ABSTRACT

A propositional attitude (PA) is a belief, desire, fear, etc., that x is the case. This dissertation addresses the question of the semantic content of a specific kind of PA-instance: an instance of a b*elief* of the form *all Fs are Gs.* The belief that *all bachelors are sports fans* has this form, while the belief that *Spain is a country in Eastern Europe* do not. Unlike a state of viewing the color of an orange, a belief-instance is *semantically* contentful because it has reference, a meaning, logical implications, or a truth-value. While the intrinsic semantics view holds that either concepts or abstract objects are the source of content for PAs, the extrinsic semantics view holds that symbols of a mental language provide this content.

I argue that a successful theory of intentionality must explain: (1) the truth-preserving causal powers of PAs, (2) the failure of the deductive principle Substitutivity to preserve truth over sentences that ascribe PAs, and (3) the truth-evaluability of PAs.

As an internalist version of the extrinsic semantics view, I first evaluate Fodor’s Computational Theory of Mind, which says the semantic ingredients of mental states are symbols governed by rules of a mental syntax. I argue that in order to meet (1), CTM would have to associate the causal patterns of each thought-type with the inferential relations of some proposition – in an arbitrary or question-begging way. I also evaluate Fodor’s causal theory, as an externalist version of the extrinsic semantics view. This view is that lawlike causal relations between mental symbols and objects determine the reference, and thus the truth-value, of a thought. I argue that Brian Loar’s circularity objection refutes the ability of this theory to meet (3); and I endorse the intrinsic semantics perspective.

I evaluate Frege’s theory of abstract, mind-external, and intrinsically semantic objects (senses), as an attempt to meet condition (2). I conclude that mind-external universals are the source of the intrinsically semantic features of concepts. Finally, I put forth a theory called ‘Bare Property Intentionality’, which describes the features of intrinsically representative and semantic concepts that connect them to these universals.

Chapter I: The Semantic Content of Propositional Attitudes

A. The Common Sense and Philosophical Notions of Intentionality

According to common sense and contemporary psychological science, people have mental states that are about things, things that are external to those states. Another popular way of describing this phenomenon is to say that the minds of people can represent other things. For instance, an idea of the plant on my dining room table can occur within my mind and the mental state I am in while it is occurring represents, or *is about,* that particular plant. I can also think about at least the following kinds of entities: non-particulars, such as the class of plants in general; non-natural entities, such as the state of Hungary; abstract entities, such as derivative rules of calculus; events or processes, such as the Vietnam War; characteristics, such as integrity; and states, such as economic recession.

Even more interesting aspects of the mind’s representative ability are exemplified by the fact that I can have thoughts whose contents are not true, such as the belief that *Henri Matisse drew sketches for decorating the ceilings of Eastern Orthodox chapels in Russia*. Finally, the most fascinating aspect of mental representation seems to be its self-referential potential, that is, the fact that people with minds can think about themselves: even about properties of their *own* minds. I can even believe, for example, that *I am currently thinking about the mind’s ability to represent other things,* or falsely imagine that *I am currently concentrating upon my performance in a cross-country athletic competition.*

Contemporary analytical philosophers use the term ‘intentionality’ to refer to the property that many mental states have in virtue of which they represent things and have “meanings”.[[1]](#footnote-1) Throughout history, many theories have attempted to explain the feature whereby a state of the mind can have a meaning or relate to something else in a way such that the mental state is about that thing. Yet, such views have faced much difficulty in explaining and providing a metaphysical account of the essential constituents of representative mental states: what it is that connects these thoughts semantically, or referentially, to objects in the world.[[2]](#footnote-2)

This essay concerns the metaphysical features of the semantic constituents of the intentional states that are often termed 'propositional attitudes' (PAs). (Actually, this discussion only addresses the constituents of instances of PAs, which are particular, concrete occurrences of a propositional attitude within a person’s mind.) PAs are beliefs, desires, fears, etc., that *x is (or be) the case*, where x is a proposition expressible, in principle, in sentential form. The belief that *there is a plant on Sharon's kitchen table* is a propositional attitude, for example, while the state of remembering a visual experience of this plant is not. The desire that *there be a plant on Sharon's kitchen table* is a PA, while a state of emotion, e.g., an experience of aesthetic pleasure that is associated with seeing the plant on the table, is not. Neither, for instance, is a merely sensory state of perceiving the plant (as it stands on the table). While intentional states are all and only those types of mental states that represent things ⎯ or purport to represent things[[3]](#footnote-3) ⎯ in an essentially *semantic* way, propositional attitudes are the category of intentional states that one can evaluate as having logical implications and truth-conditions.

The specific questions that this essay will aim to answer are: What type of semantic content do these mental ingredients have? What types of internal features or mind-external phenomena provide the constituents of instances of PAs with semantic content?[[4]](#footnote-4) To accomplish this philosophical goal, later chapters will evaluate various versions of two general competing theories of intentionality. On the one hand is the view that the constituents of intentional states are *intrinsically* semantic, i.e., that these constituents essentially have meanings, logical implications, referential features, and truth-evaluability.[[5]](#footnote-5) On the other hand is the view that these constituents are *extrinsically* semantic, i.e., that they have meanings, logical implications, referential features, and truth-evaluability only in virtue of their contingent relations with other entities. The final chapter will endorse a version of the former view, and it will provide a metaphysical characterization of the intrinsically semantic constituents of instances of PAs involving universal propositions, such as an occurrence of a belief that *all bachelors are sports fans*.

To begin, this introductory chapter will explain the intuitions that motivate the intrinsic semantics view of intentionality, by exploring their foundation in a common sense view of what it is for a person’s thoughts to have meaning and for his mind to represent something. As I will explain below, the current discussion interprets the (more or less) common idea of how thoughts represent things to mean that concrete concepts are the semantic constituents of such thoughts.Philosophical analysis of the commonsensical identification of concepts with “meanings”[[6]](#footnote-6) and “ideas” reveals that these concepts are intrinsically semantic entities.[[7]](#footnote-7)

This conceptual view of mental representation stands in opposition to the extrinsic semantics view. For in contrast to the grounding of the conceptual account in commonly held intuitions, the extrinsic semantics theory arises from a scientific and empirically based philosophical perspective. According to this perspective, mental processes are linguistic (or computational) processes, the contents of thoughts are pieces of a mental language system, and, accordingly, the elements of such thoughts are symbols. These mental symbols are, of course, extrinsically semantic entities. That is, the symbols only mean or represent what they do in virtue of their non-essential relationships with the syntactic system to which they belong or objects in the world outside the mind.[[8]](#footnote-8)

As mentioned above, after several chapters that explore and reject different versions of the symbolic theory of thought, later chapters will discuss versions of the intrinsic semantics view. Ultimately, moreover, this essay will endorse a version of the intrinsic semantics view, according to which mental states represent things in virtue of the fact that they are constituted by concrete, intrinsically semantic concepts. Moreover, I will argue that both intuitions and metaphysical considerations support the view that these concepts essentially present abstract and mind-independent objects, namely, properties, to the mind. According to this view, which I will call ‘Bare Property Intentionality’, these concepts’ presentations of properties (and instantiations of properties) to the mind provide instances of propositional attitudes of the form *all Fs are Gs* with intrinsically semantic content.[[9]](#footnote-9) One of the metaphysical conclusions of the last chapter will be that each occurrent intrinsically representative concept must include a specific type of constituent, namely, a “bare referential intention,” in virtue of which the concept directs the mind in which it occurs toward something outside the state to which the it belongs. (This claim will apply to all intrinsically *representative* concepts, independently of the type of PA-instance to which they belong.)

B. Contrasting Metaphysical Assumptions on Mental Representation

Before explaining the contrast between the intrinsic and extrinsic semantics views of intentionality, this section will first contrast the metaphysical assumptions of the common sense view of mental representation with those of the symbolic theory. To begin explaining these underlying assumptions, I will first discuss the notion of a semantic *feature*, and then I will contrast the general notion of a concept with that of a symbol as a semantic constituent of a mental state.

In the context of mental states, the term ‘semantics’ can express four general connotations, either alone or in any combination: (i) the relation whereby the mental state refers to something, (ii) the truth-value of the state, (iii) the cognitive content or “meaning” of the state, and (iv) the propositional content of the state, i.e., the proposition that it involves and the entailment relations of that proposition.[[10]](#footnote-10) Before discussing each of these ideas, it is important to note that from an analytically rigorous perspective, these are ideas of what the semantic *features* of an intentional state are, as opposed to the distinct ideas of the semantic *content* of a state and of the semantic *constituents* (or ingredients) of a state.

For example, an instance of the belief that *Albania is a country in southeastern Europe* has three semantic features. First is the characteristic of meaning something, that is, of having an internal cognitive value.[[11]](#footnote-11) Next is the characteristic of representing something, such as the mind-external objects consisting of the geographical regions of Albania and southeastern Europe. Finally, the state has the characteristic of having a truth-value, that is, (in this case) accurately representing the geo-political status of Albania. For a propositional attitude-instance to have a semantic feature is just for it to have either the characteristic *of having* a meaning, or that of representing something, or that of having a truth-value, in contrast to *not* having any such characteristics.

One can see the difference between a state’s having semantic features ⎯ that is, features in virtue of which it is evaluable as true or false, in virtue of which it may represent something, and so on ⎯ and a state’s *not* having such features by considering the difference between two kinds of mental states: a sensation and a belief. It makes no sense to say that a visual or auditory sensation of Bus 11 coming down the street, that is, the sights and sounds that this object immediately causes a person to experience, are true or false. Yet, under the assumption that the terms ‘Bus 11’ and ‘the street’ each *refer* to discrete individual objects, my belief that *Bus 11 is coming down the street* is always either true or false, and represents these individual objects.

The philosophical notion of the semantic content of a mental state is subtly distinct from the philosophical notion of a semantic feature of such a state. While the latter consists in the mere fact of a mental state’s meaning or representing something in a manner that may be evaluable as true or false (as opposed to *not* meaning or representing anything at all in this manner), the former consists in *what* a mental state actually means or represents. One can observe the distinction by considering, again, an instance of the belief that *Albania is a country in southeastern Europe*. The semantic *content* of this state is simply *what* it means, what it represents, or (from a strictly logical perspective) what it has as a truth-value. In other words, this content could be the proposition or internal idea that one may identify with this meaning, such as the proposition or idea that *Albania is a country in southeastern Europe*.[[12]](#footnote-12) Since the term ‘content’ does not necessarily refer to something internal to the mental state in this philosophical context, some might consider the semantic content of this PA-instance to be the actual objects in the world, if any, that the belief represents, such as the geo-political entity that is the state of Albania. Finally, from a strictly logical or formal perspective, the semantic content of the mental state could be simply the truth of the belief.

The semantic *features* of such a state, in contrast, are simply the *fact of* its having the property of meaning something, representing something, or having a truth-value. For example, the belief that *Albania is a country in southeastern Europe* and the belief that *Kansas is a state in the USA* both have the semantic features of meaning something, representing something, and being true. They have different semantic *contents*, however. For while the former involves the proposition or cognitive content that *Albania is a country in southeastern Europe*, and represents Albania and southeastern Europe, the latter involves the content concerning (and represents) Kansas and the U.S.

Finally, we come to the notion of the semantic *constituents* (or ingredients) of mental states. This philosophical idea is separate from both the idea of a state’s semantic features and the idea of a state’s semantic content. Consider, for example, that the name 'Albania' and the predicate 'is a country in southeastern Europe' are some of the *linguistic* ingredients of the sentence 'Albania is a country in southeastern Europe'. Analogously, the semantic constituents of a propositional attitude-instance are the actual internal phenomena that occur while the mind is in that state. These ingredients make it the case that while in that state, the mind has the semantic features of meaning or representing something and having a truth-value. They also provide the mental state with the particular semantic content that it has, i.e., its particular meaning, its particular extension (if any), and its particular truth-value.

As previously explained, one of the main projects in this essay is to evaluate the merits of two competing types of theories of intentionality: the intrinsic semantics (or conceptual) view and the extrinsic semantics (or symbolic) view. All theories of intentionality are either conceptual or symbolic. For all theories of intentionality appeal to semantic constituents to explain how a mental state has its semantic value; and these semantic constituents will either be intrinsically semantic, i.e., conceptual, or extrinsically semantic, i.e., symbolic. Having understood the philosophical difference between the notions of semantic features, contents, and constituents of mental states, we may explore the two general theories of what the actual constituents of intentional states are.

What are the pre-analytic justifications for the conceptual and symbolic theories of intentionality? As was noted earlier, although the conceptual view encourages (and perhaps demands) sophisticated metaphysical characterizations of semantic entities, basic intuitions about the mind, which are grounded deeply in common sense, actually motivate this view. The common sense view of mental representation holds that some mental states, such as beliefs, thoughts,[[13]](#footnote-13) and so forth, are about things in the world, and thus one may evaluate them as true or false. The ingredients that provide these mental states with semantic features and content, according to common sense, are something like the “meanings” or “ideas” involved in the states. For example, if a person named Sharon were to say ‘I think that Bus 11 goes downtown’, a common sense-based interpretation of this statement would be that Sharon is in a state of mind that involves the idea that *Bus 11 goes downtown*.

Moreover, if an object exists that people commonly refer to with the term ‘Bus 11’, common sense says that her state of mind represents this object in virtue of the thought’s somehow involving the idea of Bus 11. The common sense view also holds that one may evaluate this mental state as being true or false in virtue of the fact that it represents Bus 11 *as having* the property of traveling downtown. Since the intrinsic semantics view put forth in this essay partially derives from a common sense perspective on the mind, this view will depend upon a specific metaphysical assumption, which says that what it is for a mental state to be intrinsically semantic is for it to be connected essentially to some set of properties or other.

As mentioned above, the common sense view implies that the semantic constituents of PA-instances are *concept*s.[[14]](#footnote-14) This essay adopts a metaphysical interpretation of the common sense view of mental representation according to which what people commonly refer to as ‘meanings’ or ‘ideas’, namely, concepts, are intrinsically semantic entities. That is, the concept of a bus means what it does (and represents what it does) just in virtue of its own essential characteristics and independently of any contingent relation that it bears toward anything else.[[15]](#footnote-15) A later chapter will discuss different analyses of the *metaphysical* features in virtue of which a concept is intrinsically semantic. One pre-analytic intuition concerning concepts can be understood through a contrast with a pre-analytic intuition concerning symbols. This intuition is that words only have extrinsically semantic content, which they acquire solely through contingent (and often arbitrary) associations with concepts or other internal mental phenomena.

The notion of the intrinsic semantic content intuitively associated with concepts can be understood by considering the difference between the sense in which a *concept* of a bus means or represents buses, on the one hand, and the way that the *word* ‘bus’ does, for example. There is no necessary connection between the word ‘bus’ and any actual bus, for this word could easily fail to represent buses, and an infinite number of potential words could serve this very same function. A person’s utterance or inscription of the word ‘bus’ consists in her combination of the linguistic symbols (or sounds) ‘b’, ‘u’, and ‘s’ into ‘bus’ (or in making its associated sound). This act expresses the concept of a bus, represents a bus, or is true of a bus and false of a ball, only in virtue of an arbitrary convention of the language community to which she belongs. The fact that the community of English speakers agrees that utterances or inscriptions of this word refer to this kind of object constitutes such a convention.

Moreover, if one says, “point to a bus” to the English-speaking individual, and says “demonstra un [point to an] autobus” to a Spanish-speaking individual, they will both pick out the same kind of object (under normal circumstances). The intuitive explanation of this empirical fact about behavior is that both of these words express one and the same mentally-based “concept,” “meaning,” or “idea.” Thus, in contrast to utterances or inscriptions of words, when a mental state involving the *concept* of a bus occurs in one’s mind, one is automatically thinking about buses ⎯ independently of what language one speaks and independently of the social conventions in which one participates. On the one hand, the meaning of a mental state that contains the concept of a bus essentially and unchangeably involves buses, and represents possible or actual buses. On the other hand, the linguistic symbol or word ‘bus’ is *only* about buses because of the above-mentioned convention. Similarly, a sentence ⎯ a rule governed and structured concatenation of symbols ⎯ expresses and refers to what it does because of a historically based social agreement that its syntactical form will connect some names (symbols for objects) with predicates (symbols for characteristics of those objects).

The sentence ‘Albania is a country in southeastern Europe’, for example, contains the complex term ‘is a country in southeastern Europe’ and the word ‘Albania’. An agreement of the English speaking community not only determines that ‘is a country in southeastern Europe’ and ‘Albania’ represent the particular properties and objects that they do, namely, the property of being a country in a certain area and the geopolitical entity immediately northwest of Greece. This linguistic convention also determines whether each of these terms will refer to properties or objects. Thus, the sentence only expresses the complete idea that *Albania is a country in southeastern Europe* in virtue of an external social agreement involving: (i) the syntax of sentences, (ii) what kinds of logical entities (i.e. properties or objects) certain kinds of terms will refer to, and (iii) the individual referents of the terms involved in this sentence.

In contrast to the essentially *arbitrary* connection between the word ‘bus’, for example, and actual buses, intuitions say that the meaning itself that this word happens to express *cannot fail* to be related to (possible or actual) buses. For these reasons, we believe that certain mental entities, such as meanings or ideas, are related to certain sets of features essentially, such as those uniquely instantiated by buses. Thus, we believe that if there exist objects in the world that have these sets of properties, then these ideas necessarily represent such objects. This is the difference between the common sense intuition that mental phenomena like concepts have intrinsically semantic content, and the scientific view that the constituents of thought are not such “concepts,” but rather are extrinsically semantic symbols of a mental language.

According to the common sense view of intentionality, therefore, the complete semantic ingredient of Sharon’s occurrent belief that *Bus 11 goes downtown* is simply the complex *idea* that occurs (and that she affirms) inside her mind when she has that belief. This idea gives her occurrent belief the meaning that it does because of something essential to the idea itself. While this internal idea is concrete, this essay also assumes that the common sense view of the mind includes the notion of the meanings of mental states as involving abstractions (somehow). Like the notion of an intrinsically semantic entity, the notion of an *abstract* entity is defined in opposition to one that is intuitively understood, namely, the notion of a concrete, particular object.

Common sense (perhaps combined with reflection) may lead one to consider the meaning contained in a belief that *all balls are plastic* to be abstract. This is because this meaning does not solely concern any particular ball, but rather connects to all balls essentially, in virtue of the fact that it represents them. From a more philosophical standpoint, one can also consider the notion of the *meaning* of an intentional entity to suggest a connection to a non-physical but *non*-mental (mind-external) object, such as a universal or proposition. A later chapter will put forth a view of abstractions as a *source* of semantic content, which will respect the impossibility of an abstract entity residing inside the mind. According to this theory, while concepts (internal semantic constituents of mental states) get in touch with abstract entities, these universals *themselves* are not inside the mind.

A concrete entity is one that has physical or spatio-temporal properties. Given that each occurrent mental state spans a temporal interval, each mental state-instance, as well as each of its constituents, is a concrete entity. An abstract entity, in contrast, is one that has no intrinsic physical or spatio-temporal properties. A universal, in turn, is a kind of abstract entity. For while a concrete entity is particular and unrepeatable, a universal is something that is (or may be) had by more than one thing. In other words, a universal may be instantiated, and thus repeated, in more than one object. Yet, an abstract entity need not be universal, for an entity’s lack of spatio-temporal features does not necessarily imply that the entity is (or can be) instantiated by more than one thing.[[16]](#footnote-16)

There are many different versions of the intrinsic semantics perspective on intentionality. As mentioned above, later chapters will analyze and evaluate metaphysical views of non-mental abstract entities as being some of the entities that provide mental states with intrinsically semantic content along with several theories of mind-internal concepts. The theory most favored by the ancient and medieval thinkers said that our thoughts represent things in virtue of our minds being in contact with universals. According to Plato, Aristotle, and others, universals have an actual existence outside any mind; and more than one concrete thing can have the same characteristic (or stand in a certain relation) in virtue of the fact that concrete things are instances of these universals.[[17]](#footnote-17)

One can grasp the intuitions behind the ancient notion of a universal as an abstract entity by considering, for example, the set of all balls. According to Plato and Aristotle, each actual physical ball is an instance of the general category of balls, constituted by an abstract form or universal that is not *itself* a ball. In thinking about balls, moreover, one is engaging this form (or universal) with one’s mind and contemplating the essence or intrinsic qualities of “ballness” itself, such as roundness, etc. Instead of *having* properties, as concrete things do, these universals either were said to be identical to the properties of concrete things, or were said to involve those properties essentially. These properties form a kind of abstract content that can be thought, feared, desired, and known to be true of concrete things in the physical world: a content that has no intrinsic physical features and is essentially semantic.

As explained above, the symbolic (or extrinsic semantics) theory, in contrast, holds that *mental states as well as words* have only symbols, including the syntactical features accompanying the more complex symbols, as their semantic constituents.[[18]](#footnote-18) The *mental* symbols with which symbolic theories of intentionality are concerned are allegedly analogous to the symbols of written and spoken languages in that mental symbols also acquire representative content in virtue of extrinsic and non-essential relationships with other things. One should view an individual token of a type of mental symbol, such as a particular occurrence of a mental representation of buses in a particular individual’s mind, as an analogue of an individual token of a type of linguistic symbol. This mark: ‘bus’, for instance, is a particular inscription of the general type of word ‘bus’. Were the actual constituents of mental states to *be* symbols, they would be such individual tokens (according to the assumptions of this essay), just as one particular inscription of the type word ‘bus’ *is this particular inscription*: ‘bus’.

Having discussed the two general views of the semantic constituents of mental states, the next section will explain two general philosophical distinctions concerning their ontological nature and their location in relation to the mind.

C. The Atomism-Holism and Internalism-Externalism Distinctions

In contemporary philosophical discussions of the metaphysics of intentionality, there are at least two general lines drawn: that between atomism and holism and that between internalism and externalism. The first division concerns the ontological nature of whatever it is that makes a mental state have intentional features (and, more specifically, the ontological nature of whatever it is that gives the state its particular content). Each side of this debate attempts to answer the question: what very general kind of feature or ingredient is it in virtue of which a mental state is about something? Holists think that what provides a state with semantic features (and with a particular semantic content) is that state’s playing a certain role in an entire system or network of intentional states. (As a later section will explain, philosophers and cognitive scientists variously describe this role as conceptual, functional-causal, syntactic, inferential, etc.) Thus, a relationship between a state and other states determines the fact that it has intentional features (and its particular content).

Atomists, in contrast, think that that the intentional features (and thus the particular content) of a state are not determined by any relation to other *states*. Rather, for the intrinsic semantics atomist, the individual constituents of the state alone provide that state with intentional features (and a particular content). For the extrinsic semantics atomist, causal relations between the symbolic constituents of a state and objects in the mind-*external* world determine its the semantic properties (and its particular content).

The second dichotomy in contemporary philosophical discussions of intentionality is that between internalism and externalism. This debate concerns the *location*, so to speak, of whatever it is that provides a mental state with meaning or representative features. Internalists on intentionality believe that the semantic features (and the particular intentional content) of a state are constituted by something inside the mind, while externalists think that these features (and the particular content of the state) are metaphysically dependent, at least in part, upon relations to things outside the mind. Unlike the externalist, the internalist has a choice between the intrinsic and extrinsic semantics views.

According to the extrinsic semantics view, as we have seen, the semantic constituents of mental states are symbols; and a holist symbolic theory[[19]](#footnote-19) is available to the internalist, according to which the content of these symbols is completely determined by their mind-internal relations with each other. As we have also seen, under the intrinsic semantic view, the source of the semantic content of thoughts can be concrete concepts, which reside in the mind, or abstract and mind-independent mental entities, such as Platonic forms or Aristotelian universals. (Bare Property Intentionality incorporates both of these perspectives.) The internalist on intentionality may appeal to an atomist version of either of these views.[[20]](#footnote-20) Externalism on mental content holds that at least part of what determines the semantic content of a mental state is a relation between the constituents of that state and something external to the state. Usually, externalists consider this relation to be some type of perceptually and historically-based causal connection with members of a class of concrete objects.

Hilary Putnam’s famous Twin Earth thought-experiment illustrates the intuitions behind externalism on mental content, which imply that the relevant external relations are based in causal relations, and that they involve concrete objects. He asks us to imagine that there is a planet named ‘Twin Earth’, which is identical to Earth in every way except one. That is, apart from a single qualification, Twin Earth has a qualitatively identical set of objects, organisms, animals, and people, with qualitatively identical features and standing in qualitatively identical relations to each other, as Earth does. The exception is that for every H2O molecule on Earth there is a specific kind of non-H2O molecule in the corresponding place on Twin Earth, a molecule of a substance that Putnam terms ‘XYZ’; and for every XYZ molecule on Twin Earth, there is an H2O molecule in the corresponding place on Earth.[[21]](#footnote-21) In other words, Earth has no XYZ molecules, and Twin Earth has no H2O molecules. Yet, Earth has H2O everywhere that Twin Earth has XYZ, and vice versa.

Given these assumptions, for each mental state that represents an actual H2O molecule (or a collection of such molecules, as in a drinking glass) in the mind-external Earth world, there is a corresponding mental state that represents anactual XYZ molecule (or collection thereof) in the mind-external Twin Earth world, and vice versa. Consider, then, an example of such a mental state belonging to a person on Earth named ‘Eric’: an occurrent belief that *water is wet*. Eric will have a qualitatively identical counterpart on Twin Earth, that is, a person named ‘TwinEric’; and TwinEric will have an occurrent belief that is qualitatively identical to Eric’s belief that *water is wet*. (Without prejudice toward or against any position concerning what determines the semantic content of these mental states, I will call these states “water-meaning” thoughts. For since Eric and TwinEric have names suggesting that they are English-speaking men, who would use the *term* ‘water’ to refer to H2O and XYZ, respectively, it is appropriate to use this term to refer to the cognitive content of the relevant mental states.)

It is obvious to most, on an intuitive level, that while Eric’s water-meaning thought represents H2O, TwinEric’s water-meaning thought represents the chemical substance XYZ. Putnam considers the force of this intuition to support externalism on mental content.[[22]](#footnote-22) Suppose, that is, that one endorses this intuition (that the cognitively identical water-meaning beliefs of Eric and TwinEric have distinct referents). Then, consideration of Putnam’s thought-experiment raises a challenging question concerning intentionality. If the *internal* *constituents* of intentional states ⎯ the concepts or symbols that make up these states ⎯ determine what objects in the world they are about, how can mental states with qualitatively identical internalconstituents systematically represent different kinds of objects, namely, H2O molecules on Earth and XYZ molecules on Twin Earth? The externalist point is that this is incoherent. For according to the conditions of Putnam’s thought-experiment, the internal cognitive value of every H2O-directed mental state on Earth, or at least the aspect of the state that exclusively concerns H2O, is identical to the internal cognitive value of the aspect of every mental state on Twin Earth that exclusively concerns XYZ. What is going on *inside* the mind of a given person, Eric, who is thinking about or desiring H2O on Earth is identical[[23]](#footnote-23) to what is going on inside the mind of his twin, TwinEric, who is thinking about or desiring XYZ on Twin Earth.

Having used this “Twin Earth” thought-experiment to generate the intuition that the qualitatively identical mental states of Eric and TwinEric represent distinct kinds of physical entities, Putnam concludes that the *internally* accessible semantic constituents of a mental state are not sufficient to determine its semantic or intentional content. For he thinks the thought-experiment shows that what is going on inside the mind of Eric (or in the mind of TwinEric) is not *all* that determines what his mental state represents. Rather, the semantic content of a mental state, that is, what it represents and its truth-conditions, is determined *both* by the internal semantic ingredients of the mental state itself *and* by factors outside the mind, i.e., by aspects of the external context in which the thought occurs. The relevant external features are, according to the externalist, perceptually, socially, and historically-based causal connections with members of a class of concrete objects.

There is a simplistic version of an externalist-motivated conclusion concerning the Twin Earth scenario. This version is that (before the discovery of the chemical constitution of the substance to which English-speaking people refer with the term ‘water’) Eric’s water-meaning thoughts are about H2O while TwinEric’s water-meaning thoughts are about XYZ because of the distinct types of perceptual relations that have caused each man to instantiate the relevant beliefs.[[24]](#footnote-24) (This explanation obviously involves a background assumption of empiricism on concept-acquisition.)[[25]](#footnote-25) The assumption concerning Eric, on the one hand, is that his perceptual experience of H2O is the ultimate, i.e., counterfactual-supporting *cause* of his having the relevant water-concept (or symbol). From this, it follows trivially that such a perception of H2O is the ultimate cause of his instantiation of the relevant occurrent belief that *water is wet*, i.e., the belief that contains an instance of this concept. Therefore, some of Eric’s original sensations of some collection of H2O molecules, e.g., a puddle or bottle of water in infancy or early childhood, played a unique and essential role in causing his occurrent belief that *water is wet*.

Yet, on the other hand, TwinEric’s perceptual experience of XYZ is the ultimate cause of his having the relevant water-concept, or it is at least the case that XYZ molecules played a unique and essential role in causing his occurrent water-meaning belief.[[26]](#footnote-26) As previously explained, Putnam’s conclusion ⎯ that the semantic content of an intentional state is determined, at least in part, by concrete connections between the state and objects, events, etc., in his or her external environment ⎯ rests upon a certain *intuition*. This is the intuition that Eric’s belief that *water is wet* represents H2O while TwinEric’s belief that *water is wet* represents XYZ. Externalists think that the power of this intuition reflects that the notion of a state having semantic features and the notion of a state having a particular semantic content involve the implicit notion that such external connections determine this content.

The above example of an external perceptual connection between a water-meaning concept and a particular chemical in the external world is, as mentioned, oversimplified. In actuality, the externalist believes that an entire collection of complex causal connections, which often include social and historical factors, determine what our mental states mean. As a crude example of a *historically* based causal connection, consider the relationship between the American Revolutionary War and a 21st-century person’s occurrent belief that *muskets were the most popular weapons in the Revolutionary War*. While the semantic constituents of her belief may have no direct, perceptually-based causal connection with this war,[[27]](#footnote-27) there is an appropriate, historically-based causal connection between her concepts of this war and the war itself. The externalist thinks that this connection determines whether her belief actually represents this event. The above considerations reveal the basic sense in which the Twin Earth thought-experiment illustrates the intuitive philosophical underpinning for externalism on intentional content: what gives a thought the extension and truth-conditions that it has is determined, at least in part, by its relations with individuals, contexts, or processes *outside* the mind.

D. The Philosophical Possibilities for Theories of Intentionality Created by the Distinctions

So far, this essay has explained that there are two main kinds of theories of the semantic constituents of propositional attitudes, namely, the intrinsic and extrinsic semantics perspectives; and it has described two kinds of philosophical debates concerning what kind of features these constituents are and their relationship to the mind. Let us understand these debates in the context of the intrinsic and extrinsic semantics perspectives. To recall, atomists think that the individual concepts or symbols (or sets thereof) that constitute a belief (along, in the latter case, with some mind-external relation that such symbols bear) determine the meaning of a belief or a desire. Holists think, in contrast, that a network of relations between such constituents and those of other *states* give a state its semantic content. Moreover, internalists think that what gives a PA-instance its content is located completely inside the mind, while externalists hold that contextual factors in the outside environment of the thinker partially influence the meaning and extension of a PA-instance.

Thus, it would seem that there are actually *eight* relevant philosophical alternatives concerning the metaphysics of intentionality: the combination of atomism or holism with either internalism or externalism, and then the combination of each of these four positions with either the intrinsic or the extrinsic semantics views. Yet, this essay will only evaluate three of these general metaphysical positions: (1) an extrinsic semantics version of internalist *holism*, (2) an extrinsic semantics version of externalist *atomism*, and (3) an intrinsic semantics version of internalist atomism. Following are my reasons for excluding the five other (purely theoretical) possibilities.

Internalist holism is incoherent with the idea that intrinsically semantic entities are the constituents of thoughts. This is because one who endorses the view that individual concepts are the ingredients that provide propositional attitude-instances with semantic features cannot simultaneously hold that it is a network of *relations* between mental states that determines what a given state means and represents. According to the intrinsic semantics view, the concept of a bachelor, for example, essentially represents unmarried adult males. This means that it is some intrinsic feature of this concept, some property that it could not lose while retaining its identity, which makes this concept about unmarried adult males. Considering an instance of the belief that *John is a* *bachelor,* however, the internalist *holist* view would have to hold that this state has the meaning that it does because of its position in a network of inferential or causal relations with a set of other mental states {S1, S2, S3 ... Sn}. Say, for instance, that S1 and S2 together tend to cause instances of the belief that *John is a bachelor*, or that this belief-type is inferentially related to Sn (which may be the belief that *John is male*, for example). If a concept of a bachelor intrinsically means bachelor, then it contributes this meaning to any intentional state of which it is a constituent, irrespective of the causal or inferential relations between that state and other states.[[28]](#footnote-28) Therefore, the intrinsic semantics view cannot endorse internalist holism.

This discussion will not consider any internalist atomist version of the extrinsic semantics perspective on intentionality, because that type of philosophical combination would also be incoherent. This is because a symbol, by definition, only acquires the meaning that it has in virtue of its contingent and external relation to something else. Were individual symbols to constitute an instance of a belief that *all cherries are fruits*, for example, then such symbols could not be the sole determinants of what this belief represents. For on the one hand, the symbol for cherry may have its semantic content in virtue of its mind-internal relationship to a system of symbols (as the linguistic symbol ‘cherry’ does). Yet, in this case, atomism about semantic content would be false. On the other hand, this symbol may represent cherries in virtue of being uniquely associated with cherries somehow.[[29]](#footnote-29) This would be a mind-external relation, however; and to endorse this view would thus be to imply that internalism is false.

As I explained earlier, a confounding issue arises when one attempts to analyze the prospects of an externalist atomist view of mental representation in combination with the intrinsic semantics perspective. On the surface level, it may seem that such a view is inconsistent by definition. One would think that an intrinsically semantic entity would have to reside inside the mind, like a particular occurrence of a type of concept. Yet, as was also explained earlier, there is a metaphysical position according to which some intrinsically semantic entities are abstract objects that exist independently of the mind. Since the mind grasps these abstract objects through a relation to something *outside* of the mind, this may seem to be a version of externalism.

Nonetheless, I consider this view to be a version of internalism. For according to Bare Property Intentionality, intentional states are constituted by concrete concepts, which present abstract entities to the mind. It is in virtue of bearing this grasping relation toward abstractions that mental states have semantic content. Moreover, the fact that such concepts bear this grasping relation toward properties *essentially* suffices for these concepts to be intrinsically semantic.

The final group of excluded philosophical positions on intentionality includes two possible externalist holist views: the ones that would combine externalist holism with the intrinsic and extrinsic semantics perspectives, respectively. Externalist holism fits with neither of the two main views of semantic features endorsed in this essay. While some theories hold that the role that a *word* plays in the network of relations between words in a given language determines its meaning, this is externalist holism about the semantic content of *linguistic* symbols. Given that mental symbols exist inside the mind, any kind of network of relations with other mental symbols to which they *necessarily* belonged would have to exist inside the mind. Yet, the alternative possibility, i.e., that a network ofrelations among concrete entities outside the mind determine the meaning of an intrinsically semantic entity, does not appear to be coherent.

E. Three Feasible Philosophical Positions on Intentionality

A combination of internalist atomism with the intrinsic semantics view of intentionality is a feasible philosophical position. Suppose that the semantic ingredients of PA-instances are concepts, i.e., intrinsically semantic mental entities. Then according to internalist atomism, the simple,[[30]](#footnote-30) individual concepts that make up an intentional state completely determine that the state has semantic content and what (if anything) the state represents. Consider the belief that *all cherries are red*, for example. An internalist atomist of the conceptual perspective on intentionality may say that certain individual concepts determine what an instance of this propositional attitude means and represents.[[31]](#footnote-31)

Since this state definitely represents the class of all cherries, at least one of its constituent concepts must intrinsically represent each individual cherry. Yet, it is not obviously necessary, from a philosophical perspective, that some ingredient of this state intrinsically represents each individual red thing. Nonetheless, the metaphysical assumptions of the intrinsic semantics view dictate that some ingredient of this state connects the mind to the property RED, and that some ingredient connects the mind to the property CHERRY.

When combined with internalist holism, the symbolic view of mental representation is also a prima facie feasible philosophical position. How does the internalist holist define the semantic constituents of thoughts, and how does she explain how these constituents determine what the thought represents? Internalists of the extrinsic semantics perspective generally appeal to a holistic view termed“computationalism.” This view says that thoughts are pieces of mental language with an essentially syntactical kind of structure, and that they acquire their meaning, either in whole in part, through their syntactical and computational relations with other such linguistic states.

Consideration of an analogy with a linguistic entity, such as the sentence ‘This ball is red’, will facilitate a brief explanation of this idea. In the English language, ‘This ball is red’ has certain logical implications, broadly speaking, such as ‘This ball has a color that is not blue or yellow’, ‘This thing is round and not green’, ‘Something exists that is neither square nor two-dimensional’, etc. According to the computational view, the symbolic constituents of the belief that *this ball is red* have distinctive, syntactically defined shapes. The shapes of these symbols subordinate them to rules that determine their causal relations with other mental symbols ⎯ relations that are formally analogous to the logical roles that objects and properties play.

For instance, under the assumption that nothing is simultaneously a male and a female, the proposition that *John is a bachelor* entails the proposition that *John is not a woman* because of the logical relationships in both propositions between the object that is John, the property of being a bachelor, and the property of being a woman. According to computationalism, the shape of the symbols for these objects and properties are governed by syntactical rules in virtue of which the patterns of causal relations between these mental symbols is formally similar to the pattern of entailment relations among the relevant objects and properties. Since internalist holism says that the meaning of a mental state is determined by its role in a network of relations with other states, it must be interpreted to imply that the meaning of the belief that *this ball is red* is determined by such syntactically driven causal relations.

The atomic externalist perspective, in contrast, appeals to lawlike relations between members of a given class of objects in the *external* world and instances of a type of mental *symbol* to account for the symbols being about such objects. This combination (of atomic externalism with the extrinsic semantics theory) also lacks any obvious incoherence. According to the particular view that a later chapter will consider, there is a special kind of counterfactual-supporting causal relationship between actual sheep, for example, and inscriptions of a particular type of mental symbol. Actual sheep cause this symbol to be tokened, and this causal pattern is ultimately responsible for the fact that anything else causes tokenings of this symbol. This lawlike relationship, according to symbolic atomist externalism, makes it the case that that all and only symbols of that type *mean* sheep, (potentially) *represent* actual sheep, and are *true* when applied to sheep.[[32]](#footnote-32)

The particular externalist (and atomist) symbolic view that this essay will consider is Jerry Fodor’s Representational Theory of the Mind (RTM). RTM is a set of five theses concerning mental representation. (It is actually CTM without the tenet concerning the computational nature of thought processes.) The latest version of RTM, as we will see, proclaims that the semantic content of an individual inscription of a mental symbol ⎯ that it is about something and its representing the thing that it does ⎯ is constituted by a certain kind of law-like causal relation. Specifically, this is a causal relation between the symbol type and the members of the type of object that the symbol represents. By appealing to Brian Loar’s circularity objection to this view, I will argue that such a theory cannot explain false representations. In other words, if the mental symbol SHEEP may be normally caused by both sheep and goats, Fodor’s causal view of content cannot justify the claim that the symbol is true when applied to sheep and false when applied to goats.

Having discussed the basic metaphysical assumptions and general competing perspectives of philosophical theories of intentionality, below I will consider the three explanatory conditions that such theories must meet.

F. The Adequacy Conditions That Theories of Intentionality Must Meet

This essay will use three main explanatory criteria to evaluate the various versions of the conceptual and symbolic theories of mental representation. The first condition is that a theory of intentionality must account for the background assumptions of what I call the Intentional Hypothesis (IH): the thesis that actual mental states represent objects in the world. The background assumptions of this thesis are concerned with the logical and causal features of propositional attitudes: that PA-instances have an internal logical structure, and that they tend to cause each other in a way that preserves the truth of their semantic contents. The second explanatory requirement for a theory of mental representation is that the view accounts for the fact that the logical principle *substitution of terms for identicals* often fails when applied to sentences that ascribe PA-instances. Finally, the metaphysics of a theory of intentionality must explain how it is possible that one can evaluate semantically contentful states as being true and as being false. Below, I will explain these conditions further.[[33]](#footnote-33)

In attempting to define the semantic features of mental states, many contemporary philosophers of mind and language have taken up an implicitly scientific perspective that considers both the existence and nature of intentionality as a hypothesis with an extreme level of explanatory coherence and predictive reliability. Apart from the fact that the notion that mental states represent things is deeply entrenched in common sense, in other words, Fodor, Churchland, and others have argued that there are theoretical, scientific reasons to believe that mental states are often about things. The most compelling justification for the claim that intentional properties exist, according to those who hold this empirical orientation, is that using this claim as a hypothesis best allows us to successfully predict and coherently explain people’s behavior.[[34]](#footnote-34) For by assuming, on the basis of a person’s actions and surroundings, that she has beliefs and desires about certain states of affairs, one can reliably predict her future actions in the context of a coherent theory about her logical and practical rationality.

