly disjunctive. Some monetary exchanges involve strings of wampum. Some involve dollar bills. And some involve signing one's name to a check. What are the chances that a disjunction of physical predicates which covers all these events (i.e., a disjunctive predicate which can form the right hand side of a bridge law of the form 'x is a monetary exchange <--> . . .') express a physical kind? In particular, what are the chances that such a predicate forms the antecedent or consequent of some proper law of physics? The point is that monetary exchanges have interesting things in common; Gresham's law, if true, says what one of these interesting things is. But what is interesting about monetary exchanges is surely not their commonalities under physical descriptions.

(LOT, p. 15)

The point Fodor is making is that although, by brute force, an omniscient physicist might be able to construct a physical predicate that was coextensive with every instance of monetary exchange, unless

that physical predicate identified a physical kind that was lawfully coextensive with instances of monetary exchange, which seems highly unlikely, it will not be able to capture what is interesting about Gresham's law. Fodor, himself, explicitly draws a parallel between economics and physics and psychology and physics, asserting that "[e]ven if (token) psychological events are (token) neurological events, it does not follow that the kind predicates of psychology are coextensive with the kind predicates of any other discipline (including physics)" (LOT, p. 17). By construing narrow content as a function, Fodor has made content the sort of wildly disjunctive predicate that he disparages in economics. The predicate 'believes that P' has been reduced to a disjunctive predicate couched in terms of the kind predicates of ordinary language. Although Fodor does still want psychology to endorse a taxonomy of mental states that enhances its ability to explain and predict behavior (narrow content is a distinctively psychological predicate that is supposed to co-vary with causal properties), he does not seem to realize that the kind predicates of psychology are not only the basis for the individuation of mental states, but they are the basis for the individuation of objects in the world. Psychology is concerned with how mental states, in virtue of their contents, in virtue of the relations they bear to the world, influence behavior. The contents of mental states are a product of the relation between the individual and the world. The individuation of the world cannot be divorced from the individuation of content. The taxonomic types, the kind predicates, of psychology apply equally to mental state contents and to the world. The kind predicate with which psychology identifies some individual's

mental state content must also identify that in the world that is responsible for the content. The fact that what a psychological kind predicate picks out in the world may not be coextensive with what any kind predicate in any other field picks out is irrelevant when determining the kind predicates psychology should use.

The original conviction that lead Fodor to adopt methodological solipsism was that machine functionalism is the only plausible model for mental causation. The consequence of this view, which Fodor initially endorsed, is that if a scientific psychology is to explain and predict behavior in terms of intentional state, the notion of content it uses must be one that ensures that content co-varies with causal (physical) properties. For the purposes of a scientific psychology, the semantic content of mental states must be determined narrowly or individualistically. If we accept that the individuation of content must be coordinated with the individuation of mental states by causal powers, then because content is derived from a relation to the world, the individuation of the world must, therefore, also coordinate with causal powers. The true moral of methodological solipsism is that both sides of the mental state-world relation must be individuated in a way that coordinates with causal powers, but Fodor seems to espouse only half of this moral. In "Methodological Solipsism" Fodor explicitly argues for the autonomy of psychology, but he advocates its autonomy only with respect to the individuation of mental states qua physical or functional states. He is not willing to grant psychology autonomy in individuating mental states qua representational states. Psychology, according to Fodor, must be granted the right to individuate states in the head as it sees fit, but

it is not to be granted the right to individuate the world as it sees fit. Fodor wants to accept only half of the moral of methodological solipsism, whereas what we need to do if we are to vindicate commonsense psychology is to embrace both halves. Psychology must be permitted to individuate both that which is inside the head and that which is outside the head as it sees fit.

III. DRETSKE'S CONCERN FOR KNOWLEDGE LEADS HIM TO ENDORSE THE ONTOLOGICAL SUBORDINATION OF PSYCHOLOGY

Dretske is clearly in agreement with Fodor's original claim in "Methodological Solipsism" that content must be paired with causal powers if there is to be any hope of a scientific psychology of intentional states. However, in Explaining Behavior Dretske takes the position that a mere parallelism of content and causal powers, such as Fodor's theory envisions, is insufficient. We want to explain the way individuals behave by appealing to their intentional states, and Dretske claims, rightly I believe, that in order for explanations couched in terms of beliefs and desires to explain why individuals behave the way they do, the contents of beliefs and desires must be responsible, in some sense, for the types of effects on behavior to be explained. The fact that a state has the content it does must, according to Dretske, help explain why that state has the causal powers it does. Dretske argues, in Explaining Behavior, that an account of intentional state content grounded in information theory

and utilizing a notion of the function, purpose or job of a state in a system will yield a notion of content which has the requisite degree of control over causal powers.

Like Fodor and Burge, Dretske adopts a notion of content that makes a pre-existing taxonomy of what there is in the local environment the arbiter of intentional state content. However, unlike Fodor and Burge, Dretske seems to have been initially motivated to adopt such a view of content on epistemological grounds (see KFI, Ch. 5). In Knowledge and the Flow of Information, Dretske argues that in order for knowledge to be possible, it must be that a signal will carry a piece of information, P, just as long as the conditional probability of the source of the signal being P is 1, i.e., has 0 equivocation. Since evil demons, mad scientists, and Twin-Earths could introduce equivocation into a signal (just as a weakened spring could), if not somehow ruled out as genuine alternative states of the signal's source, Dretske stipulates that only possible sources of the signal in the local environment are genuine alternatives, and, therefore, only local alternatives need be considered in determining whether or not a signal is equivocal.

In Chapter 4, I argued that stipulating that only things in the local environment can affect the equivocation of a signal on the grounds that such an assumption is necessary if we are to have knowledge is unjustifiable. A scientific psychology does not require that knowledge be possible in order to predict and explain how behavior is influenced by mental states. All that it plausibly requires is belief. But, even if we grant what I have argued there are no grounds for granting, namely, Dretske's claims that "an indicator indicates what it can reliably discriminate (when functioning normally) in its

natural habitat" (ACR, p. 109), we cannot assume that it is a chemist's or a physicist's description of what it can reliably discriminate that is the correct one. Restricting indication to the local environment does not by itself authorize or justify the adoption of a pre-existing taxonomy of that environment. One cannot claim that a signal is nomically correlated with the property of being H₂O simply because all samples of H₂O give rise to the signal. Even if it turns out by chance that only samples of H₂O give rise to the signal, the signal still may not carry the information that something has the property H₂O. It may not be the H₂O-ness of the samples that is responsible for the correlation between the signal and the samples. The signal carries information about whatever property of the source is responsible for the signal being what it is. "If a structure s carries the information that t is F, it does not necessarily carry the information that t is G even though nothing is F that is not also G" (KFI, p. 172). In order for a signal to carry the information "F" it must be that the source's being F, rather than G, is responsible, in some sense, for the information carrying features of the signal (KFI, p. 64). If nomic correlation grounds information, then even if, by some cosmic accident, it turns out that all and only dingbats in the local environment are green with purple splotches, a signal can carry the information "green with purple splotches" without carrying the information "dingbat." Whether the signal carries the information "green with purple splotches," "dingbat," or both, has to be determined on the basis of what property the signal is responsive to.

Dretske, however, does not apply this criterion for what information a signal carries to the examples he gives that supposedly demonstrate

that his semantic information theory entails a wide notion of content. Dretske claims that Tommy, who lives in a world having both H₂O and XYZ, receives signals that carry the "essentially disjunctive" information "H₂O or XYZ" (See KFI, pp. 225-227; EoB, pp. 15-16). But in order to claim that Tommy receives the essentially disjunctive piece of information "H₂O or XYZ" it must be that the signal is nomically correlated with instances of either the physical property H₂O or the physical property XYZ, rather than, say, a certain set of common physical and phenomenal properties. To claim that Tommy receives the information "H₂O or XYZ", it must be that the signal he receives is one responsive to molecular structure. If the signal is responsive to the molecular structures of H₂O and XYZ, then a different signal should result from some stuff, call it 'WXY', having identical macro-physical and phenomenal properties as H₂O and XYZ but having a different molecular structure than either H₂O or XYZ. The situation with respect to H₂O/XYZ and WXY is, of course, analogous to the situation with respect to H₂O and XYZ in the Earth/Twin-Earth case, and we know that the signal does not change given a sample of XYZ rather than a sample of H₂O. Since XYZ produces the same signal as H₂O, we can conclude that a sample of WXY miraculously transported (from Twin-Twin-Earth) to Twin-Earth, would produce the same signal that H₂O and XYZ actually do. Thus, it would seem that we cannot, in fact, claim that Tommy receives signals responsive to the particular molecular structures of H₂O and XYZ. What the signal is responsive to is a set of phenomenally identified macro-physical properties, so Tommy does not receive the information "H₂O or XYZ," but rather the information "Q," where Q is some complex set of macro-physical and phenomenal properties.

Even if it were the case that the signal was responsive to the molecular structures of H_2O and XYZ, in order to claim that Tommy receives the essentially disjunctive piece of information " H_2O or XYZ," it must be that the taxonomic categories into which psychology places the stuff with which the signal is correlated in the local environment are the same (identical) taxonomic categories into which chemistry would place the stuff for its purposes. The claim that Tommy receives an essentially disjunctive piece of information presupposes that psychology uses the same taxonomic categories that chemistry does. The information that Tommy (counterfactually) receives is disjunctive only if psychology taxonomizes H_2O and XYZ as different substances for the purposes of determining Tommy's mental state content. But simply because chemistry taxonomizes H_2O and XYZ differently for its purposes, does not mean it is appropriate for psychology to taxonomize them as different stuffs for its purposes. The appropriateness of psychology adopting the taxonomy of the world of chemistry for determining mental state content must be argued for. It cannot simply be assumed or presupposed.

In Chapter 4, I indicated that Dretske's claim that his theory entails a wide notion of content could be supported if either the information theory component of content or the function component of content required such a notion of content. So far, however, I have only discussed the information component of content, arguing that it does not entail a wide notion of content. At this point I want to pick up the other half of the question, and examine whether or not the notion of function results in a wide notion of content. It will be my contention that, like the notion of information, the notion of

function does not entail or require a wide notion of content, either directly or indirectly. Consequently, it will be maintained that there are no grounds on which to claim that an information theory based account of intentionality leads to a wide notion of content.