Yet, one may consider this very justificatory claim, or more specifically, the implicit argument that it expresses, to be circular. This implicit argument is that the hypothesis that intentional features exist provides us with a high degree and quality of explanatory and predictive success; and therefore we are justified in believing that intentional features exist. Some may think, however, that to legitimately reach this conclusion, we must *assume* that our own explanations and predictions have intentional features. It may be said, for example, that we have to assume that our belief that a given action was caused by a certain set of belief and desires actually represents that action (as well as some mental states of the actor). Moreover, suppose we claim that our predictions of some future action of an individual on the basis of some hypothesis concerning her beliefs and desires ⎯ a hypothesis inferred from observations of her behavior ⎯ are successful. One may say that this claim can only be justified under the assumption that the beliefs or propositions that the predictions expressed actually represented the relevant individual, her actions, and her mental states.

In addition, one may point to a specific culprit in this circularity; and this is a biased assumption that we *can and should* approach the notion of intentionality and the issue of its existence as a hypothesis of empirical (psychological) science. For the belief that any individual (other than the self) is having (or has had) any mental state, as well as the belief that such a state causes her actions, are empirical theories constructed through inductive reasoning about observations of behavior. Yet, there may be no promise in attempting to construct theories of mental representation solely from this (purportedly) objective perspective. This is because we may not be able to adopt a third person, scientific standpoint from which we seriously consider the question whether intentionality exists. Given the circularity of the above claim concerning the justification of our belief in the existence of intentionality in other minds, one could conclude that any project of analyzing intentionality must, of necessity, concern the intentionality of our very own thought processes about this issue.[[35]](#footnote-35)

Nonetheless, we do, in fact, successfully and frequently employ IH in everyday life. More important to the current philosophical project is the fact that the actual success of such an hypothesis in explaining and predicting the actions of others depends upon several background assumptions concerning the nature of mental representation, i.e., the internal structure of propositional attitude-instances and their causal, logical, and referential features. A satisfactory theory of mental representation must explain the phenomena presupposed by this hypothesis, namely, that beliefs and desires have something like logical form and tend to cause each other in a logically valid way.[[36]](#footnote-36)

In this essay, therefore, I will disavow the (implicitly) circular claim of the empirical perspective, namely, that we are justified in believing in intentionality primarily because it allows us to explain and predict behavior. Yet, I acknowledge that any acceptable theory of intentionality must be consistent with the background assumptions that we must make about the semantic features of PA-instances ⎯ those *required* to explain and predict behavior. Most importantly, this essay will use these tenets concerning the logical and causal features of PAs as criteria for the metaphysical theories that it considers and evaluates.

The first condition for the success of a philosophical view of intentionality is that the theory is consistent with the background assumptions of IH. These assumptions support the very core common sense view that PA-instances have truth-preserving causal relations with each other (and as a result, have predictable, semantically relevant effects on behavior).

What does it mean to assume that PA-instances cause each other in a truth-preserving manner? It actually involves two specific assumptions. To explain and predict behavior in terms of intentional states, one must make several assumptions. First, one must assume that the contents of PAs may have, or mimic, a logical form, like: If P, then Q. Second, one must make a *practical syllogism* assumption. The practical syllogism assumption says that: (a) tokens of certain PA-types reliably cause tokens of other types such that the resulting set of PA contents mimics the form of a valid logical inference; and (b) this content causes predictable behavior intended to bring about some state of affairs represented by the content.

So, for example, if Sharon believes *that P* and that *If P then Q*, the assumption says that this will predictably cause her to also believe that *Q*. The behavioral aspect of the practical syllogism assumption also says that PAs of desire, PAs of belief, and actions have certain relationships of coinstantiation. These relationships are such that a person having an instance of a belief that *If P then Q* and an instance of a desire that *Q* will tend to act in a way intended to bring P about. That is, if Sharon believes that *If I board Bus 11, I will go Downtown*, and she desires that *I am Downtown,* wecan reliably predict, ceteris paribus, that she will behave in a way interpretable as attempting to board Bus 11.

The causal aspect of these background assumptions plays a very important role in justifying IH. For it cannot be a mere coincidence that causally related sets of PA tokens parallel the form of logically valid inferences. In order to utilize IH to predict behavior, we must assume that PA-instances *cause* other PA-instances in virtue of their semantic constituents. Imagine the following scenario, for example: Eric knows that only one bus (Bus 11) makes a pick-up at a certain stop and that it then goes directly downtown. After observing Sharon standing at the bus stop, Eric can use the background assumptions of IH to predict that when Bus 11 arrives, Sharon will board it. His specific versions of these assumptions are: (1) Sharon believes that *Bus 11 stops here and transports people downtown*, (2) Sharon believes that *if one boards a bus that transports people to a given place, one will be transported to that place*, (3) Beliefs 1 and 2 cause Sharon to believe that *if I get on Bus 11, I will be transported Downtown*, and (4) Sharon desires *to* *go Downtown*. In order to explain this very predictivity of IH, in turn, then, its proponents must explain the systematic causal relationship between: 1) Combinations of tokens of belief-type P with tokens of belief-type P-->Q, 2) tokens of belief-type Q, and 3) actions intended to bring about Q *if* there is a *desire* that Q.

Therefore, philosophical accounts of intentionality must provide a semantic ontology of propositional attitudes that adequately explains these truth-preserving and behaviorally effective characteristics. The ontology of the common sense view, as we have seen, involves the notion that mental states contain concepts, or ideas. It explains the practical syllogism assumption with the claim that people’s understanding of the meanings of these conceptual states provides them with a grasp of their logical implications, and thus enables and disposes them to reason about the conceptual contents of their beliefs, desires, memories, fears, etc. From this perspective, then, it is because people reason about the contents of their mental states that PA-instances reliably cause other PA-instances such that the contents of the whole set have a valid inferential form. In other words, it is because Sharon understood the contents of her belief that *Bus 11 goes downtown* and of her belief that *If I board a bus which goes to X, I will go to X* that she was caused to believe that *If I board Bus 11, I will go downtown*. The first condition upon the success of a theory of mental representation is that it explain these abilities of intentional states to cause other intentional states in a way such that the truth of their contents are preserved, which, in turn, causes predictable behavior.

A distinguishing feature of intentional states is that sentences ascribing them to people fail to support operations of the deductive inference rule often referred to as ‘Substitutivity of Identicals’ (SI). This rule says that given a sentence of the form ‘Fa’, for example, and an identity statement of the form ‘a=b’, one can infer ‘Fb’. What it means to say that the referents of two names are numerically identical is that the referent of one has exactly the same set of properties as the referent of the other. SI thus codifies the following intractable intuition: indexed to a time, a thing is always identical to itself, and since identicals share the same set of properties, a thing has any property belonging to itself.

Yet, the problem is that given a sentence ascribing an intentional state to someone, it appears possible to substitute terms for identicals in the intensional context of the sentence and still fail to preserve the truth of the sentence. For the fact that Oedipus believes that *the wife of* Oedipus is *Jocasta,* for instance, does not imply that he believes that *the mother of Oedipus is* *Jocasta*, despite the fact that ‘the wife of Oedipus’ and ‘the mother of Oedipus’ are terms for one and the same object. The second adequacy condition that this essay uses to evaluate theories of intentionality is that they must explain these inferential failures.[[37]](#footnote-37)

The final adequacy condition upon philosophical views of intentionality is that they must resolve the Semantic Evaluability (or False Representation) problem. One can consider this problem from several different theoretical perspectives, but it is probably best to present it in the context of two quite simple thought-experiments. First, imagine that there are only two objects on a table, one of which is a cherry, and the other of which is a peach. A child looks at just one of these objects and has a belief that *this is a cherry*. Under the assumption that this thought is either true or false, and without considering whether the object *really is* a cherry (that is, excluding facts about the world), one may ask a fundamental philosophical question about the mind. What is it about the content of *this* kind of mental state that makes it possible that the state *can have* a truth-value? For many other kinds of states can be directed at the object without having the capacity to be true or false. The sensory state of perceiving redness, the emotional state of feeling happy in response to seeing it, and the instinctually based state of feeling hunger in response to it are examples of such non-semantic types of states.

Another way to see the question is to consider that the object that is not a cherry is a peach. Under the assumption that the child’s belief that *this is a cherry* is true, one may ask, what is it about the content of this mental state that makes it only about that particular individual cherry, as opposed to the peach, or some other cherry, or any other object in the world?

The question posed above demands an account of how (or the conditions under which) a mental state singles out a metaphysically unique entity in exclusion of all others to be its object, so to speak, of representation. This question is philosophically quite close, as we can see from the cherries-belief example, to the question of what makes it possible for the contents of certain mental states to have the semantic property of being true or false. For one can only say that a mental state is accurate or inaccurate, true or false, in virtue of the fact that the state is about a certain entity (or group thereof) and not about any other entity. That is, it is only because the child’s thought that *this is a cherry* is about the first object and not about the second (or any other object) that it one can evaluate as true or false.

The last criterion that I impose upon theories of intentionality is that they provide a satisfactory answer to these questions, that is, that they explain the exclusiveness of the relation whereby a mental state (or some of its content) represents a particular entity in numerical distinction from every other, or how it demarcates its object. (Clearly, this does not mean that mental states can only be about individuals in the metaphysical sense, for we can also think about properties, states of affairs, events, relations, and so on, as well as about sets of these entities. But these thoughts still must be able to be about these things as individual entities, that is, as one set of things, one property or set of properties, one state of affairs or set of states of affairs, etc.) To provide an adequate account of this semantic capacity is a condition upon a satisfactory theory of intentionality.

G. A Summary of the Theories and Arguments to be Discussed

This dissertation will discuss and evaluate the ways in which several metaphysical views of intentionality attempt to meet the three explanatory conditions described above. The next chapter will first analyze the internalist (holist) symbolic theory of mental representation and its ability to account for the truth-preserving causal patterns of propositional attitudes. Supposing that one would have to explain this by an appeal to subjects' internal understanding of the semantic content of their PAs, this essay will consider BonJour's infinite regress argument against the Symbolic Theory of Thought[[38]](#footnote-38) as a challenge to its ability to make this explanation. BonJour claims that we must accept that people sometimes understand the semantic content, or meaning, of their own thoughts. Yet, since symbols are, by definition, extrinsically semantic, the purely symbolic theory of thought leads to a regress, once we accept this premise concerning internal understanding. I will then evaluate an objection by Cass Weller, who holds that this regress argument begs the question against the symbolic view, and will conclude that the regress argument is not, in fact, question-begging.

The alternative internalist symbolic view that this essay will evaluate ⎯ according to which symbols constitute PA-instances, but this need not involve any internal, conceptual understanding ⎯ is Fodor’s Computational Theory of the Mind (CTM). CTM accounts for the semantics of PA-instances in terms of causal relations between mental symbols (and in some cases, in terms of relations between such symbols and inputs and outputs to a cognitive system). This view also accounts for the reliable truth-preservation of causally related series of PA-instances in terms of computation over their syntactic properties in accordance with rules for symbol manipulation.[[39]](#footnote-39) CTM allows the contents of PAs to have something that is, at least, analogous to *logical form*, by stipulating that these contents are quasi-linguistic entities, that is, strings of symbols in a mental language.[[40]](#footnote-40)

The next chapter will argue that quasi-linguistic content can account for the logical form of, for instance, the belief that *Bus 11 goes downtown*, or the belief that *if I take Bus 11, I will be transported downtown*. It will conclude, therefore, that CTM’s use of the notions of syntax and computation can allow it to explain the truth-preserving causal relations of PA-instances in virtue of their having such logical form. For one PA-instance may be said to cause another through a mechanism for computing over its constituents, that is, for transferring the mental system from one syntactic state to another. This mechanism operates in accordance with syntactic rules for combining certain kinds of symbols with others, and something analogous to the shape of a symbol determines the kind to which it belongs. It is this system of laws for causal relations between mental states that allow us to change, in a systematic and structured way, from a given state to another with which the former is logically connected.[[41]](#footnote-41)

Yet, I will then argue that while CTM could explain mental truth-preservation under the assumption that these symbolic states have semantic content, it must fail to actually explain *how* PAs have semantic content. For CTM would have to endorse three tenets. The first tenet is that a state’s pattern of causal relations with (at least) other intentional states in a mental network constitutes its content. Next is a tenet saying that this content belongs to a PA-type in virtue of some exclusive association between the PA-type and some proposition. The final tenet is that this exclusive association is determined by the fact that the formal structure[[42]](#footnote-42) of the causal network of the PA-type mimics some portion of the formal structure of the inferential network that the proposition has with other propositions.[[43]](#footnote-43) I will argue that there is no non-circular way for CTM to account for the semantic content of a given PA-type through the creation of an exclusive association between that PA-type and some proposition. For this reason, I will conclude that CTM, as a paradigm of the internalist symbolic theory of intentionality, fails to provide symbolically constituted PAs or PA-instances with semantic content.

The externalist (and atomist) symbolic view that this essay will consider is Jerry Fodor’s Causal Theory of Content, an extension of his Representational Theory of the Mind (RTM). RTM is a set of five theses concerning mental representation. (It is actually CTM without the tenet concerning the computational nature of thought processes.) The latest version of RTM, as we will see, proclaims that the semantic content of an individual inscription of a mental symbol ⎯ that it is about something and that it represents the thing that it does ⎯ is constituted by a certain kind of relation. This is an exclusive, law-like, causal relation between the symbol and the objects it represents. Through an appeal to Brian Loar’s circularity objection to this view, the third chapter will argue that such a theory cannot explain false representations. In other words, if the mental symbol SHEEP may be normally caused by both sheep and goats, Fodor’s causal view of content cannot justify the claim that the symbol is true when applied to sheep and false when applied to goats.

If a philosophical theory of intentionality cannot provide distinct metaphysical conditions under which a mental state is true and under which it is false, then it cannot account for the semantic features of mental states. This is because to have semantic features is to represent a possible or actual set of objects, that is, to single out a metaphysically unique entity (or set thereof) in exclusion of all others. Moreover, it is only in virtue of the fact that a mental state is about a certain entity (or class thereof) and not about any other entity (or class thereof) that the state can be said to be accurate or inaccurate, true or false.

Having rejected the symbolic account of the intentional features of mental states, I will evaluate a version of the intrinsic semantics view. To reiterate, this is the position that thoughts have meanings, reference, and truth-evaluability essentially. My evaluation and endorsement of the intrinsic semantics position will be guided initially by Gottlob Frege’s discussion of the second adequacy condition of theories of intentionality, i.e., the requirement that they resolve the problem of the failure of substitution of terms for identicals in intensional sentences. Frege’s semantic and metaphysical analyses of this problem in the philosophy of language and mind appeal to the existence of non-mental and objective abstract entities termed ‘senses’. He argues that the relevant terms in intensional sentences refer to these senses. Frege’s resolution of the substitution problem is that occurrent intentional states, such as beliefs and desires, are directed toward senses, which he construes as intrinsically *semantic* objects. After analyzing a Fregean metaphysical matrix of relationships between senses, pieces of language and the objective world, the fourth chapter will conclude that the essential connection between PA-instances and *properties* explains the fundamental role that abstract objects play in determining semantic relationships between the mind and other things.

While I reject Frege’s view that abstract objects *themselves* are intrinsically semantic, I will conclude that abstract entities must exist, and that they must play an essential role in the metaphysics of intrinsic intentionality. For I will argue that the notion of x bearing a semantic relation to y includes the notion of x presenting (to the mind) some set of properties that y uniquely instantiates. For y to uniquely instantiate a set of properties {F, G, H} is for y to be the only individual that instantiates F, G, and H, or for y to be the only class each individual member of which instantiates F, G, and H. The belief that *all bachelors are sports fans*, for example, presents the set of properties {UNMARRIED, ADULT, and MALE} to the mind, and thereby represents every bachelor (or the class of all bachelors).

Despite my endorsement of an ontology of intrinsic semantics that involves mind-external abstractions, I acknowledge that it is at least partially in virtue of some concrete, mind-internal constituent of an instance of the belief that *all bachelors are sports fans* that this state exclusively and intrinsically represents the entire class of bachelors.[[44]](#footnote-44) In order for this state to be intrinsically semantic, some mind-internal ingredient of the state must be intrinsically semantic. This is because it is clear that a universal cannot be a constituent of an occurrent mental state, since universals are abstract, while occurrent mental states are concrete and particular. Therefore, I will conclude that concrete, mind-internal concepts are the only intrinsically semantic constituents of PA-instances. The final chapter will argue for a view according to which these concrete concepts essentially direct the mind toward things by presenting universals exclusively instantiated by those things to the mind.

To account for the intrinsically representative content of some intrinsically semantic concepts, such as an occurrence of the concept of the class of all bachelors, this final chapter will defend a metaphysics of intentional states according to which intrinsically representative concepts include primitive and non-propositional *intentions* of directing or pointing a mental state at something other than itself. In addition, this last chapter will argue that some intrinsically semantic concepts do not represent objects (or classes thereof) in a truth-evaluable manner. Rather, these “instantiation-concepts” merely function to present the instantiation of a property, such as SPORTS FAN, to the mind.

Chapter II: The Internalist Symbolic Theory of Intentionality

A. BonJour’s Infinite Regress Objection to the Symbolic Theory of Thought

In his 1991 paper “Is Thought a Symbolic Process?”, BonJour argues that the view that the semantic ingredients of thoughts are symbolic leads to a vicious regress. In brief, he claims that if thoughts are made solely of extrinsically semantic mental symbols, then in order to understand the meaning of her thought, a person would need another thought to interpret the meaning of the first thought. However, to understand the meaning of this latter thought, which would itself be composed of extrinsically semantic symbols, she would need a third thought, and so on, to a vicious infinite regress.

BonJour considers internal understanding of one’s own intentional mental states to be a requirement for having such states at all; and he thus may regard his infinite regress objection ⎯ that will shortly be explained in depth ⎯ to be a conclusive refutation of the symbolic theory of thought.[[45]](#footnote-45) However, this chapter

will only draw a weaker conclusion from the view that the regress argument succeeds. This is because it only concerns the first explanatory requirement for theories of intentionality: that they explain the truth-preserving causal properties of PAs. Therefore, I will only infer that since, as BonJour shows, an infinitely regressive hierarchy of higher-order states is required for subjects to understand the semantic content of their symbolically constituted PA-instances, the internalist symbolic view cannot appeal to such internal understanding to explain mental truth-preservation. I will conclude that this view must account for these logically and causally based features solely in terms of *computation* over the syntactical form of PAs*.*

It is clear that in order to explain the logical and semantic properties of an intentional state, such as the belief that *Bus 11 goes downtown*, one must hold that its complex semantic constituent[[46]](#footnote-46) is some kind of structured entity with a “meaning” and logical form. That is, one must acknowledge that such a belief has a semantic content, which, in turn, has certain entailment relations. One such relation, for example, would be the tautological implication that *it is not true that Bus 11 does not go downtown*. The previous chapter explained the intuitive reasons for thinking that the kinds of things that have semantic and logical features are concepts, symbols, or universals. As explained above, the symbolic view can only appeal to structured symbol strings to account for the semantic and logical features of mental states. A sentence is a string of symbols arranged in a syntactical form that expresses or represents something. For our explanatory purposes, the sentence-like complex constituent of an intentional state will thus have to be a string of mental symbols that has: (a) an associated meaning, (b) syntactical form, and (c) logical features, i.e., truth-preserving causal patterns in relation to other symbol strings.

A brief and simplified preview of BonJour’s objection to the symbolic theory is as follows: In order for a subject to understand and reason about the logical implications of such strings of mental symbols, she must understand their meanings. Yet, a symbol, by definition, can only have a contingent, extrinsic association with a meaning, and thus can only express a meaning in virtue of some contingent, extrinsic relation that the symbol has with something external to the symbol. A complex linguistic symbol, e.g., the English sentence-type ‘Bus 11 goes downtown’, means what it does only in virtue of the agreement of the English-speaking community to associate it with some semantic content. One can understand what such a complex mental or linguistic symbol means, therefore, only if one already has a grasp of what its constituent symbols express (and of the significance of their syntactical structure). Hence, to account for the truth-preserving effects of a PA-instance with the notion that its subject has an internal understanding of the meanings of the ingredients of the symbolic state, one must explain how the subject has a prior understanding of the semantic content of such symbols.

BonJour’s line of reasoning continues: A person’s understanding the meaning of the symbolic ingredients of her thought will, as a trivial matter, involve her standing in some kind of relation to a semantic content. The view that a PA-instance with complex symbolic constituent (T) has truth-preserving causal features in virtue of its subject’s internal understanding of the meaning of T must hence appeal to a *distinct* mental state or process (T1) involving semantically contentful constituents. According to the symbolic view, the semantic content of this latter state or process *itself* involves nothing but mental symbols, however. Thus, the subject’s understanding of the content of T1 must also be an understanding of the meaning of symbols. Given that these symbols are *also* extrinsically semantic, understanding them (or being related to them in a semantically relevant way) will constitute a distinct state or process (T2), and so on. According to BonJour, the position that people even *have* an internal understanding of the semantic content of their *symbolically* constituted PAs, therefore, encounters an infinite regress.

This discussion of BonJour’s infinite regress objection uses the term ‘symbol’ to refer to any *individual* token of a type of extrinsically semantic mental entity (or mental state ingredient). It is legitimate, for instance, to discuss the following *particular* inscription ‘bus’ as a purely physical entity or “mark” independently of its semantic function as a symbol ⎯ a bearer of an extrinsically associated content ⎯ in the English (i.e., non-mental) language. This is because the preceding physical inscription would be the same as a shaped blotch of ink even if it had no semantic function in any kind of language. Analogously, one can discuss the notion of a mental symbol as a specific, individual (and most likely physical) entity, such as a particular neurological pattern with a given shape. This discussion addresses symbols as individual entities of this kind, which, as a separate matter, function as members of some kind of mental syntactic system.[[47]](#footnote-47) In what follows, I will analyze BonJour’s argument and evaluate an objection that it begs the question against the symbolic theory of thought.

BonJour objects to the symbolic or “linguistic” theory of propositional attitudes by attempting to demonstrate that if it were true, there would have to be an infinitely regressive set of higher-order symbolic mental states, each of which was a state of understanding the semantic content of its immediate predecessor. The argument, as I reconstruct it, begins with this statement:

1) “[i]t is clearly possible to employ a set of symbols, or even to accept that such a set of symbols expresses a truth, without understanding them: merely having access to a symbol or set of symbols does not by itself yield any access to their representative content.”[[48]](#footnote-48)

From this, BonJour should be interpreted to infer that:[[49]](#footnote-49)

2) If the representative constituents or ingredients of thoughts were symbolic (in the strict sense of having their meaning, truth conditions, and extensions determined by the relation of these symbols to something else), then simply mentally inscribing such symbols would not constitute or guarantee understanding of the semantic content of the symbols.[[50]](#footnote-50)

With the next three steps, one can interpret a generation of the vicious regress:

3) Suppose that the symbolic theory were true, that is, that the representative constituents or ingredients of thoughts were symbolic (in the strict sense of having their meaning, truth conditions, and extensions determined by the relation of these symbols to something else.)

From this, BonJour concludes that:

4) Any understanding that one did have of the representative content of a given mental state T (symbolically constituted as such) would have to consist in or involve something other than mentally inscribing T.

Thus, the regress can be interpreted as follows:

5) If this understanding consisted in a set or string of mental symbols T1 other than those constituting T, then (per 2) there would need to be a further set or string of symbols T2, which constituted the subject’s understanding of T1. (To construe this understanding in any other way would be to construe it as non-symbolic in character, of course, which, in turn, would be to abandon the symbolic view of thought.

The intermediate and ultimate conclusions have this form:

6) To avoid a regress, one must acknowledge that the symbolic theory of thought cannot explain how subjects understand the meaning of their own mental states, and thus that:

7) “[T]he acceptance of the linguistic conception seems to lead inexorably to the conclusion that I have no access to the content of my thought, no internal grasp or understanding at all of what I am thinking. But, this is surely an absurd result, and thus the linguistic conception must be mistaken” (Thought, therefore, cannot be a purely symbolic process).[[51]](#footnote-51)

It may be objected that this argument begs the question in favor of the intrinsic semantics view. This is because, as Weller notes, the regress argument may be interpreted to assume that there is something more to having a thought than the mere tokening of a *symbolic* state.[[52]](#footnote-52) Yet, if this is true, he thinks, then the regress argument *assumes* the very negation of the symbolic view of intentionality ⎯ that is, it assumes the negation of the extrinsic semantics view *⎯* which says that all it *is* to have a given thought is to be in a certain *symbolic* state or other. Weller’s challenge is motivated by the main (and purportedly controversial) premise of BonJour’s objection to the symbolic theory (under my simplified interpretation): that in order to have a belief, for example, that *the cat is on the mat*, one must be able to understand what *the cat is on the mat means* from the inside. The circularity objection to BonJour’s argument is that this premise begs the question by assuming implicitly that there is more to believing, for instance, that *the cat is on the mat* than simply inscribing a set of mental symbols mentally.[[53]](#footnote-53)

The actual assumption involved in BonJour’s premise concerning understanding is that to actually have an intentional mental state requires *understanding its semantic content*.[[54]](#footnote-54) Thus, whether the regress argument begs the question against the symbolic view depends upon whether this premise involves an implicit *and* *illegitimate* assumption that there is something more to having a thought than the mere inscribing of a symbolic state. In other words, does the assumption that having a thought necessarily requires understanding the thought *unjustifiably* assume that to have a thought must be more than mentally instantiating a string of symbols? If not, then BonJour’s premise concerning understanding does not beg the question. What follows is an argument that the most plausible interpretation of the premise concerning understanding is true even under the assumption of the symbolic theory, and hence, that the regress argument is not illegitimately circular, that is, it does not beg the question[[55]](#footnote-55) in favor of the intrinsic semantics view of intentionality.

In order to object that the infinite regress argument is circular, one must argue that its conclusion is (illegitimately) contained in its premises. The conclusion of BonJour’s infinite regress argument is obviously that thought is not a symbolic process, that is, that the symbolic account of the semantic content of PAs is false. Thus, this conclusion must be contained in one of the premises in order for the infinite regress argument to be circular; and such containment must be philosophically illegitimate. Weller’s objection is based on the supposed circularity of the BonJour assumption that “I am…able to tell simply by introspection what it is I am thinking about, to reflectively grasp or understand the content of my thought.” Therefore, this claim must be shown to assume (illegitimately) that the symbolic theory of thought is false in order for BonJour’s regress argument to be shown to beg the question against the symbolic view.

Weller has pointed out an important ambiguity in BonJour’s claim that “I am apparently able to tell simply by introspection what it is I am thinking about, to reflectively grasp or understand the content of my thought.”[[56]](#footnote-56) Let us call this quoted premise (A). Weller notes that (A) could mean either of two things: (i) that when a given subject has a thought T, this entails her understanding the content of T, i.e., understanding what T means, or (ii) that when a given subject has a thought T, she must have some kind of “introspective access” to this thought in the sense of having a distinct state T1 of knowledge *that she is thinking T*. While these possibilities are not mutually exclusive, they definitely have distinct philosophical implications.

Whether or not the circularity objection succeeds depends upon which interpretation of claim (A) it takes, how it analyzes this interpretation, and how it argues that the relevant denial of the symbolic theory of thought is contained in (A). Let us call (i) the ‘mere understanding’ interpretation. This may simply amount to the seemingly trivial claim that one cannot think *that P* without understanding the content of P. As even Weller acknowledges: “...it’s hard to see how understanding what I’m thinking differs at all from my simply having a thought, if to understand what I’m thinking is simply currently to have the idea that provides my thought with content.”[[57]](#footnote-57) As we will shortly see, there are several more interpretations of what it is to “understand” the semantic content of one’s own mental state. However, the other main version of (A), namely, (ii), clearly concerns not only one’s ability to understand one’s thought (whatever that is), but also concerns one’s having “introspective access to the fact that one is thinking thus and so.”[[58]](#footnote-58)

The ‘mere understanding’ version of premise (A) clearly makes the regress argument against the symbolic theory easier to defend than the ‘introspective access’ version. This is because if, on some plausible and intuitive construal of “understanding,” understanding the meaning of one’s thought that *P* only entails that one understands the ideas involved in P, or even just simply that one actually is thinking that *P*, then it is hard to imagine how this assumption can be illegitimate.

The powerful intuition behind this interpretation is that to believe that *the cat is* on the mat is to both understand what one means in believing that *the cat is on the mat* and to believe that this semantic content is true. One can grasp the triviality of this intuition by considering the intricately related intuition that in order to understand what the cat belief means, one does not need to have a distinct mental state whose content is *about* the content of the cat belief. The strength of this intuition makes it arguable that the assumption — that one must understand the semantic content of a thought in order to have the thought — is a *conceptual truth*. In that case, it would be hard to deny that the symbolic theory is obligated to accommodate the assumption.

Yet, the introspective access version of (A) is a harder assumption for a proponent of the regress argument to justify. This is because it is quite questionable whether the mere having of a thought *that P* necessarily provides one with a distinct belief that *I am thinking that P*. This would require an act of accessing and internally affirming the contents of the first-order belief through looking inward to inspect the contents of one’s own mind.[[59]](#footnote-59) The introspective access version could also be interpreted to mean that whenever one understands the meaning of one’s belief that P, one has a higher-order belief that is meta-semantic, as well, i.e., a belief that *my belief that P means that x*.

However, if premise (A) ⎯ “I am apparently able to tell simply by introspection what it is I am thinking about, to reflectively grasp or understand the content of my thought”— meant either of these things, then the very claim that there was such a thing as internal understanding of intentional states would lead to a regress. For suppose that every time one had a thought *that P* and understood its content, one had a belief that *I am thinking that P* (or a belief that *I am thinking that P and my belief that P means that x*). Then each belief would be accompanied by an infinitely regressive series of higher-order beliefs about the occurrence and semantic content of its immediate predecessor. This is because in order to understand one’s higher-order belief that *I am thinking that P*, one would have to have a distinct state of believing that *I am thinking that I am thinking that P*, and so on. Therefore, the following analysis of the circularity charge will be based upon the ‘mere understanding’ version of premise (A) only.

The most reasonable interpretation of the circularity objection, to reiterate, is that the regress argument begs the question because it “assumes” that there is something more to understanding a (symbolically constituted) belief-state than mentally inscribing the relevant symbols. It is explicitly clear that BonJour’s regress argument does make the reasonable presumption that one can instantiate, or inscribe a set of symbols without understanding their meanings.[[60]](#footnote-60) So let us begin by considering the possibility that *this* presumption begs the question against the symbolic theory, which, of course, depends upon whether the presumption (illegitimately) denies that the ingredients of thoughts are purely symbolic in character.

Given the tenet that to have a belief-state requires understanding of its semantic content, does it beg the question against the symbolic theory for BonJour to assume that there is something more to understanding a (symbolically constituted) belief-state than merely inscribing the symbols mentally? The answer to this question depends upon what is meant by ‘understanding the semantic content’ of one’s belief-state. As mentioned above, this phrase will be considered to refer to something that is a metaphysically trivial entailment of one’s having the mental state.

Consider the belief that *the cat is on the mat*, for example. On my view, the feature of understanding the semantic content *the cat is on the mat* does not involve the propositional content of some distinct state that is about this first-order content. Elimination of the higher-order understanding interpretation, however, tells one nothing about what it *actually is* to understand the meaning of one’s belief that *the cat is on the mat*. Nevertheless, one must address this issue in order to decide whether BonJour’s assumption concerning understanding begs the question against the theory that this belief has a purely symbolic content.

BonJour himself attempts an answer to the question:

Someone may want to ask…what exactly it means for the content of one’s thought to be accessible in the way that I have been appealing to. What is it for someone to understand that about which he is thinking? At one level, the answer is easy enough: for me to understand what I am thinking about is simply for me to have an intelligible idea or conception of the object of my thought. Thus, if I think there is a dead crab on the beach, I know that I am thinking about a certain distinctive sort of marine animal and not, e.g., a furry mammal or a red fruit or a motor vehicle; I understand what it means for it to be dead rather than alive; and I have a conception of the beach as a distinctive location that differs from the fish counter or the backyard.[[61]](#footnote-61)

This passage contains a locution that could seem to suggest the introspective access interpretation of ‘understanding the semantic content of’(x), namely, “know that I am thinking about” (x). Yet, the above discussion has shown that this version is implausible and leads to a vicious regress. Therefore, I will focus on analyzing the ‘mere understanding’ interpretation in terms of BonJour’s notion of having an “intelligible idea or conception of the object of [his] thought,” a conception of an object y as distinct from other kinds of objects z, w, and so on. As we will see shortly, this notion appears to involve some logical features. For it seems to involve a grasp of at least some property that is instantiated by y; and of the fact that y instantiates this property while z and w do not; and of the fact that each of z and w instantiate some property that y does not.

The “intelligible idea or conception” view of understanding the meaning of one’s belief-state could be presumed to involve the notion of understanding as grasping some abstract and intrinsically semantic mental content. Yet, what BonJour has in mind is not so purely metaphysical. This is because while he believes that PA-instances involve intrinsically semantic entities as their constituents, his *argument* against the symbolic theory of thought is based on his specific psychological conception of the notion of “understanding what I am thinking about.” In his latest work,[[62]](#footnote-62) he describes this notion as a kind of *internal awareness* of the semantic content of one’s own thought. I will now analyze the implications of this notion for the internal understanding assumption of BonJour’s regress or of argument against the symbolic theory.

BonJour’s work, *In Defense of Pure Reason*, explicitly addresses his philosophical concept of understanding. Understanding what one is thinking about involves three defining features: (1) consciousness of what one is thinking about, (2) rational insight into what one is thinking about, and (3) there being a subjective perspective from which what one is thinking is being thought. It is reasonable to suspect that there are sub-conscious thought processes; and thus, (1) may not always be present when we are thinking. Nonetheless, (1) is an essential and undeniable feature of the thoughts whose contents are understood by the person to whom the thoughts belong. What does it mean to be conscious of what one is thinking about?

Average people, philosophers of mind and knowledge, and cognitive scientists alike generally have at least two distinct notions of what it is for a state of a system to be “conscious.” The first notion is one of a *phenomenally* conscious state, such as the sensation of burning and the perception of the color teal. All such states essentially involve a subjective experience of being in the state, or what Thomas Nagel termed a qualitative feature of “what it is like” for the subject to be in the state. The second category is that of semantically and logically contentful states, such as the belief that *the cat is on the mat*, the desire to *go downtown*, an instance of knowledge that *it is currently 2004*, and so on. All such states have a representative, cognitive, or propositional content. That is, in contrast to the phenomenal content of subjective states of sensory or perceptual awareness, the content of any intentional state of consciousness can, in principle, have deductive entailment relations, have at least a (possible) state of affairs or object as its extension, and have a truth-value.

Consideration of the average person’s proper usage of the term ‘consciousness’ reveals that such semantically contentful states ⎯ many of which involve no sensory awareness, but only cognitive awareness ⎯ also fall under the concept of consciousness. In cases where one person causes harm to another, for instance, the average person only considers the actor to be morally culpable (as opposed to legally or financially responsible) for the subsequent injury or suffering if he or she was conscious of the harm that her action was causing. This intuitive and commonly endorsed principle of moral guilt clearly depends upon a concept of cognitive consciousness, i.e., the mental state required for culpability is a state of knowledge *that* I am hurting your foot with my foot. The occurrence of such a propositional attitude is neither identical to, nor the necessary result of, any phenomenally conscious state such as the tactile sensation of one’s foot contacting another’s foot. This is because one’s foot contacting an object is an event that may cause a sensation without giving rise to any semantic state of knowing that the contacted object belongs to another sentient creature. So, a person who is numb or otherwise paralyzed may know, through a verbal report or body language, that her foot is contacting that of another without having any sensation of the contact.

Moreover, the *scientific* usage of the term ‘consciousness’ as applied to a mental system connotes the idea of the system (or its subject) being in a general state of knowing (or at least having reasonable beliefs about) what is going on in its immediate environment. This notion of consciousness is distinct from that of sensory awareness *⎯* although it is usually influenced by such phenomenal consciousness. The difference is evident from the fact that an animal or a person in an REM state of sleep, a narcotic stupor, or a coma may have sensations without having any occurrent, logically evaluable beliefs about what is going on around her. Ned Block placed such states of cognitive, propositional, and representational content under the concept of “access” consciousness, as distinct from the “phenomenal” consciousness manifested in sensory states.

Although BonJour’s conception of the conscious internal awareness that constitutes understanding the semantic content of one’s intentional state requires Block’s notion of “access” consciousness, it is *not* identical to this notion. This is because Block’s notion of access-consciousness appears to involve only the mere *fact* of a mental state’s *having* a semantic or representative content, while BonJour’s notion of conscious internal awareness of the semantic content involves not only the having of the content, but also understanding of the content. The plausible existence of sub-conscious thoughts implies that this awareness is not *always* present when we are thinking.[[63]](#footnote-63) (Moreover, the existence of incoherent, vague and confused thoughts, such as those that occur in semi-conscious states of mind, and those that one may have when one attempts, for the first time, to grasp a complex concept, appears to support the view that a semantic content may be mentally instantiated without being fully understood.)

To understand the metaphysical distinction that underlies the contrast between Block’s access-consciousness and BonJour’s conscious internal awareness, one may recall the previously explained distinction between the notions of a semantic feature and a semantic content. A state falls under Block’s concept of access-consciousness as long as it has a semantic feature, while a state does not fall under BonJour’s concept of conscious internal awareness unless its subject understands its semantic content. In other words, a belief that *Albania is a country in southeastern Europe* is “access-conscious” just because it has a meaning, in virtue of which it represents both an (actual or possible) object, namely Albania, and one of its properties, namely, its geographical location.[[64]](#footnote-64) Yet, such a belief does not involve conscious internal awareness unless its subject understands *what it means* to be a country in southeastern Europe.

One may wonder, of *what* exactly is this conscious awareness? Consideration of the second feature of BonJour’s notion of internal understanding, namely, rational insight into a semantic content, reveals that this feature of awareness is one that *accompanies* the contentful features of many PA-instances, but is not *identical to* the feature of having semantic content (or to the semantic content itself). However, the awareness that constitutes internal understanding is not a distinct semantic content, either. As mentioned above, internal understanding is not a higher-order or meta-level intentional feature that *is about* the semantic content of one’s thought. The belief that *the cat is on the mat* obviously has the semantic content, or meaning, that *the cat is on the mat*. The conscious internal awareness that, according to BonJour, normally accompanies such a belief, i.e., an awareness of thinking that *the cat is on the mat*, is not a semantic content itself, for it has no distinct meaning, extension, or set of truth-conditions. Rather, internal understanding intrinsically involves a kind of “rational insight”[[65]](#footnote-65) into the logical, analytical, and conceptual implications of the content of one’s thoughts, that is, awareness of the properties to which one’s thoughts refer. Thus, conscious awareness is simply the awareness that accompanies such insight. As BonJour tells us:

A person apprehends or grasps, for example, the properties redness and greenness, and supposedly ‘sees’ on the basis of this apprehension that they cannot be jointly instantiated. Such a picture clearly seems to presuppose that as a result of this apprehension or grasping, the properties of redness and greenness are themselves before the mind in a way that allows their natures and mutual incompatibility to be apparent.[[66]](#footnote-66)

In other words, the feature of understanding the semantic content of one’s own belief essentially involves a grasping of the logical features of the properties expressed by the belief. In grasping that redness essentially looks like *this* and that greenness essentially looks like *that*, for instance, one can immediately understand that no one object could be simultaneously red and green (all over). Yet, recall, again, that I rejected the introspective access interpretation, according to which ‘understanding the meaning of x’ involves a distinct state about one’s believing that x or about the semantic content of x. Thus, one’s grasping of logical features of a property F expressed by one’s belief cannot constitute a distinct representation of some entailment relation among propositions involving F. Understanding that nothing can be completely red and completely green at the same time must be an immediate consequence of but not identical to, grasping of the properties redness and greenness.