I think it should be obvious that the contribution that function makes to determining the content of an intentional state cannot in any direct way result in a wide notion of content. Function plays the role of determining which of the things an internal state is already an indicator of it has as its semantic content. The power of the internal state to indicate comes prior to its having the function of indicating anything. If what a state is an indicator of is determined individualistically, function cannot convert that individualistic information to wide content. For example, if the most specific piece of information that a state carries is "clear, colorless, tasteless liquid . . .," an appeal to the notion of function cannot extract from that the content " H_2O ". One cannot extract "is red" from "is colored." A type III RS state cannot acquire the function of indicating something that it has never indicated. What information a state carries limits what it can acquire the function of indicating, and since information is individualistic, then content will also have to be individualistic.

It might be claimed, although Dretske, himself, does not discuss appealing to the notion of function in this fashion, that although the notion of function does not directly result in a wide notion of content, it does so in an indirect fashion through its dependence on the notion of success or benefit. It might be claimed that since what some internal indicator has the function of indicating is determined by considering which of the things it indicates is important for the

success of the behavior it helps produce, indication should be assessed against the same background that success or benefit is assessed, i.e., the local environment. But such a claim to limit the context within which indication is determined cannot be motivated from within the theory itself, and limiting context without such theoretical motivation serves only to undermine the objectivity of information.

In order to justify claiming that we should limit the context within which we determine what something is an indicator of because the success of output is assessed in the local context, it would have to be that there was some sort of necessary connection between what a state indicates and the success of the behavior that the state plays a role in bringing about. But there is no such necessary connection. Not only is it the case that an internal indicator need not indicate "H₂O," rather than "clear, colorless, tasteless liquid . . ." for it to be recruited as a cause of some output beneficial when H₂O is present, but, in fact, an internal state need not indicate anything even related to those conditions on which some output's success is dependent for it to be recruited as a cause of that output.

If we have a system that lacks an internal indicator for condition F, a temporary solution to The Design Problem [the problem of how do we get a system to do M when and only when condition F exists] can nonetheless be reached if there is an internal indicator of some condition [G] which, through coincidence, temporary arrangement (by an experimenter, say) or circumstances of habitat, is correlated with F. . . . An internal representation of G develops because the internal indicator of G is given its job in the production of output because of what it indicates about external affairs. Depending on the degree of correlation between F and G, this will be a more or less effective solution to The Design Problem. The better the correlation, the more successful the animal will be in producing M in conditions F (and, therefore, in getting whatever reward it is that promotes that response).

If the correlation (however temporary) between F and G is

perfect, this solution to The Design Problem will (for however long the correlation persists) be indistinguishable from the original solution, the solution by a system that has an F indicator. But the explanation of the resultant behavior of these two systems will be different. Using the intentional idiom to describe this case, we say that the second animal produces M in conditions F, not because it thinks that F exists, but because it thinks G exists.

(EB, pp. 102-103)

An indicator of G can be recruited as a cause of M, whose success is dependent on condition F, even if the indicator of G carries no information that is even related to F, if there is a fortuitous correlation between G and F. The recruitment of an indicator as a cause of some output is not dependent on the indicator carrying any information about the conditions under which the output is successful or beneficial to the system, so one cannot justify limiting the context in which to assess indication to the context in which success is determined.

Even if one could justify limiting the context in which indication was determined to the local environment by some appeal to the notion of function, one would still not have wide content. According to Dretske, even if the only thing in the local environment that is a clear, colorless, tasteless liquid . . . is H₂O, and vice versa, one will still not necessarily have an indicator of H₂O where one only had an indicator of clear, colorless, tasteless liquid before.

All systems of representation, whatever type they happen to be, are what I shall call property specific. By this I mean that a system can represent something (call it s) as having property F without representing it as having the property G even though everything having the first property has the second, even though every F is G. Even if the predicate expressions "F" and "G" are coextensional (correctly apply to exactly the same things), this doesn't guarantee that an RS will represent s as F just because it represents s as G (or vice versa).

(EB, p. 75)

It may be that in some instances limiting the context does change what a state is an indicator of (for example, if one had a state responsive to the molecular structures of both H₂O and XYZ, then by limiting the context to one which has only H₂O, the state would then be an indicator of H₂O, whereas it previously had not been), but even in those cases what the state is an indicator of has to be determined by looking at what property of the source is responsible for the information carrying features of the state. Indication always has to be determined on the basis of what property is responsible for the information carrying features of the signal, regardless of how one defines the context in which to assess what information the signal carries. Even in local contexts information is individualistic.

The information a state carries must always be determined individually, regardless of whether or not one has limited the context of assessment. While the success of output may depend on biology or chemistry, one cannot on that basis claim that what an indicator indicates should be identified using the taxonomy of biology or chemistry, because the recruitment of an indicator as a cause of output does not require that the indicator indicate anything about the conditions of success of the output. What something is an indicator of is to be determined on the basis of objective relations of dependency that hold between the state and the world. One cannot presuppose that those objective relations of dependency will generate a taxonomy of the world that corresponds to the taxonomy of biology, chemistry, physics, or any other field, nor can one appeal to the notion of function to ensure the coordination of the taxonomies of the world according to internal indicator states and according to biology, chemistry, or any

other field.

Dretske's synthesis of his semantic information theory and a notion of the function of a state within a system to produce an account of intentional states that achieves the coordination of content and causal powers via a scientifically plausible mechanism (i.e., selection) represents a substantial improvement over Fodor's theory, which merely postulates that content and causal powers are coordinated. However, Dretske, like Fodor and Burge, does not recognize that taxonomic categories are discipline relative. What information a signal carries is dependent upon nomic correlation, but what a signal is nomically correlated with is partly a matter of how one has chosen to taxonomize the world. Psychology should claim that the content of Emily's belief is that Diospyros kaki have red to orange fruit only if Diospyros kaki is a taxonomic category relevant to the purposes for which psychology identifies the content of her belief. Diospyros kaki is a botanical kind, but that does not make it a psychological kind. What taxonomic categories there are in the world will depend on who you ask and what they are interested in doing. Dretske takes his theory to endorse a wide notion of content, not because of considerations internal to psychology or the theory itself, but because of his concern over preserving the possibility of knowledge, and his oversight of the interest relative nature of taxonomic categories. Notwithstanding Dretske's remarks to the contrary, neither information theory itself, nor Dretske's own account of intentional states supports a wide, as opposed to narrow, individualistic notion of content.

IV. PSYCHOLOGY MUST BE GRANTED AUTONOMY IN THE INDIVIDUATION OF BOTH
THE WORLD AND MENTAL STATE CONTENT

The rigorous scientific psychology whose conceptual framework is the subject of our concern is meant to deal with the prediction and explanation of cognitively mediated behavioral responses to changes in the environment or circumstances in which the individual finds himself. We want psychology to be able to explain why a particular individual behaved in a particular way on some particular occasion; we want it to be able to predict in principle, at least reasonably accurately, how someone will behave under a given set of circumstances; and we want it to be able to make generalizations about the relation between environmental conditions and/or changes and behavioral responses. Furthermore, while the *raison d'être* of psychology is the accurate prediction and explanation of cognitively mediated behavior, as a scientific field, psychology is subject to those methodological constraints placed on all explanatory scientific activity. Consequently, the framework within which a scientific psychology chooses to explain behavior must be one that, among other things, is consistent with accepted causal mechanisms, both those that are generally applicable and those that apply to behavior qua bodily movement. These predictive and explanatory goals of psychology, along with the general methodological constraints placed on all explanatory scientific fields are the criteria against which we should evaluate the various candidate frameworks.

In general, the basic working assumption of folk psychology is that

we can determine the content of a belief from the content-clause used to attribute the mental state: the content of a mental state is taken to be the same as the meaning of the content-clause used in attributing the mental state. Since the convention of folk psychology is that the content-clause of a belief attribution expresses the content of the belief being attributed, we assume that if the same content-clause can be used to attribute beliefs to two different individuals, then the two individuals believe the same thing, their beliefs will count as being of the same content type. On the other hand, if the same content-clause cannot be used to attribute beliefs to two different individuals, then we assume that they do not have the same belief (although they may very well have different beliefs about the same object). Nevertheless, the fact that folk psychology employs the general rule of thumb that the content of a mental state is given by the meaning of the content-clause used in attributing the mental state cannot, by itself, be taken as evidence that mental state contents are identical to the meanings of the content-clauses used in mental state attributions. The purpose of folk psychology attributions of intentional states is, I have claimed, communication. We attribute beliefs and desires to individuals in order to assist our audience in understanding -- explaining and predicting -- how the subject of the attributions has, or might, behave. Attributions of beliefs and desires -- as acts of communication -- must conform to the conventions of the means of communication. Folk psychology attributions of intentional states are made in public language, and, consequently, must conform to the conventions governing public language. One of the fundamental precepts of language is that the words we sincerely utter

are to be understood by our audience in their standard way, they are to be taken to mean what they normally do. Folk psychology attributions of content are not, cannot be, exempted from this assumption of normal or standard meaning. If folk psychology attributions were not held to this assumption of standard meaning, we, as audiences of those attributions, would be unable to interpret those attributions. Only if folk psychology attributions adhere to the assumption of normal meaning can they be used to convey information, to communicate with an audience.

In order to achieve the goal of communication we must tailor our choice of content-clauses in our intentional state attributions to the audience we are addressing, assuming that our words will be taken to mean what they normally do. We choose the content-clause we will use in a belief attribution on the basis of 1) what we know about the subject of our attribution; 2) what we know about the interests of our audience; and 3) what we know about the conventional meaning of our words. (Note that, although 1 - 3 do not ensure that the content-clauses we use in our attributions express the exact content of the intentional state we are attributing, they are likely to result in our content-clauses being reasonable approximations of the contents of the our subjects' mental states for the practical purposes of the intended audience.) Obviously, the fact that folk psychology attributions of intentional states must follow the linguistic convention of using words in their normal way in order to communicate with an audience successfully does not mean that the content-clauses we use do not express the content of the mental states we are attributing. It simply means that we cannot claim, on the basis of what we say in ordinary situations alone, that the content-clause of an intentional

state attribution expresses the precise content of the intentional state in question.