Nonetheless, the above example involves thoughts about color properties, which are essentially related to the phenomenal contents of sensory-perceptive states, such as states of seeing something as red. Note that while there may often be some kind of phenomenal *aspect* to our thought processes, BonJour’s internal awareness is not identifiable with any kind of *qualitative* *experience* of entertaining a certain semantic content. To recognize the view that to have a thought that *x* normally involves grasping the basic logical and conceptual features of the properties involved in x, it may be better to revisit BonJour’s example of the crab belief:

[I]f I think there is a dead crab on the beach, I know that I am thinking about a certain distinctive sort of marine animal and not, e.g., a furry mammal or a red fruit or a motor vehicle; I understand what it means for it to be dead rather than alive; and I have a conception of the beach as a distinctive location that differs from the fish counter or the backyard.[[67]](#footnote-67)

The final feature with which BonJour characterizes his notion of understanding the semantic contents of one’s own thought *is*, in some sense, a “subjective”[[68]](#footnote-68) awareness. This characteristic will not play a substantive role in the current analysis of the circularity charge against his regress argument as including the internal understanding assumption. Yet, it is important to note that this feature of subjectivity is not a phenomenal awareness like the awareness that constitutes sensory experiences. Rather, it should be interpreted to mean, at least, that there is a *perspectival* element involved our being conscious of the semantic content of our own propositional attitudes. In other words, for an individual to understand what her belief that *the cat is on the mat* means is not merely for her to apprehend the objective logical features of the properties of being a cat and being a mat, etc. Rather, it is also for there to be some idiosyncratic viewpoint *from which* she grasps such properties, as they are subjectively represented in her own mind. The view that objective logical and conceptual relations among properties are successfully grasped through subjective mental representations has a history in the philosophy of mind.[[69]](#footnote-69)

I argued earlier that in order for the circularity objection to be viable, it must be interpreted to make the following charge: BonJour begs the question against the view that the semantic constituents of thoughts are symbols *by* assuming that there is something more to understanding the semantic content of one’s symbolically constituted belief-state than the mere mental inscription of its symbols. Whether or not it renders his regress argument circular, BonJour’s claim that we (often) understand the semantic content of our own thoughts is inconsistent with the claim that these semantic constituents are purely symbolic. For if, this is true, then to have a belief is simply to token a set of extrinsically semantic entities in one’s mind.[[70]](#footnote-70)

According to the symbolic theory, the belief that *a crab is on the beach* is constituted by at least several atomic symbols structured in a syntactic arrangement. For reasons that will become clear, it suffices to consider such an arrangement to be at least formally analogous to the structure of the English sentence ‘A crab is on the beach’. Consider, that this complex linguistic symbol contains six primitive or atomic symbols: ‘a’, ‘crab’, ‘is’, ‘on’, ‘the’, and ‘beach’. The combination of the first two, i.e., ‘a crab’, names some individual. The combination of the last two primitive symbols, i.e., ‘the beach’, also denotes a specific object. The symbol ‘on’, in turn, represents a relation between these two objects: the relation of the crab being physically connected to the beach; and the symbol ‘is’ has the syntactical function of expressing a predicative relation, whereby the property of being on the beach is *attributed* to the object that is the crab. However, none of these linguistic symbols *essentially* represents any objects, properties, relations, or predicative connections between objects and properties. Both this physical ink mark: ‘the beach’ and the electro-magnetic pattern that constitutes the sound-waves created by *saying* “the beach” could easily represent a different object, a different kind of object, a property, or nothing at all. Mental symbols, moreover, are analogues of these physical ink marks and sound waves: they have no features that intrinsically represent any entity.

It is thus quite clear that the features of BonJour’s notion of internal understanding are quite different from the features of mental symbols. This is because to understand the meaning of one’s belief that *a crab is on the beach* is to have a *conscious rational insight* into the properties and relations (and the logical features thereof) represented by the belief. In other words, in understanding the meaning of this belief, one is consciously aware of some of the essential properties of being a crab, such as being an animal and being amphibious; and one is consciously aware that these properties are distinct from the properties of being a non-organic machine or an organic but non-sentient entity such as an apple. In addition, one is consciously aware of the predicative or attributive function of one’s belief, which is to ascribe the relational property of being on the beach to the crab.

However, if, as the symbolic theory says, the semantic constituents of thoughts are purely symbolic, then they are extrinsically semantic, just like the linguistic symbols ‘a crab’, ‘on the beach’, and so on. Given that these mental symbols do not essentially represent any entity, it is hard to imagine how the mere tokening of such symbols could provide one with BonJour’s kind of conscious rational insight into the properties that they express and the logical features of such properties. This is because the hypothetical mental symbol that constitutes thoughts of crabs only has such semantic content in virtue of some kind of extrinsic relation. As explained in the introductory chapter, the *internalist* symbolic position must consider this relation to hold between mental symbols, a relation defined, categorized and determined, in a way that I will thoroughly explain below, by a syntactico-computational system.

It thus appears impossible that one could understand, in BonJour’s sense, the meaning of one’s belief that *a crab is on the beach* without doing one of two things. One must either: (i) understand the extrinsic relation that provides the symbol with meaning and representative features (under the assumption that the state is symbolically constituted), or (ii) understand some *intrinsic* semantic content, i.e., some entity that essentially expresses the definitive properties of crabs, such as being amphibious, being an animal and so on. Given BonJour’s conception of what it is to understand the semantic content of one’s own PA-instances, it makes perfect sense for him to assume that there is something more to understanding a (symbolically constituted) belief-state than mentally inscribing the symbols. For it would seem that mentally inscribing a set of symbols, which are by definition *extrinsically* semantic, could never in itself constitute or suffice for having such conscious, rational insight into the meaning of one’s thoughts.

What does this elucidation of the inconsistency between BonJour’s notion of internal understanding and the symbolic conception of thought imply for the objection that his regress argument is circular? Under the assumption that we sometimes have such conscious rational insight into the meaning of our PA-instances, BonJour’s argument successfully generates a vicious regress from the symbolic view. There are only two ways, therefore, for a proponent of the symbolic theory to survive.

First, he or she may deny that such a feature of internal understanding *ever* accompanies intentional states, although many would consider the existence of such a feature to be intuitively undeniable. The other alternative for a proponent of the symbolic view is not to challenge BonJour’s premise that “I am apparently able to tell simply by introspection what it is I am thinking about, to reflectively grasp or understand the content of my thought.” Rather, it is to defend some other, equally legitimate view of the nature of this feature of internal “understanding.” This view would have to be that: (i) a state with this feature has no intrinsically semantic constituents, and (ii) understanding the semantic content of such a state involves no relation to any intrinsically semantic entities.[[71]](#footnote-71)

It seems that even under the interpretation according to which it does make the relevant assumption concerning understanding, the infinite regress argument against the symbolic theory of thought can survive the circularity objection. I having a thought does indeed require understanding what it means, the objection appears to establish the general conclusion that the intentional content of mental states cannot be purely symbolic in character. Yet, the specific concern of this chapter was merely to evaluate the intuition that subjects’ internal understanding of the contents of their states explains the truth-preserving causal powers of propositional attitudes.[[72]](#footnote-72) As explained earlier, to make such an explanation the symbolic theory can only appeal to the notion of computation over the purely syntactical content of PAs.

This result has profound consequences for the ability of the symbolic theory to account for the semantic content of mental states, for, as the following segment will argue, computationalism cannot support internalist semantics. This is because, as mentioned above, the syntactical (computational) explanation of thought processes must appeal to some version of functional or inferential role semantics to give an internalist account for the features of thoughts in virtue of which they have meaning, represent things, and have truth conditions. I will now present convincing arguments that these views cannot succeed. I will conclude that externalism about content is the only alternative for the symbolic view of mental representation. First, I will look at Jerry Fodor’s (now abandoned) version of computationalism as a theory of mental truth-preservation and semantic content.

B. CTM’s Account of Mental Truth-Preservation and Internal Semantic Content

In order to evaluate the prospects of the internalist symbolic view for explaining the truth-preserving causal features of PA-instances, the current segment will consider Fodor’s Computational Theory of the Mind (CTM).[[73]](#footnote-73) Fodor has now discarded this internalist view as a general theory of semantic content. Yet, it may still account for the tendency of sequences of intentional states to follow logical patterns. CTM says that logical thought processes can be described as syntactically driven causal relationships among symbolically constituted mental states. The mental states with which CTM is concerned are syntactically structured PA-instances. I will argue that under a formal notion of truth-preservation, CTM’s appeal to the (presumed) syntactical relationships between PAs can succeed in explaining their truth-preserving causal properties.

Since PA-instances are beliefs, desires, etc., that *x is the case*, they have propositionally structured entities as their ingredients; and CTM says that these ingredients are complex symbol strings. These complex symbol strings are, in turn, composed of atomic ingredients, that is, individual symbols. The complex symbol strings are pieces of a mental *language* that have a structure that mimics the logical structure of propositions. Moreover, because their atomic ingredients are combined in accordance with syntactical categories, the structure of these symbol strings mimics this logical structure. In virtue of this form, moreover, the causal relations of these symbol strings are governed (or driven) by syntactical rules. This computational structure of the resulting network of states can explain the fact that they tend to cause each other in a truth-preserving manner, i.e., that each PA-instance tends to cause (or to be caused by) PA-instances to which it is logically related. In this way, CTM claims to be able to explain the existence of seemingly[[74]](#footnote-74) logical thought processes. It can also account for the success that we achieve in explaining and predicting the behavior of other people, by hypothesizing that they have causally related PA-instances that follow an inferential form.[[75]](#footnote-75)

How is the syntactic ontology of PAs supposed to explain their truth-preserving causal powers? As mentioned, CTM allows the constituents of PA-instances to have something at least analogous to logical form by stipulating that these constituents are quasi-linguistic entities, that is, strings of symbols in a mental language.[[76]](#footnote-76) A quasi-linguistic entity can manifest (or at least mimic) the logical form of, for instance, the proposition that *Bus 11 goes downtown*. In virtue of mimicking this form, it can thus explain the truth-preserving causal relations of such a belief.

A language is a system of symbols that we interpret to ‘represent’ objects, their properties, and connections between objects and properties. The syntax of a language involves a set of symbols broken into kinds, such as symbols for objects, for properties, for relations, and so on. The semantic ingredient of the belief that *Bus 11 goes downtown* is, according to CTM, a sentence-like entity constructed from symbols for objects, like buses and the region termed ‘downtown’, actions like ‘going’, etc. The syntax of a language also involves a system of rules: imperatives for and prohibitions against combining certain kinds of symbols with others to make sentences.

For example, symbols for objects must be combined with symbols for properties or relations to express a proposition, symbols for relations must be accompanied by at least two symbols for objects, and so on. Therefore, the idea that the constituents of PA-instances have syntactic properties means that they consist of symbolic elements[[77]](#footnote-77) that belong to different kinds; and in virtue of this membership, they are organized and combined in accordance with rules. Note that prima facie, the syntactic structure of a sentence itself, will not express a traditionally conceived semantics (that is, a conceptual meaning, an external denotation, or a truth-value) for that sentence. Thus, such a symbol string would seem to require *interpretation* to account for its semantic content.CTM says that the syntax of a PA and its syntactically driven causal relationships with other PA-instances explain the fact that these relationships are truth-preserving.[[78]](#footnote-78) In an earlier stage of his theory, Fodor suggested that this means that the network of causal relations between PA-instance constituents is (at least partially) *isomorphic* with the network of entailment relations between propositions.[[79]](#footnote-79) For example, consider a set of PA-tokens:

1. The belief that *Bus 11 transports people Downtown, and if one boards a bus that transports people to a given place, one will go to that place*.
2. The belief that *If one gets on Bus 11, one will go downtown*.
3. The belief that *If I (Sharon) get on Bus 11, I will go downtown*.

Suppose that CTM is right. Then 1, 2, and 3 are symbolic states; and their constituents appear to have a deductive logical relationship that is isomorphic with the logical relationships between members of a set of propositions {P, Q, R}, where P is the proposition associated with the sentence which expresses the content of (1), Q is the proposition associated as such with (2), etc. Thus, if a subject’s having belief (1) causes her to have (2), which, in turn, causes her to have (3), the causal relationships between the constituents of PA-instances (1), (2), and (3) have a pattern whose form is isomorphic with the form of the pattern of the inferential relations between members of the set {P, Q, R}. The CTM ontology, in which PA-instances contain symbols structured in syntactical form, attempts to use the syntax of PAs to associate them uniquely with propositions.

In order to explain *how* the causal relationships between PAs are (reliably) isomorphic with the inferential relations between their propositional associates, CTM puts forth a computational mechanism. PA-instances tend to cause others in such a way that the content of each sequential PA-instances preserves “truth,” because syntactical relations between mental sentences mimic the inferential relations among propositions. Computation over the syntactic properties of these sentences follows syntactical rules that mimic the structure of inferential rules.[[80]](#footnote-80) In this way, series of PA-tokens can be said to mimic these entailment relationships in virtue of computationally based causal relations between the unique syntactic structures of their respective types.

We can hence explain the truth-preserving causal relations between PAs in terms of the syntactic properties of their symbolic contents. One PA-instance causes another because of a mechanism for computing over its constituents, that is, for transferring the mental system from one syntactic state to another. This mechanism operates in accordance with syntactic rules for combining certain kinds of symbols with others, and the kind to which a symbol belongs is determined by something analogous to its shape. Tokenings of complex symbols cause each other in virtue of: (i) the shapes, combinations, and organizations of their symbol components, and (ii) the computational rules for transforming a mental state with certain symbolic shapes, combinations, and organizations into another type of state.[[81]](#footnote-81)

For example, consider that the natural-language sentence ‘Either one does not get on Bus 11 or one goes downtown’ expresses a semantic content, that is, a meaning, with a certain logical structure. This structure can, in turn, be formalized in accordance with the rules and categories of a hypothetical syntactical system of the mind (*M*). *M* assigns “Mentalese” symbols to certain discrete classes in accordance with the types of shapes of these symbols, and just as with natural languages, these types of symbol shapes can be considered to correlate with certain discrete metaphysical classes of entities, such as objects, properties, propositions, and logical relationships between them. These will be the classes of entities and logical relationships that thought-symbols represent, just as the sentence ‘Either one does not get on Bus 11 or one goes downtown represents certain objects, such as Bus 11, certain properties, such as being transported downtown, and certain logical relationships, such as that between P and Q in the disjunction *either -P or Q*.[[82]](#footnote-82)

In accordance with these symbol-shape categories of our hypothetical system of mental syntax, the belief that *either one does not board Bus 11 or one will go downtown* can be formalized as a string of formal language symbols that represent mental symbols: (x) (-Bxb V Gxd).[[83]](#footnote-83) To understand the sense in which *M* drives this belief-state to cause other beliefs in a truth-preserving manner, we should further hypothesize a syntactical or computational rule of *M*, which determines its causal relations in terms of the shape-kinds of its symbolic constituents and their formal arrangement. This causal rule (*M1*) says thata symbol string of the syntactical structure (x) (-Pxa V Qxb) combined with a string of the structure Pxa suffices in itself to cause an instance of the mental symbol string of the structure Qxb*.* This rule mimics, among other things, the truth-conditional logical structure of the deductively valid inference-form: Either not-P or Q, P, therefore Q *(Disjunctive Syllogism)*.[[84]](#footnote-84)

C. Functionalism and CTM’s Account of Mental Truth-Preservation

In order to understand the theoretical background of this kind of causal explanation, it is important to note that the computational view of mental representation and thought processes is a species of functionalism about the mind. Functionalism, for the purposes of this essay, is the view that mental state-types can be defined in terms of what Ned Block terms their “abstract causal” properties, that is, their causal relationships with other state-types and with inputs and outputs to a mental system.[[85]](#footnote-85) In other words, if the functional properties of a state-type define that type, then the type can be defined in conditional terms. This is because a functional state S1 of a given system is defined in terms of hypothetical changes of the system, that is, what the system will do under certain conditions when it is in S1. These conditional properties are that if the system is in S1 and it receives a certain type of input *i* it will then either: (a) give a certain output type *o*, or (b) go to another state S2,[[86]](#footnote-86) or both (a) and (b), or neither (a) nor (b). If the system will give output *o* when receiving input *i* while in S1, then S1 is partly defined in terms of these inputs and outputs; and if it will go to (i.e., cause) another state S2 when the system is in S1 and receives *i*, then S1 is defined in terms of that conditional relationship, and so on.

Now it is appropriate to consider in more detail a very basic illustration of how a syntactic process is supposed, by CTM, to explain how PA-instances cause each other in a way that preserves the truth of their contents. Suppose that, as suggested above, Sharon has a belief with content either *I do not board Bus 11 or I will go downtown* (X); and one with content *I will board Bus 11* (Y). This should cause her to believe that *I will go downtown* (Z) in virtue of a mechanism for transforming the system that is her mind from one state into another, in accordance with the syntactic rule M1. The truth-preserving causal sequence of PA-instances in this intuitive example clearly mimics the form of the deductive entailment relationship between a set of propositions. These are the propositions that can be identified with the respective semantic contents of the set of beliefs {X, Y, and Z}. This is because the conjunction of the proposition *Either I do not board Bus 11 or I will go downtown* (A) with *I will board Bus 11* (B) entails the proposition *I will go downtown* (C).

Each proposition, however, entails more than one other. The proposition *Either I will not board Bus 11 or I will go downtown* entails many other propositions, such as *If I board Bus 11 I will go downtown*; and *I will not both board Bus 11 and not be transported downtown*, and so on. Thus, the entailment relations of a given proposition (A) are the relations that it has with the set of propositions {A1, A2, ... An} that logically follow from it. As this crude example illustrates, the syntactic structure of a propositional attitude (X) mimics the logical structure of a proposition (A) when: (i) instances of (X) cause instances of other PAs {X1, X2, ... Xn}; and (ii) the network of causal patterns of the set {X1, X2, ... Xn} is, in turn, at least partially isomorphic with the network of entailment relations of the set {A1, A2, ... An}. Such isomorphisms, according to Fodor’s CTM, explain the fact that when PA-instances cause each other, they seem to preserve truth. (The network of PAs caused (syntactically) by (X) include those, such as (Z) of our example, which (X) causes in combination with other PAs, e.g., in combination with (Y) of our example. These complex, syntactically driven causal relationships also mimic the entailment relations of sets of propositions, such as the deductive relationship between (C) and the set {A, B}.)

It is clear that not all entailment relations are formal in character, i.e., not all such relations are based completely upon the logical structure of the relevant propositions; and hence, it may seem that not all truth-preserving causal sequences of PA-instances can be explained in terms of the formal (i.e., syntactical) properties of their constituents. Suppose that Sharon believes that *Bus 11 is red*, for instance, and upon seeing a green bus *x* and forming the belief that *Bus x is green*, she comes to believe that *Bus x is not Bus 11*. Given that no object is simultaneously red and green (all over), this causal sequence of PA-instances is clearly truth-preserving. But, at least prima facie, it does not appear that the reason that the content of the belief that *Bus x is green* entails that *Bus x is not Bus 11* is due to any kind of *formal* characteristics of either proposition. One may, therefore, demand a justification for the claim that the *syntactical* properties of the constituents of these beliefs suffice to explain their tendencies toward truth-preserving causal relations with each other.

A reasonable answer is as follows: while there may not be any formal or logical relationship between the mental symbols for the colors red and green, there is a metaphysical relationship of exclusion between the *properties* of being red and being green. For we may accept that on any account of color properties, it is going to be a necessary or metaphysical truth[[87]](#footnote-87) that no one thing can be simultaneously red and green all over (G1). We can thus consider it to be necessary that if a given thing is red (all over), then that thing is not green (all over) (G2). The purpose of considering whether entailment relations are formalizable, of course, is to consider the prospect of explaining truth-preserving causal sequences of PA-instances. To explain why the belief that *Bus 11 is red* combined with the belief that *Bus x is green* causes the belief that *Bus 11 is not Bus x*, it would be necessary to appeal to the presence of two background beliefs: (G1) and (G2). The explanation of the existence of these background beliefs has to do with the subject’s intuition. Thus, perhaps this explanation has to do with her internal understanding of the properties red and green. This is irrelevant, however, since the current project is to use syntax to explain mental truth-preservation between beliefs that already exist.

Consider, then, a set of four propositions: {*Bus 11 is red*, *Bus x is green*, G1, and G2}. *This* set entails that *Bus 11 is not Bus x*; and this entailment relation may be formally and systematically represented in terms of symbols. One can, therefore, use syntax to represent the truth-preserving relationship between the PAs corresponding to the four propositions in the set, on the one hand, and the belief that *Bus 11 is not Bus x*, on the other hand. Under the assumption of the existence of these background beliefs, it would then make sense for the computationalist to hold that there is a syntactical explanation for the truth-preserving causal relationship between the relevant set of PAs and the belief that *Bus 11 is not Bus x*.

One may wonder how a set of syntactical categories, i.e., rules for combining various kinds of symbols, can make it the case that certain kinds of symbol strings systematically *cause* other kinds of symbol strings. For a syntactical relationship between entities that function as symbols is, by definition, an abstract, formal matter of theory, while a causal relationship between two such entities is, by definition, a brute, mechanistic matter of physics. To ask this type of question is to make what I term the ‘Syntactic Abstraction’ objection to CTM. The answer lies in the fact that the computational account of the truth-preserving causal patterns of symbolical states is a *functional* theory of intelligent thought. I discussed the basic concept of functionalism above: a state of a system can be defined in terms of if-then conditional propositions that express its causal dispositions toward other states and inputs and outputs. Every functional analysis of an information-processing system has both an abstract, symbolic level of description and a mechanistic, primitive level of description. However, common sense says that ultimately, the success of CTM as an explanation of truth-preservation depends upon there being a plausible account of how the rules of mental syntax determine patterns of causal relations between symbol ingredients of thought. As Ned Block says:

The paradigm of defining or explicating intelligence in cognitive science is a methodology sometimes known as functional analysis. Think of the human mind as represented by an intelligent being in the head, a ‘homunculus.’ Think of this homunculus as being composed of smaller and smaller and still stupider homunculi, until you reach a level of completely mechanical homunculi....

The way to discover symbols in the brain is to first map out rational relations among states of mind and then identify aspects of these states that can be thought of as symbolic in virtue of their functions.[[88]](#footnote-88)

To begin the conceptual exploration required to answer this question, one can consider a hypothetical functional system with only two states: S1 and S2. A simple example is the 1950s Coke machine (I will call it ‘CK’) that Fodor describes in “The Mind Body Problem.”[[89]](#footnote-89) CK is a 10-cent pop machine that has only two primitive inputs, namely, a nickel and a dime, and only two primitive outputs, namely, a nickel and a Coke bottle. S1 is definable in terms of the following if-then causal conditional: if CK receives a nickel, it gives no output and goes to S2; and if CK receives a dime, it outputs a Coke bottle and stays in S1; and if CK receives no input it gives no output and stays in S1. Since CK has to have received a nickel to be in S2, S2 is definable in terms of an if-then causal conditional. The conditional is: if CK receives a nickel, it gives a Coke bottle as output and goes back to S1; and if it receives a dime, it gives a nickel and a Coke bottle and goes back to S1;[[90]](#footnote-90) and if it receives no input it goes back to S1.

Fodor’s ancient pop machine is a purely mechanistic functional system. A computational system, on the other hand, is a functional system with two *syntactical* characteristics: (1) its states, inputs, and outputs are describable as symbols, that is, entities whose formal categorization in a system allow them to represent something else, and (2) the causal relations between its states, inputs, and outputs are describable as syntactical relationships. CK is also a bistable system, i.e., a system with only two states. Many electronic computational systems, in contrast, have the potential for billions of discrete symbolic states. But, each has only has two primitive physical states, of which all of its other physical states are composed; and each has only two primitive symbolic states of which all its other symbolic states are composed.[[91]](#footnote-91)

The two primitive physical states of an electronic computational system S, i.e., S1 and S2, may be simplistically defined in terms of the presence or absence of macro-level electrical charge: S1 has a 1-volt charge and S2 has a 0-volt charge.[[92]](#footnote-92) Suppose that there are two atomic symbols in S, from which all other symbols of the system are constructed, and these are the binary-based symbols ‘1’ and ‘0’.[[93]](#footnote-93) This hypothetical system is engineered so as to instantiate the symbol ‘1’ in correlation with the 1-volt charge and instantiate the symbol ‘0’ in correlation with the 0-volt charge.[[94]](#footnote-94)

An electronic computational system is essentially a functional system, as well. A given state S1 of a functional system is, as explained above, defined in terms of ‘if-then’ conditionals. These conditionals describe two things: (a) what combination of inputs and other states (if any) will cause the system to go to state S1, and (b) what outputs and transformations to other states S will be caused to accomplish, once in S1. Electronic gates physically implement such discrete causal conditionals. A gate is a mechanism that is structured to prevent and allow specific quantities of electrical charge past a physical barrier only under rigidly specified conditions. Electrical charges, their presence and volume (in terms of voltage) are the primitive *physical* inputs and outputs, as meditated by gates, to a computational system, while binary symbols ‘1’ and ‘0’ are the primitive *symbolic* inputs and outputs to such a system.

Electronic gates are physical circuit structures, but their symbolic or abstract identity is analogous to that of certain truth-functional connectives in deductive logic. For example, certain kinds of gates are “and” gates, that is, they accept two inputs and produce one output based on the electrical, and thus the symbolic, quality of the inputs. The primitive computational symbols ‘1’ and ‘0’ determine the symbolic output of the “and” gate in the same way as the primitive semantic values *true* and *false* determine the truth-value of a conjunctive proposition. Just as *P & Q* is only true when both *P* and *Q* are true, the “and” gate only outputs a ‘1’ symbol when both of its inputs are ‘1’ symbols; and only emits a ‘0’ when at least one is a ‘0’. A matrix of the “and” gate thus shows its symbolic input-output function, i.e., it represents how the gate’s symbolic output is determined by the four alternative functions for assigning either the symbol ‘0’ or ‘1’ to each of its two inputs. In direct analogy, the truth-table matrix of the logical connective ‘&’ shows how the truth-value of a proposition *P & Q* is determined by which of the four possible functions it uses to assign the primitive semantic values true and false to its conjuncts.

Aside from the obvious formal analogy between the binary matrix of the “and” gate and the truth-functional matrix of the logical connective ‘&’, we have seen no sense in which a computational system can be described as a semantic system. This is necessary in order to understand how the syntactical rules of a computational mental system can be responsible for truth-preserving causal processes among its strings of mental symbols. Having taken note of the distinction between a number as an abstract mathematical entity and a numeral as a symbol that refers to such an entity, Block argues that the functional processes of the computer’s “and ” gate are physical manifestations of a syntactical system whose symbolic ingredients are interpretable as having semantic extension or reference:

The inputs and outputs of the adder must be seen as referring to numbers... But, once we go inside the adder, we must see the binary states as referring to the symbols themselves...Typically, as we functionally decompose a computational system. We reach a point where there is a shift of subject matter from things in the world to the symbols themselves.[[95]](#footnote-95)

There is a semantic matrix of mathematical relations that directly parallels the syntactical or symbolic matrix of relations among numerals, i.e., the matrix of input-output functions realized by the primitive, electrically driven “and” gate of the computer in Block’s example. This parallel allows the functional or causal processes of the “and” gate to mimic a mathematical or logical, i.e., truth-preserving deductive inference. For at the physical level, the combination of a 4-volt charge input and a 7-volt charge input to the “and” gate causes the gate to output a 4-volt charge. Yet, at the symbolic level, the combination of a binary ‘1’ symbol input with a binary ‘0’ symbol input causes the gate to output a binary ‘0’ output. This syntactical process of binary addition directly mimics the inference that the addition (or conjunction) of the number 1 to the number 0 is equivalent in semantic value to the number 1. This is how syntactical rules govern certain causal processes. These processes are those that mimic truth-preserving causal processes among the contents of PA-instances:

The symbolic manipulations correspond to useful rational relations among the meanings of the symbols⎯namely, the relations of addition. The useful relations among the meanings are captured by the semantic function...and the corresponding symbolic relations are the ones described in the symbolic function...It is the correlation between these two functions that explains how it is that a device that manipulates symbols manages to add numbers....

The idea is that we have symbolic structures in our brains, and that nature has seen to it that there are correlations between causal interactions among these structures and rational relations among the meanings of the symbolic structures. The primitive mechanistic processors “know” only the “syntactic” form of the symbols they process...and not what the symbols mean. Nonetheless, these meaning-blind primitive processors control processes that “make sense”⎯ processes of decision, problem solving, and the like. In short, there is a correlation between the meanings of our internal representations and their forms. And this explains how it is that our syntactic engine can drive our semantic engine.[[96]](#footnote-96)

Such is the sense in which a physical, causal system can express an abstract, syntactic, and semantic system.

So far, I have been focusing on the ability of CTM to explain mental truth-preservation, and I have not discussed any way in which this theory could explain how intentional states, symbolically construed, have any cognitive meaning or represent objects and states of affairs in the world. It could be thought that CTM eliminates or at least rejects the semantic character of a PA in favor of its symbolic, syntactical, computational, and causal properties, for the claim that the belief that *Bus 11 goes downtown* has purely symbolic ingredients apparently tells us nothing about how that state means or represents Bus 11 or its property of going downtown. A view called Functional Role Semantics (FRS) does, however, hold that the general line taken by Fodor’s syntactical explanation of mental truth-preservation can also account for the *semantic* properties of PAs.

D. The Computational Theory in the Context of Functional Role Semantics

The purpose of a computationalist’s endorsement of Functional Role Semantics is to create a reasonable *association* between a symbolically constituted mental state and a semantic content. He or she would attempt to do this by showing a parallel between the causal role of the state and the inferential role of a proposition.[[97]](#footnote-97) CTM appeals to FRS as a way of justifying the claim that symbolic propositional attitudes have semantic content.

What, then, is Functional Role Semantics? As mentioned above, to define a state S1 of a system in terms of its functional properties is to define it in terms of what the system will do under certain conditions when it is in S1. Some functionalists about mental state-type-individuation also believe that defining propositional attitude types in terms of their functional roles (i.e., their roles within a causal network) can account for the *semantic* properties of those types. These are Functional Role Semanticists.[[98]](#footnote-98)

Symbolically constituted PA-instances are presumed by CTM to have truth-preserving causal patterns. It is perfectly consistent, therefore, for CTM to appeal to FRS to identify PA-instances as having certain kinds of causal relations with each other, i.e., syntactically governed state-state transformations that mimic the form of deductive inferences, while nonetheless maintaining that these states can be associated with individual propositions as their semantic content.[[99]](#footnote-99)

To understand the specific conception of FRS that this essay assumes, consider that FRS may say that the meaning of a mental state with content *John is a bachelor* is defined in terms of this state’s *causal* relations to perceptual and symbolic inputs and outputs and other mental (including intentional) states. The belief *John is a bachelor*, for example, may tend to cause the logically related beliefs that *John is unmarried*, that *John is an adult*, and that *John is a male*. It may also tend to cause the stereotypically (but not deductively) related beliefs that, for instance, *John lives alone in an urban apartment*, and that *John is under 35*. Yet, it has non-inferential causal relationships with non-intentional states and even with behavior, as well. The belief that *John is a bachelor* may tend to be caused by visual perceptions of John (looking young and male) and by auditory states of hearing someone else say ‘John is a bachelor’. It also may tend to cause certain emotions and behaviors toward John, such as affection or condescension, and to cause associative and memory based states, such as mental images of other bachelors whom the subject knows. Finally, it may have causal relationships with symbolic inputs and outputs; e.g., being caused by a verbal question, ‘Is John married?’ and causing statements to the effect that John is male.

FRS would say that the semantic content of the belief that *John is a bachelor* is reductively analyzable in terms of *all* of these causal relationships.[[100]](#footnote-100) However, since the notion of attaching meaning to purely *symbolic* states is currently under consideration, one cannot assume any of their particular causal patterns to be inferential patterns. Rather, this section will consider them as purely causal relationships. This is because to ascribe inferential characteristics to relations among states would be to presume those states to have intentional contents. Yet, the purpose of a computationalist’s endorsement of FRS is to *argue* that there is a formal and unique isomorphism between the purely *causal* network that each PA has with others, on the one hand, and the inferential network that some proposition has with others, on the other hand. Most importantly, he or she wants to use this conclusion to justify the claim that such purely symbolic states have an internal semantic content.

Whatever the ontological status of propositions, this essay has explained that some of their undeniable intrinsic (and defining) characteristics are their deductive relationships with other propositions. So, for instance, the proposition *Bus 11 goes downtown* (D) has, as an essential property, its inferential relationship with the set of propositions {(a) *If one boards a bus which goes to X, one goes to X*, and (b) *If one boards Bus 11, one goes downtown*}. The conjunction of (D) with (a) logically entails (b). The notion of the meaning of a PA-type connects this meaning to the notion of the proposition with which the PA is intuitively associated. Since PAs are reliably truth-preserving, FRS thus explains their semantic content by connecting some of the intrinsic causal properties of a PA-type with the intrinsic inferential properties of the relevant proposition. The functional role of a PA-type is supposed to provide its meaning in the sense that the node (in the causal network) that defines a PA-type is (at least partially) isomorphic with the node (in the entailment network) that presumably is a unique characteristic of a proposition.

A profound problem with FRS concerns its need to incorporate some *ideal* notion of the causal relations of propositional attitudes. As Fodor describes the source of this idealization problem, “[t]he pattern of causal dispositions actually accruing to a given mental state must surely diverge very greatly from the pattern of inferences characteristic of its propositional object.”[[101]](#footnote-101) That is, people don’t always, or even usually, think in logically valid inferences; and even when the semantic content of a set of causally related PA-instantiations by an individual does constitute a logically valid inference, she has only made one of the infinite number of valid inferences that she could have made from the semantic content of a given PA. Therefore, for any instance of a PA-type, the contents of the PAs that it actually causes will not necessarily (or even probably) be derivable from its content. Moreover, there will be an infinite number of logically derivable PAs that it *could* have caused, but that it did not cause.[[102]](#footnote-102)

However, according to FRS, the *fact* of a PA-type having semantic properties at all, and not just its truth-preserving tendency, is reducible to the parallel between its role in the causal network of PAs and the role of some proposition in some portion of its inferential network. Thus, Fodor says, in order to account for the semantic content of a PA-type through an isomorphism between its causal network and some portion of an inferential network, one must *idealize* its causal patterns. What this means is that one must create an ideal *schema* of its causal network, which accomplishes two theoretical tasks: First, it highlights and manifests the relevant truth-preserving patterns with other PAs, by eliminating the non-truth preserving causal relations from its representation of these patterns. Second, it shows the structure of the resulting pattern to be symmetrical with the structure of (at least some portion of) the unique network of entailment relations that some proposition has with other propositions.

Idealization of the causal patterns of a PA-type would consist in creating a schema that only represents causal relations between the individual instances of that PA and instances of logically related PAs.[[103]](#footnote-103) The resulting pattern of the relations between the focal PA-type and other PAs in such a schema would, according to the present proposal, have a structure that could be directly mapped on to the structure of some portion of the network of inferential relations belonging to a certain proposition. Yet, construction of such an abstract schema from the foundation of the *actual* set of causal dispositions belonging to a given PA-type would have to be *justified* in order to prevent the obvious charge of question-begging. How can this be justified?

One way to justify the construction of such an abstract schema from the actual causal patterns of a PA-type is to create a *syntactical* model of its network of causal relations with other PA-types. By appealing to the notion of a network among mental states which merely *lacks* any syntactically “invalid” causal relations, this assumption avoids the question-begging assumption (which is implied by the creation of an ideally or completely truth-preserving network of causal relations) that mental states have a certain content semantic content. This is because a causal sequence of PA-instances can be considered to preserve truth when it is completely governed by syntactical rules for state-state transformations, rules that are implemented by computational mechanisms and which completely determine the set of PA-tokens that it causes (and that it is caused by).

In order for the idealization of causal networks of PAs to succeed, it would have to be theoretically justifiable to idealize the syntactic relations that a given type of symbolic state has with other types.[[104]](#footnote-104) Recall, however, that the idealization objection is based in the indubitable claim that in reality, the causal patterns of the tokens of a given PA-type are not parallel to the inferential patterns of any proposition. For instances of a given PA-type often cause other PA-instances in such a way that the sequence of PA contents does not follow a valid inferential (or syntactical) form. Many times, in fact, a person has a belief and fails to infer anything from it all, or infers something that doesn’t follow, or has random or vaguely associative series of thoughts that follow a ‘stream of consciousness’ pattern. For instance, Sharon may first think *Bus 11 goes downtown*, then think *how beautiful was that bus ride through Paris when I was* 17. Then she may think *it would be awful if the “Quebecois” seceded from Canada just because they speak French*. Next, she may think something else about the Parisian bus; and it's possible that only after all these mental states, does she think *If I board Bus 11, I will get downtown*.

This presumption of a system of syntactically-determined causal relations thus contrasts with the concrete psychological reality that a belief often causes other PAs whose semantic contents have no deductive logical relation to, as opposed to a personal, subjective association with its own content. It also, therefore, contrasts with the fact that a symbolic state should also cause other states whose internal structure has no syntactically determined relation with its internal structure. This means that syntactical laws do not determine the real causal relations between many actual instances of beliefs and desires. This is because the syntactic structure of a PA is a formal organization of its symbols in terms of the metaphysical categories of, and logical relationships between, the objects and properties that these symbols represent.

Thus, the assumption of syntactic ideality would have to incorporate a *statistical* assumption about the overall causal tendencies of tokens of a given PA-type.[[105]](#footnote-105) Any individual instance of the belief that *either I do not go downtown or I board Bus 11*, for example, may cause PA-instances that are not logical consequences of its content. Nonetheless, this *type* of belief could be assumed to have a law like and formalizable relationship, that is, a probabilistic relationship, with the belief that *If I board Bus 11 I will go downtown*.[[106]](#footnote-106) For overall, instances of the former type probably do tend to cause instances of the latter type (and other deductively related beliefs) more often than they cause tokens of any type of non-deductively related belief. The syntactical rules governing (or influencing) these beliefs could be assumed, moreover, to *explain* this disproportionate probability.

Intuition says that there is a parallel between the internal logical structure (or form) of the proposition that *either I do not board Bus 11 or I will go downtown* and the syntactical structure or form of the ingredients of the belief that *either I do not board Bus 11 or I will go downtown*. So that analogous to the deductive principle in virtue of which this proposition entails that *If I board Bus 11 I will go downtown*[[107]](#footnote-107)would be a syntactical rule that, along with a physical mechanism for implementation of its dictates, governs the causal relations of any symbolically constituted mental state of the form *either not P or Q*. One of these would, of course, be rule *M1*, discussed above, which says that a symbolic state whose ingredients are structured in this form can, just in virtue of syntax alone, cause a transformation to a symbolic state whose ingredients are structured in the form *if P then Q*. CTM says that there is a causal rule that governs such a transformation in virtue of the *intrinsic* syntactical features of states of the type *either not P or Q*. Thus, it might be reasonable to suppose that this type of state will cause those of the form *if P then Q* (and those of other deductively related forms) more often that it will cause those of other forms.[[108]](#footnote-108)

Therefore, given the assumption that the syntactic properties of a PA-type provide its instances with specific *probabilistic* causal properties, one could say that PA-*types* have *abstract* networks of relations, which parallel inferential relations among propositions (at least partially). Consider a given PA-type T, e.g., the belief that *Either I do not board Bus 11 or I will go downtown*, and any one of its instances *i*. According to the syntactic idealization proposal, any given PA-instance *i*2 (of a distinct belief-type T2) that *i* causes is more likely than not to be syntactically related to *i* in a manner formally analogous to an inferential relation between T1 and T2. Due to influence by associative, evaluative, preferential and emotional factors, any instance of the belief that *Either I do not board Bus 11 or I will go downtown* may cause an instance of a syntactically unrelated PA, such the belief that *It will take 25 minutes for me to go downtown* (T3). Yet, *i* is *more likely* to cause an instance of a syntactically related PA, such as the belief that *If I board Bus 11 I will go downtown*.

Based on this reasonable statistical presumption, one may think that it is justified to represent the overall network of relations that a given PA has with other PAs in terms of the syntactically driven causal relations that instances of this PA have with instances of others.The statistical assumption concerns two classes of PA-instances. The first class (C1) is the set of all PA-instances that *i* causes. The second class (C2) is the set of all instances of all PAs that are deductively related to T, such as the belief that *If I board Bus 11 I will go downtown*, the belief that *if I do not go downtown I will not board Bus 11*, and so on. The statistical assumption is that a member of C1 is more likely to be a member of C2 than it is to not be a member of C2. The pattern of relations that T has with other PAs is accounted for, for our theoretical purposes, in terms of the statistical features of the set of causal relations that all the instances of T have with instances of other PAs. In other words, this is defined both in terms of the kinds of PAs that instances of T actually cause and in terms of the kinds of PAs that instances of T are caused by. The kinds of PA-instances that instances of T usually cause are syntactically related to T in the sense of being connected to T through a syntactico-causal rule.

The computationalist holds that the syntactico-causal rules governing state-state transformations are those mimicking the logical form of deductive inference rules. Thus, the Functional Role Semanticist may use such an abstract, statistical schema of a causal network to justify the assumption that the causal pattern of each PA-type is at least *partially* isomorphic to some portion of the network of inferential relations belonging exclusively to some proposition.

One serious problem with this way of associating symbolically constituted propositional attitudes with semantic content is that the true set of inferential connections between a given proposition and others to which it is deductively related is so infinite and complex that its structure could not be parallel to the syntactical structure of any PA-type. This is because PA-types are types of human belief-states, and the finitude of the human mind prevents such a parallel from existing. Take the complex proposition (S1): *Sharon is a redhead or (Sharon is not a redhead or Sharon plays baseball)*, for example. This is logically connected to the proposition (S2): *Sharon is a redhead or (((Sharon is not a redhead or Sharon plays baseball) or Sharon does not play baseball) and John is a plumber)*. In fact, the former entails the latter. However, this latter proposition entails another proposition (S3). S3 makes a disjunction with *or John is not a plumber or John is an engineer*, and so on, to the former proposition. This entails an infinitely regressive tree of disjuncts. Nonetheless, the most important objection to Functional Role Semantics as a way of providing computational states with content is that a proposition itself cannot even be *individuated* by the pattern of its network of deductive entailment relations with other propositions. Therefore, even if the syntactical-causal idealization project could succeed, a *partial* isomorphism between the syntactical role of a PA in a network and the network of entailment relations belonging to a given proposition would still fail to associate that PA-type *uniquely* with that proposition. The following section will explain this ‘Propositional Individuation’ problem.