Certainly, the success of folk psychology should count heavily in favor of adopting its constructs, and we should not discard lightly those concepts and constructs that are central to the folk psychology view and contribute significantly to its success. But, on the other hand, neither should we accept all of the constructs of folk psychology prior to careful examination. A scientific psychology that can/will vindicate commonsense folk psychology need not accept and endorse the particular attributions of intentional states that are made under folk psychology. The vindication of commonsense folk psychology requires only that a scientific basis be found for the folk psychology practice of explaining and predicting behavior in terms of beliefs and desires. The fundamental claim of folk psychology is that what people believe and want influences how they behave. A scientific psychology that endorses the basic framework of explaining behavior in terms of beliefs and desires will have succeeded in vindicating commonsense psychology regardless of whether it endorses a wide or a narrow/individualistic notion of content.

In the vast majority of commonsense attributions, it is not even possible to determine whether the attribution was based on a wide notion of content or a narrow notion. Given the limited number of terms we have readily available, in the vast majority of cases our folk psychology attributions of mental states will be the same regardless of whether they are based on a wide notion of content or a narrow notion. It is only in the case of unusual circumstances (typically dreamed up with the explicit intention of causing

difficulty) that we would make different attributions using a narrow notion of content than using a wide notion of content, and in those cases intuitions are divided over which attribution is "correct." Even if folk psychology does, in fact, use one notion of content over the other (although to expect that level of consistency in a "theory" that consists mainly of a set of aphorisms, maxims, and rules of thumb seems rather unreasonable), nevertheless, it is only to the extent that the constructs and taxonomic categories of folk psychology facilitate the achievement of the goals of a rigorous psychology that they should be accepted by a scientific theory of psychology.

Psychology is primarily concerned with the explanation and prediction of cognitively determined behavior. Its goal is to formulate generalizations about how the behavior of individuals is determined by what they believe and want. The taxonomy of the world (which should be understood to include not only states internal to the individual, but the external environment as well) that psychology should be using is one that serves its purposes, just as other endeavors use taxonomies that serve their purposes. Concerns that fall outside of the province of psychology cannot legitimately influence the kind predicates in terms of which a psychological theory conducts its business. A scientific theory of psychology, unlike commonsense folk psychology, is not, I submit, concerned with communication. The goal of folk psychology attributions of beliefs and desires is to convey to an audience information which they will then use to make their own predictions and explanations. Consequently, folk psychology attributions of mental state contents will be influenced by how the speaker feels the message can best be communicated to the particular

audience to which the attribution is addressed. The attributions of particular intentional states by a scientific psychology, on the other hand, are for the prediction and explanation of the behavior of the particular individual in question using the generalizations of psychology. The generalizations or "laws" of a scientific psychology are the "audience" of its attributions of mental states. Consequently, there are no grounds for the assumption that the attributions of intentional states of a scientific psychology will even be made using common everyday language, and it is clear that such scientific attributions will not need to conform to the conventions governing communication in language.

V. A SCIENTIFIC PSYCHOLOGY WILL EMPLOY A RADICALLY INDIVIDUALISTIC NOTION OF CONTENT

If what we want from a scientific psychology are generalizations about how behavior is dependent on intentional states, the principled starting point for psychology in determining how to individuate the world for the purposes of identifying the content of intentional states, is a mapping between internal state-types, identified on the basis of causal powers, and external conditions. Any two particulars that map to the same internal causal state-type should be placed in the same taxonomic category for determining the content associated with that causal state-type. Conversely, any two particulars that map to different causal state-types should be placed in different

taxonomic categories for determining the content of those causal state-types. I want to suggest that the content of those causal state-types would then be that property or set of properties that all members of the taxonomic set share and by virtue of which they are members of that set, in other words, those properties that are necessary and sufficient for inclusion in the taxonomic set. The hope for a scientific psychology of intentional states is based on the assumption that content does not merely parallel (either by stipulations or by epiphenomenal accident) causal powers, but that there is some mechanism that is responsible for the coordination of content and causal powers. Thus, the only content that a scientific psychology can legitimately attribute to a specific intentional state will be determined by those properties that are necessary and sufficient for the underlying mechanisms to effect the mapping of particulars to causal state-type that has been empirically determined.

Clearly, if a scientific psychology must base its attributions of content on properties of particulars necessary to and sufficient for a mechanism coordinating content and causal powers, it will not be satisfactory for psychology to make its attributions of content on the basis of those properties that are coordinated with the causal state-type in the local environment. It must base its attributions of content on those properties that are relevant to the functioning of the mechanism, without regard to the peculiarities of the local environment. One cannot arbitrarily restrict the context of the generalizations and kind predicates of psychology any more than one can arbitrarily restrict the context of the generalizations and kind predicates of physics. Even though all and only A's are B's, in the

local environment, the B-ness of A's may be irrelevant to the mechanism that coordinates content and causal powers. In order to determine what content is associated with a certain causal state-type, it may be necessary to look outside of the local environment. For example, stickleback fish display certain typical behavior in the presence of conspecifics. In order to determine the content of the internal state that causes this behavior, we need to determine what property or properties in the environment those causal states are coordinated with. One might be inclined to say that the content of those states is "stickleback" because the state is coordinated with the property of being a stickleback. However, in the environment in which sticklebacks naturally occur, (roughly) all and only sticklebacks have a bright red underside, so the causal state may in fact be correlated with having a red underside rather than with being a stickleback. In order to determine what the state has been coordinated with we have to look outside of the local environment for instances in which the various candidates are separate from each another. From Tinbergen's studies⁵ we know that it is the property of having a bright red underside that is coordinated with the causal state type, not the property of being a stickleback. Thus the content of that state, according to a scientific psychology, might be something more akin to "thing with a red underside," rather than "stickleback," although, for reasons that will become apparent, even this is (probably) an overspecification of the content of the stickleback's belief. ⁶

In the case of stickleback fish, noctuid moths, marine bacteria, and, in general, any system whose behavioral responses to the environment are instinctive or genetically determined, the mechanism

responsible for the coordination of causal state-types and properties involves natural selection. Individuals with a certain pairing of content and causal powers were reproductively more successful, for any number of reasons, and therefore influenced the evolution of the species to a greater extent than those individuals with a different, less useful, pairing, who were reproductively less successful. Consequently, there is little variation across individuals of the same species in the properties they respond to, and their responses to those properties. All sticklebacks distinguish between things with a bright red underside, and things without a bright red underside, and (relative to their sex) behave in the same way towards those things having a red underside. "Thing with a bright red underside" is a taxonomic category for all sticklebacks, and all sticklebacks have causal state types with the content "thing with a bright red underside." Cases involving systems that learn, on the other hand, are somewhat more varied. The mechanism responsible for the coordination of properties and causal state-types operates internal to the individual. Granting that the general type of mechanism an individual has may be determined by natural selection, nevertheless, since the mechanism of coordination operates at the level of the individual rather than the species, there is a much greater chance for variation in the taxonomic categories different individuals of the same species employ. Unlike in the case of sticklebacks, noctuid moths, and marine bacteria, one cannot conclude from the fact that "object of such and such size and appearance" is a taxonomic category for some particular individual that it is a taxonomic category for all individuals. Thus, if content is to be determined by what a state has

been reliably correlated with, the content of the mental states of individuals will have to be determined on a case by case basis.

A mechanism that coordinates properties and causal state types can only effect such coordination with respect to properties it can detect. (Only if a mechanism has a means of detecting when the grass needs watering can it turn on the sprinklers when, and only when, the grass needs watering.) All things that fall into the same "detection category" will be of the same kind from the point of view of the mechanism. To put it in another fashion, everything that is stimulus equivalent will count as of the same kind from the point of view of the mechanism. Thus, in the most primitive perceptual cases, for example, stimulus equivalence from the point of view of the individual might be the basis on which content is determined. Figure 1 on page 236 is a schematic view of the relation between the way the world is taxonomized on the basis of stimulus equivalence for an individual who has only perceptual beliefs, and the taxonomies of various other fields. All of the things that fall within the wedge are stimulus equivalent for the individual (located at the center of the concentric circles). Holy water, mundane water, H_2^1O , H_2^2O , H_2^3O , and XYZ are all of the same type from the point of view of the individual. He cannot detect any differences between these things. On the basis of commonsense stereotypes, all of these things are also typed together. Chemistry, however, discriminates two things, H_2O and XYZ, where both the individual and commonsense stereotypes perceive only one. From the point of view of physics, we can recognize three different kinds (H_2^1O , H_2^2O , H_2^3O) where chemistry recognizes only one (H_2O).