E. The Propositional Individuation Objection

The Propositional Individuation problem challenges the attempt, in principle, to associate mental states with propositions through networks of truth-preservation and entailment relations. The problem is based in the fact that many propositions of the same logical type will have identical *patterns* of formal deductive relations with other propositions, although the actual set of individual propositions to which a given proposition is deductively related will differ for each proposition. The proposition *John is a bachelor* (J1), for instance, is of a different logical type from the proposition *John is not a molecular biologist* (J2). This is obviously because the former expresses the instantiation of (or is an abstraction of the instantiation of) one property by one object, while the latter expresses the negation of the proposition that one object instantiates one property. There are an infinite number of propositions belonging to the former type; and propositions of this logical type belong to a larger kind, i.e., the kind that categorizes a proposition in terms of the number of properties or objects that it involves. There is also another category of propositions: those that are constituted by a certain number of simpler propositions. Belonging to this latter kind is a proposition that conjoins the two simpler propositions *John is a bachelor* and the negation of John *is 30 and John is not 30*: *John is a bachelor* *and* *it is not true that* *John is 30 and John is not 30* (J3).

It is clear that the entailment relationship between the proposition *John is a bachelor* and the proposition *John is male* is determined by the intrinsic character of some of the properties that, in some sense, are expressed by these propositions. These are the properties of being a bachelor and being male. One can see, however, that the most specific logical *type* to which a proposition belongs, and not the properties that it intrinsically expresses, will determine the formal character of its network of entailment relations with other propositions. This is because the formal character of entailment *networks* has to do with the patterns created by deductive inference rules that apply to general types of propositions, and not with the logical relations between individual properties that propositions involve.

Consider, for instance, the proposition *John is either a bachelor or a   molecular biologist* (J4). Because this is a simple disjunction, the formal character of its network of entailment relations with other propositions will be identical to the formal  character of the network of entailment relations had by any other proposition of this simple disjunctive type. Deductive inference rule governing its logical type necessitate that J4 entails that *it is false that John is not a bachelor and not a molecular biologist*, that *if John is not a bachelor then he is a molecular biologist,* etc*.* Yet, there are an infinite number of propositions that are logical analogues of J4 above, for there is no limit to the number of propositions belonging to its logical type, such as, *Mary is either a scientist or an attorney* (J5), *Eric is either a sportscaster or an artist* (J6), and so on. Thus, deductive inference rules *alone*, and not the character of any of the properties involved, determine that each of these propositions entails other specific kinds of propositions. In other words, deductive inference rules determine that J5 entails that *it is false that Mary is not a scientist and not an attorney*; and that J6 entails that *if Eric is not a sportscaster then he is an artist, and so on*. Hence, J4 has a pattern of entailment relations that is identical to the entailment patterns had by an infinite number of propositions.

From this, it is clear that a proposition cannot be individuated in terms of the formal character of the network of entailment relations that it has with other propositions, since each distinct pattern of entailment relations attaches to an infinite number of propositions. Neither can a proposition be individuated in terms of the specific set of propositions that it entails or by which it is entailed. This is because while the propositions J1 and J3 above are of distinct logical types (and even belong to different *kinds* of logical types), it is nonetheless clear that J1 entails and is entailed by the exact class of propositions as J3.[[109]](#footnote-109) (Everything, including every tautology, entails every tautology; and a tautology entails all and only other tautologies.) Therefore, inferential role necessarily *underdetermines* proposition individuation. This is because a proposition P has the exact set of inferential connections as (at least) the conjunction of P and any tautology, and Yet, we nonetheless preserve the intuition that the two — such as *John is a bachelor* and *John is a bachelor and it is not true that* *John is 30 and John is not 30 ⎯* are distinct propositions. For they have different “meanings,” given that J1 expresses only a contingent fact, while J3 expresses the conjunction of that state of affairs and a tautology.

Thus, it appears that even were it possible to analyze the semantic content of a given symbolic state reductively, in terms of some partial isomorphism between the causal network of the state and some inferential network of some proposition, this would not suffice to explain how the state has a *unique* meaning, reference, and set of truth conditions. The Propositional Individuation objection thus exposes a profound problem with the appeal to Functional Role Semantics as a way of associating the purely symbolic states of CTM with intentional content. This is that the choice of which proposition to associate with a given PA−type would have to be either circular or arbitrary. This is because as explained, the proposition intuitively associated with a given PA is merely one out of an infinite number that share the same network of entailment relations. Hence, to succeed in associating a PA exclusively with a proposition P would require the *assumption* that its causal network was (at least partially) isomorphic with the entailment relations of P, as opposed to being isomorphic with any of the other propositions of the same logical type that have the same network of entailment relations.[[110]](#footnote-110)

Yet, the point of FRS was to account for the semantic content of symbolically constituted PAs *through* the (presumably already existing or already warranted belief in) isomorphism between the abstract structures of the causal networks of individual PAs and the abstract structures of the inferential networks of individual propositions.[[111]](#footnote-111) It was *from* such a parallel structure that the association of a PA-type with a given proposition was to be justified. One therefore cannot justify the claim that any particular truth-preserving causal pattern belongs to any particular PA-type, Fodor and others have said, without already *assuming a prior association* between the PA and a proposition.[[112]](#footnote-112) The Propositional Individuation objection, therefore, appears to refute Functional Role Semantics as an account of the semantic content of symbolically constituted intentional states.

F. Conclusion

For the reasons I have explained above, the internalist-symbolic theory fails as an account of the semantic content of intentional states, even though it may succeed as a view of (apparently) truth-preserving causal relations among those states, when these relations are conceived in a purely syntactical-causal manner. The implicit holism of the internalist symbolic view says that: (1) mental representations are constituted by symbols *qua* extrinsically meaningful entities; and (2) the network of computational-causal relations among these symbols (somehow) constitutes their content. As this theory solely concerns relations between symbols inside the mind, and is unable to associate mental symbols with propositions, it cannot connect these symbols semantically with properties or with objects outside the mind. Clearly, this result prevents the holistic symbolic theory from explaining intentionality, which is, if nothing, the feature that mental states represent things outside the mind in virtue of *some* structured connection to their properties.

The syntactical-computational view, in other words, can account for why the belief that *John is a bachelor* tends to cause the belief that *John is male*,but cannot explain how this belief *mean* that John is a bachelor, etc., or how they bear the relation of *aboutness* to the actual individual in the world that is John. The implicit Functional Role Semantics that I have criticized as being required by CTM, however, assumes that the features in virtue of which states have semantic content are completely internal.[[113]](#footnote-113) Thus, a proponent of the general symbolic view of thought may still appeal to *externalism* on content, in order to explain the semantic content of mental symbols. To recall, the division between internalism and externalism concerns the location, so to speak, of whatever it is that provides a mental state with content. The computationalist, for instance, is an internalist holist: he or she thinks that the semantic contents of intentional states are functional-syntactical relations between mental symbols. These relations are determined by their formally defined shape. Externalism, in contrast, holds that what makes a state represent something is constituted, at least in part, by something outside the mind.

Having rejected the internalist version of the theory that the semantic ingredients of PA-instances are symbols, the next chapter will explain and evaluate Fodor’s externalist symbolic theory of the content of intentional states.[[114]](#footnote-114) Fodor inherits this atomist and informationalist conception from Fred Dretske; and in this tradition he provides a law-like causal condition that is sufficient for a mental state to be about something: “Dog thoughts are about dogs because they are the kinds of thoughts that dogs can be relied upon to cause.”[[115]](#footnote-115)

Chapter III:Fodor’s Causal Theory of Mental Content

A. Fodor’s Causal Theory of Intentional Content and Semantic Evaluability

Intentionality is the property that a mental state has in virtue of which it has a meaning, in virtue of which it represents an object; and thus, in virtue of which it has a truth-value. When a belief that *all sheep have white hair* occurs in a person’s mind, for instance, this is an intentional mental state. The feature of intentionality allows this mental state to have a meaning, namely, that *all sheep have white hair*, which implies (among other things) that no sheep fail to have white hair. The fact that this occurrent state has the feature of intentionality also allows the person’s mind to represent each member of the class of all sheep ― that is, to represent each individual sheep ― as having white hair. Finally, the fact that this mental state is intentional means that one can evaluate it as being true or false, with its actual truth-value being that of falsehood.

Philosophers have two general views of the metaphysics of the intentional features of these types of mental states: the intrinsic semantics view and the extrinsic semantics view. According to the intrinsic semantics view, the constituents of intentional states have semantic features independently of their relation to anything else; but according to the extrinsic semantics view, thought processes are computational and their ingredients are symbols of a mental language. That is, the extrinsic semantics position is that these mental symbols are extrinsically semantic entities, which, like words and sentences in a spoken language, refer to objects only in virtue of contingent relationships born by these symbols to other things.

For instance, a proponent of the intrinsic semantics view would say that some ingredient of a person’s occurrent belief that *all sheep are white*, such as an instance of a *concept* of sheep, essentially connects her mind to sheep. This kind of philosopher may believe that this semantic connection takes place because that concept is related intrinsically to the *property* of being a sheep, which each individual sheep instantiates. In contrast, one who endorses the extrinsic semantics view would say that when a belief that *all sheep are white* occurs in a person’s mind, the ingredients of her mental state are a complex set of mental symbols, which, by definition, do not intrinsically represent anything. A written or spoken symbol, such as an utterance of the word ‘sheep’, only happens to represent actual sheep because speakers of the English language accept the convention that proclaims it to do so.

Analogously, a *mental* symbol is a member of a syntactical system or mental language. A symbol’s membership in this language and its relation to other symbols within the language may give it a contingent ― that is, a non-necessary and non-essential ― relationship with sheep. An extrinsic semantics philosopher who holds this view endorses internalist holism. Internalist holism says that this type of contingent relation between the symbol and the whole mental language makes it the case that a token of the symbol within a person's mind represents sheep.

As an alternative to internalist holism, a proponent of the extrinsic semantics perspective on intentionality may hold that mental symbols represent objects in virtue of *causal* relationships that they have with objects of those types. Since causal relationships between two events can neither be essential to either event nor necessary in a philosophical sense, these connections between symbols and objects are contingent. Given that relationships between instances of one kind of mental symbol and members of classes of objects, such as sheep, are connections between individual instances of mental symbols and the mind-external world, this causal view of intentionality is called ‘externalist atomism’.

One may wonder how to evaluate the merits of these competing general perspectives on mental representation. One of the criteria for a successful metaphysics of intentionality is that it must explain what makes it the case that a mental state may be false.[[116]](#footnote-116) When a person comes upon an individual sheep in a field, for example, he or she may have a thought such as: *this object is a sheep*. In order for this mental state to be objectively truth-evaluable, some ingredient of the state must represent sheep only, and not goats, pencils, blades of grass, or any other type of object. This is because there is no basis from which to evaluate this thought as being *true* if the state does not represent the object as being a member of an exclusive class, namely, the class of all sheep. Likewise, there are no grounds upon which to consider a thought that *this object is a goat* as being false when a person applies it to the very same animal, if some ingredient of the state does not represent goats *only*. Therefore, a philosophical view of mental representation must explain the exclusivity of the semantic relation between a type of thought and a specific class of objects. Due to this exclusivity, occurrences of that thought are true only when applied to sheep, and are false when applied to any other class of objects, such as goats.

Jerry Fodor has put forth an externalist-atomist version of the symbolic theory of intentionality, with which he attempts to explain what makes it the case that a mental state may be objectively evaluated as true or false. In *Psychosemantics*[[117]](#footnote-117)and *Concepts: Where Cognitive Science Went Wrong*,[[118]](#footnote-118) he argues that a mental symbol exclusively represents a particular kind of object if instances of that symbol have specific kinds of causal relations with objects of that kind.[[119]](#footnote-119) A mental symbol represents only sheep, for example, if sheep normally cause tokens of that type of symbol in accordance with a law that Fodor defines with a counterfactual proposition. This law determines that only events in which sheep independently cause the symbol to occur in someone’s mind, and not intermediate or dependent events in which other objects, such as goats, cause it to occur, meet the criteria for the symbol to have an exclusive semantic connection to sheep. A mental symbol that satisfies these conditions would be false if it were applied to any object that is not a sheep. With this reductive causal theory, Fodor attempts to provide a naturalistic and *non-semantic* account of the semantic evaluability of mental states, that is, their capacity to be true when applied to some objects, but false when applied to others.[[120]](#footnote-120)

Consider again a mental symbol for sheep, which may occur as a constituent of a person’s belief *that this is a sheep*,[[121]](#footnote-121) for example*.* Fodor thinks that instances of this symbol represent actual sheep exclusively if they are caused in a specific way. That is, a person’s mind connects exclusively with actual sheep if an individual *event* in which this symbol occurs in her mind meets a set of complex causal criteria. Specifically, Fodor gives three causal conditions for a person's symbolic thought to represent one class of objects exclusively, such as sheep.[[122]](#footnote-122)

To begin understanding Fodor's causal theory of symbolic mental representation, let us refer to a person's mental symbol for sheep with the linguistic term 'SHEEP'. This mental symbol is a type, which has particular instances, or tokens. When a person has an occurrent mental state, such as an instance of the belief that *this is a sheep*, a numerically unique event takes place in time, namely, an *occurrence* of the mental symbol-type SHEEP within her mind. Fodor's first causal condition says that such an instance of SHEEP represents sheep and no other type of object if sheep generally[[123]](#footnote-123) cause instances of SHEEP to occur in her mind.

Yet, Fodor acknowledges that the fact that sheep normally cause instances of SHEEP to occur in a person’s mind will not suffice to determine that these thoughts represent only sheep. This is because goats or other types of objects may occasionally cause a person to have thoughts such as *this is a sheep*; and these thoughts would use the symbol SHEEP to refer to goats (or other types of objects).[[124]](#footnote-124) It must be possible for a person to use the mental symbol SHEEP to represent an object inaccurately; because any time a person believes that an object is a sheep, it must be *possible* that this belief is false. A thought is false when it applies a mental symbol to an object that is not a member of the class to which that symbol accurately refers. Therefore, Fodor must provide additional causal conditions to answer the question: if non-sheep cause people to apply sheep-thoughts to them, in what (non-trivial) sense are these thoughts wrong or false?

Another way to conceive of the question is as follows: given the hypothetical case that both sheep and goats cause sheep thoughts, how is it that such thoughts do not actually refer to the *disjunctive* class sheep *or* goat? Since sheep and goats are distinct types of objects, it must be wrong to think SHEEP in response to goats. There should be some reason that the mental symbol-type SHEEP represents sheep exclusively, that is, there should be some condition under which mental states involving this symbol refer only to sheep. Thus, Fodor faces the challenge of what philosophers have called the "false representation problem:" he needs to provide conditions that demarcate the class of things to which a given type of mental symbol refers exclusively.

Fodor describes the intuition behind his resolution of the false representation problem:

It’s an old observation—as old as Plato, I suppose—that falsehoods are ontologically dependent on truths in a way that truths are not ontologically dependent on falsehoods. The mechanisms that deliver falsehoods are somehow parasitic on the ones that deliver truths. In consequence, you can only have false beliefs about what you can have true beliefs about (whereas you can have true beliefs about anything you can have beliefs about at all).[[125]](#footnote-125)

This historic philosophical idea — that the notion of falsehood is definable in terms of the notion of truth — motivates Fodor’s position on the metaphysics of mental representation. Since false applications of any given mental symbol must be possible and might be frequent, Fodor adds another condition to his causal theory. With this condition, he attempts to demarcate one class of objects as the class to which a type of mental symbol accurately refers. His goal is to define what it is for a person's occurrent thought to be false in terms of a kind of causal *dependence* that the thought has upon true thoughts.

Consider, for example, an event in which a person sees a goat and the thought that *this is a sheep* occurs in her mind. Fodor’s view, of course, is that this mental state contains an occurrence of SHEEP as an ingredient. To recall, his first condition says that mental events in which a person applies this symbol to *sheep* are true if sheep cause occurrences of this symbol. This would appear to imply that a person’s occurrent belief that *this is a sheep* is true if her sighting of a goat (or some other non-sheep object) causes this mental state to occur. To prevent this implication, he gives a condition for occurrences of SHEEP to be false.

Fodor’s second condition says that an occurrence of a mental symbol, such as SHEEP, is false only if it is causally dependent upon true occurrences of that symbol. That is, if a person applies SHEEP to non-sheep objects, such as goats, this mental state is false if and only if events in which she applies SHEEP to sheep *cause* events in which she applies SHEEP to goats. Since the problem Fodor addresses with this condition assumes that goats (or other non-sheep objects) may be the immediate cause of a person having a thought that *this is a sheep*, the second condition actually requires a more precise description. It says that if a non-sheep object, such as a goat, causes the thought that *this is a sheep* to occur in a person’s mind (with an instance of the mental symbol SHEEP as an ingredient), this mental state is false only in certain cases. The cases are those in which the immediate causal relation between a goat and the mental state involving SHEEP is itself caused by the fact that *sheep* cause her to have mental states involving SHEEP.

Fodor puts forth the second condition of his causal theory in an attempt to determine that all instances of the belief that *this is a sheep* ― that involve occurrences of the mental symbol SHEEP ― refer exclusively to sheep. In this way, he intends to use the assumption that sheep bear an exclusive causal relationship to occurrences of SHEEP to define the exclusive *semantic* relationship between SHEEP and actual sheep. (Obviously, it is in virtue of this exclusive semantic relationship that all applications of SHEEP to non-sheep are false.) He assumes that the causal relationship between sheep and mental states involving SHEEP is ultimately *responsibl*e for all causal relationships between non-sheep and mental states involving SHEEP.

At this point in the discussion, it is clear that Fodor has a program for resolving the philosophical problem of false representation in terms of a naturalistic notion of causal dependence. As explained above, an example of the false representation question would be: Why do mental states involving SHEEP not refer to goats also, since goats may cause mental states involving this symbol to occur. That is, why do occurrences of the mental symbol SHEEP not refer to the disjunctive class sheep *or* goats?

Fodor’s answer is that even if another type of object, such as goats, meets his first causal condition, then his theory does not imply that the mental symbol SHEEP refers to sheep *and* goats. Nor does his theory imply that mental states involving this symbol, such as the belief that *this is a sheep*, are true when goats cause them to occur in a person’s mind. This is because non-sheep may cause thoughts involving the SHEEP symbol only *because* sheep cause thoughts involving that symbol. In other words, given a case in which a person’s encounter with a goat causes her to have a thought like *this is a sheep*, Fodor intends his second condition to determine that this mental state is false because the causal relationship between the goat and her mind is dependent upon the causal relationship between sheep and her mind. This dependence relation can be described as follows: goats would not cause the mental symbol SHEEP to occur in her mind *unless* sheep caused it to occur in her mind.

However, the notion of causal dependence characterized by Fodor’s second condition does not suffice to fully distinguish true thoughts from false thoughts. As he recognizes, some events in which goats cause occurrences of the mental symbol SHEEP may be causally *interdependent* with events in which sheep cause occurrences of this symbol. This is because features that are common to both sheep and goats may be the cause of a given occurrence of the thought that *this is a sheep*, which, by assumption, would include an instance of the mental symbol SHEEP. This possibility raises the question again: how can Fodor’s theory determine that this belief-instance is *false* if a person applies it to a goat, but *true* if she applies it to a sheep?

There are probably events in which a feature that both sheep and goats have, such as white hair for example, causes SHEEP to occur in a person’s mind. Some of these events would involve goats and not sheep. Nonetheless, at least some of these events involving goat hair would meet Fodor’s second causal criterion, because they would be causally *interdependent* with events in which sheep cause occurrences of this symbol in a person’s mind. In such cases, Fodor’s second criterion fails to demarcate sheep from goats as the *proper* referents of sheep-thoughts.

To save his causal theory of symbolic mental representation from the implication that at least some occurrences of the belief that *this is a sheep* are true when caused by either goats or sheep, Fodor adds a third causal criterion. He intends this last condition to ensure that instances of the mental symbol SHEEP are true only when applied to sheep, and false when applied to any object that is not a sheep. SHEEP refers exclusively to sheep only if all events in which non-sheep, such as goats, cause SHEEP to occur in a person’s mind have an *asymmetrical* causal dependence upon events in which sheep cause SHEEP to occur in her mind.[[126]](#footnote-126) This dependence relation can be described as follows: goats (or other non-sheep) would not cause SHEEP to occur in a person’s mind unless sheep caused it to occur in her mind; but sheep would cause SHEEP to occur in her mind *even if* no non-sheep caused it to occur in her mind. Thoughts that use occurrences of SHEEP to represent objects, such as instances of the belief that *this is a sheep*, are true only when applied to sheep only if all events in which sheep cause this symbol to occur in a person’s mind are causally *independent* of events in which any other type of object causes this symbol to occur in her mind.[[127]](#footnote-127)

In this way, Fodor’s causal theory of mental content is supposed to guarantee that sheep thoughts are true when applied to sheep and only when applied to sheep. Thus, in keeping with the Platonic intuition that false beliefs are dependent, in some way, upon true ones, Fodor defines the intentional content (or reference) of symbolic thoughts in terms of their *counterfactual* causal relations.[[128]](#footnote-128) For a person's application of the thought SHEEP to a non-sheep object to be *false* is for it to be true that: If there were no causal relation between SHEEP thoughts and sheep, there would be no causal relation between SHEEP and non-sheep objects, but there would be a causal relation between SHEEP and sheep even were there none between SHEEP and non-sheep.[[129]](#footnote-129)

In “Can we Explain Intentionality?” however, Brian Loar states: “[W]e will not find the fully naturalistic treatment Fodor seeks, for no causal condition can explicate reference except in conjunction with further intentional conditions.”[[130]](#footnote-130) Loar’s point is that a philosopher cannot successfully use a causal relationship that a concept has to reduce the semantic content of that concept to a purely naturalistic, *non*-semantic relationship.[[131]](#footnote-131) This is because in order to determine what a concept represents, a naturalistic-causal theory would have to appeal to some non-naturalistic *semantic* relationship the concept has. Therefore, Fodor’s attempt to reduce the semantic content of mental states to their causal relationships falls into a vicious regress.

B. Brian Loar’s Guiding Conception Objection to Fodor’s Causal Theory

Brian Loar’s specific view on Fodor’s theory is that his causal conditions are not sufficient to distinguish the conditions under which occurrences of a mental symbol are true from the conditions under which they are false. This is because in order to explain why an occurrence of a mental symbol has an asymmetrical causal dependence upon occurrences of another symbol in a person’s mind, Fodor would have to appeal to some distinct intentional or semantic event in her mind. (Correspondingly, Fodor would have to make the same type of appeal to explain why an occurrence of a mental symbol is counterfactually independent of occurrences of other symbols.) Suppose, for example, that a goat causes a person to have an instance of the belief that *this is a sheep*, which includes an occurrence of SHEEP*.* Loar thinks that Fodor would need to appeal to a semantic and non-causal event in her mind to justify the claim that it is false *in virtue of* meeting the asymmetrical dependence condition. That is, he must use some distinct intentional feature to explain why the goat would not have caused SHEEP to occur unless sheep caused it to occur in her mind, while sheep would cause it to occur in her mind *independently* of whether goats did.

Loar’s objection centers around the following line of argument: Fodor’s theory must account for the semantic evaluability of a given mental symbol-type, such as SHEEP. That is, he must explain why an occurrence of SHEEP is true only when applied to sheep and false when applied to any other kind of object. To account for the fact that one can evaluate a thought objectively as being true or as being false, Fodor puts forth the asymmetrical causal dependence condition, which determines that a thought is false ― and he puts forth a corresponding counterfactual condition to determine that a thought is true. However, Loar thinks that one can explain why an occurrence of a mental symbol within a person’s mind satisfies one of these conditions only if one assumes that some distinct intentional or semantic event in her mind causally influences that occurrence of the symbol.

This distinct event is an occurrence of the person’s higher-order concept of the class of object to which her SHEEP symbol refers. Appropriately, this semantic and intentional phenomenon occurs within an instance of a meta-semantic belief that the person has ― a belief *about* the class of object to which SHEEP applies. Thus, a meta-semantic belief must have a causal relationship with an occurrence of the mental symbol SHEEP in order for that symbol’s occurrence to satisfy Fodor’s asymmetrical causal condition for being false (or to satisfy his counterfactual causal condition for being true).

Loar concludes from the above argument that Fodor’s theory leads to a vicious regress. This is because it reduces the semantic feature whereby a concept, such as SHEEP, represents a class of objects in a truth-evaluable manner to a complex causal relationship, but in order for this relationship to obtain, a distinct concept must have a semantic feature whereby *it* represents a class of objects in a truth-evaluable manner.[[132]](#footnote-132) The occurrence of the latter concept is a meta-semantic or meta-intentional event: a person’s belief that an instance of her SHEEP concept represents a certain class of objects, such as sheep, to the exclusion of all others.[[133]](#footnote-133)

To demonstrate his view, Loar first argues that in principle, Fodor’s causal theory could only account for the intentional features and truth-conditions of *one* general type of concept. These are what Loar calls “individual, demonstrative” concepts. A person’s demonstrative concept expresses her subjective recognition of a given type of thing as “ ‘a thing of *that* sort’.”[[134]](#footnote-134) A person may have a demonstrative concept of a concrete, physical type of object, which directly causes her to have a certain type of perceptual experience (or group of qualia). For instance, if a person perceives sheep as being white and four-legged, to make a sound like “baaa,” and to feel rough or wooly when she touches them, her individual, demonstrative concept of sheep would be her subjective mental representation of sheep as being the type of objects that causes her to experience “that” kind of qualitative phenomenon.[[135]](#footnote-135)

Loar believes that non-demonstrative concepts are “socially deferential” or “socially mediated;” and he argues that Fodor’s theory cannot account for the truth-conditions of socially mediated concepts without abandoning naturalism and falling into a vicious regress.A person’s concept is socially deferential when she would, in principle, defer to some community to define the concept’s intentional content and truth-conditions.[[136]](#footnote-136) For a simple example, a person has an individual, demonstrative concept of sheep when she herself defines its intentional content and truth-conditions in terms of subjective phenomena, such as the way sheep look to her. In contrast, a person has a socially deferential concept of sheep when she allows some community, perhaps of average people with agricultural experience, or of zoologists, etc., to determine what type of object that concept represents accurately.

Loar’s categories of demonstrative and socially deferential concepts play a central role in his objection to Fodor’s theory. Loar argues that when one uses Fodor’s causal conditions to determine the truth or falsity of members of either category of concept, one must appeal to a distinct semantic event to explain why those causal conditions are met. Thus, with either type of concept, Fodor’s theory leads to a regress as described above. Yet, Loar thinks that Fodor’s theory *automatically* falls into this regress when he uses it to account for the truth-values of socially mediated concepts. This is because by definition, a person’s meta-semantic belief that *my concept refers to that which my community considers it to refer* determines the class of objects that any of her socially mediated concepts represents.[[137]](#footnote-137) He concludes that ultimately, Fodor’s view fails as a naturalistic, non-semantic account of the intentional features of demonstrative concepts, as well.[[138]](#footnote-138)

To support his argument, Loar discusses Fodor’s thought-experiment concerning the theoretically sophisticated concept PROTON. In this experiment, Fodor acknowledges that a person’s use of this concept depends upon a set of background concepts that her community constructs. Yet, he argues, her visual experiences of protons play a role in causing her to apply the concept PROTON in some cases, which means that causal conditions can determine what these thoughts represent. Loar objects that PROTON *must* be a socially deferential concept, and that the perceptually-based causal relationship between protons and some of a person’s usages of the concept PROTON does not determine the class of objects that these thoughts represent:

Consider what [Fodor] says about ‘proton.’ Suppose a person ascribes it in response to visible traces of protons, depending on theoretical beliefs to the

effect that these traces are caused by protons in such-and-such physical ways. According to Fodor, ‘proton’ can come to refer to protons by virtue of such a causal connection... But... ‘proton’ does not express a demonstrative that acquires its reference from those discriminations, as becomes clear by considering another case. You observe the same visible phenomenon and with no theory of protons in mind you form the conception ‘that sort of thing,’ the conception of that, whatever it is. If you succeed in referring to an underlying property or kind, the referential link is in a general way demonstrative. But that is not what happens in Fodor’s case, where ‘proton’ is already embedded in a background physical theory and has a cognitive role independent of and presupposed by its employment in the current discrimination of traces, which apparently makes it not a demonstrative whose reference derives from that discrimination.[[139]](#footnote-139)

Here Loar illustrates that a specific visual experience a person has of protons can be the immediate cause of an event in which she applies the concept PROTON to these particles. Fodor acknowledges that such a person would have pre-existing beliefs concerning protons as being the kind of objects that cause these qualia ― beliefs that are the *ultimate* cause of her *applying* PROTON to the particles. These beliefs do not arise from her perceptual experiences of protons: they are grounded in scientific theories.

Loar’s argument appears to be that since the person’s PROTON concept has developed within a theoretical structure prior to the occurrence of these specific visual phenomena, its pre-existence within the background theory is what causes her to apply PROTON, as opposed to any other concept, to the specific visual phenomena. To understand this point, suppose that our culture has a socially-constructed concept of a proton as a subatomic particle that resides in the nucleus of an atom, which has a positive electromagnetic charge, which creates certain visual traces under technologically enhanced visual conditions, and so on.Since the existence of the background theory and her PROTON concept are prior to and independent of the specific event in which she has the proton-like qualia, the “cognitive role” of PROTON within such a theory would cause her to use PROTON to *refer* to the particles as a theoretical explanation of the experience. Moreover, this cognitive role provides the semantic content of this occurrence of PROTON. Thus, the immediate, perceptually-based causal relationship between the relevant protons and the occurrence of PROTON does not determine what this occurrence represents.

Recall that Fodor’s goal is to use a naturalistic theory to define the semantic content of a type of mental symbol, including its truth-conditions, in terms of causal relations between objects and occurrences of the symbol. From his analysis of the PROTON thought-experiment, Loar makes the general conclusion that Fodor’s causal theory cannot apply to socially deferential concepts. This is because when a person would defer to a community-based theory to determine the semantic content of her concept, that theory, and not any naturalistic causal relationship, determines that occurrences of that concept represent a certain type of object, such as protons, exclusively. But Loar’s objection to Fodor’s causal theory as applied to socially deferential concepts is not merely that a pre-existing set of concepts (and no perceptually-based causal relationship) determines their semantic content. Loar also argues that Fodor’s theory falls into a vicious infinite regress when applied to occurrences of socially deferential concepts.

For example, the theory says that an application of PROTON to an object, such as an electron, would be false if it bears an asymmetrical causal dependence to applications of the symbol to another type of object, i.e., protons. According to Loar, a person’s pre-existing, meta-semantic belief that *my PROTON concept refers to that which my community intends it to refer*, or that *my PROTON concept refers to a subatomic particle with a positive electromagnetic charge and residing in the nucleus of an atom*, actually causes her to apply the concept PROTON to an object in response to visual phenomena. Fodor’s causal theory must explain why the causal relation between an electron and an occurrence of the PROTON concept would be asymmetrically dependent upon the causal relation between protons and occurrences of the PROTON concept. Yet, Loar thinks that one can only explain this dependence the person’s meta-semantic belief that her community (or that a certain scientific definition that this community constructs) determines what her PROTON thoughts represent.

Obviously, the causal relationship between this occurrence of this meta-semantic belief and the person’s application of PROTON to an electron involves the occurrence of a *semantic* event that is distinct from the event in which the person applies PROTON to an object. Were Fodor to attempt to reduce the semantic content of a person’s belief that *my PROTON concept refers to a subatomic particle with a positive electromagnetic charge and residing in the nucleus of an atom* to a naturalistic, causal relation like those described by his theory, he would fall into a vicious regress. This is because he would have to define the truth-conditions of her higher-order concept of *that to which* her PROTON concept refers in terms of the influence of another higher-order concept: her concept of that to which the first higher-order concept refers. Socially deferential concepts are thus, “not candidates for fitting Fodor’s condition,” according to Loar, for Fodor’s causal criteria will necessarily imply that these concepts have an infinitely regressive set of truth-conditions. Thus, the only concepts to which Fodor’s account could apply must be demonstrative, that is, they “must acquire reference as a function of [an individual’s] non-socially mediated [presumably perceptually and memory based] discriminations.”[[140]](#footnote-140)

To recall, a demonstrative concept is a concept of “a thing of *that* sort,” usually a concept of a thing that causes “that” kind of perceptual experience. Given an individual’s demonstrative concept of sheep, for example, the ‘that’ in the “thing of that sort” locution can be understood to pick out the properties of the object that causes *that* quale, e.g., “that” fuzziness or “that” ‘baaa’ sound, etc. To understand Loar’s objection to Fodor’s causal theory as applied to demonstrative concepts, suppose that a person applies a demonstrative concept, SHEEP, to a goat. Fodor’s theory says that this mental state is false only if it is asymmetrically dependent upon events in which she applies SHEEP to sheep. Loar argues that in order to justify the claim that goats would not cause SHEEP to occur in her mind unless sheep did, but that sheep would cause SHEEP to occur in her mind independently of whether goats did, Fodor must appeal to the causal influence of a distinct semantic event. This event is the occurrence within her mind of a meta-semantic belief, such as a belief that *my concept refers to that kind of thing*. Within the intensional sentence that ascribes this belief to the person, ‘that’ generally points out the kind of object that causes her to have a certain kind of qualia.[[141]](#footnote-141) What does the fact that this event is required imply? Fodor’s attempt to use a causal theory to reduce the truth-conditions of demonstrative causal relations encounters a vicious regress of meta-intentional concepts.

To exemplify his argument, Loar asks us to consider a case where a person is informed that he has been applying a demonstrative (or “recognitional”) concept to more than one natural kind of object. He describes the case of “Armand, a novice bird watcher who learns that the birds he has been identifying under a certain recognitional concept [B] are of two not especially closely related kinds.” If one asked Armand what, if anything, his B concept represents, he could give three mutually exclusive answers. He could think that occurrences of B in his mind represent (1) something indeterminate, (2) the disjunctive class of the two kinds of birds, or (3) “a resemblance in configuration,” which would mean that B represents the disjunctive class of the two kinds of birds and any other kinds of objects that share the relevant perceptual resemblance.

Loar concludes that in each case, only Armand’s idea of that to which he intends to refer with his B concept can determine its reference. That is, in case (1) B would not refer, in case (2) it would refer to the disjunctive class, and in case (3) it would refer to the configuration-resemblance class.[[142]](#footnote-142) Thus, in controversial cases, a person’s *guiding* *conception* of that to which he *intends* to refer determines that to which his concept does refer. Yet, according to Loar, Fodor’s asymmetrical causal condition does nothing to distinguish or adjudicate between these mutually exclusive candidates for being the extension of Armand’s B concept.

In the case where more than one kind of object tends to cause a person to apply the same concept-type, Loar thinks that it is not necessary that applications of the concept to one kind of object are causally dependent, in any way, upon applications of the concept to the other kind of object. There may be no sense in which the causal relationship between one type of bird and Armand’s bird thoughts is dependent upon, or caused by, the causal relationship between the other type of bird and his bird thoughts. Moreover, Armand’s *intention* to refer exclusively to a certain type of object would determine this asymmetrical causal dependence, if it obtained. Of course, this intention itself involves semantically contentful concepts.

For a simpler example, imagine that a person is informed that he has been applying his demonstrative sheep-concept (SHEEP) to two distinct kinds of objects, e.g., sheep and goats. Fodor’s theory says that the person’s belief that *this is a sheep* is true only when applied to sheep if he would not use SHEEP to refer to non-sheep unless he used it to refer to sheep, but he *would* use SHEEP to refer to sheep even if he did not use it to refer to non-sheep. According to Loar, one would be able to account for the truth of this counterfactual only by assuming that the person has some *concept* of sheep *qua* distinct from goats, *prior* to this occurrence of the mental symbol SHEEP. (He may, for example, conceive of the thing as that which causes “that” kind of fuzzy, white, four-legged impression; and this impression would *not* include perceptions of horns or beards, as may be the case if goats were the true referents.) This “guiding conception” *of what* he means by SHEEP is necessary to determine what SHEEP represents. Moreover, the person must have some *intention* to refer to sheep and not to goats (or to any other non-sheep); and this intention *involves* his guiding conception. Since this intention has meta-semantic and representative content, Loar concludes that Fodor’s causal theory requires a concept to have objective representative properties and truth-values in order to *acquire* them.[[143]](#footnote-143)

When a person has a mental state involving a demonstrative concept, Loar thinks that one can only explain asymmetrical dependence condition for falsity if one assumes that the person’s thought is influenced by some guiding conception of the type of object to which she intends to refer. (Nor can one explain why the state meets Fodor’s counterfactual independence condition for truth without making this type of assumption.) Therefore, Fodor’s theory cannot demarcate the class of objects that a person’s occurrent belief represents from all other classes of objects without appealing to a distinct, meta-intentional belief that the person has. This is a belief that *my concept refers to* *that kind of thing*, where ‘that’ expresses a guiding conception that is a constituent of the mental state.

In the context of the current discussion, for a mental phenomenon to have an intentional feature is for it to represent some object (or class thereof) to the exclusion of all others, and for it to be evaluable as true or false. Thus, an occurrent concept must have truth-conditions in order for it to have intentionality. Fodor defines a concept’s referent and its truth-conditions in terms of causal conditions of counterfactual independence and asymmetrical dependence. Loar says that a concept cannot meet either of these conditions unless the occurrence of a meta-semantic (or meta-intentional) guiding conception causes the first concept to occur. Thus, such a concept cannot have intentional features without being caused by a distinct intentional phenomenon.

Moreover, this distinct intentional phenomenon, i.e., a guiding conception, determines what an occurrent concept represents exclusively. This act (of determining the extension of the occurrent concept) is, itself, an intentional event. Therefore, Loar concludes that Fodor’s causal theory fails to provide naturalistic and non-semantic conditions for a mental state to have intentional content. In addition, he concludes that Fodor’s theory implies a vicious regress, in which the causal conditions that determine the reference and truth-value of one concept cannot be met without the reference-determining causal influence of a distinct concept, which, in turn, has its own extension and truth-conditions. In other words, what it is for a mental state to represent something includes conditions under which that representation is false. But since one must appeal to the causal influence of representative features of a different state to explain how the first state meets Fodor’s conditions for falsity, his view fails as a *naturalisti*c explanation of intentionality: one which accounts for the intentional properties of thoughts in purely *non*-intentional terms.

C. Fodor’s *Mere Nomic Relation* Response

In *Psychosemantics,* in his reply to Loar in *Meaning in Mind*, and in his latest work, *Concepts: Where Cognitive Science Went Wrong*, Fodor makes what I call the ‘Mere Nomic Relation’ response to Loar’s objection. He bases this metaphysical argument upon a distinction between two kinds of relations involving occurrences of mental symbols: causal relations *per se*, and “nomic” (or “lawlike”) relations held by causal properties. Causal relations *per se* are just those whereby dogs, for example, cause occurrences of the mental symbol DOG in a counterfactual-supporting way. The relevant kind of nomic relation, in contrast, is a law that relates the *property* of *being* a dog to the property of *being* an appropriate cause of occurrences of DOG. (An appropriate cause is counterfactually independent of causal relations between any other type of object and occurrences of DOG.) Fodor also describes the nomic relation that makes it the case that DOG represents dogs exclusively as a relation that holds if the property of being an occurrence of DOG is *co-instantiated* with, or co-variant with, the property of being caused by dogs (appropriately).[[144]](#footnote-144) This is because for obvious theoretical purposes, the property of being a dog that causes occurrences of the mental symbol DOG appropriately must be coinstantiated with the property of being an occurrence of DOG that is caused by dogs appropriately.

According to Fodor, the nomic relations of his theory manifest certain mechanisms that are merely contingent. These mechanisms allow there to *be* the lawlike pattern (between the property of being a dog and the property of being an appropriate cause of an occurrence of the mental symbol DOG). Only the fact that this pattern of coinstantiation *obtains* makes up a semantic relation between dog concepts and actual dogs. Having the concept DOG “is constituted by there being the appropriate meaning-making lawful relations between instantiated doghood and one’s neural-cum-mental states. It’s *that*, not *how,* your mental structures contrive to resonate to doghood, that is constitutive of concept possession.”[[145]](#footnote-145) From this distinction, Fodor concludes that there is a sense in which Loar is right to conclude that intentional features are required to explain how an occurrence of a mental symbol meets [Fodor’s] causal conditions for truth or for falsity. Yet, the explanation requires these intentional features only at the level of the relevant causal mechanisms, and not at the *nomic* level.

For these reasons, Fodor rejects Loar’s charge that the causal theory of intentional content is viciously regressive. Fodor thinks that the *law* that the mental symbol meets his asymmetrical causal condition (or his counterfactual independence condition)— and not any causal or intentional *mechanism* that sustains or *explains* the truth of this law — determines the intentional features and truth-conditions of that mental symbol. In *Concepts*, Fodor concedes that “as often as not, the mechanisms whereby semantic access is achieved themselves involve the operation of intentional processes,” for “it’s patent that applying some concepts mediates applying others wherever semantic access is sustained by gossip, theoretical inference, expertise, deployment of instruments of observation, and the like.”[[146]](#footnote-146) In the case of gossip, for example, Fodor says, “Somebody may tell me things about dogs...and that too may cause me to think dog.” In other cases, mechanisms are *deliberately* engineered to make one’s concepts correlate with previously conceived classes of objects. The case of “dog bells,” for example, is set-up “so that a bell goes off when the dog shakes his head. If I know how things are rigged, hearing the bell may reliably cause me to think *dog*.”[[147]](#footnote-147)

It is clear that a set of distinct semantic relations are involved in the various beliefs, memories, and so on, which Fodor has in response to the gossip or alarms that he hears. It is also clear that these semantic relations affect the semantic relation in which Fodor’s thought represents dogs. This is because the presence of these thoughts causes Fodor to think about dogs. Nonetheless, Fodor thinks that the distinction between nomic relations and causal mechanisms prevents his theory from generating a vicious regress of intentional features. This is because the issue of how the relevant mechanisms achieve and sustain the nomic relation is a mere question of *engineering*, and not one of metaphysics. Therefore, this engineering issue does not prevent him from creating a reductive (or naturalistic) definition of the semantic relation between a mental symbol and a type of object, such as dogs.

In what follows, I argue that Loar’s circularity objection does stand; and that one can show that Fodor’s theory implies a vicious regress. This is because one must appeal to a specific, higher-order intentional feature of a person’s mind to explain why a causal relation between a kind of object and a mental symbol that occurs in her mind is counterfactually independent (or asymmetrically dependent). Thus, one must appeal to a distinct, meta-intentional state that a person has to justify the claim that any of her beliefs are true in virtue of being counterfactually independent (or false in virtue of being asymmetrically dependent). As explained earlier, a metaphysical definition of mental representation must include a definition of false mental representation, that is, an account of the truth-conditions of an intentional mental state. I will conclude that the naturalistic, non-semantic nomic relation with which Fodor defines mental representation can obtain only if a distinct, non-naturalistic, and *semantic* nomic relation obtains. This latter nomic relation includes an intentional property as a relatum; and it co-varies with, sustains, and explains the obtaining of the nomic relation with which Fodor defines intentionality.

Another way to look at the vicious regress that Fodor’s “Mere Nomic Relation” argument implies is as follows: one should consider the nomic relation between the property of being a sheep and the property of being an occurrence of the mental symbol SHEEP (that is caused in a counterfactually independent way) to be a *three* *place* nomic relation. This relation obtains between the property of being a sheep, for example, the property of being an occurrence of SHEEP that is caused in the appropriate way, and a higher-order intentional property of believing that *my SHEEP concept refers to sheep exclusively*.

My argument has the following form: Fodor claims to have reduced the intentional relation between a mental symbol and a kind of object to a naturalistic nomic (or lawlike) relation of coinstantiation between property X and property Y. Yet, in order to explain the fact that this naturalistic (or non-intentional) nomic relation of coinstantiation (N) obtains between property X and property Y, one must appeal to the causal influence of an intentional (meta-semantic) event upon property X. The necessity of the causal influence of this intentional event means that an intentional property Z must be instantiated if property X and property Y are coinstantiated. In turn, this implies that in order to account for the fact that a naturalistic nomic relation obtains between property X and property Y, one must appeal to the fact that a distinct nomic relation (N1) obtains between the intentional property Z, property X, and property Y.

Fodor’s view implies that the fact that nomic relation N1 obtains *makes it the case that* the mental symbol represents the kind of object exclusively. Also, the instantiation of the intentional (meta-semantic) property Z makes it the case that a person’s mind represents the intentional relation between the mental symbol and the kind of object. In order to define *this* higher-order intentional relation with a counterfactual causal independence condition, Fodor would need to explain the nomic relation between property Z and the first-order intentional relation between the mental symbol and the kind of object. He could only accomplish this through an appeal to a third nomic relation (N2). N2 would obtain between a third-order meta-intentional property Z1 and another relation. Therefore, Fodor’s attempt to define intentional relations reductively through counterfactual (and asymmetrical) causal conditions leads to a vicious infinite regress of higher-order intentional events and relations, which, obviously, are not naturalistic.

Why must one appeal to a person’s prior intentional states to justify the claim that any of her beliefs are false in virtue of being caused in an asymmetrically dependent manner? Why must one appeal to a person’s prior intentional states to justify the claim that any of her beliefs are true in virtue of being caused in a counterfactually independent manner? Fodor’s only examples of mechanisms that implement mental representation are of either perceptual processes or of those, such as gossip and theoretical inference, which are clearly driven by intentional features, i.e., prior beliefs and desires about the kind of thing that the mechanism will induce the subject to represent with her mental symbols. As he acknowledges, the processes through which these non-perceptual mechanisms bring about a thought clearly involve intentional states. This, I conclude, prevents the relevant *causal* relations between mental symbols and objects from being defined reductively or naturalistically.

I argue that intentional conditions are also required to explain the falsity of a thought in the case where the nomic relation (that constitutes an intentional relation) is implemented by *perceptual* mechanisms. The ultimate reasons that sheep, for instance, cause occurrences of the mental symbol SHEEP are going to be perceptual; and thus, some qualia that an individual has of sheep is the ultimate reason, i.e., the proximate cause, of her applying SHEEP to sheep. In cases where she applies the thought to *non*-sheep, therefore, some similarity between the qualia that she has of sheep and those that she has of non-sheep causes the false thought. Yet, since similarity is a *symmetrical* relation, this same principle would imply that some similarity between the two kinds of qualia causes her to apply SHEEP to sheep. It follows that goats and sheep cause occurrences of SHEEP for the same underlying perceptual reason; and neither relation is *asymmetrically* dependent upon the other.

Moreover, the person must have some belief that her SHEEP thought refers to some class of objects exclusively, in order to determine that the qualia she experiences in response to sheep (as opposed to those she has in response to goats or other non-sheep) are the cause of her SHEEP thoughts. It appears that in order to justify the claim that sheep and not goats are the proper referents of her thought, this class should be some natural kind, such as the kind that causes some qualia, or the biological class of sheep. As such, the naturalistic nomic relation that Fodor thinks makes up an intentional relation must co-vary with the instantiation of a higher-order intentional (and thus, non-naturalistic) feature.

It may be objected, however,that this line of reasoning about similarity relations among perceptual experiences only shows that given some *particular set* of thoughts, consisting, for example, of one application of SHEEP to a sheep and one application of SHEEP to a goat, the two applications are not related asymmetrically. Yet, it does not show that the two *general* types of causal relations, i.e., between qualia of sheep and SHEEP-thoughts and that between qualia of goats and SHEEP thoughts, are not related asymmetrically. One may think this because the reason that sheep as a class of objects *tend* to cause some particular individual to apply SHEEP to them involves both the features of the quale induced by sheep *and* some mental module that the experience activates. The characteristics of this module make the general type of causal relation between goats and SHEEP asymmetrically dependent upon the general type of causal relation between sheep and SHEEP.[[148]](#footnote-148)

D. Normal Conditions Constraints and the Perceptual Template Hypothesis

One way to make the above kind of argument (concerning similarity relations among perceptual experiences) is to assume that a person’s memories of the perceptual experiences that she has of a kind of object create perceptual templates. The differences between the relationships that these templates have with different kinds of objects create the relevant type of asymmetrical causal dependence. In other words, her initial perception of a sheep may cause a perceptual template to be established in her brain; and later triggerings of this template may cause the mental symbol SHEEP to occur in her mind. Subsequently, her visual experiences of sheep trigger the template under all normal conditions, but goats only trigger it occasionally (or rarely). One could conclude from this that sheep would cause SHEEP occurrences even if goats did not, but goats would not trigger SHEEP *unless* sheep did. In this way, one could explain the asymmetrical causal dependence of the goat-SHEEP relation upon the sheep-SHEEP relation at the perceptual level, without appealing to any prior intentional states of the person, such as her guiding conceptions. Thus, one may conclude, the asymmetrical causal dependence condition can explain the falsity of SHEEP-thoughts applied to goats without generating a vicious regress of intentional states.

My position, however, is that one who endorses the perceptual template response must define misrepresentation either in terms of some notion of “normal conditions” for triggering an occurrence of a mental symbol, or in terms of some statistical qualification. That is, one must either: (1) specify the characteristics and parameters for the perceptual conditions under which triggerings of the template will count as normal, or (2) define the proper referent of the mental symbol as being the most specific or exclusive class of objects that *usually* triggers the template. One could support a normalcy constraint with the intuition that certain *kinds* of perceptual environments are substandard, such as bad lighting or auditory conditions, blizzards, long-range viewing distances, and so on. In this case, the definition of normal conditions for the triggering of SHEEP would be conditions in which none of these substandard environments obtain. It follows that to justify the claim that goat-triggered SHEEP occurrences are asymmetrically dependent upon sheep-triggered SHEEP occurrences solely with the normal conditions qualification, one must claim that goats only trigger SHEEP occurrences under abnormal perceptual (or cognitive) conditions.

Yet, the problem with this approach centers on the question whether only non-sheep, such as goats, trigger the SHEEP occurrence under abnormal conditions. It appears to be impossible to specify the relevant normal causal conditions in a manner that succeeds in falsifying all applications of SHEEP to non-sheep without begging the question. The question would be begged by infecting the definition of normal conditions with preconceived notions of the distinction between the conditions under which an occurrence of the mental symbol is true and those under which it is false. Specifically, one might create a circular account by defining normal conditions as those under which sheep cause a mental SHEEP symbol to occur (through triggering the template), while defining abnormal conditions as those under which non-sheep cause it to occur. The fact remains that both sheep and goats may trigger the template under abnormal perceptual circumstances, since it cannot be that sheep would not usually trigger the template under these circumstances.

The assumption that sheep would not usually trigger the template under abnormal circumstances would imply that SHEEP thoughts were false when applied to sheep under these circumstances, which is impossible. Under such substandard conditions, a person’s belief that a given sheep *is* a sheep certainly meets the semantic standard of *truth*, although it may not meet an epistemological standard of knowledge. Therefore, normal conditions for triggering the perceptual template cannot be normal causal conditions. They would have to be conditions under which by definition, only sheep would trigger the template.

As mentioned above, one may try to use a statistical superiority qualification to supplement the “normal conditions” account of asymmetrical causal dependence. This would be to argue that goats and other non-sheep trigger the perceptual template for SHEEP thoughts only *rarely* and *because* of substandard perceptual (or cognitive) conditions, while sheep trigger the template the overwhelming majority of the time and under *any* perceptual conditions. Yet, in order to succeed, this proposal must appeal to the combination of some intrinsic feature of the template and some intrinsic feature of sheep, an essential connection which allows the relationship between sheep and this template to be the *independent* reason for the relationship between goats and the template. The mere statistical predominance of triggerings of the template (and thus, of the symbol SHEEP) is an arithmetical, and not a causal fact; and therefore, it cannot justify the claim that triggerings by non-sheep are caused by triggerings by sheep. Hence, this predominance cannot justify the claim that only applications of SHEEP to sheep are true. This is because even though sheep are more highly correlated than non-sheep with triggerings of the template (and thus, with the symbol SHEEP),[[149]](#footnote-149) the fact that goats *ever* trigger SHEEP implies that the kind of object most highly correlated with the template’s triggering is that of the disjunctive class sheep *or* goats.

In other words, suppose that sheep trigger ninety-eight percent of SHEEP symbols, while goats only trigger two percent. Then the disjunctive property *being either a sheep or a goat* is one hundred percent correlated (or coinstantiated with) the property of being a cause of SHEEP occurrences. The property *being a sheep*, in contrast, is correlated with the property of being a cause of SHEEP occurrences only to a degree of ninety-eight percent.[[150]](#footnote-150) Under such an assumption, SHEEP thoughts applied to either sheep or goats would be true. In order to determine that applications of a SHEEP symbol to goats are false, therefore, the perceptual template defense of Fodor’s theory must appeal to some unique feature of the causal connection between the template and sheep to *explain* the statistical predominance of SHEEP applications to sheep. Such a feature would also allow this connection to be the cause of the connection between the template and goats.

A proponent of the perceptual template hypothesis must appeal to some unique connection between sheep and some essential feature of the relevant perceptual template. He or she must argue that sheep somehow *fit* the template better, and much better, than any non-sheep (including goats).[[151]](#footnote-151) Goats may only trigger the template when substandard perceptual conditions make them look like sheep to the subject, for example. That is, goats only trigger it when these conditions induce goats to cause a perceptual experience in the subject that triggers the perceptual template for SHEEP because it is sufficiently similar to the experience that sheep cause. Yet, this quite plausible suggestion does not prevent an intentional feature from being an essential aspect of the nomic relation with which Fodor defines intentionality. For an intentional feature must still be present to explain why applications of SHEEP to goats are asymmetrically dependent upon such applications to sheep.

To understand this connection, consider what it is for an object to fit the perceptual template. It is ultimately some feature of the template, namely, the memory that it encodes of a specific set of phenomenal properties, which determines whether an object fits the template (and thus, causes a occurrence of SHEEP). Hence, an object fits the template through having the *extrinsic* feature of causing a certain kind of quale to which the template intrinsically responds.[[152]](#footnote-152) Since sheep and goats have this extrinsic causal feature, whatever it means for the former to fit the template “better” than the latter must imply that the causal relation between sheep and the template is the *reason* for the causal relation between goats and the template. The asymmetrical dependence of one causal relation upon another implies, in turn, that there is a different explanation for these causal relationships. Moreover, it implies that this explanatory difference *justifies* the claim that the former would occur independently of the latter, but not vice versa.

However, the only difference between the reasons that sheep and goats have such a causal property lies in the perceptual conditions under which they are induced to cause these qualia. The important issue is whether the difference between their external perceptual conditions justifies the claim that the former relation is (counterfactually) independent of the latter, but that the latter relation is asymmetrically *dependent* upon the former. Yet, the fact that sheep trigger the template under normal perceptual (and cognitive) conditions, while goats only trigger it under substandard perceptual (or cognitive) conditions does not mean that goats trigger the template only *because* sheep trigger the template.

Even were sheep never to have existed (or were they to cease existing), goats could still have the extrinsic property of triggering the perceptual template for SHEEP symbols under certain substandard perceptual conditions. This is because a perceptual experience with a goat under substandard conditions could bring about the initial *formation* of such a template. Subsequently, only sheep would trigger it under normal perceptual conditions. All that is required to initiate or trigger the template is to induce a certain set of qualia in the subject, namely, the ones for which the template encodes memories of phenomenal properties. The template is a purely perceptual and mechanical module which sets forth no evaluative criteria for its own triggering, such as standards for the quality of the lighting conditions under which it is triggered. Therefore, the proper referent of the SHEEP symbol in the perceptual template thought-experiment remains the class of objects that trigger the template by inducing the relevant phenomenal experience, which is the disjunctive class sheep *or* goat.

Ultimately, then, the only way to justify the claim that the causal relation between goats and SHEEP symbols is asymmetrically dependent upon that between sheep and SHEEP symbols is to say that sheep meet some *pre-conceived criterion* for being the true referents of the SHEEP. Even an appeal to the standard of normal perceptual conditions cannot succeed in establishing the falsity of SHEEP thoughts applied to goats unless the thinker has an implicit *conception* of the class of objects to which her SHEEP thoughts properly refer. An interpretation of this conception would determine the truth-value of these thoughts. She may believe, that her SHEEP thoughts refer to the natural kind of object that usually causes the relevant set of qualia, e.g., the biological species of sheep. In this case, the subject’s mistaken belief that a given perceptual experience was caused by an object belonging to the natural kind of object that usually causes it would be the reason that she applied a SHEEP symbol in response to a goat.

Most importantly, this same type of belief would be the reason for her to apply SHEEP in response to sheep, as well. The same kind of belief and the same kind of perceptual experience are jointly responsible for each application of SHEEP to *any* kind of object. Therefore, the truth or falsity of any of her SHEEP thoughts would be determined by whether the object to which she applies the thought has the properties that she conceives of herself as representing with that thought. Whether she applies SHEEP to something with the property of belonging to the natural kind that usually causes her to have “that” perceptual experience determines the truth-value of this thought.

E. Conclusion

In conclusion, Loar’s guiding conception objection appears to show that Fodor’s causal theory of content cannot provide a successful naturalistic reduction of intentionality. Specifically, Fodor cannot reduce the semantic relationship whereby one type of mental symbol represents one type of object exclusively to a causal relationship between occurrences of that type of symbol and members of the relevant class of object. This is because in order to determine what a concept represents, a naturalistic-causal theory would have to appeal to some non-naturalistic, *semantic* relationship the concept has. Therefore, Fodor’s attempt to reduce the semantic content of mental states to their causal relationships falls into a vicious regress.

I discussed Fodor’s Causal Theory of Content in order to consider a version of the externalist atomist theory of symbolic perspective on mental representation. Based upon my conclusion that this type of view leads to a vicious regress, I argue that the symbolic perspective on intentionality cannot satisfy the third explanatory requirement for successful theories of intentionality: to account for the ability of a mental state to be evaluated as true or false.

Chapter IV: Frege’s Theory of Senses

1. Frege’s Theory of Senses as the Source of Intrinsically Semantic Content for Propositional Attitudes

In the introductory chapter of this essay, I explained that there are two general kinds of theories of the intentionality of propositional attitude-instances: the intrinsic semantics and the extrinsic semantics perspective. In the preceding two chapters of this essay, moreover, I evaluated two versions of the latter perspective. I explored the Computational Theory of Mind, as a paradigm of the internalist holist theory of the intentionality of propositional attitudes. Specifically, I concluded that there is no way for this theory to associate symbolic constituted PA-types exclusively with a proposition as its semantic content, such that each PA-type can be said to have a unique set of logical implications and a unique and demarcated class of objects in the world as its extension. I also evaluated Fodor’s Causal Theory of Content as a paradigm of the externalist atomist theory of the intentionality of symbolic PA-instances. Yet, I concluded that this theory cannot explain how a type of mental symbol exclusively represents a demarcated set of objects in the world, such that the symbol is true only when applied to that kind of object, and false when applied to any other kind of object. Since the third requirement that I imposed upon a satisfactory theory of intentionality is to explain the ability of a mental state to be evaluated as true or false, no version of the symbolic theory of intentionality can succeed.

For this reason, in this chapter I begin to explore the intrinsic semantics position on intentionality. This general position is that the constituents of propositional attitude-instances provide those mental states with semantic content in virtue of their own essential features, and not in virtue of any contingent relation between these constituents and any thing else. The version of the intrinsic semantics perspective that I will evaluate in this chapter is Gottlob Frege’s theory of senses. According to Frege, senses are abstract, mind-independent, and intrinsically semantic objects, which provide intrinsically semantic content to mental states. I will first evaluate the ability of Frege’s theory to meet the second explanatory requirement that the introductory chapter imposed upon successful theories of intentionality: to explain the inferential failures of sentences that ascribe propositional attitudes to people.

The rest of the discussion will assume the general linguistic-semantic point of Frege’s resolution of the Substitution of Identicals puzzle. This resolution says that terms in the intensional contexts of sentences that ascribe propositional attitudes to people refer to intrinsically semantic entities. The discussion will also assume Frege’s general metaphysical point that mental states have intrinsically semantic content through connections with abstract objects.Yet, Frege’s metaphysical theory of *senses* as abstract, mind-independent, *and* intrinsically semantic objects fails to explain how intrinsically semantic contentcan belong to *concrete* mental states, which are internal to the mind. Thus, in the final chapter, I will put forth a metaphysical theory of concrete, mind-internal concepts as the only intrinsically semantic objects. I will argue that these intrinsically semantic constituent*s* of intentional mental states are essentially directed toward properties. This essential connection between concrete concepts and non-semantic abstract objects provides intentional mental states with intrinsically semantic content.

In “On Sense and Reference”[[153]](#footnote-153) and in “The Thought: A Logical Enquiry,”[[154]](#footnote-154) Gottlob Frege puts forth a semantic and metaphysical theory of a kind of abstract entity that he calls a ‘sense’. The philosophical motivation for his theory of senses is to resolve certain puzzles in the philosophy of language that have important metaphysical implications for the philosophy of mind. In this essay, I will analyze and evaluate Frege’s use of this theory to resolve a particular question in the philosophy of language, namely, the substitution of identicals (SI) puzzle. SI is the deductive inference rule that says that substituting co-referential terms in a sentence never affects the truth-value of the sentence. The puzzle concerning SI is that in what Frege called ‘oblique contexts’ of sentences, substituting one term for another term with the same referent may fail to preserve the truth-value of the sentence. While ‘the wife of Oedipus’ and ‘the mother of Oedipus’, for example, are co-referential terms, one may substitute the latter for the former in a certain kind of sentence and fail to preserve its truth. From ‘Oedipus believes that the wife of Oedipus is Jocasta’, for example, one cannot infer, by SI, ‘Oedipus believes that the mother of Oedipus is Jocasta’.

At the beginning of this essay, I argued that a successful theory of intentionality must give a satisfactory explanation of this phenomenon. Frege’s analysis of the SI problem is that one and the same term may have a different type of semantic function in different sentential contexts, such as oblique contexts. A term that refers to an individual thing under normal circumstances, that is, may not refer to that same individual when it occurs in an oblique context. According to this position, the term ‘Jocasta’, for instance, refers to *distinct* objects in the sentences ‘Jocasta is the Queen of Thebes’ and ‘Oedipus believes that Jocasta is the Queen of Thebes’. Frege introduces the notion of a sense to play two explanatory philosophical roles: First, senses function as the referents of terms in the “that” clauses of sentences of the latter kind, e.g., as the referent of the term ‘Jocasta’ or ‘the Queen of Thebes’[[155]](#footnote-155) in the sentence ‘Oedipus believes that Jocasta is the Queen of Thebes’. Second, senses function as the semantic constituents of the instances of propositional attitudes that such sentences ascribe, e.g., as the ingredient in virtue of which Oedipus’s belief that *Jocasta is the Queen of Thebes* is about Jocasta.

The basic *semantic* thesis of Frege’s theory is that the sense of a term roughly amounts to the *meaning* or conceptual content of the term, i.e., what is expressed by normal usage of the term. Such a meaning is not identical to the individual to which the term refers under normal use. Rather, the sense of a term is that which *determines* its referent in normal use; and there is thus a fundamental distinction between the sense and the referent of a term. The basic *metaphysical* thesis of Frege’s theory is that senses exist as abstract and mind-independent entities: things that have no physical constituents yet do not exist in any mind. In this chapter, I will explore Frege’s semantic and metaphysical arguments for these theses. Frege begins “On Sense and Reference” by posing the question why some kinds of identity sentences are informative, while other kinds are not. He asserts that a sentence of the form ‘a = a’, for instance, conveys no information, while one of the form ‘a = b’ does; and he uses his theory of senses as the conceptual contents (or meaning) of terms to explain this cognitive difference. Because we can attach different senses to co-referential singular terms, it is possible (and it often happens) that we fail to know that something that is picked out by one concept can also be designated by a second concept.

A way of understanding the cognitive difference between a sentence of the form ‘a = a’ and one of the form ‘a = b’ is to recall that identity statements of the form ‘a = a’ are trivial logical truths, so that any sentence with this form must be true. The sentences ‘Oedipus = Oedipus’, ‘The author of “Sense and Reference” = the author of “Sense and Reference”’, and so on, are examples of such trivialities. Presumably, then, Frege means that one could never learn something that one did not already know from reading or hearing such sentences, and thus, one can gain no information from them. Sentences of the form ‘a = b’, however, such as ‘Oedipus is the husband of Jocasta’ and ‘Frege is the author of “On Sense and Reference” ’ can, and do convey information. These are sentences with different terms on either side of the identity sign, e.g., ‘Oedipus’ and ‘the husband of Jocasta’. If such a claim is true, one often learns a previously unknown fact upon reading or hearing it.

Frege proposes to use the notion of a term’s having a sense to explain why we may learn something from sentences of this form, but nothing from sentences of the form ‘a = a’. More specifically, he uses the notion of a sense to answer two interrelated questions: (1) why can ‘a = b’ express non-trivial and previously unknown information; and (2) what kind of non-trivial information does it convey, i.e., what is the *difference* between the information provided by sentences of the form ‘a = b’ and those of the form ‘a = a’? The following passage expresses his view of how the notion of the sense of a term helps to explain why we can fail to know that a term applies to a given object:

The sense of a proper name is grasped by everybody who is sufficiently familiar with the language…to which it belongs; but this serves to illuminate only a single aspect of the reference, supposing it to have one. Comprehensive knowledge of the reference would require us to be able to say immediately whether any given sense belongs to it. To such knowledge we never attain.[[156]](#footnote-156)

Frege’s point here can be grasped by considering a proper name ‘a’ within a given language L. He assumes that every competent speaker of L grasps the sense of this proper name, and therefore, one should not interpret him to consider terms such as ‘John’ or ‘Erica’, which are considered “proper names” under current linguistic terminology. This is because such terms have no universally understood meanings in any language. Rather, Frege must intend terms such as ‘the Queen of Thebes’, which represent an individual through expressing a unique property.[[157]](#footnote-157) For the sense of such a term both guarantees that the term denotes a unique individual *and* that it is understood by everyone competent in a language.

The view that the sense of ‘the Queen of Thebes’, for example, “serves to illuminate only a single aspect of the reference” means that if the term has a referent, then the sense of the term only expresses information about some properties of that object. For example, the term does not convey the information that the object is also the mother of Oedipus, or the former wife of Laius, or five feet and nine inches tall. Since one never has knowledge of every single property belonging to the object, moreover, another term, such as ‘the mother of Oedipus’, can refer to it in virtue of expressing *another* one of its properties. This is why one may learn something new from claims of the form ‘a = b’, such as ‘The Queen of Thebes = the mother of Oedipus’, while one cannot learn anything non-trivial from claims of the form ‘a = a’, such as ‘The Queen of Thebes = the Queen of Thebes’.

Above, the cognitive and semantic intuitions behind Frege’s resolution of the puzzle of the informativity of some identity statements were explained. What of the metaphysical thesis that terms ‘a’ and ‘b’ in claims of the form ‘a = b’ are actually associated with real entities called senses? Frege justifies this view by answering the explanatory question: what is the *difference* between the information provided by sentences of the form ‘a = b’ and those of the form ‘a = a’? He has established that only the former type of sentence can convey non-trivial information, because only in that case is it possible for us to be ignorant of the fact that the terms one each side of the identity symbol have the same extensions. Yet, this ignorance is only possible, in turn, *because* we can think about one and the same object in different ways, and thus can attach different conceptual or cognitive contents to co-referential terms. It appears to follow that the difference in the information conveyed by sentence-forms ‘a = a’ and ‘a = b’ is a function of the difference in senses associated with the different terms contained in the latter type of sentence. Frege thus resolves the puzzle of the informativity of identity statements by claiming that two terms may refer to one and the same object, and yet do so by means of different senses or conceptual contents.

Frege uses his notion of a sense to resolve the Substitution of Identicals puzzle in the Philosophy of Language. The validity of SI as a rule is based in the logical intuition that if two terms, ‘a’ and ‘b’, refer to one and the same object, the substitution of ‘b’ for ‘a’ in a sentence must preserve the truth-value of the sentence. The question why sentences of the form ‘a = b’ can be informative, moreover, clearly involves sentences that use co-referential terms ‘a’ and ‘b’ to express an identity relation between object a and object b. This leaves us with a puzzle why the substitution of ‘a’ and ‘b’ in such contexts fails to preserve truth. Frege’s solution is to claim that normally co-referential terms ‘a’ and ‘b’ can refer to distinct senses in intensional contexts, including those ascribing beliefs that *a = b* or that *a is b*.

The notion of an intensional context pertains to certain kinds of complex (but non-truth functional compound) sentences that contain simpler sentences as subordinate clauses.[[158]](#footnote-158) These clauses create a sentential context within the larger sentence, and are governed by ‘that’ operators, such as ‘is possible that’, ‘believes that’, ‘desires that’, and so on. Intensional sentences include modal sentences, such as ‘It is necessary that 1 is the smallest prime number’; and sentences ascribing propositional attitudes to a subject, such as ‘Oedipus believes that Jocasta is the wife of Oedipus.’ The portion of the intensional sentence that follows ‘that’ is its *intensional context*.

Having explored the intuitive basis for Frege’s use of the semantic notion of a sense to resolve the SI problem, I will now consider the complete semantic theory of indirect sense and reference that underlies his position that terms refer to their senses in intensional contexts. (I will later discuss his further arguments for this view, as well as his metaphysical arguments for construing senses as abstract and non-mental objects.) Frege constructs this semantic theory around his distinction between the *customary* and the *indirect* senses of a term, and, correlatively, between its customary and its indirect referents. Frege proposes that intensional sentences are a species of indirect speech, and he claims that the terms involved in indirect speech refer to the senses that they have in normal speech. That is, terms in indirect speech refer to their customary (i.e., normal or direct) *senses*. Yet, these terms must also have *indirect* senses: the senses that they have, i.e., their meanings, in *indirect* speech.

Frege’s notions of the indirect sense and reference of a term derive from his conception of the role of terms (as proper names) in direct speech. His functional definition of a proper name is that which, in direct speech, “expresses its sense, stands for or designates its reference.”[[159]](#footnote-159) When you use a proper name (or a term in general), Frege holds, there are three kinds of things to which you can use it to refer: (1) You can use a word to speak of what the word refers to, or (2) you may use it to talk about the word itself, or (3) you can use it to speak of the sense of the word, i.e., its meaning. In other words, by using the term ‘ball’ you may refer to the referent of the word ‘ball’, namely, balls, as in, for example, ‘This ball is red’. Alternatively, you may refer to the word ‘ball’ itself, e.g., “‘Ball’ has four letters.” Finally, you may refer to what the word ‘ball’ means, i.e., its sense. This last kind of semantic action may be accomplished through indirect, i.e., reported or quoted speech, or through discussion of this meaning itself.[[160]](#footnote-160)

According to Frege, sentences as well as terms function as proper names (for reasons I explore below), and thus a sentence itself (or its sense) can be discussed in indirect speech. An example of this kind of indirect speech is the following sentence: ‘Pericles said that all good things flow to the city’. Words in reported speech, such as those of the component sentence ‘All good things flow to the city’, do not have their ordinary (or customary) reference. That is, in such indirect contexts, words denote their customary senses, and not their normal extensions, such as good things and cities.

The customary sense of a term is its conceptual content in ordinary speech, while the customary referent of a term is the object that it is semantically directed toward — the unique object that it denotes ⎯ in ordinary speech. Thus, in the direct speech sentence, ‘Jocasta is the wife of Oedipus’, the term ‘the wife of Oedipus’ has its customary sense. This is the idea that the term normally expresses: perhaps the concept of an individual who uniquely has the property of being married to the individual termed ‘Oedipus’. The customary *referent* of ‘the wife of Oedipus’, in contrast, is the individual entity that it uniquely singles out in the world, namely, Jocasta. Terms occurring in intensional contexts refer to their customary senses, according to Frege. In this sentence, therefore, the term ‘the wife of Oedipus’ does not refer to Jocasta, but rather to its customary sense.

Frege infers that terms in intensional contexts must not only have indirect referents, but must also have indirect senses. The indirect sense of a term *t* (when *t* is used in indirect speech) is the direct meaning, or normal sense of a distinct term *t*1, where *t*1 is a term that refers to the normal sense of *t*. Terms in intensional contexts must have indirect senses, because if such terms refer indirectly to their *customary* senses, then these senses cannot be identical to the senses of the terms *as they are used* in such intensional contexts. In other words, if the referent of ‘the wife of Oedipus’ in ‘Oedipus believes that Jocasta is the wife of Oedipus’ is the concept associated with the term ‘the wife of Oedipus’, then the concept that the term ‘the wife of Oedipus’ expresses in this intensional sentence must be the concept expressed by a term like ‘the concept (or sense) Oedipus associates with the term ‘the wife of Oedipus’’. This is because Frege holds that the sense of a term is *distinct* from its referent. Yet, if, in the sentence ‘Oedipus believes that Jocasta is the wife of Oedipus’, the term ‘the wife of Oedipus’ did not have an indirect sense distinct from its direct or normal sense, then it would both refer to *and* express its normal sense. The problem is that this would mean that its sense and its referent were identical

B. Russell’s Alternative to Frege’s Semantic Arguments for Senses

In “On Denoting,”[[161]](#footnote-161) Bertrand Russell argues that there is no need to appeal to the existence of senses in order to resolve the problem of the inferential failure of substitutivity in intensional contexts. Russell argues that the true logical and semantic analysis of sentences containing singular terms reveals that such terms do not function as distinct semantic units, and thus do notdenote individuals:

[T]he principle of the theory of denoting I wish to advocate...[is] that denoting phrases never have any meaning in themselves, but that every proposition in whose verbal expression they occur has a meaning. The difficulties concerning denoting are, I believe, all the result of a wrong analysis of propositions whose verbal expressions contain denoting phrases.[[162]](#footnote-162)

Russell puts forth an analytical reduction of singular terms, such as definite descriptions like ‘the mother of Oedipus’, which we intuitively consider to have the semantic function of referring to individual things in the world. Contrary to appearances, he says, such (presumably) “denoting phrases” do not occur in the true logical and semantic analysis of sentences that seem to contain them, like ‘The wife of Oedipus is a mother’. (This holds independently of whether the sentences contain intensional contexts.) Russell says the following concerning such terms:

By a ‘denoting phrase’ I mean such as any one of the following: a man, some man, any man, every man, all men, the present King of England... the centre of mass of the solar system at the first instant of the twentieth century.... Thus, a phrase is denoting solely in virtue of its form. We may distinguish three cases: (1) A phrase may be denoting, and yet not denote anything; e.g., ‘the present King of France’. (2) A phrase may denote one definite object; e.g., ‘the present King of England’ denotes a certain man. (3) A phrase may denote ambiguously; e.g., ‘a man’ denotes not many men, but an ambiguous man. [[163]](#footnote-163)

When Russell analyzes sentences containing denoting phrases logically and semantically, he finds that these phrases do not refer to any individual thing. Rather, a sentence in which such phrases occur actually expresses the proposition that a certain set of properties is uniquely instantiated. This is the set of properties through which a definite description would normally be considered to denote a unique individual. In Russell’s analysis of the sentence ‘The wife of Oedipus is a queen’, for example, there is no occurrence of the singular term ‘the wife of Oedipus’ to refer directly to any individual object, like Jocasta. Rather, there is a conjunction of propositions that jointly assert that the property of being a wife of Oedipus is uniquely instantiated by an object that is a queen.

Specifically, sentences containing (implicit) definite descriptions have the logical and semantic import of expressing a conjunction of three kinds of propositions: The first is an existential proposition, which says that a given property is instantiated. Hence, such existential propositions (indirectly) say that something *exists* which has that property, e.g., that *something is a wife of Oedipus*.[[164]](#footnote-164) The second is a uniqueness proposition, which says that the relevant property is instantiated by exactly one thing, e.g., that *at most one thing is a wife of Oedipus*. Finally, Russell thinks that sentences containing (implicit) definite descriptions also express a kind of proposition that *subsumes* the class of things that instantiate the relevant property under another class. This proposition says that everything that has the property of being a wife of Oedipus also has another property, such as the property of being a queen: *everything that is a wife of Oedipus is a queen*.[[165]](#footnote-165) The ultimate semantic import of this conjunction of propositions, and thus the ultimate import of the original sentence, is to express an existential quantification such as: *there exists exactly one thing that is a wife of Oedipus, and that thing instantiates the property of being a queen*.[[166]](#footnote-166)

Russell objects to Frege’s resolution of the SI puzzle concerning intensional sentences with two interrelated positions. First, singular terms (including definite descriptions) do not have referents if they are disguised or concealed, and this includes descriptions in intensional contexts. Second, any sentence of sufficient logical complexity ⎯ whether intensional or not ⎯ that contains such a description is semantically ambiguous. For it can be interpreted to express two logically *distinct* conjunctive propositions: one in which the implicit definite description has what Russell calls “primary occurrence” and one in which it has “secondary occurrence.”

Russell’s distinction between the primary and the secondary occurrence of a definite description in a complex sentence can be understood as analogous to the distinction between the wide and narrow scope of quantifiers in predicate logic. To understand this distinction, consider the non-intensional sentence

(P) The wife of Oedipus is not Roman.

One can interpret ‘the wife of Oedipus’ to have either primary or secondary occurrence in P. Which type of occurrence this term is interpreted to have determines *which* conjunction of existential, uniqueness, and subsumption propositions P is understood to express. On one interpretation of P, it expresses this conjunctive proposition:

(P1) Something is a wife of Oedipus, there is at most one wife of Oedipus, and nothing that is a wife of Oedipus is a Roman.

The term ‘the wife of Oedipus’ has primary occurrence in P, because the ultimate import of P1 is to express an existential quantification such as: *there exists exactly one thing which is a wife of Oedipus and that thing instantiates the property of not being a Roman.*

On another interpretation of P, the definite description ‘the wife of Oedipus’ has a secondary occurrence. In this case, P expresses a different proposition:

(P2) It is not the case that: something is a wife of Oedipus, there is at most one wife of Oedipus, and everything that is a wife of Oedipus is a Roman.

Clearly, P1 and P2 can have different truth-values. While P1 implies P2, there is a possible case in which the latter could be true while the former is false. For P2 expresses the negation of a conjunction: the claim that Oedipus has a unique mother *and* this mother is Roman. The negation could be true if Oedipus had no mother (and thus no unique mother); and yet this state of affairs would make P1 false. For P1 asserts that there *exists* exactly one object that is Oedipus’s mother.

The reason that the definite description above may have primary or secondary occurrence is that it may be subordinated to another clause, i.e., ‘is not Roman’, which contains a sentential operator, i.e., ‘not’. The presence of operators often makes a sentence logically complex. ‘Not’ is obviously a logical operator for negation, but there are other kinds of sentential operators, including the intensional operator ‘believes that’ in ‘Oedipus believes that the wife of Oedipus is a queen’, for example. The governance of implicit definite descriptions by such intensional operators, therefore, also provides each intensional sentence of sufficient complexity with two logically distinct interpretations. Moreover, note that one such interpretation (P2) does not imply that there is any referent of the term ‘the wife of Oedipus’. I will show below that one interpretation of an intensional sentence will fail to imply that there is a referent of an (apparent) singular term in its intensional context.

The puzzle of the substitutivity of co-referential terms in intensional contexts depends upon such terms having reference. For it is only the assumption that ‘the wife of Oedipus’ and ‘the mother of Oedipus’ denote one and the same object that generates the appearance of a logical problem with our inability to infer ‘Oedipus believes that the mother of Oedipus is a queen’ from ‘Oedipus believes that the wife of Oedipus is a queen’. Clearly, Russell thought that his reductive theory of descriptions could solve the SI puzzle:

A logical theory may be tested by its capacity for dealing with puzzles, and it is a wholesome plan, in thinking about logic, to stock the mind with as many puzzles as possible, since these must serve the same purpose as physical experiments in physical science. I shall therefore state three puzzles which a theory as to denoting out to be able to solve; and I shall show later that my theory solves them...

He then continues to list the first puzzle, a clear instance of the substitutivity problem:

If *a* is identical with *b*, whatever is true of the one is true of the other and either may be substituted for the other in any proposition without altering the truth or falsehood of that proposition. Now George IV wished to know whether Scott *was* the author of *Waverley*. Hence, we may substitute *Scott* for *the author of ‘Waverley’*, and thereby prove that George IV wished to know whether Scott was Scott. Yet, an interest in the law of identity can hardly be attributed to the first gentleman of Europe. [[167]](#footnote-167)

Russell’s theory provides a way to resolve the puzzle by denying the truth of one of its assumptions, namely, that terms in intensional contexts *have* reference. For singular terms refer to nothing in the relevant sentential contexts, because they never refer to anything. Russell puts it as follows:

The puzzle about George IV’s curiosity [the SI puzzle about the sentence ‘George IV wished to know whether Scott was the author of *Waverley*’] now seems to have a very simple solution. The proposition ‘Scott was the author of *Waverley*’... does not contain any constituent ‘the author of *Waverley*’ for which we could substitute ‘Scott’. [[168]](#footnote-168)

Russell begins to tell us that the substitution of a term for an identical does not even appear to fail to preserve the truth of this primary occurrence interpretation, because all that is being substituted is the *term*, not its meaning: “This does not interfere with the truth of inferences resulting from making what is verbally the substitution of ‘Scott’ for ‘the author of *Waverley*’, so long as ‘the author of *Waverley*’ has what I call a primary occurrence in the proposition considered.”[[169]](#footnote-169) Yet, in order to work, Russell’s theory must establish that definite descriptions do not refer, but rather occur only as parts of sentences with no independent semantic content.

So far, I have explained the difference between the primary and secondary occurrence of definite descriptions; and I have explained how these different semantic interpretations provide a non-intensional sentence with two distinct sets of truth-conditions. The following will examine the details of Russell’s solution.