Moving outward to the next circle, theology recognizes two

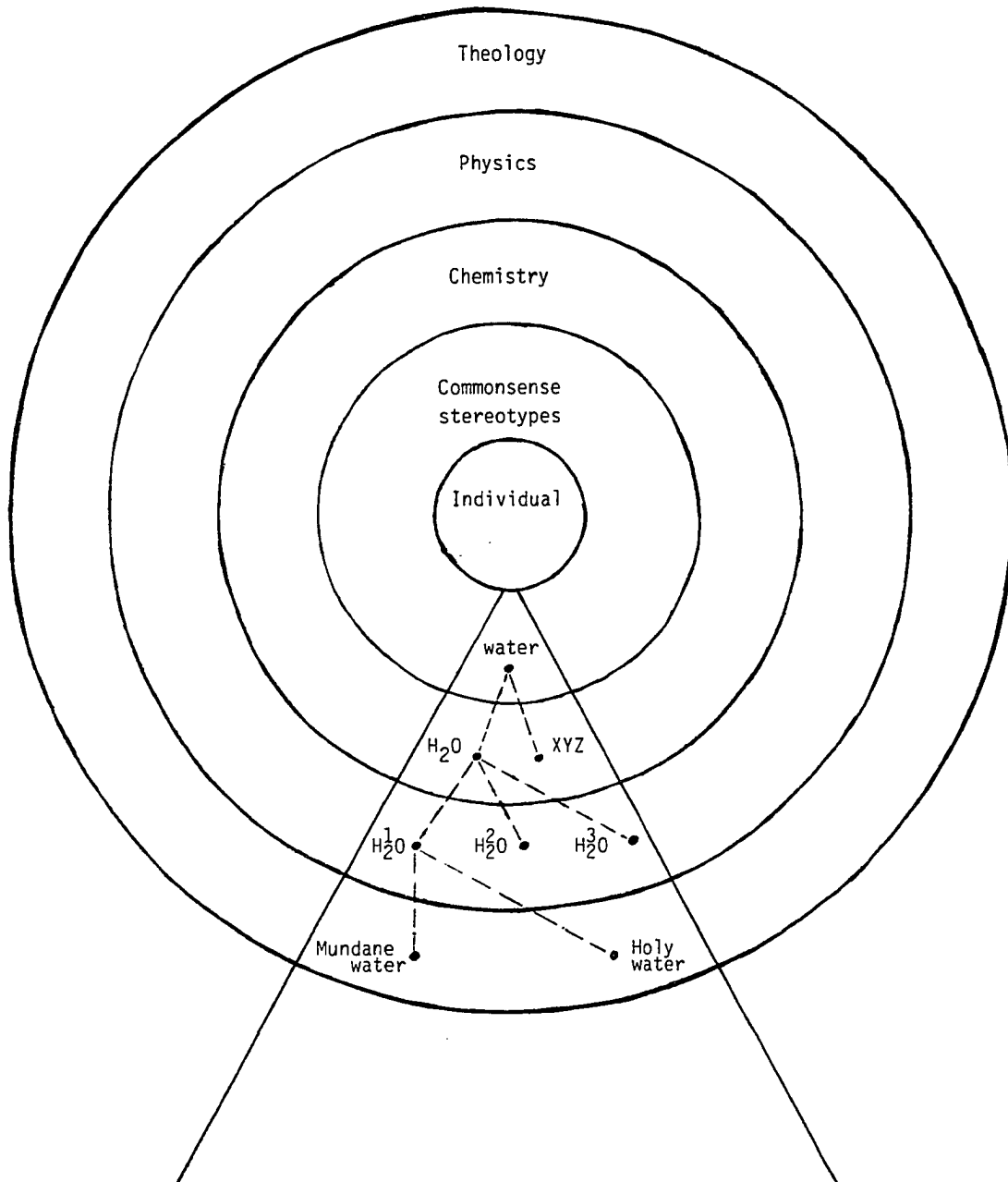


Figure 1: A Schematic View of the Relation Between a Psychological Taxonomy of the World for a Hypothetical Individual and the Taxonomies of the World of Various Other Fields

different kinds (holy water and mundane water) where physics could only detect one. Holy water from the point of view of catholic theology is as different from mundane water as H_2O is from CO_2 . Holy water has different effects than mundane water does. I am not suggesting that the difference between holy water and mundane water is not metaphysically and epistemologically problematic, but simply that taxonomic categories are relative to one's interest. If one is troubled by the theological basis for the difference between holy water and mundane water, one might instead consider cosmology. Cosmology wants to distinguish between "big bang" water and "continuous formation" water, a difference that physics does not recognize. The point is that different viewpoints require different distinctions to be drawn, and yield different taxonomic schemes. The fact that the individual cannot detect differences that physics and chemistry can does not mean they do not exist for physics and chemistry. It just means that they are irrelevant when determining the content of the individual's mental state. In general, if a distinction is not, in fact reflected in the cognitive system of the individual then the distinction does not exist for the purposes of determining the content of the individual's mental state. It is a distinction without a difference, for the psychology of that individual.

Two points about the way Figure 1 is drawn need to be clarified. First, Figure 1 shows the taxonomy of each field neatly embedded inside that of another, commonsense stereotypes within chemistry, chemistry within physics, etc. But the actual situation is much more complex and much less neat than the figure indicates. Not only can the space-time continuum be divided up in any number of different

ways, all of which do not neatly reduce to some other way, but vast portions of it can be simply ignored. How we divide up the world depends on what we are trying to do.

Second, although the individual is located at the center of the figure "looking out" on the world, and I have claimed that the individuation of the world for the purpose of determining content must be done on the basis of what is reflected in the cognitive system of the individual, the question of whether psychology should individuate mental state contents narrowly or widely is not a question about internal versus external individuation. The question is really about what distinctions are relevant for determining content. There are distinctions internal to the individual that I want to claim are irrelevant for the purposes of determining mental state content. Imagine that Twin-Earth is just like Earth in every respect, including having H₂O running in streams and falling from the sky. Now, imagine that the only difference between myself and my doppelganger is that we have relatively inverted spectra. The quale she experiences when gazing at a patch of blue is what I experience when looking at a patch of green, and vice versa. Since the inversion is total and complete, there will be no difference between the way we taxonomize the world. The difference in qualia between us makes no difference in the way the world is correlated with our (identical) causal state-types. All of our beliefs have the same content, irrespective of our inverted spectra, because the distinction in our qualia is not one that can be detected by the mechanism that coordinates properties and state-types. Distinctions of qualia, or differences of qualitative feel, may be more fine grained than can be detected by the mechanism of coordination.

Such distinctions, I claim, are immaterial for the determination of mental state content. The point that should be emphasized here is that the debate over using a wide notion of content or a narrow notion of content is not about whether things external to the individual should be considered when determining content, or only things internal to the individual. The wide-narrow debate is about what is to count as relevant to the individuation of the world for the purposes of psychology. To argue for a wide notion of content in psychology is, in essence, to argue that psychology must adopt the taxonomic scheme of another field. Narrow content is based on an individuation of the world determined by the needs of psychology. Wide content is based on an individuation of the world determined by physics, chemistry, economics, the semantics of language, etc., irrespective of whether or not that individuation facilitates the ability of psychology to construct viable theories.

A scientific psychology that determines the content of a mental state on a case by case basis using a criterion of psychological or cognitive relevance will entail a profusion of mental state contents. But this profusion of different mental state contents is not, in any way, a disadvantage. It is, in fact, of significant advantage to a scientific psychology that wants to be able to predict and explain behavior accurately using generalizations that abstract across the contents of beliefs and desires. One of the goals of a scientific psychology is to construct generalizations roughly of the form

If X believes that not-P unless Q, and wants that P,
then, other things being equal, X will bring about Q,

that can accurately predict the behavior of individuals. The more

accurately we can specify the contents of the subject's beliefs and desires (those things that are substituted for 'P' and 'Q' in the formula), the more accurately we will be able to predict what the subject will do. There is nothing about the sorts of "laws" or generalizations that a scientific psychology will rely upon that militates for a restriction of what mental state contents 'P' and 'Q' can range over. Claiming that the representational contents over which the variables of psychological generalizations can range are limited to those that can be captured using some currently existing taxonomic scheme is like claiming that the speeds at which a car can travel are limited to those that our digital speedometer can represent. If we want to know exactly how fast the car was going, we do not simply claim that it was going exactly 47mph because that is what the speedometer says. If we do, we might find ourselves faced with explaining why one car got to the finish line before a second, even though they were both going "exactly" 47mph according to our digital speedometer. What we do is get a speedometer with higher fidelity. Similarly, if we want to predict more accurately how someone is going to behave we may need to use a more accurate method of specifying the contents of their beliefs and desires than the taxonomies of physics, biology, economics, etc. permit. Individuating mental state content individualistically gives you that higher fidelity.

Any and every difference in the way two individuals classify or taxonomize the world will entail a difference in the contents of those of their mental states that are dependent on the sets that differ will also differ. Every change, for example, in the membership of a stimulus equivalency set will entail a change of the corresponding

mental state content. Identity of mental state content is determined by identity of taxonomic category. Two mental state tokens have the same content if and only if they have the same extensions, where extension is determined by psychological/cognitive equivalence for the particular individual. For example, if you treat broad overstuffed chairs as taxonomically equivalent to sofas, but I do not treat such chairs as taxonomically equivalent to sofas, then the mental state contents that you and I have when we think about such things will be different, even when we are thinking about some particular object that is in both of our sets of taxonomically equivalent objects. Mental state content is determined on a case by case basis by how each individual divides up the world and, consequently, content will be radically individualistic, radically idiosyncratic.

Obviously, if we grant that a scientific psychology must base its attributions of mental state contents on the way the particular individual in question taxonomizes the world, then we are also going to have to accept that commonsense language is inadequate for the specification of mental state content. We do not, and in practice cannot, have enough words to express all the possible mental contents that individuals can have. Mental state contents are much more varied and numerous than we have any need of words for in everyday situations. But the claim that we cannot in English, or any other natural language, express with perfect accuracy what everyone thinks is hardly a startling or revolutionary view. In fact, it would seem to be supported by common intuitions about our perceptions of the world. Certainly, when we are dealing with colors, feels, or scents we are perfectly happy accepting that language is inadequate to

express all of the discriminations that we in fact make. Simply because our language, as it stands, is insufficiently rich to express the differences between two shades of emerald green does not force us to deny that there is a difference in our perception or experience of the two shades. We identify two paint chips as both being royal blue even though we can distinguish between the colors. The fact that we say they are both royal blue does not make them the same exact color. Similarly with scents: we say of two flowers that they both smell like a rose, but we accept that they have somewhat different scents. Why should we not accept that the vocabulary of the natural languages we have are not adequate to express all mental contents. Mental contents are simply more finegrained (or sometimes less finegrained) than language.

If we view mental contents as being strictly defined scientific states, the case for allowing wide variation in how the terms are used in casual conversation is even stronger. Just as in casual conversation we identify any number of samples as being samples of water even though they may vary very greatly in their actual chemical composition, so too, it would seem, that we would identify any number of belief tokens as being the belief that sofas are pieces of furniture even though the actual contents of the belief tokens vary. In geometry there are countless different shapes that can be defined by geometric functions. We do not have words for the vast majority of those shapes, and the words we do have for shapes are used, in common situations, to refer to any number of similar shapes. The word 'circle' is used in common situations to mean anything that is roughly circular. In practice, all we really require is that whatever is identified as a

circle look more like a circle than any of the other common shapes (e.g., square, triangle, rectangle, pentagon, ellipse). The same will hold with respect to content in psychology. What content a particular mental state token has will be specified precisely in terms that psychology defines for this purpose. It may very well be that a scientific psychology's specification of content will be done in terms of functions or formuli. Common language will not have terms for all of the different contents a science of psychology must recognize, but it does have enough terms for the practical purposes of folk psychology. To make the sorts of rough and ready predictions and explanations of behavior that we use folk psychology for we do not need to know the exact content of the individual's mental state, any more than we need to know the exact shape of an object to have a good idea of whether or not it will roll off a table.

The vindication of commonsense belief/desire psychology requires that there be a plausible, scientific explanation of how the causal powers of intentional states contrive to respect the logical form of their semantic contents. It does not require, however, that a scientific psychology accept the individuation of mental state contents that commonsense psychology uses. Nevertheless, we would hope that a rigorous scientific psychology would be able to account for or accommodate many of our everyday folk psychological practices, while having as few unintuitive consequences as possible. A scientific psychology that endorses a fully or radically individualistic individuation of mental state contents is, I believe, the best hope we currently have for the vindication of mental causation and commonsense belief/desire psychology.

A fully individualistic individuation of mental state contents is, as I have argued, not only compatible with but required by a function-cum-information theory account of intentional states such as Dretske has suggested. The sort of theory that Dretske proposes is currently the only theory that provides an account of how content can play a causal role in the explanation of how and why intentional state have causal powers that follow their semantic properties. Content and causal powers co-vary because the indicator properties of a state are relevant to the process that results in intentional states having the causal powers they do, and, so mental causation is accounted for in a scientifically plausible way. Thus, the vindication of commonsense psychology would seem to require a fully individualistic individuation of mental state contents regardless of whether or not such an individuation conflicted with the way we individuated mental state contents using commonsense psychology.

While I have claimed that a scientific psychology should adopt an individualistic individuation of mental state contents regardless of how well it agrees with commonsense psychological practices, I believe that such an individuation is, in fact, better able to account for our common practices and ordinary intuitions than are the notions of Burge and Fodor. A fully individualistic individuation of mental state contents permits a more general psychological theory than can be constructed using the sort of socio-linguistic individuation that Burge advocates. If we adopt Burge's method of individuating mental state content, in which the the linguistic community is the arbiter of the mental states that a linguistically competent member of that community can have, we are forced to claim that our attributions of

mental states to animals and pre-linguistic children do not and cannot mean the same thing as the same attributions made to language using adults. According to Burge's theory attributions of propositional attitudes to animals, etc., are of a fundamentally different nature than are attributions of propositional attitudes to competent users of language. Psychology will have to have a special branch to treat attributions of beliefs to animals. Not only does bifurcating psychology in this fashion clearly run counter to our common practice of attributing beliefs to animals and children, but it also increased the complexity of psychological theory, giving it a somewhat ad hoc appearance. If, however, we adopt a fully individualistic individuation of mental state content we can account for our attributions of beliefs and desires to animals and young children within the same framework that we use for language users.

One of the problems that is often mentioned in connection with the attribution of beliefs and desires to animals is that we have no way of knowing if the way they taxonomize the world is anything like the way we or the semantics of our language does. If, however, we acknowledge that how any given individual taxonomizes the world may vary from how the semantics of our language does, then attributions of beliefs and desires to animals are not different in this respect than are attributions to anyone else. A scientific psychology that individuates mental state contents individualistically will be able to handle animal beliefs in exactly the same fashion as it does the beliefs of linguistically competent individuals. The contents of all intentional states will be determined in the same fashion, and a single psychological theory will be able to deal with attributions of beliefs

and desires to anyone.

Furthermore, by explicitly acknowledging that there can be gaps between the language an individual speaks and the contents of that individual's mental states, we can accommodate perfectly legitimate concerns about the "cognitive fit" between ourselves and those around us, while avoiding the implausible skepticism about our own minds that Burge's theory entails. Wondering whether the belief that you have and express with the content-clause "spatulas are useful in the kitchen" is the same belief that other express with the same content-clause is a straight forward concern about whether you taxonomize the world in the way stipulated by the semantics of the word 'spatula' in your linguistic community. Recognizing that one may not fully understand all of the words one is a 'competent' user of does not lead to the conclusion that one never knows what one believes. Language and thought are certainly intertwined for those that speak a language, but language does not determine what thoughts we can have. An individualistic notion of mental state content allows language to develop from meaningful thought while still remaining distinct from it.

Fodor's shift from the original notion of narrow content in "Methodological Solipsism" to the new narrow content of Psychosemantics, was motivated, in part, by a concern over how to account for two beliefs with the same narrow content having different truth conditions, without severing the connection between content and extension. This problem of reconciling same narrow content with different extensions only exists, however, if one allows psychology autonomy in individuating mental states qua physical states, but denies psychology autonomy in individuating the world. If one grants

psychology autonomy in individuating both mental states and the world, there is no conflict between narrow content and truth conditions. In a fully individualistic, fully autonomous psychology mental states with the same contents will have the same truth conditions. A fully individualistic psychology will view Oscar1's belief that there is water nearby, and Oscar2's belief that there is water nearby as not only having the same narrow content, but as having the same truth conditions as well.

What I want to stress is that when I claim that a fully individualistic psychology will identify Oscar1's belief and Oscar2's belief as having both the same content and the same truth conditions, I do not mean that both XYZ qua XYZ and H₂O qua H₂O will render their beliefs true. Rather it is that H₂O and XYZ are, as far as the psychology of Oscar1 and Oscar2 is concerned, taxonomically the same stuff. The taxonomy of the world according to a psychology specifying the belief contents of Oscar1 and Oscar2 has some stuff, call it 'bwater,' but does not have either H₂O or XYZ. H₂O and XYZ are in the taxonomy of the world for the psychology of the two Oscars, like holy water and mundane water are in the taxonomy of the world for physics. They simply do not exist. In an individualistic psychology, beliefs with the same contents have the same truth conditions because content and extension are determined together.

Finally, it should be noted that a fully individualistic notion of content is compatible with a computational model of the mind. Using the fully individualistic notion of content that I have outlined, indication will be determined by mapping causally individuated internal state-types to those things in the environment with which

they are nomically correlated. Indication and causation will move in tandem. Two beliefs will be content identical if and only if they are functionally identical. However, psychology will still need to couch its generalizations in terms of the semantic properties of propositional attitudes. If psychology were to couch its generalizations in terms of the physical or functional properties of mental states then those could be appropriately applied to states that lacked semantic properties altogether. A "psychology" that couches its generalizations in other than semantic terms would not be psychology, because it would fail to capture the causal role of representational content in the behavior of the system.

We taxonomize the world in a wide variety of ways depending upon what we are doing. Every other field, scientific as well as non-scientific, has its own way of individuating the world. We individuate the world in different ways for different purposes. Commonsense basically divides the world into water, elms, squares, fortnights, and colors. Physics, on the other hand, recognizes an entirely different taxonomy. It divides the world up into things like protons, electrons, positrons, quarks, charm, and strangeness, but certainly does not recognize as taxonomic categories such things as elms, arthritis, fortnights, or squares.

The meaning of words, unlike mental state content, is fundamentally and essentially a product of communal activity. Words come to have the meanings they do by virtue of the fact that they are used by various individuals to say something meaningful. Language is derived from individual practice. The meaning of a word is a consensus, if not actually a compromise, among numerous individual idiolects, and,

at best, is capable of fully and accurately expressing that which is expressed by only a small number of those idiolects. Mental state contents, on the other hand, are not the product of communal activity. Certainly, the activities of the community in which an individual finds himself can influence the mental state of the individual just as anything else in his environment can, but the mental state an individual is in at any given moment is not determined by the activities of some sort of committee. The actual mental state an individual is in is arrived at in a highly individualistic way. It may be indirectly influenced by the community, but it is not any sort of consensus or compromise. If anything, an individual's mental state contents would have to be considered to be analogous to his idiolect, rather than his community language.

What there is in the world, that is, what the ontology of the world is, according to psychology has to be determined by what makes a difference to the individual in ways that are relevant to mental causation. Mental state content and the taxonomy of the world for the purposes of psychology must be tailored to the individual. In "Methodological Solipsism" Fodor claimed that a naturalistic psychology will not be possible until all the other sciences are complete (MS, p. 249), and that, therefore, we will have to make due with a computational psychology. What I hope is, by now, clear is that there is no dichotomy between a computational psychology and a naturalistic psychology. A naturalistic psychology is concerned with those relations between mental states and the world that determine the content of mental states. But since for the purposes of determining mental state content, what there is in the world is a matter for

psychology to decide, a naturalistic psychology is not dependent upon the other sciences. Both sides of the mental state-world relation must be taxonomized by psychology. A computational psychology that individuates mental states on the basis of causal powers will effect a similar individuation of the world. Content will determine extension, and a naturalistic psychology need not wait.

ENDNOTES

Chapter 1

1. See, in particular, "Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology," in Fodor, Representations, Cambridge, Mass.: MIT Press, 1981.
2. Ibid.
3. Fodor, Jerry, Psychosemantics, Cambridge, Mass.: MIT Press, 1987.

Chapter 2

1. This is not strictly true. One of the arguments for machine functionalism is, precisely, that it allows that different types of systems might implement the same psychology, and that different systems will use different symbols (states) to represent the same thing. (Consider the sequence of states of an abacus when it computes $5 + 7$ verses the sequence of states of an electronic pocket calculator when it computes $5 + 7$.) The different system-types will thus be in formally different states even though they represent the same thing. However, Fodor says that in such instances there will be translation tables that lead from the symbols of one system-type to the symbols of the other(s), so that the representational states of the systems will be formally isomorphic with one another, though not formally identical:

[I]t may be empirically possible that there should be creatures that have the same propositional attitudes we do (e.g., the same beliefs) but not the same system of internal representations. . . . Suppose, for example, it turns out that Martians, or porpoises, believe what we do but have a very different cost accounting. We might then want to say that there are translation relations among systems of internal representations (viz., that formally distinct representations can express the same proposition). Presumably which proposition an internal representation expresses -- what content it has -- would be completely determined by its functional role in the organism's mental life, including, especially, the way it is connected to stimulations and responses. Functional identity of internal representations would then be criterial for their intertranslatability.

(PA, p. 202)

However, for the present discussion, this is a fine point that need not be of concern. What is the case is that within a single system-type all states that have the same content must be formally identical, and those states that are formally identical must have the same content.