How does Russell’s theory of descriptions apply to the SI problem? Intensional sentences often contain both sentential operators and definite descriptions. For example, the sentence ‘Oedipus believes that the wife of Oedipus is Jocasta’ contains both the sentential operator ‘believes that’ and the term ‘the wife of Oedipus’.[[170]](#footnote-170) These intensional contexts thus pose problems of ambiguity. Russell’s theory is intended to solve these problems *without* invoking the Fregean distinction between the sense (or the ordinary conceptual meaning)[[171]](#footnote-171) and the denotation of a term.

The intensional sentence ‘Oedipus believes that the wife of Oedipus is a queen’ (O) appears to be composed, in part, of a distinct sentence embedded within the intensional context of O. This distinct sentence is ‘The wife of Oedipus is a queen’ (M). M involves both the definite description ‘the wife of Oedipus’ and the predicate ‘is a queen’. According to Russell, the true logical analysis of M is the conjunction ‘There is exactly one wife of Oedipus and any wife of Oedipus is a queen’. Because O involves not only these descriptions, but an intensional sentential operator ‘believes that’, O is of sufficient logical complexity to have two distinct semantic interpretations. In one interpretation, the definite description has primary occurrence, and in the other interpretation, it has secondary occurrence. The secondary occurrence interpretation of O is simply that Oedipus believes the semantic content of M to be true:

(O2) Oedipus believes that both (a) there is exactly one wife of Oedipus and (b) any wife of Oedipus is a queen.

The primary occurrence interpretation of O, in contrast, is:

(O1) There is exactly one wife of Oedipus and Oedipus believes that she is a queen.

Clearly, O1 is an interpretation under which the definite description of the embedded sentence M has wide scope. Under O2, however, this definite description has narrow scope, and therefore, the propositional attitude operator ‘believes that’ governs it. In the case where the description has a secondary occurrence, Russell thinks that the intensional operator of the larger sentence O, i.e., ‘believes that’, includes the predicate ‘*x*-is-a-wife-of-Oedipus’ within its scope.

The fundamental point to recognize about Russell’s resolution of the substitutivity problem with intensional sentences concerns the truth-conditions of their secondary occurrence versions. The conjoined existential, uniqueness, and subsumption propositions which Russell considers to constitute the true content of the embedded sentence M are not being asserted as *true* in O2. For O2 does not say that there actually *exists* such a thing. O2 can be true even if there exists no object that satisfies the embedded definite description ‘the wife of Oedipus’. Its essential semantic function, therefore, cannot be to *refer* to such an object. Therefore, according to Russell, it makes no sense in such a context to attempt to substitute a *co-referential* term, e.g., ‘the mother of Oedipus’, for the embedded definite description ‘the wife of Oedipus’. This is because such descriptions refer to no individual. This is why the substitution of an (apparently) co-referential term in such a context appears to fail to preserve truth, because of the false intuition that such terms actually have reference.

In the case where the implicit definite description is interpreted to have primary occurrence, however, the intensional sentence must be interpreted to say that some object actually exists which is believed to have a certain property. Thus, O1 asserts the *existence* of a thing, which uniquely instantiates certain properties, i.e., the properties *x*-is-a-wife-of-Oedipus and *x*-is-believed-by-Oedipus-to-be-a-queen.

It is clear that under O1, applications of SI to O will preserve its truth-value. For suppose that one substitute the co-referential term ‘the mother of Oedipus’ for ‘the wife of Oedipus’ in O. This results in the sentence:

(O3) Oedipus believes that the mother of Oedipus is a queen.

Yet, if the description ‘the mother of Oedipus’ has a primary occurrence in O3, then it is to be analyzed as:

(O4) There is exactly one mother of Oedipus and Oedipus believes that she is a queen.

However, since the wife of Oedipus = the mother of Oedipus, and it is she, according to (O1) and (O4), whom Oedipus believes to be a queen, these two sentences will not differ in truth-value. Therefore, SI is valid when applied to a definite description having a primary occurrence in an intensional sentence. Moreover, only intensional sentences that contain secondary occurrences of definite descriptions pose a problem for SI.

Russell’s solution is that applications of SI in these cases are illegitimate in *principle*. For secondary occurrence interpretations do not even express the existence of an *apparent* referent, e.g., the actual wife of Oedipus, for the relevant term that is a candidate for substitution, e.g., ‘the wife of Oedipus’. The *apparent* inferential failure of SI is illusory and spurious, according to Russell. He concludes that there is no need to appeal to the metaphysical notion of senses to explain the inferential failure of substitutivity in intensional contexts. For the problem is generated by the false assumption that singular terms, e.g., the mother of Oedipus’ and ‘the wife of Oedipus’ may refer to one and the same individual thing, such as Jocasta. Thus, abstract objects, or senses, are not required to resolve of the problem applying SI in intensional contexts.

Nonetheless, in my view, one still needs to appeal to the existence of some type of abstract object to resolve the SI puzzle; and one still needs to hold that terms in intensional contexts of propositional attitude-ascriptions refer to intrinsically semantic entities. I do not endorse Frege’s *metaphysical* view that such a term refers to a type of object that is *both* abstract and intrinsically semantic. Yet, I will argue that Russell’s semantic theory of descriptions and of intensional sentences fails to show that the SI puzzle can be resolved without appeal to Frege’s *semantic* view that terms in intensional contexts refer to their (ordinary) conceptual meanings. For intensional sentences ascribe PA-instances to people; and thus the truth-conditions of these sentences involve the existence of the semantic *constituents* of these PA-instances. Intuition as well as analysis reveals that these constituents are concepts that the subject considers to apply to some set of objects, and that these constituents express, in some sense, properties that the subject considers to be instantiated by some set of objects. (The ordinary conceptual meanings of these constituents accomplish this semantic feat.)

The truth-conditions of intensional sentences thus involve the existence of mind-internal concepts, which, according to the assumptions of this essay, are intrinsically semantic. Even if, as Russell says, terms in the intensional contexts of these sentences do not function by referring to individuals, some of them must function by referring to these conceptual meanings or by expressing the properties that these meanings involve. These terms must be the predicates that constitute definite descriptions, and the function of these terms must be relevant to the solution of the SI puzzle.

It must be the fact that normally co-referential terms express different sets of properties that creates the apparent failure to substitute one such term for another in an intensional sentence and preserve its truth. For example, it must be the fact that ‘the wife of Oedipus’ expresses the property of being a wife of Oedipus, while ‘the mother of Oedipus’ expresses the property of being a mother of Oedipus, which explains why the substitution of one for the other in the sentence ‘Oedipus believes that the wife of Oedipus is a queen’ appears to fail to preserve its truth. Yet, Russell’s analysis of the SI puzzle fails to address the question of the mind’s grasping of these properties, which is of profound and central importance, since a PA-ascription says that some mind involves some concepts that express such properties. This question must be answered in order to account for the truth-conditions of such sentences. Therefore, this answer is required to justify any view concerning the success or failure of applications of SI to preserve their truth.

In the absence of any argument to the contrary, Frege’s view that terms in intensional contexts refer to their ordinary conceptual meanings meets every intuitive and philosophical requirement for providing this explanation ⎯ even should these turn out to be the meanings of predicates instead of the meanings of terms for individuals, and even if these meanings are not identifiable as “senses,” which are abstract, mind-independent, *and* intrinsically semantic. As has been explained, Frege’s semantic view is that through referring to a sense, a term in an intensional context of a sentence is semantically connected to the object (or class thereof) to which that term ordinarily refers. Moreover, a later section will explain the metaphysical view that this semantic connection occurs through an exclusive relation between an intrinsically semantic entity and some set of properties uniquely instantiated by the object (or class thereof).

Russell has successfully established that from a semantic perspective, there are two distinct interpretations of the meaning of an intensional sentence: distinguished by whether the definite description has primary or secondary occurrence. Yet, the former is logically irrelevant to my philosophical concern since it implies, for example, that there exists a wife of Oedipus, while this is not implied by interpretations under which SI appears to fail. Russell may also have established that terms in intensional contexts do not refer to individuals;[[172]](#footnote-172) from which it follows that the relevant propositions do not imply the existence of any individual or of any *single* sense (or intrinsic conceptual meaning) corresponding to descriptions such as ‘the wife of Oedipus’. In order to succeed in resolving the SI puzzle without the Fregean view that some terms in intensional contexts refer to senses (or intrinsic conceptual meanings), his theory must be able to interpret the truth-conditions of these sentences without the existence of senses (or intrinsic conceptual meanings). To show that Russell cannot succeed, I will give an intuitive argument and a metaphysical argument concerning the truth-conditions of intensional sentences.

Intuitively, what it means to say that *x believes that Fa* is that *x has some conception involving a property F being instantiated by an individual a*. Since Russell’s (secondary) interpretation of ‘*x* believes that the F is an H’ is:

(Φ) *x* believes that something is a unique F and is (an) H

from an intuitive perspective, Φ cannot be true unless *x* has in mind a conception involving the properties F and H.

Suppose, as is required to generate the SI puzzle, that the unique F and the unique G are one and the same object *y*. Intuition says that the reason that Φ does not *entail* that *x believes that something is a unique G and that thing is an H* must be that *x* does not *conceive* of *y* as having the property G. Let Ψ be the secondary interpretation of ‘*x* believes that the G is an H’; the only difference between Ψ and Φ is that the latter includes the predicate ‘F’, while the former includes the predicate ‘G’. Thus, the reason that what Russell terms the “verbal”[[173]](#footnote-173) substitution of ‘the G’ for ‘the F’ in Φ does not preserve the truth of Φ is that the predicate ‘G’ refers to *or* expresses a distinct property than the predicate ‘F’. Note that this intuitive resolution does not require that the terms ‘the F’ and ‘the G’ refer to *individuals* in the relevant sentences. Therefore, it does not require that they function as discrete semantic units at all; and it is consistent with Russell’s theory of descriptions as non-referring linguistic entities.

The intuitive conclusion above was that the difference between the semantic content of Φ and Ψ is the semantic content of the terms ‘F’ and ‘G’. Now the component terms ‘F’ and ‘G’ occurring in the complex terms ‘the F’ and ‘the G’ do not refer to individuals, but these components do contribute semantically to the sentences in which they occur. In fact, Russell says that such terms are “broken up” in the true semantic analysis of such sentences.[[174]](#footnote-174) If these components referred to individuals, then the complex terms of which they were components would have referred to individuals. Therefore, they must have the semantic function of referring to properties or of referring to some conception belonging to the subject *x*. This conception involves properties ‘F’ and ‘G’, e.g., ‘wife of Oedipus’ and ‘mother of Oedipus’. The existence of and the distinction between these referential or connotative relationships must be the explanation of the fact that the “verbal” substitution of ‘the G’ for ‘the F’ in an intensional sentence fails to preserve its truth.

According to Russell, the semantic content of an intensional sentence containing a definite description is a complex proposition. The content of O2, for example, is that *Oedipus believes that something is a unique wife of Oedipus and that thing is a queen*. While Russell does not think that the non-intensional sentence ‘The wife of Oedipus is a queen’ occurs as a discrete, unified embedded sentence in O, it is clear that the proposition *something is a unique mother of Oedipus and that thing is a queen* is a discrete, unified component of O2. For Oedipus must bear the *believes that* relation to this smaller proposition if O2 is true. Thus, considering our earlier analyses, O2 seems to imply either that the mind of Oedipus has contact with the properties *is-a-mother-of-Oedipus* and *is-a-queen*, or that he has some conception of the predicates that express or refer to these properties. Therefore, Russell’s analysis of the content of intensional sentences seems to imply that the existence of senses of predicates (as intrinsic conceptual meanings) is part of the truth-conditions of these sentences.

Senses as intrinsic meanings of predicates are not merely, as a metaphysical fact, part of the truth-conditions of intensional sentences. They must also be part of the semantic function of *some* terms in intensional contexts (albeit not unified definite descriptions) to refer to such senses or meanings. For even Russell’s resolution of the SI problem requires that such senses or meanings have a semantic relationship with terms in intensional contexts. Even though such terms do not (at least directly) refer to mind-external individuals, it is clear that their “broken up” components are semantically connected to properties, i.e., they either express, in some way, or refer to, properties.

Let us assume, from an intuitive perspective, that an individual’s sense of a term is her conception involving some set of properties or features that she associates with the referent of the term. Therefore, Oedipus’s sense of the combined components of the term ‘the wife of Oedipus’, i.e., ‘wife of’, ‘Oedipus’, and ‘the’ is his conception involving some set of properties {p1, p2, p3,  ...pn} that he considers something to instantiate uniquely.

On the one hand, O2 could be interpreted to imply that Oedipus’s mind gets directly in contact with the properties *is-a-wife-of-Oedipus* and *is-a-queen*. Whatever this metaphysical relationship (of a mind directly contacting a property) would be, O2 implies that this grasping relationship is instantiated. This implies that terms in intensional contexts refer to these properties. In other words, if O implies that Oedipus’s mind gets directly in contact with the property *is-a-wife-of-Oedipus*, then it expresses this grasping relationship semantically. It can only do so by way of the term ‘the wife of Oedipus’ referring to the property *is-a-wife-of-Oedipus*.

If the components of the term ‘the wife of Oedipus’, i.e., ‘wife (of)’ and ‘Oedipus’ succeed in referring to or expressing the property *is-a-wife-of-Oedipus* in the intensional context of O, however, then such reference is relevant to the true resolution of the SI problem. For if this referential relationship occurs, then the components of term ‘the mother of Oedipus’ ⎯ ‘mother of’ and ‘Oedipus’⎯ in the intensional sentence ‘Oedipus believes that the mother of Oedipus is a queen’ must combine to stand in the same type of referential relationship with the property *is-a-mother-of-Oedipus*. If the components of the terms ‘the mother of Oedipus’ and ‘the wife of Oedipus’ combine to bear referential relationships toward these properties in the relevant intensional sentences, then this fact simply *must* be relevant to the question why the merely “verbal” substitution of former term for the latter in the sentence O may fail to preserve its truth. (The obvious explanation is that the terms combine to refer to distinct properties and not to any individual, let alone one and the same individual (Jocasta)).

On the other hand, suppose that the components of the term ‘the wife of Oedipus’ combine to refer to the property *is-a-wife-of-Oedipus* through Oedipus’s *sense* of these component terms, that is, through his conception involving these properties. Note that a person’s occurrent concept of x is not identical to the *Fregean* sense of a term through which he or she refers to x. For a sense of a term for x is a mind-independent and abstract entity, while an instance of a concept of x is a mind-internal and concrete entity, i.e., a constituent of an occurrent mental state. Yet, according to Fregean metaphysical assumptions, Oedipus’s occurrent concept of the property *is-a-wife-of-Oedipus* must, nonetheless, represent this property through the direction of this concept toward a mind-independent sense; and thus, if the term ‘the wife of Oedipus’ refers to a set of properties through a sense it must also do so through a mind-dependent concept, and vice-versa.

Thus, it nonetheless follows that the components of the term ‘the mother of Oedipus’ combine to refer to the property *is-a-mother of Oedipus* through his conception of these properties, i.e., through Oedipus’s sense of the predicate ‘mother’, etc.[[175]](#footnote-175) In this case, the combined effect of each set of component terms is to refer to distinct senses of distinct predicates. For it is just as legitimate to consider a concept of a property to be a constituent of an intentional mental state as it is to consider a concept of an individual to be such a constituent. Analogously, it is just as legitimate to consider a sense of a predicate to be involved in the truth-conditions of a propositional attitude ascription as it is to consider a sense of a term for an individual to be so involved. Under this interpretation, more importantly, one component of the term ‘the wife of Oedipus’, i.e., ‘wife’, is referential; and one component of the term ‘the mother of Oedipus’, i.e., ‘mother’, is referential. However, these terms are not co-referential, for they refer to different properties. The fact that these terms have this referential difference in the intensional contexts simply must be relevant to the why substitutivity of the latter for the former in ‘Oedipus believes that the wife of Oedipus is a queen’ fails to preserve truth.

O2 essentially ascribes a semantic relationship between the mind of Oedipus and the properties of being a wife of Oedipus and of being a queen. That is, the sentence intrinsically states that Oedipus has some concept that essentially involves these properties.

The intuitive account of an individual’s sense of a term just is his concept involving the properties to which the term refers. Therefore, even the view according to which the fragmented components of singular terms refer directly to unique properties implies two things about the truth-conditions of PA-instances. First, they involve the existence of intrinsically semantic entities that essentially connect with properties, and second, they involve a semantic relationship between these intrinsically semantic entities and the components of definite descriptions in intensional contexts. Russell’s logical analysis of the semantic value of intensional sentences ascribing PAs fails, therefore, to explain the SI puzzle independently of the Fregean view that (at least) *some* terms in intensional contexts refer, in some fundamental way, to intrinsically semantic entities.

At this point in the discussion, it seems clear that intrinsically semantic entities are needed to explain the inferential failure of SI in intensional sentences. Nonetheless, the question remains, what *types* of things are these intrinsically semantic entities? While it is intuitive to think that the sense of a term is the mental *concept* that an individual associates with her use of the term, Frege gives good reasons for thinking that senses are abstract entities that exist *outside* the mind. In what follows, I will evaluate Frege’s arguments for this metaphysical view of senses as intrinsically semantic, abstract, and mind-independent objects.

C. The Metaphysics of Fregean Senses

What are the intuitions behind Frege’s view of senses as abstract, mind-independent, and intrinsically semantic entities? So far, I have worked with a rough, intuitive notion of the sense that attaches to a term, such as ‘the wife of Oedipus’. This notion is of something like the conceptual content of the term, or the idea that it expresses, e.g., the concept of what it is to be a wife of Oedipus. Yet, in “The Thought,” Frege makes it clear that a sense is not a *mental* entity, in the way that (under current terminology) an *idea* is. The sense of the term ‘the wife of Oedipus’ is not, for Frege, something internal to the mind of a particular individual who uses the term.

Prior to considering Frege’s metaphysical arguments concerning senses, it is important to establish a preliminary and intuitive distinction between the notions of sense, concept, and idea. Frege uses the term ‘sense’ to express his notion of a kind of entity that exists independently of the mind but is not a physical object. He uses the term ‘idea’, in contrast, to express the notion of a kind of entity whose existence is dependent upon the existence of a mind, and which is essentially connected to (or constituted by) *subjective* imagery, associations, memories, sensations, or emotions. The sense of the term ‘the wife of Oedipus’, for Frege, is thus something like an abstraction or universal that intrinsically means *an entity that bears the relation of being a wife of to Oedipus*. There is another intuitive way to think about this sense and the way that, according to Frege, it provides words and PA-instances with semantic content: As I will later explain, a sense can be considered to involve the logical implications of the property of being a wife of Oedipus essentially, such as its relations to *being a female* and *being a spouse of Oedipus*.

As mentioned above, Frege’s notion of an *idea* of the property of being a wife of Oedipus, in contrast, is one of something internal to the mind of an individual. This idea essentially involves the memories of his own sensory images, experiences, etc., that he associates with the terms ‘Oedipus’ and the predicate ‘is-a-wife-of’. Moreover, particular features of this individual’s perspective may cause his idea of the term ‘the wife of Oedipus’ to involve specific (but logically unrelated) concepts such as that of being a king, of noble character, of cultural rites, of opinions about marriage, and so on. While there may be ideas that accompany a person’s use of a term, such mental entities are not identical to the *sense* that attaches to it.

The intuition that the sense of a term is something like the conceptual content associated with it makes it reasonable to think that senses are something like concepts.[[176]](#footnote-176) Yet, unlike Fregean senses, concepts are usually considered mental entities, that is, things that exist in the mind. Therefore, senses exist independently of the existence of any mind, while concepts depend for their existence upon their instantiation in some mind. A Fregean sense is an abstract entity that constitutes (or somehow contains) the objective meaning expressed by a term.[[177]](#footnote-177) Concepts, on the other hand, are non-physical but concrete constituents of mental states. Because of the historical ambiguity concerning the philosophical meanings of ‘idea’ and ‘concept’, I reserve the term ‘concept’ for mental entities with cognitive or objective logical content (as opposed to subjective sensations, emotions, etc.); and I reserve the term ‘idea’ for mental entities that involve no such logical content, such as images, etc.[[178]](#footnote-178) I will next explore Frege’s arguments for the metaphysical distinction between senses and ideas.

In “The Thought,” Frege offers several arguments to establish that the senses that attach to terms must be abstract but non-mental objects.[[179]](#footnote-179)The arguments center on the assumption that *ideas*, in contrast to senses, essentially involve the subjective, associative, and idiosyncratic features of individuals’ thought processes. It is perhaps best to analyze Frege’s description of the metaphysical difference between a sense and an idea, and to then evaluate his arguments in the context of such a characterization. In “Sense and Reference,” he tells us:

The reference of a proper name is the object itself which we designate by its means; the idea, which we have in that case, is wholly subjective; in between lies the sense, which is indeed no longer subjective like the idea, but is yet not the object itself. [[180]](#footnote-180)

The argument in “The Thought” concerning senses is based upon a metaphysical matrix containing four sets of opposing general characteristics which may apply to any object: (1) immateriality or materiality, (2) mind-independence or mind-dependence, (3) graspability or non-graspability, and (4) objectivity or subjectivity.[[181]](#footnote-181)

Frege tells us that senses, like ideas, are immaterial:

So I can say: the thought is the sense of the sentence without wishing to say as well that the sense of every sentence is a thought. The thought, in itself immaterial, clothes itself in the material garment of a sentence and thereby becomes comprehensible to us. We say a sentence expresses a thought. .... A thought is something immaterial and everything material and perceptible is excluded from this sphere of that for which the question of truth arises. [[182]](#footnote-182)

Unlike ideas, however, senses are mind-independent. For an entity to be mind-dependent is for its existence to depend upon the existence of (some) mind. Presumably, this dependence consists in the fact that the entity can only exist as a constituent of a state of a mind. Thus, a thing is mind-independent when it exists independently of the existence of any mind. It is clear that Frege thought that senses were mind-independent when he said that thoughts do not require a bearer. For according to him, any mind-dependent entity *would* require a bearer. As he says:

[I]deas need a bearer. Things of the outer world are however independent … thoughts are neither things of the outer world nor ideas.... A third realm must be recognized. What belongs to this corresponds with ideas, in that it cannot be perceived by the senses, but with things, in that it needs no bearer to the contents of whose consciousness to belong. [[183]](#footnote-183)

It could be true that an abstract entity existed outside the mind while all of its instantiations depended upon the mind for its existence. Alternatively, it could be that an abstract entity existed outside the mind and nothing instantiated it. A particular feeling of anger in an individual or an *instance* of a belief that *roses are red*, for example, is mind-dependent, since neither can exist except as a state of some individual’s mind. However, physical objects, such as boxes, quarks, planets and so on, are clearly mind-independent, because their existence is not determined by the existence or content of any mental state.[[184]](#footnote-184) Yet, whether there exist any immaterial things that are *also* mind-independent is another philosophical question.[[185]](#footnote-185)

A thing’s being graspable is the feature that it has in virtue of which some mind may stand in a relation to it or entertain it, as in the case where one is said to grasp the proposition expressible by the English sentence ‘Some roses are red’. Clearly, then, material objects themselves are not graspable by a mind. Rather, ideas and senses are graspable or apprehendable. Frege goes on to explain this apprehending relation, as follows:

The apprehension of a thought presupposes someone who apprehends it, who thinks. He is the bearer of the thinking but not of the thought. Although the thought does not belong to the contents of the thinker’s consciousness yet something in his consciousness must be aimed at the thought. But, this should not be confused with the thought itself. Similarly Algol itself is different from the idea someone has of Algol. ... The thought belongs neither to my inner world as an idea nor yet to the outer world of material, perceptible things.[[186]](#footnote-186)

Kluge translates Frege’s term for ‘apprehend’ as ‘grasp’, so we know that Frege is discussing what it is for a non-physical entity to be grasped in this passage. When Frege tells us that something in the consciousness of an individual must aim at the thought in order for the individual to grasp the thought, we know that the subjective ideas that Frege distinguishes from thoughts (and which do belong to the contents of individuals’ consciousnesses) are often media through which our mental states grasp the objective, logical content of thoughts.[[187]](#footnote-187) This, then, is what it is for a non-physical thing to be graspable: to exist independently of the mind, but to be graspable by the mind through its ideas ⎯ what Kluge calls “representations.”[[188]](#footnote-188) Therefore, the passages in which Frege says that ideas are possessed by someone, that is, that they “belong to the content of his consciousness” and that they “need a bearer,” refer to the features of subjectivity and mind-dependence and not to “graspability.”[[189]](#footnote-189)

Finally, while ideas are subjective, according to Frege, senses are objective. Some things appear to depend upon the conscious experience of an individual both for their essential features *and* for their existence, and these are subjective. A state of toothache pain, for instance, cannot exist unless there is some subject who is consciously experiencing the pain; but of equal importance is that the intrinsic properties of the pain itself are in part constituted by that experiential character. As Frege says:

[I]deas are had. One has sensations, feelings, moods, inclinations, wishes. An idea which someone has belongs to the content of his consciousness.... The fields and the frogs in it, the sun which shines on them are there no matter whether I look at them or not, but the sense-impression I have of green exists only because of me, I am its bearer. It seems absurd that a pain, a mood, a wish should rove about the world without a bearer, independently. [[190]](#footnote-190)

It is theoretically possible that the experiential features of every (physically identical) state of toothache pain, i.e., *what it is like* to feel the toothache pain, are identical for each individual. Yet, Frege’s conception of the subjectivity of ideas also seems to involve the notion of an idea’s being unique to an individual’s perspective:

My companion and I are convinced that we both see the same field; but each of us has a particular sense-impression of green. I notice a strawberry among the green strawberry leaves. My companion does not notice it, he is color-blind. The colour-impression, which he receives from the strawberry, is not noticeably different from the one he receives from the leaf. Now does my companion see the green leaf as red or does he see both as of one colour with which I am not acquainted at all? These are unanswerable, indeed really nonsensical, questions. For when the word ‘red’ does not state a property of things but is supposed to characterize sense-impression belonging to my consciousness, it is only applicable within the sphere of my consciousness. [[191]](#footnote-191)

An entity (such as a particular person’s belief that *roses are red*) may be mind-dependent without any of its essential constituents being unique to the perspective of this person. This uniqueness and perspectival, experiential character, therefore, distinguish the subjectivity of an entity from its mind-dependence. As an empirical matter, it is reasonable to think that only one individual may have any given particular type of deja vu memory experience, for example, with all of its unique constituent emotional nuances, sensory images, and the conceptual associations that these generate. Yet, the belief that *roses are red*, in contrast, may be mind-dependent but qualitatively identical in each person who instantiates it.

Frege’s arguments for the objectivity of senses are based in the impossibility of their being subjective. What he appears to mean by the (hypothetical) subjectivity of the sense of a *sentence,* as opposed to that of an idea associated with it,is that its truth-value would be determined by, or variant with, the perspectives of individuals who token the sentence. His account of a subjective entity may be interpreted to be that *x* is a subjective entity if: (i) *x* is a mind-dependent entity and (ii) *x* exists only when there is a conscious perspective upon *x*, and at least one of either (iii) only one person may have this conscious perspective, or (iv) the truth of a proposition about the intrinsic features of *x* is determined by this perspective. A physical object, such as a ball, for example, is obviously objective, since it lacks any of (i) through (iv). The philosophical question that Frege attempts to answer in “The Thought” is whether the sense of a sentence (and thus the senses of its component terms) is an *objective* non-physical object.[[192]](#footnote-192) He attempts to answer this question by considering a presumably absurd consequence of the subjectivity of senses, namely, that the truth-value of a sentence expressing a mathematical proposition would be relative to the person who utters the sentence.

To summarize, Frege thinks that the difference between the abstract objects that are senses of terms and the mental entities that he calls “ideas” is this: although both senses and ideas are immaterial and mentally graspable, senses are mind-independent and objective (like physical objects), while ideas are mind-dependent and subjective.

D. Frege’s Metaphysical Arguments for the Distinction between Senses and Ideas

I will now consider Frege’s two main arguments (in “On Sense and Reference” and “The Thought”) for the conclusion that the sense of a term cannot be identical to any mental entity, i.e., to any idea. I label the first line of reasoning the ‘Commonality of Senses’ argument; and I label the second the ‘Intersubjectivity of Truth’ argument.[[193]](#footnote-193) Frege bases the Commonality of Senses argument (CS) upon two assumptions: (i) that no two persons associate the same idea with any given term, and (ii) that the sense of a term is held in common by more than one person. In “On Sense and Reference,” Frege asserts that the idea that an individual associates with a term for a perceivable object[[194]](#footnote-194) is an internal image grounded in her emotions and memories of sensations and actions. As an earlier section explained, Frege holds that an idea has an unclear subjective content, which depends upon the unique, perspectival idiosyncrasies of her past experiences as processed and remembered by her own mind.[[195]](#footnote-195) Therefore, he concludes, no two people have the same idea. It also follows that no such idea is identical to the *sense* of a term, which belongs to people in common: “one can hardly deny that mankind has a common store of thoughts that is transmitted from one generation to another.” Given this commonality of senses, *no* such sense is “a part or a mode of the individual mind," and this “constitutes an essential distinction between the idea and the sign’s sense.”[[196]](#footnote-196)

The ultimate purpose of the CS argument is to establish that the sense of a term is not mind-dependent by proving that is not identical to any mental entity. I have laid out a preliminary, intuitive notion of mind-dependence, namely, that an entity is mind-dependent when it depends for its existence upon the existence of some mind. Yet, this implicit conclusion is based upon the explicit claim that the sense of a term is “held in common by many,” and there is more than one interpretation of this concept of the commonality of senses. This multiplicity of accounts of commonality, in turn, generates more than one metaphysical interpretation of the implied conclusion that the sense of a term is mind-independent, i.e., that it is not a “part or mode” of any individual mind (M).

Frege’s view that the sense of a term is held in common by many (C) has two possible meanings, one of which says that all the tokens of a given term share one and the same sense, and one of which says that each token of a term is numerically *distinct* from, although qualitatively identical to, the other. The first version of C is motivated by consideration of two of Frege’s positions: that the sense of a term t determines its referent, and that a sentence containing t refers to its own truth-value. This interpretation of Frege’s view that the sense of a term is held in common by many, namely, C1, is that (i) more than one individual tokens t; and that (ii) every such act of reference is determined by the association of *one and the same* abstract entity *s* with that particular tokening. The other interpretation of C, namely, C2, is that the sense s of a term t is a type, which is “held in common by many” in that many different people have manifested s by instantiating it in their own mental states.

C1 supports an interpretation of the implied conclusion M, which endorses complete mind-independence for the senses of terms (M1). For if each particular token of a term *t* has *one and the same* abstract entity s as its semantic ingredient, then since each particular token of *t* is an expression of a different individual mind (or at least of a different time stage of some individual), the sense of a term *t* cannot exist in any individual mind. C2, in contrast, supports two different interpretations of Frege’s implied conclusion M, each of which only endorses some form of *partial* mind-independence for senses.[[197]](#footnote-197) The first partial mind-dependence interpretation of M is that each particular token of a term *t* expresses a semantic ingredient (or sense) *s* that is qualitatively identical to every other, but that these particular senses are not instantiations of an existent general or universal sense (M2). The other partial mind-dependence interpretation is that each particular token of a term *t* expresses a semantic ingredient (or sense) *s* that is qualitatively identical to every other; and that these distinct particular senses *are* instantiations of a universal sense *S*, which itself is an abstract entity that exists independently of the existence of any particular mind (M3).

In order to understand the CS argument, it is necessary to interpret the principles that ground Frege’s view that senses are held in common by many, i.e., the principles which, when combined with his conception of the unshareability of an individual’s ideas, makes it impossible for senses to be identical to ideas. It is also necessary to understand the metaphysical notions of mind-independence and objectivity from which he concludes that the commonality of senses prevents them from being mind-dependent.[[198]](#footnote-198) The CS argument in “Sense and Reference” does nothing to elucidate these issues. In “The Thought,” however, Frege constructs a reductio *ad absurdum* of the view that when a person utters or inscribes a sentence, the sense of the *sentence* ⎯ that is, the complete thought that it expresses ⎯ belongs uniquely to that person.[[199]](#footnote-199) The falsity of this view implies that more than one person, at least, grasps the sense of a sentence. I label this extensive passage of metaphysical reasoning the ‘Intersubjectivity of Truth’ argument (IT). Evaluation of IT will help to reveal the concepts and lines of reasoning behind Frege’s view that senses are mind-independent and objective entities.

In “The Thought: A Logical Enquiry,” Frege argues that the sense of a sentence cannot be identical to any idea that an individual associates with that sentence. In the process of constructing a reductio against the identity of senses and ideas, he constructs a reductio against the view that the sense of a sentence-token is unique to the individual who tokens it. His complex line of reasoning is based on twoassumptions: (i) that whether (at least) some kinds of sentences, i.e., mathematical sentences, are true or false is an objective matter; and that (ii) the idea that an individual associates with a sentence is private and unique. IT begins as follows: [[200]](#footnote-200)

I now return to the question: is a thought an idea? If the thought I express in the Pythagorean theorem can be recognized by others just as much as by me then it does not belong to the content of my consciousness, I am not its bearer; yet I can, nevertheless, recognize it to be true.

Frege begins IT by considering whether the thought expressed by an individual’s token of a sentence, such as his own utterance of the Pythagorean theorem, is intersubjectively accessible. In other words, can others recognize what he means with an utterance of ‘a2 + b2 = c2’ (PT)? With this question, Frege constructs a conceptual opposition between a semantic entity’s being graspable by more than one person, on the one hand, and its being a unique aspect of an individual’s consciousness, on the other. Thus, an implicit logical dilemma grounds his reasoning about this question, and this dilemma is his first premise of his argument:

1. Either others can recognize the sense expressed by an individual’s sentence-token of PT, or they cannot.

The next premise simply asserts the above-mentioned conceptual opposition:

2. If others can recognize the thought expressed by an individual’s utterance of a sentence, then it does not belong solely to the content of that individual’s consciousness.

Premise (2) clearly means that if the sense of an individual’s token of a *sentence* PT can be understood or apprehended by others, then it is not identical to any idea that she associates with PT.[[201]](#footnote-201) (We should interpret his notion of the idea that an individual associates with a sentence to be a notion of a *type* of idea that is unique to her and thus is dependent for its existence upon having been instantiated in her “consciousness,” or mind.)[[202]](#footnote-202) Frege supports the inference from (2) to the conclusion that senses are not identical to ideas by establishing that the sense that an individual expresses with a sentence must be recognizable by others. He infers this proposition by constructing a reductio of the proposition that the thought expressed by a sentence-token is *not* held in common by more than one individual:

However, if it is not the same thought at all which is taken to be the content of the Pythagorean theorem by me and another person, one should not really say ‘the Pythagorean theorem’ but ‘my Pythagorean theorem’, ‘his Pythagorean theorem’ and these would be different; for the sense belongs necessarily to the sentence. Then my thought can be the content of my consciousness and his thought the content of his.... [[203]](#footnote-203)

This passage addresses the negative horn of the dilemma of premise (1), namely, others cannot recognize the sense expressed by an individual’s token of a sentence. Frege argues that the negative horn implies that the sense he expresses by tokening PT is different from the sense that anyone else expresses by tokening PT. If the sense of his own utterance of PT were not intersubjectively recognizable, there would be no one individual entity that is the sentence-*type* PT. Therefore, there would be no one type PT that different individuals could token by uttering or inscribing ‘a2 + b2 = c2’. Rather, there would be a multitude of such sentence-types, one for each individual who utters (or inscribes) them; and describable only with respect to such individuals, as “his PT,” “her PT,” “my PT,” etc.

This argument is based on the semantic assumption that a sentence is defined, in part, in terms of the semantic content, or sense, that it expresses. Thus, it implies that if every token of a sentence expresses a different sense, then there is no objective semantic content to a mathematical proposition or thought.[[204]](#footnote-204) If Frege’s token of PT expresses a sense that can only be grasped by him, then each other individual’s token of (what seems to be) that same sentence-type must actually be a token of his or her own individual sentence-type. Therefore, these individuals’ tokens must also be expressions of their own theorem-*senses*. This would imply an extreme relativism about mathematical propositions.

If the semantic content of a mathematical sentence is defined relative to its individual utterers, however, this means that the Pythagorean theorem, for example, has no objective truth-value. Frege thus establishes the falsity of this semantic indexicality implication by pointing out its intuitively absurd consequence concerning the *truth* of mathematical claims. Recall that for Frege, the sense of a sentence determines its referent, and the referent of a sentence is its truth-value. Thus, the semantic content of a sentence determines its truth-value. If the sense or meaning of a sentence-type, such as ‘a2 + b2 = c2’, could only be defined relative to each individual who tokens it, then the *truth* of such a mathematical sentence-type could only be defined in the context, or “sphere,” of an individual’s mind. This implies, of course, that there is no objective, non-indexical truth to the Pythagorean theorem. Yet, this entails, in turn, that one person’s belief that *a2 + b2 = c2* could be *true* while that of another person was *false*:

Could the sense of my Pythagorean theorem be true while that of his was false? I said that the word ‘red’ was applicable only in the sphere of my consciousness if it did not state a property of things but was supposed to characterize one of my sense-impressions. Therefore, the words ‘true’ and ‘false’, as I understand them, could also be applicable only in the sphere of my consciousness if it did not state a property of things but was supposed to characterize one of my sense impressions. Therefore, the words ‘true’ and ‘false’, as I understand them, could also be applicable only in the sphere of my consciousness, if they were not supposed to be concerned with something of which I was the bearer, but were somehow appointed to characterize the content of my consciousness. Then truth would be restricted to the content of my consciousness and it would remain doubtful whether anything at all comparable occurred in the consciousness of others. [[205]](#footnote-205)

In the passage above, Frege continues the reductio of the proposition that his token of a sentence expresses a sense that only he can grasp. He considers himself to show that this implies that his utterance of the Pythagorean Theorem could only be considered *true* relative to the context of his individual mind. It is important to analyze exactly what notion of relative or contextual truth Frege is using in this inference, however, and this analysis must concern his definition and examples of relative truth. For the truth of the sense of a sentence to be relative to an individual mind is for the concept of truth to be legitimately applied to that sentence only within the sphere, or context, of that individual’s consciousness.

In discussing the issue of the relative, contextual truth of the sense of a sentence, Frege first considers an implicit example of a truth that applies (or ascribes) a sensory property to an object. Since a sentence ‘Fa’ is true just in case it is legitimate to predicate ‘F’ to the subject ‘a’, he considers this example from a linguistic perspective, i.e., in terms of the question of the applicability of the *predicate* ‘red’. The conditions under which the predicate ‘red’ is applicable, according to Frege, are those under which he, as the utterer, intends to refer to a property of his sense impressions. Therefore, ‘red’ is only applicable in the context or sphere of his own consciousness, given that the phenomenal features of his sense impressions are features of his own consciousness.

The important philosophical point to recognize about this example is that the relative contextual truth of a sentence predicating ‘red’ is determined by the fact that the property of being red is not a “real property of things.” ‘Red’ is not an intrinsic property of any object that exists independently of the mind. Therefore, only a conception of truth as relative to an individual can apply to sentences ascribing red. This is because in Frege’s philosophical view of color-properties, ‘*x* is red’ is true just in case the utterer of the sentence (or conceiver of the proposition) normally has a red-like sense experience of *x*.[[206]](#footnote-206) It follows that the sentence ‘*x* is red’ could be true for one individual and simultaneously be false for another individual. For example, it could be false for a person who was color-blind, one whose cognitive wiring processed the light spectrum differently, and so on. Frege appears to conclude (by analogy) from this thought experiment that if the truth of the sense of a *mathematical* sentence, such as ‘a2 + b2 = c2’, for example, were relative to an individual, then the truth of the Pythagorean Theorem would not be a real property of the theorem.

In other words, Frege reasons that if the *feature*[[207]](#footnote-207) of truth were only applicable to a mathematical proposition (or “sense”) in the context of an individual’s consciousness, then this would be because truth and falsity were not real properties of such propositions. Some would consider this result in itself to affect a reductio of the assumption that the sense of ‘a2 + b2 = c2’ is not intersubjectively accessible. Yet, what Frege considers the intolerable result of this assumption is that the sense of one individual’s entertainment of the Pythagorean theorem could be true while that of another individual was false. For he has endorsed the mathematical *necessity* of the truth of the theorem, that is, its being “timelessly true, true independently of whether anyone takes it to be true.”[[208]](#footnote-208) Obviously, this necessity entails that the theorem cannot be false; and he thus considers the claim of its possible falsehood to effect a reductio of the antecedent assumption that the sense that he himself expresses with a given sentence, such as ‘a2 + b2 = c2 ’, is not recognizable by others.

This intermediate reductio, in turn, makes it possible to apply modus tollens on the original assumption (for reductio) of the IT argument, namely, the assumption that the senses of sentences are identical to ideas. Although the denial of *this* proposition is the ultimate (albeit implicit) conclusion of IT, I explored the reasoning of IT in order to understand the concepts and principles that support Frege’s view that senses are “held in common by many.” For he makes this latter claim in the CS argument; and I interpret the statement to express his metaphysical view that the sense of a term (and of a sentence) is a mind-independent (and objective) entity.