2. Fodor's claim that the opaque construal of the content clauses of propositional attitude ascriptions pick out or identify those semantic features of mental representations that a scientific psychology can use in its generalizations ultimately requires an account of what the relation between the semantic features of mental representations and the semantic features of content clauses opaquely construed is, such that there is a strict correspondence between the semantic properties of sentences in a natural language and the semantic properties of mental representations that are the objects of propositional attitudes ascribed with those sentences. There must be a one to one correspondence between the semantic properties of a natural language sentence and the semantic properties of the mental representations that we use that sentence to identify. The connection between the opaque construal of a sentence and the semantic properties of mental representations must be such that every individual of whom we can say that he believes that P will necessarily have a mental representation whose "opaque" semantic properties are identical to the opaque construal of "P".
3. It should be noted that there is an ambiguity in common usage that arises when beliefs and desires are specified in this fashion. "Leonora believes (that) the cougar is an endangered species" can be used to indicate either that Leonora has a belief about cougars, that they are an endangered species, or that she believes the proposition "the cougar is an endangered species." The general assumption in common usage is that in the vast majority of propositional attitude ascriptions it will be the case that both interpretations are true for the individual in question. However, it need not be the case that both interpretations are true for any given individual.
4. Putnam, Hilary, "The Meaning of 'Meaning'," in Language, Mind, and Knowledge: Minnesota Studies in the Philosophy of Science, Vol. VII, ed. Keith Gunderson, Minneapolis: Univ. of MN Press, 1975.
5. Fodor makes a distinction between opaque and what he calls 'fully opaque', but the distinction need not concern us here.
6. Although Fodor (see MS, p. 244), as well as others, attribute the phrase "methodological solipsism" to Putnam, it was, in fact, Rudolf Carnap who first coined the phrase in Section 64 of his Der Logische Aufbau der Welt (Berlin, 1928). While there is no doubt that Putnam's usage differs from Carnap's, just as Fodor's differs from Putnam's (terms and phrases once coined do tend to take on lives of their own as they are picked up by others and used in ways that diverge from the original coinage), the phrase does owe its genesis to Carnap, not Putnam.

7. Fodor, Jerry, "Something in the State of the Art," in Representations, Cambridge, Mass.: MIT Press, 1981.
8. Fodor, PWTCCF, p. 6.
9. Fodor, SWOBCD, p. 431.
10. Stich, FFPCS, pp 188-189.

Chapter 3

1. While Burge does at one point state that his objective is merely to demonstrate that some parts of psychology are not individualistic ("In questioning the view that psychology is individualistic, I am not thereby doubting whether there are some sub-parts of psychology that conform to the strictures of individualism. I am doubting whether all of psychology as it is currently practiced is or should be individualistic." (IP, p. 10)), he elsewhere indicates that his objection to individualism applies to all areas of psychology that attribute intentional states ("What I have to say, throughout the paper [IP], will bear on all parts of psychology that attribute intentional states. But I will make special reference to explanation in cognitive psychology." (IP, p. 3)). Since individualism is a view about how to distinguish intentional content kinds, it is unclear what portions of psychology he views as potentially being individualistic but outside the scope of his doubt.
2. Burge, Tyler, "Individualism and Psychology," Philosophical Review, Vol. 95, No. 1 (1986), pp. 3-45.
3. The text actually reads "On the few occasions where, in the actual case, P misperceives shadows as cracks, P is counterfactually confronted with cracks" [emphasis added]. But the example makes no sense if we accept what the text actually says as correct. For P to misperceive shadows as cracks, P would have to have a representation of a crack, but it was stipulated that on the few occasions on which P was confronted with a crack, P represented it as a shadow. P has no representation of a crack. P represents not only shadows as shadows, but all the cracks P has ever encountered as shadows as well. If, however, we replace "P misperceives shadows as cracks" with "P misperceives cracks as shadows" the example is at least consistent and understandable. Consequently, I am assuming that there is an error in the text, and have taken the liberty of inserting the correct phrase in my quotation.
4. Burge, Tyler, "Individualism and the Mental," in Midwest Studies in Philosophy, Vol. IV: Studies in Metaphysics, ed. French, Uehling, and Wettstein, Minneapolis: University of Minnesota Press, 1979.

5. Burge, Tyler, "Other Bodies," in Thought and Object, ed. A. Woodfield, Oxford: Oxford University Press, 1982
6. Burge, Tyler, "Intellectual Norms and the Foundations of Mind," Journal of Philosophy, Vol. 83, No. 12 (1986), pp. 697-720.
7. Burge, Tyler, "Belief and Synonymy," Journal of Philosophy, Vol. 75, No. 3 (1978), pp. 119-138.
8. Fodor, Jerry, "Cognitive Science and the Twin-Earth Problem," Notre Dame Journal of Formal Logic, Vol. 23, No. 2, (1982).
9. Of course, Jones would seem to be very unlikely to say that he denies that verbal contracts bind since he doesn't think there are any. He would simply deny that there was such a thing. However, this point doesn't alter the basic charge against Burge.
10. Burge, Tyler, "Two Thought Experiments Reviewed," Notre Dame Journal of Formal Logic, Vol. 23, No. 3 (1982).
11. Burge, Tyler, "Self-Reference and Translation," in Meaning and Translation, ed. F. Guenther and M. Guenther-Reutter, New York: New York University Press, 1978.
12. Burge, Tyler, "Belief De Re," Journal of Philosophy, Vol. 74, No. 6 (1977).
13. It should be mentioned that I am not alone in wanting a theory of content that can account for our attributions of beliefs to animals in the same fashion that it accounts for our attributions of beliefs to individuals that use language. Fodor, Stich, and Dennett, just to mention three prominent individuals who write in the general area of philosophy of psychology, all provide accounts of belief attributions that treat attributions of beliefs to animals in the same way as they treat attributions of beliefs to people. While there may be some criticism of Burge's theory on these grounds implicit in the positions of Fodor, Stich, and Dennett, none of their discussions of animal beliefs constitutes a particular criticism of Burge's theory.
14. This is not to say that every content clause will necessarily pick out two mental state contents, since there may be some content clauses that we might be unwilling to use to specify a non-linguist's belief, e.g., 'that 37 is a prime number,' 'that electrons have a negative charge', 'that 'fortnight' begins with the sixth letter of the alphabet', etc. However, this fact does not affect the basic point that psychology will have to have two notions of content given Burge's view of content.
15. It should be noted that my argument against Burge's theory of content on the grounds that it requires psychology to construct an additional notion of content for dealing with non-linguists, is independent of the arguments that Bilgrami and Loar have made about

his theory requiring two notions of content.^{15a} Bilgrami's argument is that Burge's notion of content requires that we construct a second notion of content because his, Burge's, cannot satisfactorily fill the explanatory role of content in psychology because "it yields individuations of content too course-grained to always capture the inferences of failures of inference in agents to whom they are attributed" (EAPC, p. 195).^{15b} The charge is made that if we attribute content on the basis of the literal interpretation of content clauses, then we cannot use content to explain the behavior of individuals. The argument runs roughly as follows: Suppose that Simon believes that a fortnight is ten days, and upon learning that Max will arrive in ten days, he says to himself "Max will arrive in precisely a fortnight. I must remember to pick him up at the airport." If we attribute to Simon the literal belief that Max will arrive in a fortnight, then that belief alone cannot account for why Simon goes to the airport ten days later expecting to pick up Max. Consequently, we need another notion of content that will allow us to explain why Simon went to the airport when he did. While it is certainly the case that if we attribute the literal belief that Max will arrive in a fortnight, further explanation will be necessary to account for why Simon went to the airport when he did. However that fact, in and of itself, cannot be counted against Burge's theory, because if we use a notion of content that does explain why he arrived at the airport when he did we will not have accounted for why he said what he did, i.e., "Max will arrive in a fortnight". If we attribute a belief to Simon on the basis of such a notion of content, that is, one that explains why he went to the airport when he did, we will require further explanation in order to account for why he said what he did. Given Burge's notion of content, in order to explain why Simon went to the airport ten days later even though he thinks that Max will arrive in a fortnight, we need to add that Simon thinks that a fortnight is ten days. On the other hand, if we deny that Simon believes that Max will arrive in a fortnight, holding that he believes only that Max will arrive in ten days, we need to attribute to Simon the belief that a fortnight is ten days (or at least the belief that 'fortnight' means ten days, although this does not seem to be what Simon has said he believes), in order to account for why he said that Max will arrive in a fortnight.

When there is a mismatch between what a word means and what some individual thinks the word means, no attribution of any single belief, regardless of whether the content is based on a literal interpretation of what the individual says or on what the individual thinks what she says means, will be adequate to explain all of the relevant behavior of the individual, both verbal and non-verbal. When there is no mismatch, that is, when what an individual thinks a word means and what it actually means are the same, we can predict both the verbal and the non-verbal behavior of the individual equally well regardless of whether we attribute a belief on the basis of what she says or what she thinks she has said. If Leonora believes that a fortnight is fourteen days and Max tells her that he will be arriving in two weeks, we can, by attributing to Leonora the belief that Max will arrive in a fortnight, explain why she said

"Max will arrive in a fortnight", and why she went to the airport when she did. The fact that Burge's method of attributing belief content cannot explain all of an individual's behavior when there is a linguistic error involved, is not unique to Burge's theory. Any way of specifying content will find itself unable to account for some aspect of the individual's behavior, verbal or non-verbal, when the individual has made a linguistic error. Burge's account of content fails in this sort of situation to no greater or lesser extent than alternate accounts of content. Burge's theory simply views verbal behavior as the primary focus of psychological explanation rather than non-verbal behavior when dealing with members of a linguistic community.

- 15a. See Akeel Bilgrami, "An Externalist Account of Psychological Content," *Philosophical Topics*, Vol. XV, No. 1, (1987), pp. 191-226; and Brian Loar, "Social Content and Psychological Content," in R. Grimm and D. Merrill (eds.) Contents of Thoughts, University of Arizona Press, 1988, pp. 99-110, 121-139.
- 15b. It is interesting to note that Burge makes a very similar comment in "Belief and Synonymy," only in support of his position. He claims that one must interpret content clauses literally, permitting no substitution of synonyms because "[t]he denotations of that-clauses in de dicto propositional attitudes are, from a purely logical point of view, more fine-grained than ordinary linguistic meaning" (BS, p. 136).
16. Burge may want to claim that people defer on the content of their beliefs, not merely the truth or falsity of them, but their behavior and ours often runs counter to this view. For example, if Rachel says that she believes that Milton made slanderous remarks about George, and is later told that he didn't (because slanderous remarks are false utterances that defame or damage someone's reputation), she can quite easily deny that she ever believed that Milton made slanderous remarks, only that he made false and malicious --though not defaming -- remarks about George. When Rachel says this, we do not necessarily hold her to having had the belief that Milton slandered George, claiming that she simply did not know that she had that belief. We take her remarks to be simply clarifying what, in fact, she did believe. Rachel has not deferred to the linguistic community and admitted that she was wrong in what she believed. She has claimed that we have misunderstood her views, not because of any failing on our part, but because she described her belief incorrectly.

Chapter 4

1. Armstrong, D. M. Belief, Truth, and Knowledge. Cambridge: Cambridge University Press, 1973.
2. Ramsey, F. P. The Foundations of Mathematics, and Other Logical Essays. London: Routledge and Kegan Paul, 1931.

3. Communication theory measures information using a base 2 logarithmic function, but it could, of course, be measured using various other functions. A base 2 logarithmic function is simply the most convenient for the uses of communication theory. Since, as we will see, Dretske will not actually be making use of the amounts of information themselves, the question of whether or not a base 2 logarithmic function is appropriate for a semantic theory of information is immaterial.
4. In addition to equivocation, which is the extent to which the source can vary without having a corresponding change in the receiver, there is another type of independence between source and receiver called 'noise.' Noise is the extent to which events at the receiver can vary without there being a corresponding change in the source. Together equivocation and noise are the degree of independence of events occurring at the receiver from those occurring at the source.
5. What Dretske actually says is

"Informational content: A signal r carries the information that s is F = The conditional probability of s 's being F , given r (and k) is 1 (but, given k alone, less than 1)."

However, his use of the phrase "the information that s is F " here, as well as through out KFI, is potentially misleading. The signal does not carry the information that s is F . It only carries the information F or is F , about s . While Dretske is careful to mention at one point that he is talking about de re content, where the object is not identified by the content, his continual references to the information that s is F is unnecessarily confusion. Therefore, I have, whenever it has seemed necessary, and could be done without doing violence to Dretske's meaning, rephrased what Dretske says to remove this possible source of confusion.
6. Although Dretske claims that relativizing the information a signal carries to what the recipient already knows "accurately reflects our ordinary way of thinking about such matters" (KFI, p. 79), it seems more natural to say that I come to know your exact address not from being told that you live at 44 Horatio Street, but rather from the combination of this new piece of information with one I already have received, i.e., that you live in New York City. The information conveyed to me when I am told that you live at 44 Horatio Street is the same regardless of what else I know about your address. I simply can do more with it in the one case than I can in the other.

Dretske's way of looking at it seems to result in an unintuitive overlap of information content. If I know nothing about where you live and am told that you live at 44 Horatio Street (r_1), the information I receive is simply that you live at 44 Horatio Street. On the other hand, if I had been told that you live in New York City (r_2), the information I would have received would have been simply that you live in New York City. However, if I had already received r_2 , r_1 would carry not only the information that you live at 44 Horatio Street, but also your exact address. Similarly, if I had already received r_1 , r_2 would carry not only the information that you live in New York City, but also your exact address. Thus, once

I have received both r1 and r2, r1 and r2 carry some of the same information, i.e., your exact address. Dretske has chosen to define the information content of a signal in a way that seems unnatural (to my mind), and that unnecessarily conflates the content of different signals. Nevertheless, whether or not one takes into account what the recipient already knows when determining the information content of a signal does not have a significant impact on the general plausibility of Dretske's theory.

7. I assume that the context conditions of the voltmeter are things like being surrounded by air rather than, say, water or molten tin, and being in a relatively normal - at least for Earth - magnetic field. Since the effect of changing the context conditions of the voltmeter is less obvious, my contrast case will not use the voltmeter, though such a contrast case could certainly be constructed.
8. This is not to say that (cognitive) mental phenomena, e.g., beliefs and desires, are nothing more than information carrying states, but merely that they possess intentionality in virtue of being derived from signals. Dretske says that in addition to having intentionality, a structure must also have an executive or functional role in determining the behavior of the system of which it is a part if it is to qualify as a (cognitive) mental phenomenon. However, Dretske views this executive or functional character of beliefs as being distinct from, although related to, their intentional character (KFI, p. 197).
9. Even though, according to Dretske, all signals are intentional states, I will continue to use "intentional states" to refer to beliefs, desires, hopes, etc. -- what I have called "cognitive mental states."
10. Dretske achieves this task in a different way in Knowledge and the Flow of Information than in Explaining Behavior, and, as the two are incompatible, I will discuss only the latter view. For those who are familiar with KFI, in EB the notion of the function of a signal within a system replaces both the notions of primary representation and of digitalization in determining what one piece of information a signal carries that is to be considered its content. It should be noted that the notion of primary representation that Dretske used in KFI is inadequate for his purposes. The notion of primary representation would make the fuel gauge of a car, at least an electrically operated one, an indicator of something like the torque on the pointer armature.
11. Natural signs are signals, and the Gricean "natural" meaning of such signs is the information they carry.
12. Lewontin, Richard, "Darwin's Revolution," New York Review of Books, Vol. 3 (1983), pp. 21-27.

13. Sober, Elliot, The Nature of Selection, (Cambridge, Mass.: MIT Press, A Bradford Book, 1984), pp. 147-155.
14. Cummins, Robert, "Functional Analysis," Journal of Philosophy, Vol. 72, No. 20 (1975), pp. 741-765.
15. Ruth Garrett Millikan has suggested a way of distinguish between a moth's representation of an approaching bat and my representation of a person at the door that is at least intuitively appealing. She suggests that moth representations differ from my representations in that the moth's are imperative in character, while mine are indicative in character. An activated 'bat approaching' representation in a moth has the force of an imperative. The moth will invariably and inevitably, assuming it is physically able, execute a series of evasive flight maneuvers. It cannot "choose" to do otherwise. On the other hand, an activated 'person at the door' representation in me has no such imperative force. It only indicates that something is the case. As a result of having that representation activated I may go to the door, but I may not. Knowing only that a 'person at the door' representation has been activated in me, one can draw no definitive conclusions about what I will do.
16. Dretske seems to be making a distinction here between what someone means and what concept he has, but since he does not provide an elaboration on this point, it is unclear what force this caveat is supposed to have.
17. The analog/digital distinction that Dretske makes so much of in KFI is essentially a distinction between the information available in a signal and the information a receiver of the signal extracts from that signal and uses (See Chapter 6, particularly pp. 135-153).

Chapter 5

1. See, in particular Ch.1 of Language, Thought, and Other Biological Categories, (Cambridge, Mass.: Mit Press, 1984) and "Biosemantics," Journal of Philosophy, Vol. 86, No. 6 (1989).
2. By claiming that we need to know what there is in the world before we can determine how to individuate mental content, I am not endorsing the view that psychology will have to wait until all the nonpsychological sciences, e.g., physics and chemistry, are complete. In "Methodological Solipsism" Fodor argues that the distinction between wanting a naturalistic psychology, a psychology of organism/environment interactions, and not wanting any psychology at all, although real, is academic (MS, pp. 244-251). A naturalistic psychology will not be possible until physics, chemistry, and the like, have determined what there is in the environment.

Generally, then, a naturalistic psychology would attempt to specify environmental objects in a vocabulary such that environment/organism relations are law-instantiating when so described. But here's the depressing consequence again: we have no access to such a vocabulary prior to the elaboration (completion ?) of the nonpsychological sciences. "What Granny likes with her herring" isn't, for example, a description under which salt is law-instantiating; nor, presumably, is "salt." What we need is something like "NaCl," and descriptions like "NaCl" are available only after we've done our chemistry.

(MS, p. 249)

But contrary to Fodor's claim, psychology, even a psychology of organism/environment transactions will not have to wait until all the other sciences are complete.

3. Undoubtedly, our willingness to attribute intentional states to non-users on language, such as dogs, dolphins, and young children, is based, in part, on the anthropomorphic assumption that they are basically like us. But by assuming that they are basically like us we are hardly committing ourselves to the position that they have a semantics like ours.
4. Note, however, that narrow content is not, under this interpretation, a semantic property, whereas in "Methodological Solipsism" narrow content was a semantic property.
5. Tinbergen, N., "The Curious Behavior of the Stickleback," Scientific American, Vol 187, No. 6 (1952), pp. 22-26.
6. Accepting that the content of the stickleback's "belief" is "thing with a red underside" does not mean that we have to try to make sense of the behavior as behavior towards a thing with a red underside. We can still understand the stickleback's behavior as behavior towards a conspecific, because it is the fact that things with a red underside are, in the local environment, conspecifics that has "rewarded" (via natural selection) the coordination between things with a red underside and the typical behavior we witness today.

BIBLIOGRAPHY

The letters in parentheses immediately following an author's name in a entry are the abbreviation used to refer to that particular work when documentation (for quotations, etc.) is given parenthetically. Abbreviations have been assigned only to those works that are cited in this fashion.

- Armstrong, D. M. Belief, Truth, and Knowledge. Cambridge: Cambridge University Press, 1973.
- Bach, Kent. "Burge's New Thought Experiment: Back to the Drawing Room." Journal of Philosophy, Vol. 85 (1988), pp. 88-97.
- Bach, Kent. "De Re Belief and Methodological Solipsism." In Thought and Object, ed. A. Woodfield. Oxford: Oxford University Press, 1982.
- Bach, Kent. "Thought and Object: De Re Representations an Relations." In Representation of Knowledge and Belief, ed. M. Brand and R. Harnish. Tuscon: University of Arizona Press, 1986.
- Bilgrami, Akeel. (EAPC) "An Externalist Account of Psychological Content." Philosophical Topics, Vol. 15, No. 1, Spring 1987.
- Boden, Margaret. "What is Computational Psychology?" Proceedings of the Aristotelian Society, Vol. 58, Supplement (1984), pp. 17-53.
- Burge, Tyler. "Belief De Re." Journal of Philosophy, Vol. 74, No. 6 (June 1977), pp. 338-362.
- Burge, Tyler. (BS) "Belief and Synonymy." Journal of Philosophy, Vol. 75, No. 3 (March 1978), pp. 119-138.
- Burge, Tyler. "Cartesian Error and the Objectivity of Perception." In Contents of Thoughts: Proceedings of the 1985 Oberlin Colloquium in Philosophy. Tuscon: University of Arizona Press, 1987.
- Burge, Tyler. (IM) "Individualism and the Mental." In Midwest Studies in Philosophy, Vol.IV: Studies in Metaphysics, ed. French, Uehling, and Wettstein. Minneapolis: University of Minnesota Press, 1979, pp. 73-121.