In the final chapter, I will argue that concrete, mind-internal concepts are the only intrinsically semantic entities, not abstract, mind-independent senses. However, I will defend the Fregean metaphysical position that abstract objects play an essential role in providing intrinsically semantic content to propositional attitude-instances.

Chapter V: Bare Property Intentionality

A. The Metaphysical Assumptions of Bare Property Intentionality

The previous chapter analyzed a version of the intrinsic semantics position on intentionality, namely, the Fregean view that there exist intrinsically semantic entities. Frege called these entities ‘senses’, and he considered them mind-independent, abstract objects. Since earlier analyses in this essay rejected the extrinsic semantics perspective, this chapter will adopt the intrinsic semantics position on intentionality. More specifically, the discussion will assume the general linguistic-semantic point of Frege’s resolution of the Substitution of Identicals puzzle. This resolution says that terms in the intensional contexts of sentences that ascribe propositional attitudes to people refer to intrinsically semantic entities. For example, the sentence ‘Eric believes that all bachelors are sports fans’ ascribes an instance of a propositional attitude to Eric: it says that he has the belief that *all bachelors are sports fans*. Frege’s resolution of the SI puzzle says that this sentence refers to an entity that makes Eric’s mental state essentially mean that *all bachelors are sports fans*. Whatever their metaphysical nature turns out to be, these entities are the source of the intrinsically semantic content of PA-instances.

From this Fregean linguistic assumption, this discussion will draw an important, but *non*-Fregean metaphysical conclusion on the intentionality of PA-instances. I will argue that for a mental state to have intrinsically semantic content is, at least in part, for the state to include concrete concepts that present properties to the mind essentially. These concepts’ mental presentations of particular properties provide the state with its particular semantic content. Eric’s occurrent belief that *all bachelors are sports fans*, for instance, includes a concrete concept of the class of all bachelors. This concept essentially presents the property BACHELOR to his mind. In this sense, properties, as existent universals, are the *source* of the particular intrinsically semantic content of a PA-instance, while concepts are the only intrinsically semantic *entities*. Moreover, many of these concepts are also intrinsically *representative*.

Thus, I will conclude that there is an essential relationship between the intrinsically semantic content of mental states and abstract objects (properties). Yet, my metaphysics of intentionality, which I call ‘Bare Property Intentionality’,strongly diverges from Frege’s view, according to which abstract objects (senses) are, *in themselves*, intrinsically semantic.

The current discussion of intentionality will depend upon the assumption that there is a distinction between the feature of being intrinsically representative and the feature of being intrinsically semantic. While all intrinsically representative mental phenomena are intrinsically *semantic*, not all intrinsically semantic mental phenomena are intrinsically *representative.* According to Bare Property Intentionality, Eric’s occurrent belief that *all bachelors are sports fans* contains an intrinsically representative concept of the class of all bachelors and an intrinsically semantic (but not intrinsically representative) instantiation-concept of the property SPORTS FAN.

An occurrence of the concept of the class of all bachelors is an intrinsically representative mental phenomenon. This concept-instance is intrinsically semantic, because its content presents the property BACHLEOR to the mind essentially. An occurrence of this concept is also intrinsically representative, because it intrinsically refers to a particular possible (or actual group of entities) exclusively. Moreover, this mental reference takes place in such a way that the concept may be an ingredient in a propositional attitude-instance that is truth-evaluable. This is because when one applies an occurrence (or instance) of an intrinsically representative concept to an object (or to a class thereof) this representation of that object (or class) is either accurate or inaccurate. Whether an occurrence of such a concept is accurate or inaccurate determines whether the mental state of which it is a constituent is true or false.[[209]](#footnote-209) In contrast, a mental phenomenon that is merely intrinsically semantic presents some property to the mind essentially, but it does not essentially *refer* to any possible or actual entity (or set thereof).

A simple way to understand this distinction may be to consider the difference between merely having an unrealistic idea or concept, such as the idea of a golden mountain, presented to one’s mind, and actually applying this concept to something in the external world. That is, the concept of a golden mountain may occur in one’s mind without one having an occurrent belief (or mentally asserting) that any particular object is a golden mountain. This concept may also occur in someone’s mind unaccompanied by any occurrent belief (or mental assertions) that any golden mountains do (or do not) exist.[[210]](#footnote-210)

Moreover, such a concept can occur in a person’s mind even if he neither has an occurrent belief that anything is true of any possible or actual golden mountains, nor mentally asserts anything about any possible or actual golden mountains. This occurrence presents the property GOLDEN MOUNTAIN to his mind without mentally representing anything in any possible world *as being* a golden mountain (or as being a golden mountain of which any given proposition is true). These types of intrinsically semantic mental states involve some mental presentations that are not mental *re*presentations, because they have no reference. Moreover, such states are not truth-evaluable. Eric’s occurrent belief that *all bachelors are sports fans* includes an intrinsically semantic but non-representative concept, which presents the property SPORTS FAN to his mind as being instantiated without *representing* any object as being a sports fan. (When combined with the concept of the class of all bachelors, however, the entire mental state does represent each and every bachelor as being a sports fan.) \

Bare Property Intentionality says that only the concrete, internal constituents of mental states, and not any abstract entities, such as properties or Fregean senses, have intrinsically semantic content. These internal constituents are concepts, some of which are merely intrinsically semantic, and some of which are intrinsically representative, as well. In contrast, Frege’s view of the internal constituents of mental states is that these “mental representations” are subjective, which means that of necessity, they are not objectively evaluable as true or false. (Nor can they be semantic *components* of a mental state that is objectively truth-evaluable.) [[211]](#footnote-211) To recall, the background metaphysical assumptions of this essay require a successful theory of intentionality to account for the capacity of intentional mental states to be evaluated as objectively true or false. This means that Frege’s view does not allow the internal constituents of intentional states to *be* intrinsically representative. Since their subjective content prevents them from having deductive force, Frege’s view does not allow them to be intrinsically semantic, either.

Moreover, while it is clear that Frege considers senses (mind-external, abstract entities) to be intrinsically semantic, he appears to provide no account of the features of senses that make them so. Bare Property Intentionality, in contrast, does account for the intrinsically semantic (and representative) content of mind-internal concepts in terms of specific features. For according to Bare Property Intentionality,concepts of classes of objects essentially direct the mind to a (possible or actual) class by presenting a specific property, such as NUT, to the mind. The property is one exclusively instantiated by the class that the concept represents.

An instance of the belief that *all* *nuts are fruit*,[[212]](#footnote-212) for example, contains an intrinsically semantic and representative concept of the class of all nuts. This concept essentially represents each individual nut because it has specific features, which make it mentally present the property NUT in a way that connects the mind semantically to each individual nut. Such a state also contains a concept of the *instantiation* of the property FRUIT. While this “instantiation-concept” is intrinsically *semantic*, in that it allows the mental state to present the property FRUIT to the mind as being instantiated by each individual nut, it does not *represent* the property FRUIT intrinsically. (This discussion concerns instances of PAs of the form *all Fs are Gs* only, i.e., whose contents are universal propositions.[[213]](#footnote-213) Thus, this chapter will only discuss the kinds of concepts included in this form of PA-instance.)

How do these mental ingredients play their semantic roles? Both class concepts and instantiation-concepts involve mental presentations of properties. For example, the concepts of the class of all bachelors and of the class of all spheres, as well as the instantiation-concepts of the properties ROUND and SPORTS FAN, must each present different particular properties —BACHELOR, SPHERE, ROUND, and SPORTS FAN — to the mind. Bare Property Intentionality only concerns the mental presentation of *some* of the physical properties with which humans currently have sensory contact. These are the simple, perceptible properties of our every day experience, such as FRUIT, CAR, NOSE, etc. This view does not apply to instantiation-concepts of properties of which we may have no direct perceptual experience, such as sub-atomic properties, grand cosmological properties, and so on. Instantiation-concepts of these types of properties are complex, because they develop from logical, mathematical, and theoretical concepts.[[214]](#footnote-214) (Of course, the properties that humans either think about or have sensory contact with are a miniscule fraction of all the physical properties that exist.)

A proponent of Bare Property Intentionality argues that in addition to a presentation of a property to the mind, however, each *class* concept must have an ingredient that makes it intrinsically semantic *and* representative. This ingredient is a ‘bare referential intention’ (a BRI). All intrinsically representative concepts must contain a bare referential intention. Yet, a BRI in itself is not necessarily associated with any particular property. No such ingredient in itself can refer to the class of all fruit intrinsically, or be essentially connected with the property FRUIT.

A BRI is a general referential event (an act committed by the mind) without any particular semantic content.[[215]](#footnote-215) Within a class concept, it is a primitive and non-purposeful event of directing the mind toward some given class of objects through mentally presenting some property. Yet, a BRI does not have a *propositional* structure, as would a mental state, such as an occurrent belief that *all nuts are fruit*, in which a person consciously connects their mind with (or to thinks *about*) a class of objects. The intrinsically representative *feature* of a concept, which a bare referential intention provides, is that which makes the occurrence of the concept a semantically *representative* as opposed to a *non*-representative phenomenon. Thus, an occurrence of the concept of the class of all nuts is able to intrinsically represent something as opposed to representing nothing, because it contains a BRI. The fact that a classconcept also essentially presents a particular property to the mind, such as NUT, provides the concept with its particular semantic *content*: a relation whereby it concerns a specific class of objects intrinsically and exclusively.

An instance of the belief that *all nuts are fruit* contains a concept of the class of all nuts. This concept both involves a bare referential intention and presents the property NUT to the mind. Yet, a class concept is not enough to make the belief as a whole express that each member of the concept of the class of all nuts instantiates the property FRUIT. The belief must do this in order to be evaluable as true or false (and in order to have a logical or propositional structure); and merely including the class concept will not suffice to make the belief accomplish this semantic feat. Therefore, an instance of a PA of the form *all Fs are Gs* must also contain a distinct type of concept, which merely expresses the instantiation of the property G. Later, I will explain the metaphysical features and semantic function of these “instantiation-concepts,” and how they develop in a person’s mind.

B. The Relationship Between Properties and Intrinsically Semantic Content

The intuitive notion of a property is one of a universal: something that is (or that may be) had by more than one thing. While a concrete entity is particular and unrepeatable, a universal may be instantiated, and thus repeated, in more than one object. To understand this, consider two objects that, as our intuition tells us, have something in common. An orange and a basketball, for example, intuitively appear to share the property of being spherical; and intuition also says that because it may be instantiated, or had, by more than one thing, this feature must be universal.[[216]](#footnote-216) From a philosophical standpoint, there may exist some properties, such “abstract particulars,” which are *non*-universal, *non*-repeatable, *and* non-concrete. This chapter will argue that what it is for a mental state to be intrinsically intentional is for that state to be connected to in a certain way properties *qua* universals. The questions of whether and how abstract particulars are involved in the relationship between the mind and properties *qua* universals is outside the scope of this essay.

Bare Property Intentionality’s account of mental representation depends upon a specific *a priori* metaphysical tenet, called the ‘Property Principle’. This principle says that if x *represents* a class of objects *C* in an intrinsically semantic manner (in a manner that can be an ingredient of a truth-evaluable proposition), then x presents to the mind some property F that all and only members of *C* instantiate. [[217]](#footnote-217) One can see that this principle implies that the properties that intrinsically representative concepts of classes of objects present to the mind must be universals, since these properties are instantiated by more than one object belonging to the class. (While this essay does not address intrinsically semantic representations of universals, the principle implies that if x represents an individual abstract object *a* in an intrinsically semantic manner, then x presents to the mind some property F that is exclusively instantiated by *a*.)[[218]](#footnote-218)

What is it for a semantic entity to represent something in a manner that can be an ingredient of a truth-evaluable proposition? The answer to this question, according to Bare Property Intentionality, involves an intuitive connection between the intrinsically semantic content of mental states and properties as universals. It appears that properties are involved in what it is for a concept to refer (objectively and exclusively) to a class of objects, such as the class of all nuts. What it is for an individual to belong to this class is for it to instantiate a property, i.e., NUT, which all and only members of that class instantiate. Therefore, what it is for a concept to intrinsically refer to that class must involve a relationship between that concept and the property. Analogously, it is intuitive that properties are involved in what it is for a mental state to have an objective truth-value. For in order for a mental state, such as an instance of the belief that *all nuts are fruit*, to be truth-evaluable, it must single out a class of objects (or an individual object), such as the class of all nuts, as its exclusive extension.

For a theory of intentionality to explain what it is for a PA-instance to be truth-evaluable, if one recalls, is one of the three criteria for success put forth in an earlier chapter. The discussion above shows how reflection on this first criterion points, intuitively, toward the Property Principle. The other two requirements are, as previously described, for the theory to resolve the Substitution of Identicals puzzle, and for the theory to explain how an instance of a PA can have truth-preserving causal powers. I will now argue that the assumption that a propositional attitude of the form *all Fs are Gs* must present a property to the mind in order to represent a class of objects also helps Bare Property Intentionality to meet the second explanatory criterion. [[219]](#footnote-219) Since the explanatory efficacy of this assumption strongly supports a connection between properties as universals and intrinsically semantic content, this efficacy also strongly supports the Property Principle.

To show how Bare Property Intentionality meets the second explanatory condition, I will analyze the role of properties in the Fregean resolution of the SI puzzle. To recall, Frege concludes from his analysis of the SI puzzle that terms in intensional contexts of sentences refer to intrinsically semantic, abstract, and mind-independent objects termed ‘senses’. He addresses the question of the identity of a sense in terms of its linguistic function. A sense is that in virtue of which a term refers to an object (or that in virtue of which a sentence refers to a truth-value [[220]](#footnote-220)). This is the sense in which Fregean senses are intrinsically semantic: while a word is a mere symbol that refers to something solely in virtue of its sense, a sense provides meaning (and thus, reference) to a mental state, word, or sentence.

In what follows, I argue that a Fregean sense must be related to a set of properties essentially, and that the specific set of properties to which a Fregean sense is related determines its intrinsically semantic content. It is obvious that words, terms, and sentences express semantic contents, or meanings. According to Frege, a word, term, or sentence expresses a metaphysically real abstract object called a ‘sense’. Through expressing a sense, a word refers to whatever object(s) that sense determines. Specifically, I conclude that a word’s expression of a sense allows it to ascribe a certain property (or set thereof) to the relevant object (or class thereof). Since Frege addresses the question of the identity of a sense in terms of its linguistic role, the property that a term (or sentence) ascribes to its referent is the property that that determines the semantic content of the sense of that term. As explained above, according to Bare Property Intentionality, mind-internal concepts, and not Fregean senses, are the only intrinsically semantic objects. Nonetheless, from my analysis of Frege’s view, I draw an important conclusion about the metaphysics of intrinsically semantic mental states: that an essential connection between their ingredients and a set of properties determines the semantic content of these states.

The Substitution of Identicals rule (SI) says, to recall, that if an object *a* is identical to an object *b*, ‘Fa’ and ‘Fb’ will always have the same truth-value in the actual world. The U.S. Secretary of Defense in February 2002 A.D. is identical to Donald Rumsfeld, in the actual world. Thus, ‘The U.S. Secretary of Defense in February, 2002 A.D. lives in Washington, D.C.’ will always have the same truth-value (in the actual world) as ‘Donald Rumsfeld lives in Washington, D.C.’. However, consideration of two propositional attitude ascriptions will reveal that the sentences do not have the same *sense*. The sentence: ‘Eric believes that *the U.S. Secretary of Defense in February, 2002 A.D. lives in Washington, D.C.*’ may have a different truth-value than the sentence: ‘Eric believes that *Donald Rumsfeld lives in Washington, D.C.*’. This shows that the *intensional content* of ‘The U.S. Secretary of Defense in February, 2002 A.D. lives in Washington, D.C.’ may be distinct from that of ‘Donald Rumsfeld lives in Washington, D.C.’. Thus, the two sentences do not express identical senses.[[221]](#footnote-221)

If two senses (s1 and s2)have the internal structure of propositions, it is safe to conclude that one can determine whether they are identical by determining whether two sentences that express them are logically equivalent. (It is debatable whether the concept of logical equivalence actually applies to the notion of a Fregean sense itself. For it is unclear that senses themselves may *have* truth-values, as opposed to merely conferring truth-values, as referents, upon sentences that express them.) If two sentences S1 and S2 express identical senses s1 and s2, then S1 and S2 must have the same truth-value in *all* possible worlds. That is, a difference that makes two propositionally structured senses non-identical entails a difference, in at least one possible world, between the truth-values of sentences that express these senses.

Another way of looking at this point is to consider that the intensional content of a sentence S partially determines its truth-value. For one can consider this content to be the sense or proposition that S expresses; and the truth of this proposition determines whether S is true. The sense expressed by S, therefore, has an intrinsic feature through which it (partially) determines the truth-value of S (in a given possible world). Since a sentence says some object (or class) instantiates some property, the other factor that determines the truth-value of a sentence is whether the object (or class) to which it refers actually instantiates the property that the sentence ascribes to it.

Therefore, the intrinsic feature through which a sense partially determines the truth-value of a sentence is also the feature through which the sense determines what the sentence represents (and how it characterizes this referent). For the referent of a sentence also determines its truth-value. It is safe to conclude from this that the intrinsic feature through which a sense contributes to the truth-value of a sentence is one through which the sense determines that the sentence refers to an object (or class) as instantiating a certain property. Thus, one can also conclude that some connection between a Fregean sense and a property provides the sense with its intrinsically semantic function. Moreover, as Frege defines a sense in terms of its function within a sentence, this connection determines the identity of the sense.

As mentioned above, Bare Property Intentionality is not any version of Frege’s metaphysical view that senses, qua *mind-independent* abstract objects, are intrinsically semantic. Rather, the mind-internal constituents of a PA-instance, i.e., concepts, provide the mental state with intrinsically semantic content. Yet, a proponent of Bare Property Intentionality does accept the essential semantic point of Frege’s resolution of the SI puzzle: that the terms in the intensional context of a sentence that ascribes a PA function to refer to the intrinsically semantic content of that PA.[[222]](#footnote-222) The above section concluded that if Fregean senses were to exist, some connection between a sense and a property would both provide the sense with its intrinsically semantic function and determine the sense’s identity. However, a proponent of Bare Property Intentionality holds that terms in the intensional context of a PA-ascription refer to intrinsically semantic entities that are not mind-external and abstract. Rather, they are mind internal and concrete concepts. Therefore, it is safe to conclude that some essential connection between a concrete concept and a property (or set thereof) provides an instance of a propositional attitude with its particular semantic content.

Assume, for example, that a person named Eric has an occurrent belief that *all nuts are fruit*. Bare Property Intentionality says that under the appropriate conditions, the term ‘all nuts are fruit’ within a token of the sentence ‘Eric believes that *all nuts are fruit*’ stands for the intrinsically semantic constituents of his mental state. That is, ‘all nuts are fruit’ refers to the set of internal concepts that function (intrinsically) to provide his occurrent belief with its *particular* semantic content. These concepts carry out this semantic function, in turn, by presenting properties to Eric’s mind. With this view of intrinsically semantic entities as mind-internal concepts, Bare Property Intentionality rejects the (relevant) Fregean metaphysics, in which mind-external, abstract entities are the sole source of the intrinsically semantic content of mental states. From my analysis of the Fregean resolution of the SI puzzle, I conclude that some connection between a concrete concept and a property both provides that concept with its intrinsically semantic function and determines its identity.

As we saw above, a token of the sentence ‘Eric believes that *the U.S. Secretary of Defense in February 2002 A.D. lives in Washington, D.C.*’ may have a different truth-value than a token of the sentence: ‘Eric believes that *Donald Rumsfeld lives in Washington, D.C.*’. Bare Property Intentionality says that within these sentence-tokens, the terms ‘the U.S. Secretary of Defense in 45 February 2002 A.D.*’* and ‘Donald Rumsfeld’ refer to intrinsically semantic concepts that Eric’s occurrent beliefs respectively include. Bare Property Intentionality resolves the Substitution of Identicals puzzle in a manner that is analogous to Frege: co-referential terms may have different referents in the intensional contexts of sentences ascribing propositional mental states, for they may refer to distinct intrinsically semantic concepts.

Based on the intuitive and theoretical connections that I have discussed between intentionality and properties, I conclude that what it is for a mental state to have intrinsically semantic content is for its concepts to bear essential relations to properties, relations through which those properties are presented to the mind. This conclusion is manifested in my endorsement of the Property Principle, which says that if x represents a class of objects *C* in an intrinsically semantic manner, then x presents to the mind some property F that all and only members of *C* instantiate. A later section will explain how the ingredients of an intrinsically representative mental state give it internal logical or propositional structure. The following section, however, will give a more specific account of the metaphysical relationship between class concepts and the properties that they present to the mind. [[223]](#footnote-223)

C. Some Metaphysical Pre-requisites for a Relationship between a Concept and a Property

How, then, do *class* concepts connect the mind connect with properties? As stated above, a concrete concept of a class of objects must essentially involve a presentation of a property to the mind, a property that is instantiated by all members of that class. Therefore, the question becomes how a concept of a class of objects presents the relevant property to the mind. On the one hand, such a concept (or some element or ingredient of it) could intrinsically *represent* that property in the same way that the concept intrinsically represents the relevant *class* of objects. On the other hand, the concept could present the property to the mind in a different way (as opposed to *re*presenting it). In the former case, the Property Principle would apply to the concept’s presentation of the property to the mind, which would imply that the concept would present to the mind some distinct property that only the original property instantiates. In the latter case, the concept’s presentation of the property to the mind would not be an intrinsically semantic *representation* of the property; and the Property Principle would not apply to this presentation. Thus, the concept would not need to present to the mind some distinct property that only the original property instantiates. I will now explain what is wrong with the former alternative; and a later section will provide an account of a non-intrinsically representative, but intrinsically semantic type of mental presentation.

Consider, for example, the concept of the class of all bachelors, as a constituent of the belief that *all bachelors are sports fans*. The present question is whether this concept intrinsically represents the class of all bachelors by representing the property BACHELOR in the same way, i.e., in the way that, under the assumptions of this essay, the Property Principle would dictate. Following is an argument that a vicious infinite regress follows from the assumption that the concept functions like this. Its conclusion is that a concept of a class of objects can only intrinsically *represent* that class of objects, and not any property. On a metaphysical level, this relation essentially *involves* a presentation of a particular property, e.g., the property BACHELOR, to the mind, but this relation is not one in which the concept *represents* the property in an intrinsically semantic manner.

Suppose that a concrete concept intrinsically represents a class of objects by virtue of representing, in the way dictated by the Property Principle, a property F, which is exclusively instantiated by members of that class. For example, a concept represents the class of all bachelors intrinsically in virtue of representing the complex property ADULT & UNMARRIED & MALE in the manner dictated by the Property Principle. (It is possible for a concept to represent a property in such a way because the principle applies to intrinsic representation of individual abstract objects as well as to intrinsic representation of classes of concrete objects.) According to the supposition, in order for the concept of the class of bachelors to represent the property BACHELOR in this manner, the concept must also represent, in this manner, some second-order property *exclusively* instantiated by the first-order property ADULT & UNMARRIED & MALE. An example of such a second-order property might be, for example, NOT POSSIBLY INSTANTIATED BY CHILDREN OR BY SPOUSES OR BY FEMALES. Yet, on the alternative view, the Property Principle must govern the relationship between the concept of the class of all bachelors and this *second*-order property, as well. In order to represent this second-order property as such, however, the concept would then have to bear such a relation to some third-order property that the second-order property (exclusively) instantiates.

Thus, an infinite regress, in which the concept of a bachelor intrinsically represents an infinite number of higher-order properties of lower-order properties, ensues. Under the relevant assumptions, an intrinsically representative relation, i.e., a relation governed by the Property Principle, can obtain between the concept of the class of all bachelors (x) and the class of all bachelors (y) only under one condition. This condition is that a distinct intrinsically representative relation obtains between some *component* of x (x1) and the property BACHELOR (z). This must be a component of the concept of the class of all bachelors, for the entire concept cannot simultaneously represent this class *and* represent just some second-order property of the property BACHELOR.[[224]](#footnote-224) This second relation can obtain, in turn, only if a third intrinsically representative relation obtains between some component of x1 (which would be x2), and z1 (some property *of* the property BACHELOR), and on to infinity.

The question then becomes whether this infinite regress is vicious. While there does not appear to be any traditional or formal *definition* of a vicious infinite regress as distinct from a merely infinite regress, one can begin to consider the question whether a regress is vicious from a logical standpoint. One reason to think that the regress may be vicious may be that the infinitude of the regress *negates* the relevant supposition, i.e., the assumption that the concept of the class of all bachelors intrinsically represents bachelors in virtue of intrinsically representing the property BACHELOR. The argument for this “negation” notion of viciousness would be the following:

1) Assume the following conditional (A): If a concept C intrinsically represents a class of objects x (in the manner dictated by the Property Principle) then (i) there is a property F that is exclusively instantiated by members of x; and (ii) concept C (or some component thereof) represents property F in the manner dictated by the Property Principle.

2) Proposition (A) implies this infinitely regressive proposition (B):

The relation whereby a concept C intrinsically represents a class of objects x includes the obtaining of a relation whereby C intrinsically represents y, where x ≠ y.

3) If A implies B, then in order for a concept to intrinsically represent a class of objects, an infinite series of requirements must be met. This is because there must be an infinite series of intrinsically representative relations between that concept and some property F. Let us call this series ‘Φ’.

4) Under the assumption that human beings do have concepts that intrinsically represent classes of objects, proposition A implies that the mind of a human being engages in an infinite series of intrinsically representative relations with a property each time he has a concept that intrinsically represents a class of objects.

5) Since it cannot be true that the mind of a human being engages in an infinite series of intrinsically representative relations with a property each time he has a concept that intrinsically represents a class of objects, proposition A cannot be true.

The above argument appears to reveal that the assumption in question implies an infinite regress that falls under the “negation” notion of viciousness. The assumption, to recall, is that if a concept C intrinsically represents a class of objects x (in the manner dictated by the Property Principle) then: i. there is a Property F that is exclusively instantiated by members of x; and ii. C (or some component thereof) represents F in the manner dictated by the Property Principle. Yet, the reasoning above demonstrates that this assumption directly implies an infinitely regressive proposition, which, in turn, implies the falsehood of the relevant assumption.

A non-formal reason to consider the regress vicious has to do with the particular philosophical issues involved in the supposition that a concrete concept must bear an infinite number of semantic relations to an infinite number of higher-order properties, in order to bear an intrinsically representative relation toward one class of mind-external objects. Consider again an occurrence an instance of the belief that *all nuts are fruit*, which is a concrete and common sense-based mental state. Within this state is a concrete, intrinsically semantic concept that represents the class of all nuts. Due to the conscious and finite nature of such a cognitive act, it appears absurd to hold that such a concept bears an infinite number of intrinsically representative relations toward an infinite number of higher-order properties. Thus, it is safe to conclude that the infinite regress implied by the supposition that this concept intrinsically represents the class of all nuts by representing the property NUT in the way that the Property Principle requires is vicious.

The implication of the vicious regress is that if the Property Principle is to govern the intrinsically representative relationship between a concept and a class of objects, then the concept must present to the mind a property exclusively instantiated by that class in a way *other* than the way dictated by the Property Principle. That is, given the above-described regress, Bare Property Intentionality cannot hold that a concept C intrinsically represents a class of objects x through intrinsically *representing* a property that is exclusively instantiated by this class. In an instance of the belief that *all nuts are fruit*, the concept of the class of all nuts presents the property NUT to a person’s mind, but the concept does not intrinsically *represent* this property.

Under the assumption of the Property Principle, the relationship whereby an intrinsically semantic concept of a class of objects presents a property to the mind is not, therefore, an intrinsically representative relationship between the concept and the property. This reveals that there must be some other way for a concept to present a property to a person’s mind. This is because the concept does not represent the *property* exclusively instantiated by that classin a way that may be an ingredient in a truth-evaluable proposition.[[225]](#footnote-225) Yet, such a class concept does have some type of intrinsic feature (or set thereof) in virtue of which it presents that property to the mind, and in virtue of which it represents the relevant class of objects in a manner that may be an ingredient of a truth-evaluable proposition.I will now explain what makes concepts intrinsically semantic and representative, including their presentations of properties to the mind.

D. The Features of Concepts as Intrinsically Semantic Constituents of PAs

In order to represent all objects that are nuts as objects that are pieces of fruit, an instance of the belief that *all nuts are fruit* must first (somehow) represent each member of the class of all nuts. Yet, this cannot mean that whenever a person believes that *all nuts are fruit*, his or her mind represents each individual nut on its own.Rather, it means that his or her mind-internal concept of the class of all nuts must connect *semantically* with all and only nuts. Moreover, there must be an exclusive semantic relationship between the concept and the class of all nuts, in order for an instance of a belief that *all nuts are fruit* to be objectively evaluable as true or false.

A metaphysical relationship between this concept and some entity that all and only nuts have in common should determine the exclusivity of this semantic relationship between the concept and the class of all nuts. Since the property NUT is one thing that all individual nuts have in common, the concept of the class of all nuts only needs to connect semantically with one individual thing, namely, the *property* NUT, in order for such a semantic connection to obtain. Through this intrinsic connection, the concept bears an intrinsically semantic and representative relationship to the class of all nuts.

However, the fact that a classconcept essentially presents the property NUT to the mind is merely what provides it with its particular semantic *content*: a relation whereby it exclusively concerns the class of all nuts, as opposed to some other class of objects. For this reason, a proponent of Bare Property Intentionality says that a class concept must also include a bare referential intention. Why does the mere presentation of a property to the mind not suffice to make a concept represent a class of objects intrinsically? Because there are mental states with intrinsically semantic content that do not intrinsically represent any class of objects in a manner that one may evaluate as true or false. While the *propositions* that form the semantic contents of such states have truth-values, the *mental states* that contain these semantic contents are not *beliefs*, which could be evaluated as true or false. For these mental states are assertorically neutral: a person’s mind does not “say” anything about any object just by being in this type of state.

There are two main reasons to think that there are intrinsically semantic states that are not intrinsically representative. First, a proposition may be presented to a person’s mind (and thus a concept may be presented to a person’s mind) even if he does not affirm the truth of that proposition, i.e., even if he does not believe that the proposition is actually true. For a proposition to be presented to one’s mind is for a propositionally structured semantic content to be presented to one’s mind. According to the assumptions of this discussion, this involves the presentation of *properties* to one’s mind. It appears to follow that there is such a thing as having a semantic content in one’s mind, that is, having properties presented to one’s mind, without intrinsically *representin*g anything external to that state. [[226]](#footnote-226)

Moreover, since the deductive and inductive logical implications of a proposition can be presented to a person’s mind even if the person does not affirm or deny the proposition, one can be in a semantic and logically structured mental state without having an occurrent belief *that* the proposition is true or having an occurrent belief that the proposition is false. This means that one can have properties presented to one’s mind in a logically structured manner without intrinsically representing anything through those mental presentations. Such states are intrinsically semantic, without being intrinsically *representative*, because they are not beliefs in the truth or falsity of a proposition. Propositional attitudes, of which beliefs that *x is the case* are a subset, are obviously attitudes, in some sense, toward propositions. These can be cognitive attitudes, i.e., affirmations or negations of the truth of a proposition, or they can be emotional attitudes, i.e., desires that *x be the case*.

One may entertain a proposition, therefore, without having a propositional *attitude* (with that proposition as its content). This means that there are mental states that are not PA-instances, which nonetheless have logically structured semantic contents. Such states should have intrinsically semantic concepts, some of which *concern* entire classes of objects. For example, one could entertain the proposition that *all professional boxing fans are violence-prone*, without believing that this proposition was true or believing that it was false. One could also conclude that if this were true, then *all professional boxing fans are dangerous* would be true. In order to draw this conditional conclusion, moreover, one would have to entertain the proposition that *all professional boxing fans are dangerous* without affirming or denying it.[[227]](#footnote-227) It seems that the states involved in such a thought process are assertorically neutral, but they have internal propositional structures. While these contents are intrinsically semantic, they do not actually *represent* classes of objects or individual objects. According to Bare Property Intentionality, therefore, a class concept requires a BRI to direct the concept externally toward actual objects (or classes).

One may object that a lack of mental assertion, and not a lack of mental representation, constitutes the difference between a mental state of merely entertaining a proposition and a propositional attitude consisting of a belief that *x is the case*. However, the framework of Bare Property Intentionality requires that an intrinsically representative mental state be evaluable as a true or false mental assertion. If a propositional attitude-instance, such as Eric’s belief that *all bachelors are sports fans,* is to be *intrinsically* representative, then:

1. In virtue of the essential features of some occurrent phenomenon within Eric’s mind, Eric’s mind must represent all bachelors as instantiating the property SPORTS FAN.

(2) In principle, we must be able to evaluate Eric’s mental state as being true or false, i.e., we must be able to say that Eric currently believes something that is true, or that Eric currently believes something that is false.

The objection depends upon the assumption that a mental state may intrinsically represent a class of objects without asserting anything of that class of objects. A case that would exemplify this distinction may be a case in which Eric looks at a coffee mug and entertains the proposition that *this coffee mug was made in Korea*. Having no knowledge or evidence of the origin of the cup or of any facts about the circumstances under which it was manufactured, the objection implies that Eric can, nonetheless, mentally *represent* the cup as having been created in Korea without believing that it was or that it was not created in Korea.

While such a case may belong to a category of some type of mental representation, it is not a case of intrinsically semantic mental representation. To be intrinsically representative, according to Bare Property Intentionality, is to be a mental *state* that is objectively evaluable as true or false, not merely to be in a mental state whose propositional *contents* are evaluable as true or false. Intrinsic mental representation necessarily involves mental assertion. This is because for a mental state, or belief, to be objectively evaluable as true or false is for the owner of that state to assert something, or to mentally endorse the truth of a proposition. This does not mean that the subject must have an occurrent higher-order or meta-semantic belief, such as the belief that *my belief that all Hyundais are made in Japan is true*. Specifically, what it is for a PA-instance of this form to have an objective truth-value is for the subject of the belief to mentally assert that a particular class of objects, namely, the class of all Hyundais, has a certain property, namely, the property of being created in Japan.

In other words, the metaphysics of truth are distinct from the metaphysics of intentionality. While one genre of philosophical issues concerns the types of entities that must exist and the types of relations that must obtain in order for a proposition to be true, the other genre concerns the types of entities that must exist and the types of relations that must obtain in order for a belief to have an objective truth-value. One can entertain the proposition that *all Hyundais are made in Japan* without believing either that *all Hyundais are made in Japan* or that *some Hyundais are not made in Japan*. According to the current assumptions, this state of mere entertainment is intrinsically semantic; and the proposition that makes up its semantic content has the objective truth-value of being false. Yet, one cannot say that the state itself is false, because the person who is in the state has not mentally ascribed the property of Japanese origin to the class of all Hyundais. Thus, whatever type of mental representation occurs within the state of merely entertaining the proposition that *all Hyundais are made in Japan*, it is not the type of relationship between the mind and the world that is, in itself, true or false. Given the fact that such mental states of mere entertainment do occur, belief-instances, as assertorically representative mental states, must have some additional, distinctive feature that makes them evaluable as true or false. This feature provides PA-instances with intrinsic intentionality, as opposed to intrinsic semantic content alone.

To recall, one piece of evidence that some mental states entertain propositions but are not truth-evaluable PA-instances is the fact that certain types of inferential thought-processes occur. In these mental processes, one reasons that if X were true Y would be true, one neither believes that X is true nor that Y is true, and one believes neither that X is false nor that Y is false. For example, one may infer that if *all professional boxing fans are violence-prone then all professional boxing fans are dangerous* without affirming or denying that all professional boxing fans are violence-prone or dangerous. It is intuitive that within such a process, an internal mental association of the properties VIOLENCE-PRONE and DANGEROUS occurs; and that the inference depends upon such an association. A person of no philosophical education whatsoever, however, may go through such an inference easily. He or she would begin by entertaining the proposition that *all professional boxing fans are violence-prone* without mentally representing the *property* VIOLENCE-PRONE as a universal.

What is the explanation for this evidence that states of merely entertaining propositions exist? That is, what features must belong to the mental states involved in inferential thought-processes, such as that concerning the dangerous character of professional boxing fans? Bare Property Intentionality says that concepts included in these states mentally presentproperties, e.g., PROFESSIONAL BOXING FAN and VIOLENCE-PRONE,in a way that tends to bring about a reasoning process. A mere mental entertainment of the proposition *all professional boxing fans* *are violence-prone* includes such concepts. These concepts make a state intrinsically semantic but not intrinsically representative, because they do not *represent* any object in a manner that may be an ingredient of a truth-evaluable proposition.

The motivation for exploring this question is to begin to demonstrate the necessity, within Bare Property Intentionality, for an intrinsically representative mental state to have a BRI as an ingredient. Having established the existence of intrinsically semantic but non-representative mental states, the following section will address the question *why* a BRI is the feature that provides a state with intrinsically representative content. Both merely intrinsically semantic and intrinsically representative states must essentially present properties to the mind. Thus, a way to consider the question of the metaphysical role of BRIs is to ask what types of ingredients the latter type of state must have that the former type of state lacks.

In order for a state of merely entertaining the proposition that *all Fs are Gs* to lead one to conclude that the proposition that all *Fs are Hs* would follow from it, the concepts within the state must play a certain role. That is, their mental presentations of the properties F and G must present the logical structure and semantic content of these propositions in a way that suffices to bring about the relevant inference. Yet, the requirements are stronger for the concepts of an instance of an intrinsically semantic (and representative) *belief* that *all* *Fs are Gs*. For these concepts must direct the mind toward (possible or actual) objects in the world, i.e., the concept of the class of all professional boxing fans must make the belief represent each individual member of that class. Both this belief and a state that merely entertains the proposition *all Fs are Gs* involve mental presentations of F. Thus, the class concept cannot represent each professional boxing fan solely through mentally presenting the property F. An intrinsically representative class concept, therefore, must contain an ingredient in virtue of which it intrinsically directs the mind toward a class of objects. This constituent, in my view, is a bare referential intention.

The semantic function and metaphysical features of a BRI are distinct from those of mental presentations of properties. Each type of concept must have a different semantic content, e.g., a concept of the class of all nuts has a different content than a concept of the class of all bachelors. The fact that these concepts present distinct properties to the mind, i.e., the property NUT and the property BACHELOR, respectively, determines their distinct contents. Since every intrinsically semantic concept must present some given property to a person’s mind, and since a BRI is the type of ingredient that all class concepts have in common, all occurrent BRIs must be metaphysically and semantically associated with *some* given property. Yet, no particular BRI is *necessarily* associated with a mental presentation of a particular property. In other words, every BRI occurs in a concept that intrinsically represents some entity, such as the class of all bachelors. Yet, any given BRI could occur in a concept that represents any given property.

One can understand the logical and semantic role of a bare referential intention through an analogy with the logico-semantic role of a bare variable in a formal sentence. Consider the function of a bare variable (‘x’) in the sentence ‘(x) (Bx → Sx)’, which may express that any object that is bachelor is also an object that is a sports fan. [[228]](#footnote-228) The sentence cannot succeed with only predicate symbols, such as B and S, which merely stand for properties, since symbolizing the properties BACHELOR and SPORTS FAN (or some relationship between them) will not successfully say that any object that is a bachelor is also a sports fan. Thus, the sentence requires a different type of symbol, such as the bare variable ‘x’ that functions by representing objects.

A proponent of Bare Property Intentionality says that the logico-semantic role played by a BRI in an instance of the belief that *all bachelors are sports fans* is analogous to the function of this bare variable ‘x’. For ‘x’ is a symbol-type that must stand for objects (in contrast to properties), but it cannot essentially stand for any particular object (or kind). In this sense, the symbol-type ‘x’ plays a semantically *generic* role within a formal sentence, just as a BRI plays a semantically generic role within an intrinsically representative mental state. Within a syntactically sound formal sentence, such as ‘(x) (Bx --> Sx)’, ‘x’ represents every (possible and actual) object. This is because its function in the sentence is to facilitate the expression of: *any* thing *that is a B is an S*. Since a BRI is the ingredient in virtue of which any mental state intrinsically represents *something* as opposed to representing nothing, a BRI cannot essentially represent, in itself, any *particular* class (or individual).[[229]](#footnote-229)

This analogy also supports the view that to provide a mental state with an intrinsically representative feature, a BRI must be separate from a mental presentation of a property. For there is a syntactical basis for the semantic detachment of a bare variable from any particular object (or class), which is analogous to the basis for the semantic detachment of a BRI from any particular object (or class). A bare variable is not essentially attached to any particular predicate-symbol type or token; and a BRI is neither essentially attached to the presentation of any particular property to the mind, nor to any particular presentation of a property to the mind. For example, ‘x’ can be attached to a symbol for any type of property, such as in the sentence ‘(x) (Mx --> Ax)’, which may say express that *all objects that are mothers are objects that are adults*. Similarly, a BRI may be connected to a mental presentation of any property, such as in an instance of the belief that *all mothers are adults*.