- Burge, Tyler. (IP) "Individualism and Psychology." Philosophical Review, Vol. 95, No. 1 (1986), pp. 3-45.
- Burge, Tyler. (IN) "Intellectual Norms and the Foundations of Mind." Journal of Philosophy, Vol., 83, No. 12, December 1986, pp. 697-720.
- Burge, Tyler. (OB) "Other Bodies." In Thought and Object, ed. A. Woodfield. Oxford: Clarendon Press, 1982.
- Burge, Tyler. "Self-Reference and Translation." In Meaning and Translation, ed. F. Guenther and M. Guenther-Reutter. New York: New York University Press, 1978.
- Burge, Tyler. (TTER) "Two Thought Experiments Reviewed." Notre Dame Journal of Formal Logic, Vol. 23, No. 3, July 1982.
- Cummins, Robert. "Functional Analysis." Journal of Philosophy, Vol. 72, No. 20 (1975), pp. 741-765.
- Cummins, Robert. "Inexplicit Information." In Representation of Knowledge and Belief, ed. M. Brand and R. Harnish. Tuscon: University of Arizona Press, 1986.
- Cummins, Robert. The Nature of Psychological Explanations. Cambridge, Mass.: MIT Press, A Bradford Book, 1983.
- Cummins, Robert. "Social Content and Psychological Content." In Contents of Thoughts: Proceedings of the 1985 Oberlin Colloquium in Philosophy. Tuscon: University of Arizona Press, 1987.
- Cummins, Robert. "Why Adding Machines Are Better Examples than Thermostats: Comments of Dretske's 'The Explanatory Role of Content' ." In Contents of Thoughts: Proceedings of the 1985 Oberlin Colloquium in Philosophy. Tuscon: University of Arizona Press, 1987.
-
- Dennett, Daniel. "Beyond Belief." In Thought and Object, ed. A. Woodfield. Oxford: Oxford University Press, 1982.
- Dennett, Daniel. "Intentional Systems." Journal of Philosophy, Vol. 68 (1971), pp. 87-106.
- Dennett, Daniel. "The Logical Geography of Computational Approaches: A View From the East Pole." In The Representation of Knowledge and Belief, Ed. Myles Brand and Robert M. Harnish. Tuscon: University of Arizona Press, 1986.
- Dretske, Fred. (ACR) "Aspects of Cognitive Representation." In The Representation of Knowledge and Belief. Ed. Myles Brand and Robert M. Harnish. Tuscon: University of Arizona Press, 1986.
- Dretske, Fred. "Contrastive Statements." Philosophical Review, Vol. 81 (1972), pp.411-437.

- Dretske, Fred. (EoB) "The Epistemology of Belief." Synthese, Vol. 55 (Apr. 1983), pp. 3-20.
- Dretske, Fred. (EB) Explaining Behavior: Reasons in a World of Causes. Cambridge, Mass: MIT Press, A Bradford Book, 1988.
- Dretske, Fred. "The Explanatory Role of Content." In Contents of Thoughts: Proceedings of the 1985 Oberlin Colloquium in Philosophy. Tuscon: University of Arizona Press, 1987.
- Dretske, Fred. (KFI) Knowledge and the Flow of Information. Cambridge, Mass: MIT Press, A Bradford Book, 1981.
- Dretske, Fred. "Misrepresentation." In Belief. Ed. R. Bogdan. Oxford: Oxford University Press, 1986.
- Dretske, Fred. "Precis of Knowledge and the Flow of Information." Behavior and Brain Sciences, Vol. 6, No. 1 (1983), pp. 55-63.
- Enc, Berent. "Intentional States of Mechanical Devices." Mind, Vol. 91 (1982), pp. 161-182.
- Fodor, Jerry (CSTP) "Cognitive Science and the Twin-Earth Problem." Notre Dame Journal of Formal Logic, Vol. 23, No. 2, April 1982.
- Fodor, Jerry. "Fodor's Guide to Mental Representations." Mind, Vol. 94 (1984), pp. 76-100.
- Fodor, Jerry. "Information and Association." In Representation of Knowledge and Belief, ed. M. Brand and R. Harnish. Tuscon: University of Arizona Press, 1986.
- Fodor, Jerry. (LOT) Language of Thought. Cambridge, Mass.: Harvard University Press, 1975.
- Fodor, Jerry. (MS) "Methodological Solipsism Considered as a Research Strategy in Cognitive Science." Behavioral and Brain Sciences, 1980. (Reprinted in Fodor, J., Representations, q.v., page references are to that volume.)
- Fodor, Jerry. (PA) "Propositional Attitudes." Monist, Vol. 61, No. 4 (1978). (Reprinted in Fodor, J., Representations, q.v., page references are to that volume.)
- Fodor, Jerry. (Psycho) Psychosemantics. Cambridge, Mass.: MIT Press, A Bradford Book, 1987.
- Fodor, Jerry. (PWTCCF) "Psychosemantics or: Where Do Truth Conditions Come From." (a circulated mimeo quoted by Stich).
- Fodor, Jerry. Representations. Cambridge, Mass.: MIT Press, 1981.

- Fodor, Jerry. (SWOBCD) "Searle On What Only Brains Can Do." Behavioral and Brain Sciences, Vol. 3 (1980).
- Fodor, Jerry. "Semantics, Wisconsin Style." Synthese, Vol. 59 (1984), pp. 231-250.
- Fodor, Jerry. (SSA) "Something on the State of the Art." In Fodor, J. Representations, q.v.
- Fodor, Jerry. (TCPA) "Three Cheers for Propositional Attitudes." In Fodor, J. Representations, q.v.
- Fodor, Jerry. (TSPG) "Tom Swift and his Procedural Grandmother." Cognition, 1978. (Reprinted in Fodor, J., Representations, q.v., page references are to that volume.)
- Grice, Paul. "Meaning." Philosophical Review, Vol. 66 (1957), pp. 377-388.
- Harman, Gilbert. "Wide Functionalism." In Representation of Knowledge and Belief, ed. M. Brand and R. Harnish. Tuscon: University of Arizona Press, 1986.
- Haugeland, John. "Semantic Engines: An Introduction to Mind Design." In Mind Design, ed. J. Haugeland. Cambridge, Mass.: MIT Press, A Bradford Book, 1981
- Kripke, Saul. Naming and Necessity. Cambridge, Mass.: Harvard University Press, 1980
- Lewontin, Richard. "Darwin's Revolution." New York Review of Books, Vol. 3 (1983), pp. 21-27.
- Loar, Brian. Mind and Meaning. Cambridge: Cambridge University Press, 1981.
- Loar, Brian. "Reference and Propositional Attitudes." Philosophical Review, Vol. 81, No. 1 (January 1972), pp. 43- 62.
- Loar, Brian. "Subjective Intentionality." Philosophical Topics, Vol. 15, No. 1 (Spring 1987), pp. 89-124.
- Loewer, Barry. "From Information to Intentionality." Synthese, Vol. 70 (1987), pp. 287-317.
- Marr, David. Vision. New York: W. H. Freeman and Co., 1983
- McGinn, Colin. The Character of Mind. Oxford: Oxford University Press, 1982.
- McGinn, Colin. "The Structure of Content." In Thought and Object, ed. A. Woodfield. Oxford: Oxford University Press, 1982.

- Millikan, Ruth Garrett. "Biosemantics." Journal of Philosophy, Vol. 86 (June 1989), pp. 281-297.
- Millikan, Ruth Garrett. "In Defense of Proper Functions." Philosophy of Science, Vol. 56 (June 1989), pp. 288-302.
- Millikan, Ruth Garrett. Language, Thought, and Other Biological Categories: New Foundations for Realism. Cambridge, Mass: MIT Press, A Bradford Book, 1984.
- Millikan, Ruth Garrett. "Naturalist Reflections on Knowledge." Pacific Philosophical Quarterly, Vol. 65 (Oct. 1984), pp. 315-334.
- Millikan, Ruth Garrett. "Thoughts Without Laws: Cognitive Science with Content." Philosophical Review, Vol. 95 (Jan. 1986), pp. 47-80.
- Millikan, Ruth Garrett. "Truth Rules, Hoverflies, and the Kripke-Wittgenstein Paradigm." Philosophical Review, Vol. 99 (July 1990), pp. 323-353.
- Putnam, Hilary. "Computational Psychology and Interpretation Theory." In Philosophical Papers Vol. III: Realism and Reason. Cambridge: Cambridge University Press, 1983.
- Putnam, Hilary. "The Meaning of 'Meaning'." In Language, Mind and Knowledge: Minnesota Studies in the Philosophy of Science, Vol. VII, ed. Keith Gunderson. Minneapolis: University of Minnesota Press, 1975.
- Putnam, Hilary. "Minds and Machines." In Dimensions of Mind, ed. S. Hook. New York: New York University Press, 1960.
- Pylshyn, Zenon. Computation and Cognition. Cambridge, Mass.: MIT Press, A Bradford Book, 1984
- Quine, W. V. O. "Propositional Objects." In Ontological Relativity and Other Essays. New York: Columbia University Press, 1969.
- Quine, W. V. O. "Quantifiers and Propositional Attitudes." Journal of Philosophy, Vol. 53, No. 5, (March 1956), pp. 177-187.
- Quine, W. V. O. Word and Object. Cambridge, Mass.: MIT Press, 1960.
- Ramsey, F. P. The Foundations of Mathematics, and Other Logical Essays. London: Routledge and Kegan Paul, 1931.
- Ryle, Gilbert. The Concept of Mind. London: Hutchinson, 1949 (Reprinted, Chicago: University of Chicago Press, 1984).
- Searle, John. "Minds, Brains and Programs." Behavioral and Brain Sciences, Vol. 3, No. 3 (1980), pp. 417-457.
- Searle, John. "What is an Intentional State?" Mind, Vol. 88 (1979), pp. 74-92.

- Sober, Elliot. The Nature of Selection. Cambridge, Mass.: MIT Press, A Bradford Book, 1984.
- Sosa, Ernest. "Propositional Attitudes De Dicto and De Re." Journal of Philosophy, Vol. 67, pp. 883-896.
- Sterelny, Kim. "Fodor's Nativism." Philosophical Studies, Vol. 55 (1989), pp. 119-141.
- Stich, Stephen. (FFPCS) From Folk Psychology to Cognitive Science: the Case Against Belief. Cambridge, Mass.: MIT Press, A Bradford Book, 1983.
- Stich, Stephen. "On the Ascription of Content." In Thought and Object, ed. A. Woodfield. Oxford: Oxford University Press, 1982.
- Tinbergen, N. "The Curious Behavior of the Stickleback." Scientific American, Vol. 187, No. 6. (1952), pp. 22-26.
- Woodfield, Andrew. "On Specifying the Contents of Thoughts." In Thought and Object, ed. A. Woodfield. Oxford: Oxford University Press, 1982.
- Woodfield, Andrew (ed.). Thought and Object, ed. A. Woodfield. Oxford: Oxford University Press, 1982.