Thus, the *semantic* role of a bare referential intention within a class concept is to allow the mind to represent any given class of objects through the medium of some property. Just as the relevant kind of formula requires the linguistic symbol ‘x’ in order to represent at least *some* object as opposed to no object, a mental state requires a BRI in order to represent *something* intrinsically, as opposed to representing nothing intrinsically. Just as the variable symbol-type ‘x’ is what all generalizations in first-order predicate logic must have in common, moreover, a BRI is what all instances of intrinsically representative PAs of the form *all Fs are Gs* musthave in common.[[230]](#footnote-230) This is because a bare referential intention has the most general type of semantic content. A mental state with intrinsically representative content requires an ingredient with this highest level of semantic generality.

One may object that one can entertain a semantic content or proposition of the form (x) (Bx → Sx) without mentally asserting it. The semantic function of a BRI is to provide a concept with an intrinsically representative feature, which, by assumption, includes an intrinsically assertoric feature. Thus, the logical role of a bare variable in the formula is not analogous to the semantic role of a BRI within a concept; and the syntactical necessity of a bare variable fails to justify the claim that a BRI is necessary for a concept to intrinsically represent objects. However, the claim that one can mentally entertain a proposition of the form (x) (Bx → Sx) without asserting this proposition can only mean two things:

1. One can mentally entertain the proposition that *all bachelors are sports fans*, for example,without having an instance of the belief that *all bachelors are sports fans.*

2) One can mentally entertain the more abstract proposition that (x) (Bx → Sx).

The possibility described by the first interpretation does not change the logico-syntactical function of the bare variable ‘x’ within the formula (x) (Bx → Sx). The relevant analogy is drawn between the necessity of the semantic role that this variable plays when such a formula *is* used to represent actual objects, on the one hand, and the necessity of the semantic role that a bare referential intention plays within a concept that is used to represent actual objects. Thus, the fact that one can mentally entertain the proposition that *all bachelors are sports fans* without believing that *all bachelors are sports fans* does not eliminate the necessity for BRIs. That is, an intrinsically representative concept still needs an ingredient that plays a logico-semantic role that is analogous to the role of a bare variable in a certain type of formal sentence. This function is to make the concept essentially represent an object (or class) by asserting something of that object (or class).

The possibility described by the second interpretation exists because one can mentally entertain the abstract logical form that belongs to all universal propositions, as interpreted by first-order predicate logic. Under this interpretation, one does not mentally entertain any first-order properties, such as BACHELOR or SPORTS FAN. When one entertains the proposition that *for any object x, if that x has a certain property B, then x has a certain property S*, it would seem that one is entertaining its abstract logical form. This, in turn, would mean that one is entertaining the logical implications and equivalences of this proposition without mentally asserting that it is true. This is because under (2), the symbols ‘B’ and ‘S’ do not have do not refer to any particular properties; and no particular first-order properties are presented to the mind. One cannot have a mental state describable as ‘an instance of a belief that *all Bs are Ss*’, unless ‘B’ and ‘S’ refer to properties presented to the mind within that mental state.

As a view of the intentionality of propositional attitudes, what is important about bare referential intentions for Bare Property Intentionality is their functional role within PA-instances. Therefore, this discussion will not provide a more extensive account of the features of BRIs themselves. Yet, reasons have been given above to think that BRIs are semantically generic and primitive ingredients. First, we have seen that some ingredient of a belief-instance must direct the mental state toward something external to itself; and this ingredient cannot be associated *necessarily* with any particular property.[[231]](#footnote-231) Therefore, some states must include a semantically generic constituent that is not necessarily associated with the mental presentation of any particular property. Next, we have seen that there is an analogy between the logico-semantic role of bare variables in formal sentences and the semantic role of BRIs in propositionally structured mental states. Finally, we have seen that PA-instances must have some type of constituent that provides them with intrinsically representative features; and we may conclude that these constituents must accomplish this feat through their intrinsic characteristics, in turn. A bare referential intention is the type of entity that satisfies all three of the above conditions.

Based upon the last conclusion, one can assume that bare referential intentions have one specific intrinsic feature: this is to be semantically *primitive*. This is because they cannot be associated necessarily with any particular presentation of any particular property to the mind. According to the metaphysics of Bare Property intentionality, this implies that a BRI can have no internal propositional or predicative content. Rather, a BRI is a concrete mental event that makes an occurrent intentional mental state directed toward something other than itself (or any of its own semantic ingredients).[[232]](#footnote-232) Thus, the functional definition of ‘bare referential intention’ is: “an ingredient in virtue of which a mental state may be intrinsically representative.” A BRI combines with a mental presentation of a property (or set of properties) to determine that towards which the state is directed. This is the sense in which BRIs are required for mental states to have intrinsic intentionality. Having explained the nature of intrinsically representative class concepts and their role in propositional attitude-instances, the following section will discuss the metaphysical features and intrinsically semantic roles of instantiation-concepts.

E. Intrinsically Presenting the Instantiations of Primary Properties to the Mind

The previous sections explained that an instance of the belief that *all nuts are fruit* involves a concept, which intrinsically represents the members of the class of all nuts as instantiating the property NUT by intrinsically presenting the property NUT to the mind. Yet, these discussions did not explain how this concept presents the property to a person’s mind. This question poses a challenging philosophical problem, because a mental state and its constituents are concrete entities, while a property is a mind-external universal. How can such a universal be presented to the mind? How does a concrete entity, such as a concept, present such a universal to the mind *intrinsically*? A proponent of Bare Property Intentionality must also account for how an instance of the belief that *all nuts are fruit* would intrinsically represent this class *as instantiating* the property FRUIT. To answer the latter question, he or she says that the mental state includes a concept of the instantiation of the property FRUIT: an “instantiation-concept.” This basic concept [[233]](#footnote-233) presents FRUIT to a person’s mind as being instantiated, so that when this concept is combined with the concept of the class of all nuts, the resulting mental state represents each individual nut as instantiating FRUIT. It is important to recognize, moreover, that an instantiation-concept is not an intrinsically semantic *representation* of the property whose instantiation it mentally presents.[[234]](#footnote-234) An instantiation-concept is intrinsically semantic, but it is not intrinsically representative.

The notion of an instantiation-concept does not, in itself, explain how the concept of the *class* of all nuts presents the property NUT to a person’s mind. For instantiation-concepts are not class concepts, and class concepts do not present properties to the mind as being instantiated. However, Bare Property Intentionality says that mental presentations of perceptible physical properties, such as NUT, originally develop from instantiation-concepts of those properties.A person’s mind *originally* gains contact with a perceptible physical property through the same intuitions concerning perceptual experiences that cause the development of an instantiation-concept. Thus, the notion of an instantiation-concept indirectly accounts for the ability of the class of concept of nuts to present the property NUT to a person’s mind.

Since instantiation-concepts function solely to present properties to a person’s mind as being instantiated by a class of objects, they do not merely present a particular, concrete instantiation of a property to the mind. Rather, they present both the property *qua* abstract object *and* its concrete instantiation to a person’s mind. Consider that the phrase ‘Jesus, child of Mary’ expresses the notion of Jesus as bearing the relation CHILD OF to Mary; and it therefore also represents *Mary* (as being a parent of Jesus). The phrase does not merely represent a child of Mary. By a loose analogy, an instantiation-concept of the property FRUIT does not merely present an instantiation of this property to the mind, but rather, it also presents the property to his mind *as being* instantiated. The term ‘presents the instantiation of’ is therefore interchangeable with ‘presents the property to the mind (or to a person’s mind) as being instantiated’ in this discussion.

From a more metaphysical perspective, all instantiation-concepts must share the same form. For all PA-instances (of the form *all Fs are Gs*) must share types of ingredients that allow them to have the appropriate internal logical structure. Yet, there is no one single component that all instantiation-concepts share. As a section that follows will explain, there actually need to be two types of instantiation-concepts: primitive and basic. Primitive instantiation-concepts (PICs) present the instantiation of a property, such as RED, to the mind only as some mind-external thing that tends to cause a certain type of perceptual experience. Thus, PICs belong to instances of PAs such as the belief that *all cherries are red*.[[235]](#footnote-235) In contrast, basic instantiation-concepts (BICs) actually succeed in presenting the instantiation of a particular mind-*independent* property in-itself, such as FRUIT, to the mind. As mentioned, BICs belong to instances of PAs such as the belief that *all nuts are fruit*. Moreover, a BIC develops from a more abstract mental process than that from which a PIC develops, and is a more abstract type of concept.

A proponent of Bare Property Intentionality holds that instantiation-concepts that present perceptible physical properties (such as FRUIT) to the mind develop from *intuitions* concerning perceptual experiences of objects, and are constituted by intuitions of these properties as mind-independent abstract objects that are manifested in the physical world. An *intuition* of x is a mode of presenting x to the mind without the aide of logically structured mental processes. Despite this paucity of representative and logical features, an instantiation-concept is intrinsically *semantic*, because its essential semantic function is to present a property to the mind as being instantiated. This function means that the features of this type of concept are fundamentally dissimilar to those of a *class* concept. Whatever accounts for this difference, moreover, must also account for the fact that the class concept, and not the instantiation-concept, allows the mind to represent concrete objects in the mind-external world.

Within an instance of the belief that *all nuts are fruit*, both types of concepts present properties to the mind: NUT and FRUIT. But the instantiation-concept does not pick out a class of objects in the world. Thus, its features must allow the mental state to present FRUIT to the person’s mind as the property that is merely instantiated by each nut, but whose instantiation does not exclusively single out pieces of fruit as a class. Otherwise, the state could mean that *all pieces of fruit are nuts*, or that *some things are nuts that are pieces of fruit*. In addition, the class concept does not present the property NUT to the mind as being instantiated. To see this, consider that in order for a concept of the class of all nuts to present the property NUT to the mind as being instantiated within an instance of the belief that *all nuts are fruit*, it would have to be a component of something that expressed an embedded proposition of the form *all Fs are nuts*. This would lead to a vicious infinite regress.

In other words, in order for a concept of the class of all nuts to present the property NUT to a person’s mind as being instantiated, this concept would have to be a component of an instance of a PA of the form *all Xs are nuts.* This would require that the concept of the class of all nuts combine with a distinct class concept, namely, the class of the concept of all Xs, to create a belief-instance of the form *all Xs are nuts*. Since the concept of the class of all Xs would also have to present the property X to the person’s mind as being instantiated, the belief that *all Xs are nuts* would actually have the form *all (all Xs are Y) are nuts*. The complex embedded ingredient of the form all *Xs are Y* would also have an embedded proposition, which would make the proposition have the form (*all* (*all* (*all Xs are Y*) *are nuts*)). This is an absurdity, and it would obviously lead a to a vicious infinite regress. Thus, the metaphysical assumptions of Bare Property Intentionality preclude the class concept contained in an instance of a belief of the form *all Fs are Gs* from presenting the property F to the mind as being instantiated.

The fact that instantiation-concepts present instantiations of properties to a person's mind is not the only difference between these concepts and class concepts. An additional distinction lies in the fact that the relationship between an instantiation-concept and the property that it presents to the mind is not analogous to the relationship between a class concept and the property that it presents to the mind. As was explained previously, every class concept shares one type of constituent, because each type of class concept is composed of qualitatively identical bare referential intentions. Thus, an occurrence of a type of class concept differs in its essential features from an occurrence of another type only with respect to the particular *property* that it presents to the mind. For example, the concepts of the class of all nuts, of the class of all spheres, and of the class of all bachelors each has a bare referential intention as an ingredient.

In contrast, while all instantiation-concepts present *some* property to the mind as being instantiated, and while some instantiation-concepts present one and the same property to the mind as being instantiated, no one type of *constituent* is held in common by all. This is because a mental presentation of the instantiation of a property is not a component or ingredient of a concept. It is the event of an instantiation-concept occurring, i.e., an event of a person having an intuition. For instance, the instantiation-concepts of FRUIT, of SPORTS FAN, and of GOLDEN MOUNTAIN share no one type of ingredient. Having understood the general differences between instantiation-concepts and class concepts, the next section will answer the most important question: what features allow an instantiation-concept (whether primitive or basic) to present a property to the mind as being instantiated?

F. The Semantic Roles of Basic Instantiation-Concepts

Like most propositional attitude-instances, an instance of the belief that *all nuts are fruit* concerns the instantiation of a primary property. The property FRUIT is a non-perceptually based (or non-dispositional) primary property, since an object’s instantiation of FRUIT is independent of any tendency that object has to cause any particular type of perceptual experience. For what makes an object a piece of fruit is *not* any tendency that it has to cause any kind of qualitative phenomena in humans.[[236]](#footnote-236) A thing is a fruit solely in virtue of having certain intrinsic, primary physical properties; and these properties can be called organic, biological, or reproductive in nature. While most types of fruit tend to cause certain types of perceptual experiences in humans, such as tangy-sweet, juicy, and fleshy tastes, these secondary qualities are mere epiphenomena, which are not intrinsic to the organic identity of fruit as a class of objects. The following section will describe the metaphysical features in virtue of which a BIC of FRUIT, such as that involved in an instance of the belief that *all nuts are fruit*, intrinsically presents a particular *primary* property, i.e., FRUIT, to the mind as being instantiated.

How does a concept present a property to the mind? Specifically, how can a basic instantiation-concept *intrinsically* present the instantiation of FRUIT to the mind as the instantiation of a particular primary property?This way of presenting FRUIT to the mind is quite different from presenting it merely as something in virtue of which objects cause certain qualia. For while pieces of fruit do tend to taste certain ways, what it is *to be* a fruit is not determined by this contingent and extrinsic feature. Therefore, a BIC must present this property as being distinct from any disposition that an object has to cause qualia.

Bare Property Intentionality says that the exclusive metaphysical relationships that instantiation-concepts have with properties allow them to present those properties (and their instantiations) to a person's mind intrinsically and exclusively. A BIC presents an instantiation of the property FRUIT to that person’s mind as an individual manifestation of something that is held in common by more than one object, that is, something that is present in all pieces of fruit. Thus, this concept functions to present the instantiation of FRUIT to the mind as a manifestation of an abstract object.

On the one hand, a BIC must present an instantiation of FRUIT to a person’s mind as a manifestation of something that is not identical to any disposition to cause any specific kinds of perceptual experiences. On the other hand, a BIC involves a person’s intuition of this property as being the mind-independent source of this tendency. For each particular class of middle-sized perceptible physical objects, such as the class of all pieces of fruit, one may assume that there is a specific set of perceptual experiences that members of that class exclusively tend to cause humans to have at the same time (under normal perceptual conditions). The tangy-sweet, juicy, and acidic qualia are the ones that *only* members of the class of all fruit exclusively tend to cause people to experience simultaneously. [[237]](#footnote-237) Therefore, the property FRUIT, which is exclusively instantiated by members of that class, has a special type of relationship with that set of qualia. For under normal conditions, the only thing outside of a person’s mind that causes these qualia is the instantiation of FRUIT in objects. The property FRUIT is thus the unique and ultimate *source* of the causal relationship between its own instantiations and the relevant sets of perceptual experiences.

To fulfill these philosophical responsibilities, a proponent of Bare Property Intentionality must give a basic instantiation-concept of the property FRUIT a specific metaphysical character and semantic function. But what are these features? First, the concept presents the property FRUIT to the mind as something mind-independent and as *distinct* from any entity that *causes* the tangy-sweet, juicy, and fleshy perceptual experiences. Next, a BIC of FRUIT is an intuition of FRUIT as a mind-independent entity that is not a concrete object, but which is *present* in concrete objects (through its instantiations). In this sense, a BIC presents instantiations of the property FRUIT to a person’s mind as instantiations of an *abstract* object, i.e., as concrete instances of a primary, mind-independent, physical property.

One may wonder how the qualia that fruit cause can be divorced from the essential semantic content of a BIC, that is, from the way that the concept essentially presents the property FRUIT to the mind. How does the BIC present the property to the mind as separate from the qualia it causes, which are not essential to its identity? The intuition that a basic instantiation-concept of the property FRUIT involves is a mental *abstraction*, in the sense that it presents an abstract object to a person's mind as being instantiated, and in the sense that it is the result of a cognitive process of abstraction. This process has abstracted the notion of a mind-independent universal from the notion of the ultimate source of the fruit-related qualitative experiences. Thus, the qualitative experiences that pieces of fruit cause are divorced from the essential semantic content of a BIC, that is, from the way that the concept presents the property FRUIT to the mind. Although the constituents of a BIC of a perceptible physical property, such as FRUIT, must originate with intuitions concerning perceptual experiences, the *intrinsically semantic* content of the BIC is not connected essentially to internal qualitative phenomena. The next section will provide an account of the features of a BIC.

G. The Metaphysics of Intrinsically Semantic Basic Instantiation-Concepts

Within an instance of a belief that *all nuts are fruit*, a basic instantiation-concept presents the property FRUIT to the mind as being instantiated (by each individual nut). This section will provide an account of this BIC and its intrinsically semantic character. To facilitate this discussion, assume that under normal human perceptual conditions, pieces of fruit exclusively instantiate these three secondary properties at the same time:

G1: A disposition to cause a particular tangy-sweet taste,

G2: A disposition to cause a particular juicy tactile sensation, and

G3: A disposition to cause a particular fleshy taste.

According to Bare Property Intentionality, a BIC of the property FRUIT is an immediate awareness, or intuition, of the instantiation of one particular property that is, in turn, instantiated by all objects that tend to feel juicy and to taste tangy-sweet and fleshy.[[238]](#footnote-238)

Bare Property Intentionality’s account of the semantic content of BICs is based upon certain assumptions concerning the relationship between perceptible physical properties and sets of dispositional properties, such as G1 throughG3 above. Suppose that when a particular person encounters a piece of fruit in a certain way, that is, by tasting it, he experiences these perceptual phenomena simultaneously: a particular tangy-sweet taste, a particular juicy tactile sensation, and a particular fleshy taste. As a result, a pattern of correlation develops between occurrences of these qualia in a person’s mind. The immediate cause of this pattern is the pattern of simultaneous instantiation of the three *secondary* properties, G1 ─ G3 which are obviously dispositions to cause these perceptual experiences. While concrete, individual instantiations of the property FRUIT actually cause these correlated qualia, the abstract object FRUIT itself is the ultimate, mind-independent *source* of both patterns of correlation.

The property FRUIT has certain logical and metaphysical relationships with the set of secondary properties G1 throughG3 above. These relations are exclusive because under normal human perceptual conditions, only an object that instantiates FRUIT instantiates each member of this set of properties simultaneously. A person's BIC of FRUIT is an immediate awareness of these logical and metaphysical relationships. When a BIC of the property FRUIT occurs in such a person’s mind, he has a basic metaphysical intuition of a particular mind-independent property as being the source of the pattern of simultaneous instantiation of secondary properties G1 ─ G3. He also has a basic logical intuition of the exclusive relationship between the property FRUIT and this set of secondary properties. That is, he is immediately aware of objects that instantiate FRUIT as being the only objects that simultaneously instantiate each of these perceptual dispositions.[[239]](#footnote-239) Moreover, a BIC presents FRUIT to a person’s mind as being the *exclusive* source of these causal tendencies.

A proponent of Bare Property Intentionality must explain how the content of a basic instantiation-concept involves the presentation of the instantiation of a particular, mind-independent property, such as FRUIT, to a person's mind. The above discussion addressed the ways in which the content of this BIC is connected to the property FRUIT, the relevant set of perceptual experiences, and the relevant set of secondary properties. An advocate of this view of intentionality must also answer the question: what makes this mental presentation *intrinsically semantic*? The answer is that a BIC of FRUIT has a special relationship with two things: the property FRUIT itself as an abstract object and a higher-order property that FRUIT instantiates. This relationship accounts for the intrinsically semantic content of the concept.

A BIC of FRUIT occurs when an individual has an intuitive and exclusive awareness of the instantiation of FRUIT as a *particular* primary property, as opposed to being aware merely of the instantiation of *some* property. As I will explain later, a mental process of abstraction causes the initial occurrence of this concept in a person’s mind. This process begins when the person becomes immediately aware of a pattern of correlations between specific types of internal qualitative phenomena: tangy-tastes, juicy tactile sensations, and fleshy tastes.

However, the exclusive and intrinsically semantic relationship between the property FRUIT and a BIC of FRUIT goes beyond fruit-related perceptual experiences and dispositions to cause those experiences. This is because while FRUIT is a first-order property, it exclusively instantiates a certain second-order property. Since, by assumption, all objects that are disposed to cause the relevant set of perceptual experiences simultaneously (under normal conditions in contemporary humans) instantiate FRUIT, only FRUIT instantiates: INSTANTIATED BY ALL OBJECTS THAT INSTANTIATE THE SET {G1, G2, ... G3} (UNDER NORMAL CONDITIONS). Thus, there is a metaphysical connection between the property FRUIT and this set of secondary properties**:** under normal human perceptual conditions,when the set is instantiated, FRUIT is instantiated.

One may wonder how this connection between the property FRUIT and the set of dispositional properties G1 throughG3 makes it the case that a person’s mind may have an *intrinsically* *semantic* connection with the property FRUIT (or with its instantiations). The answer lies in the connection between a person’s mind and the second-order property described above. This connection allows a BIC to have an abstract semantic content, a content that presents an abstract object (through its instantiation) to a person's mind. As I will explain in more detail later, a proponent of Bare Property Intentionality believes that humans are disposed toward having abstract intuitions concerning the sources of their perceptual experiences. Based upon this assumption, it is reasonable to believe a person can become intuitively aware of the relationship between sets of fruit-like perceptual experiences and a set of external dispositions to bring them about. From this, he develops a basic, *immediate* awareness—not an analytical or philosophical understanding—of the connections between this set of dispositions and a particular, mind-independent, primary physical property.

While this awareness is immediate, it is also *relatively* abstract, for several reasons. First, it is an awareness of a relationship between things, and not merely an awareness of a thing (or things). Second, it is an awareness of a higher-order relationship: a connection between one relationship and something else. For the property FRUIT is the source of the relationship between a set of perceptual experiences and the set of secondary properties. Finally, it is not merely an awareness of something concrete, such as an object, a particular perceptual experience, or an individual instantiation of a property. To understand the source of the exclusive and intrinsically semantic connection between the mind and FRUIT, consider that only pieces of fruit have secondary property F (under normal conditions), while the property FRUIT itself has property F1:

F: TENDS TO CAUSE TANGY-SWEET, JUICY, AND FLESHY QUALIA (SIMULTANEOUSLY)

F1: INSTANTIATED BY ALL OBJECTS THAT TEND TO CAUSE TANGY-SWEET, JUICY, AND FLESHY SENSATIONS (SIMULTANEOUSLY).

Suppose that a person has a tendency to develop an immediate awareness of the instantiation of a mind-external thing, which is held in common by all things that cause her to experience tangy-sweet, juicy, and acidic qualia at the same time. Having become aware of this, she then has an intuition that is more abstract from her immediate awareness of something that tends to cause her to experience certain simultaneous perceptions. This is because it is an intuition of the instantiation of a property (F1) that is *not* defined in terms of its causal responsibility for perceptual phenomena.

In simpler terms, once a person has an immediate awareness that something is held in common by all objects that taste tangy-sweet, juicy, and fleshy — something in virtue of which they tend to cause her to have those perceptual experiences — she becomes immediately aware of the instantiation of something objective in those pieces of fruit. This awareness is an intuition of the existence (and instantiation) of something held in common by those objects that is not merely *identical* to their tendency to create these tastes in her mind, but which, nonetheless, is the *source* of this causal tendency. By definition, a person’s intuition says that this “something held in common” by all objects that tend to cause tangy-sweet, juicy, and fleshy qualitative phenomena is something independent of the mind. The primary and first-order physical property FRUIT is the only object that instantiates this primary but higher-order property, namely, F1. Therefore, her awareness of the instantiation of F1 creates an exclusive, intrinsically semantic connection between her mind and FRUIT.

It is important to recognize that this intuition of the instantiation of F1 is not an analytical or theoretical conceptualization of this second-order property (or of its instantiation), but rather an immediate, basic level awareness. Moreover, while the abstract object FRUIT is the ultimate source of the tendency of pieces of fruit to cause the relevant qualia, concrete instantiations of this property are *causally* responsible for this tendency. Under normal conditions, only physical objects that actually are pieces of fruit cause those sensations simultaneously in human beings, so that these sensations have no connection with the abstract object FRUIT except through its manifestations in real pieces of fruit. For this reason, the connection between a person’s mind and the property FRUIT in these circumstances is actually an intrinsically semantic concept of FRUIT as being instantiated. An occurrence of this connection is an occurrence of a basic instantiation-concept.

H. Intuitions and Higher-Level Processes that Create Instantiation-Concepts

The previous discussions of the metaphysics of propositional attitude-instances may be considered to leave several questions unanswered: In exactly what sense are instantiation-concepts intuitive? How do they develop? When these concepts occur, in what concrete way does the mind connect concrete, mind-internal perceptual experiences with mind-independent abstractions (and their instantiations)? The philosophical purpose of this section is to explain the sense in which instantiation-concepts are intuitions, to explain the intuition-based mental process of abstraction from which they develop, and to explain the role that perceptual experiences play in their development. However, this discussion of their origins is not an account of their intrinsically semantic character. That is, the causal explanation of their development and occurrence is not a *metaphysical* view of how they intrinsically present the instantiations of abstract objects to the mind.

As an earlier section mentioned, instantiation-concepts may be primitive or basic. Primitive instantiation-concepts (PICs) are ingredients of mental states like an instance of the belief that *all cherries are red*. This is because they present the instantiation of a secondary property, such as RED, to a person's mind as some mind-external (but not mind-*independent*)entity that tends to cause a certain type of perceptual experience, such as red-like visual phenomena. Basic instantiation-concepts (BICs) are ingredients of mental states like an instance of the belief that *all nuts are fruit*. They connect a person's mind with the instantiation of a particular mind-independent property in-itself, such as FRUIT. When a BIC of FRUIT occurs with a class concept of fruit, a person's mind intrinsically represents each nut as being a piece of fruit. Both BICs and PICs develop from a mental process of abstraction over a person's intuitions concerning the source of his perceptual experiences. Yet the process that creates a BIC occurs at a higher level of abstraction than that which brings about a PIC; and the content of a BIC is more abstract than that of a PIC. Since most properties that objects instantiate are primary properties, this discussion will focus primarily on the abstraction process that causes *basic* instantiation- concepts.

To recall, while some intuitions cause instantiation-concepts, others actually *constitute* instantiation-concepts. As explained above, an *intuition* of x is a mode of presenting x to the mind without the aide of logically structured mental processes. The mental states in the process through which an instantiation-concept develops are intuitive because they are states in which a person is immediately aware of something without the aid of inference or reasoning. In addition, these states can occur without explicit or conscious thought; and in many cases, their contents involve fundamental, axiomatic principles that can be considered as prerequisites for rational thought. Awareness of these basic principles is immediate, since one does not arrive at them through deductive or inductive inference. Finally, while one can describe contents of these states in propositional or sentential form, they are not PA-instances because they are intrinsically semantic, but not intrinsically *representative*.

The process through which instantiation-concepts develop is an *abstraction* process because during these processes, a person applies his innate awareness of basic logical, arithmetical, and metaphysical relations to intuitions concerning the source of his perceptual experiences.[[240]](#footnote-240) Ultimately, these intuitions concern the instantiation of an abstract object as being the mind-external source of red-like qualia, for example, or as being the mind-independent source of fruit-like qualia, for example. A person’s mental tendency to process perceptual experiences in a certain way is the original cause of the instantiation of the secondary property RED (or the primary property FRUIT) being presented to his mind.[[241]](#footnote-241)

One could look at the development of a PIC of the property RED, for example, as follows: There is an initial time that a child goes through this abstraction process concerning red-like visual experiences. The process may begin when he sees something red and has an intuition of there being other things in existence that are not identical to his or her inner qualia of redness. The content of this state is intuitive, even though it can be described in propositional form, i.e., as ‘an intuition “that” *things exist that are not identical to my inner qualia of redness*’. For during its occurrence, its content is presented to the child’s mind without the aide of logically structured mental processes. That is, he does not undertake an inference to arrive at a *conclusion* that there are things that are not the same as his internal red-like qualia.

From this first state, the child may develop an intuition of the existence of things that are somehow *outside* of his seat of consciousness, as opposed to being merely distinct from his qualia. That is, he becomes aware of there being things external to and distinct from his self. At some point, he becomes aware of something external to himself as being the *cause* of red-like visual phenomena. His encounter with red-like visions (along with other types of qualia) may then lead him to intuit the existence of *more than one* entity that causes red-like qualia.

For example, in one situation, he might encounter a cherry, and experience a red-like sensation along with a sensation of tangy-sweetness. On another occasion, he may encounter a droplet of blood, and experience a red-like sensation along with a visual or tactile sensation of liquidity; and on still another occasion, he may encounter a red leaf and experience a red-like vision in combination with a dry, solid tactile experience. These experiences (and the memories through which they become associated) lead him to have an intuition of the existence of more than one external thing that brings about red-like perceptual experiences. Whatever causes him to sense the existence of more than one source of red-like qualitative phenomena will eventually cause him to have an intuition of there being one mind-external entity that is held in common by *all* of the individual external things that bring about red-like phenomena. This is the secondary property RED. Finally, he has a primitive instantiation-concept of the secondary property RED. This is an intuition of the instantiation of some property as being the mind-external source of his red visual experiences.

When the child's mind moves from the intuition of several individual causes of individual red-like qualia to the intuition of all of these causes as being instantiations of one thing, it is undergoing a process of *abstraction.* A person’s immediate awareness of an instantiation of redness in a particular object as being the immediate *cause* of some particular red-like visual experience that he has undergone is an awareness of something concrete. Similarly, his intuition of several different red-like perceptual phenomena as being caused by several individual instantiations of redness is an intuition of several different concrete things as being in existence. However, when he begins to intuit that all of these distinct rednesses are manifestations of one entity, namely, the secondary property RED, he has become aware of the existence and instantiation of something abstract.

Having explained how a person develops a PIC, I will now discuss more precisely the series of intuitions through which he develops a *basic* instantiation-concept. Within aninstance of a belief that *all nuts are fruit*, a BIC presents the property FRUIT to a person’s mind as being instantiated (by each individual nut). Earlier, I made the assumption that pieces of fruit exclusively instantiate three dispositional properties simultaneously, which are tendencies to cause a particular tangy-sweet taste (G1), a particular juicy tactile sensation (G2), and a particular fleshy taste (G3). While acknowledging that some objects that are non-sweet, non-juicy, and non-fleshy, such as peanuts, are also pieces of fruit, I made the assumption for discursive purposes that a person’s initial encounter with a piece of fruit should produce a correlation of these types of perceptual experiences. Each of these qualia manifests one of the secondary properties G1 through G3. In the mental process that he underwent to develop a PIC, this type of interaction caused the child to have an intuition of the instantiation of *some* mind-external entity, i.e., some property, as being the source of this correlation. Yet, a BIC results from his undergoing a higher-level abstraction process. The consequence of this process is his immediate awareness of the instantiation of a particular, mind-independent physical property.

Just as in the case of a PIC, the first intuitions of the child's BIC development process concern the cause of his perceptual experiences. While intuitions in the PIC development process only concern more than one occurrence of the *red*-like visual experience, the intuitions leading to a BIC concern the cause or source of a *pattern of correlations* between different typesof qualia. The first three states in this series are:

1) An intuition of there being a pattern of correlations among his fruit-like perceptual experiences,

2) An intuition of there being a mind-external cause of this pattern of correlations, and

3) An intuition of there being a set of tendencies to cause these three perceptual experiences (the set of secondary properties G1 ─ G3), which are co-instantiated in a mind-external entity.[[242]](#footnote-242)

The first state concerns a pattern in which tangy-sweet tastes, juicy tactile sensations, and fleshy tastes occur simultaneously in the person’s mind, while the second concerns there being something outside his mind that *causes* this correlation.[[243]](#footnote-243) The third intuition is his immediate awareness of a pattern in which several different “things” that are co-present outside his mind cause the pattern in which the three types of sensations simultaneously occur.

As mentioned earlier, the states that occur in the mental process through which instantiation-concepts develop involve intuitions of certain fundamental logical, metaphysical, and arithmetical relations. Specifically, these relations are: (i) x IS IDENTICAL TO y, (ii) x is EXTERNAL TO y, (iii) x CAUSES y, and (iv) x IS A SUCCESSOR OF y. One can understand how a person’s basic awareness of these relations manifests itself in the three intuitions mentioned above. For example, when a person undergoes a mental state such as (2), he has an insight of something that is highly correlated with perceptual experiences as being external to and distinct from his self and his perceptual experiences. This awareness involves an intuition of the relations of externality and numerical or logical identity. M Intuition (2) also involves an awareness of the relation of causality: the person having the intuition has an insight of there being something outside of his own mind that brings about the perceptual correlations. Similarly, when he has a mental state like (3) above, he has a basic awareness of a set of tendencies in external objects, which are causally responsible for a certain pattern. This is the pattern in which he experiences a tangy-sweet taste, a fleshy taste, and a juicy tactile experience simultaneously.

It is important to note that a person’s awareness of these causal relations does not make up a BIC. As mentioned before, the causal origins of a BIC do not establish its metaphysical features or its intrinsically semantic content. While there is certainly some bio-chemical property (or set thereof) that grounds the disposition of pieces of fruit to have a certain kind of tangy-sweet taste, for example, this is not identical to the property of *being* a piece of fruit. Thus, a person’s BIC of FRUIT is distinct from an awareness of the instantiation of a property as that which merely forms the basis for a piece of fruit to cause a certain group of qualia. Appropriately, the fourth mental state that occurs in the process of causing a BIC of FRUIT is his immediate awareness of something outside his mind as being the *source* of the tendency of different things to taste tangy-sweet, juicy, and fleshy. This is a higher-order causal intuition, which is on a higher level of abstraction than his mere awareness of the instantiation of a set of dispositions to cause the relevant qualia. The tendency of pieces of fruit to cause these qualia makes these qualia occur simultaneously. Thus, the fourth mental state intuits the existence of a source for this tendency, which is distinct from the tendency itself. While state (3) merely concerns the existence and co-instantiation of G1 through G3 in pieces of fruit, the fourth state is an awareness of the instantiation of this set of properties *itself* as having a source in something outside the mind.

Mental state (4) also involves immediate awareness of relations (i), (ii), and (iv) above. This state develops from a person’s intuition of some mind-external thing that is not identical to any particular thing as being the source of the three causal-perceptual tendencies G1 through G3. For while the source of such tendencies being held by *one* piece of fruit could be concrete and particular, nothing that is the source of this tendency in *all* of the pieces of fruitthat have it can be concrete and particular. This intuition manifests a person’s sense of the difference between the relation whereby x is the source of a feature of one object and the relation in which x is the source of a feature held by *more* than one object. This involves a logical and numerical sense of the identity relation, as well as a sense of the arithmetical difference between the number one and any given number that is more than one.

Two serious questions may arise about intuition four. An empirical question asks whether its semantic content is true, while a philosophical question asks why this state is necessary for a basic instantiation-concept of the property FRUIT to develop. There are good reasons to believe that the property FRUIT is the ultimate *source* of the tendency of pieces of fruit to taste tangy, sweet, and fleshy (while its instantiations are the *cause* of this disposition). However, from the philosophical perspective of Bare Property Intentionality, what matters is that this semantic content occurs in the mind of a person who is developing a concept of the instantiation of the property FRUIT. Moreover, the development of a BIC of FRUIT requires an intuition like the fourth mental state because it presents FRUIT to the mind as being distinct from the relevant secondary properties (or from any particular co-instantiation of these properties). The simple, intuitive way for the mind to do this is to have an intuition of the instantiation of a property as the *source* of *any* object’s being disposed to cause the relevant perceptual experiences simultaneously.

Once a person becomes aware that no secondary property is the source of the pattern in which pieces of fruit tend to taste tangy-sweet, fleshy, and juicy simultaneously, he intuits that this source is a mind-independent primary property. While the abstract object FRUIT cannot cause anything, its particular, concrete instantiations can cause perceptual experiences, and the source of these instantiations is obviously the primary property itself. Therefore, the fifth mental state in the process of creating a BIC of FRUIT involves a person’s intuition of the existence and instantiation of some primary property. While state (4) is only an awareness of a mind-independent source of the occurrence of all of these dispositions, mental state (5) is an awareness of something *abstract* as being instantiated in many different objects that taste tangy-sweet, juicy, and fleshy.

Finally, after a person becomes aware of the instantiation of *some* mind-independent primary property as the source of a pattern of correlations between certain of his qualia, he becomes aware of FRUIT as a *particular* property. Specifically, this mental state is an intuition of the instantiation of the one particular, mind-independent property that is instantiated by all objects that tend to taste tangy-sweet, juicy, and fleshy.

What reason is there, one may ask, to think that the mental process through which an instantiation-concept develops is a process of *abstraction*, which results in the presentation of an *abstract* object to the mind? The first reason, as was explained earlier, is that the content of the third, fourth, and fifth intuitions clearly concern logical, arithmetical, and metaphysical intuitions. I propose two additional such reasons: First, the content of each state in the series is both more abstract, in the common sense of the term, than the preceding intuition and the relevant fruit-like phenomenal experiences. This is because each successive mental state in the series comes closer to being an intuition of the instantiation of a property as an abstract object *per se*, as apart from any application of that property to a particular object. Second, each successive state in the series comes closer to being an exclusive intuition of the instantiation of a *particular* property.

In the process that creates a basic instantiation-concept of FRUIT, for example, the content of each mental state from (1) to (3) is more abstract than both the state that immediately precedes it and the phenomenal experiences from which the intuition originates. Intuition (2) concerns the mind-external *cause* of the pattern of correlations between tangy-sweet, fleshy, and juicy qualia, while the first mental state only concerns this pattern. Intuition (3), in turn, concerns a set of external *tendencies* to cause these perceptual correlations. Moreover, while intuition (3) is about the set of secondary properties as the cause of the pattern of correlated perceptual experiences, intuition (4) is about the *source* of this cause. Thus, (4) moves even closer to being a mental presentation of an abstract object. A person in this state is aware of the existence of some entity that is the source for any object’s tendency to taste tangy-sweet, juicy, and fleshy at the same time. This entity can exist in any number of objects; and thus state (4) comes closer to being an intuition of a property *per se,* as apart from any application of that property to a *particular* object. The fifth mental state, in turn, is an awareness of some particular property as the source of the coincidence between members of a groupof dispositions.Finally, intuition (6) is a mental state that specifically concerns the instantiation of the property FRUIT.

I. Objections to Bare Property Intentionality

One may raise several objections against the view of intrinsically representative propositional attitude-instances that this chapter has proposed. First, one may ask why, according to the view given in section C, it is more plausible that properties can be presented to the mind in a way other than that which the Property Principle dictates than it is plausible that individuals (and classes of them) can be so presented. In response, it is important to note that given the Property Principle — which is the guiding principle of this theory of intentionality — concepts must present classes of physical objects to the mind *through* properties. This principle is supported, independently, by analysis of the ways in which a view of intrinsic intentionality should meet the three explanatory criteria that were imposed earlier. The regress argument revealed that the assumption of the Property Principle means that PA-instances *must* present certain properties to the mind in a different way than the principle dictates. On the other hand, analysis of the ways that a view of intentionality should meet the three explanatory conditions does not appear to imply any corresponding “Object Principle.” Moreover, there are no intuitions concerning an *a priori* connection between classes of *objects* and intrinsic intentionality, while there are such intuitions concerning an *a priori* connection between properties and intrinsic intentionality.

To recall, the first explanatory requirement that I imposed upon a successful theory of intentionality is that the theory account for the tendency of an instance of a propositional attitude-instance to cause other PA-instances in a truth-preserving manner. One may argue that Bare Property Intentionality’s claim that class concepts and basic instantiation-concepts combine to form PA-instances of the form *all Fs are Gs* does nothing to account for truth preservation among such mental states. In response to this objection, I point out that the logical structure that a BIC provides to such a mental state only must suffice to account for causal relationships among series of such states whose semantic contents have this inferential form:

a) All Fs are Gs.

b) All Gs are Hs. Therefore,

c) All Fs are Hs.

Following is an application of this pattern to propositional attitude-instances of the form *all Fs are Gs*:

a) An instance of the belief that *all nuts are fruit*, and

b) An instance of the belief that *all fruit are cherries*, which together cause:

c) An instance of the belief that *all nuts are cherries*.

One way to begin to address this issue may be to ask whether humans are innately disposed to go through certain types of inference-patterns. For example, the occurrence of an instance of a belief of the form *all Fs are Gs*, in combinationwith an instance of a belief of the form *all Gs are Hs*, may manifest an innate tendency to cause an instance of a belief of the form *all Fs are Hs.* This suggestion does not seem to succeed because if the semantic content of an intuition involves a logical principle, it is reasonable to assume that this principle must be *axiomatic* in order for the content of this intuition to be *innate* to the mind. Yet, there is nothing axiomatic, from a logical perspective, about the relevant deductive inference rule.

An alternative may be to interpret the semantic content of a proposition of the form *all Fs are Gs* in terms of class inclusion and exclusion. Under this interpretation, the truth-preserving causal patterns of an instance of any PA of this form might be explained through the metaphysics of representing classes of objects as standing in certain relationships of inclusion and exclusion. This way, one might not need to appeal to the notion of the innateness of certain deductive inferential patterns. The class-oriented semantic perspective says that a sentence (or mental state) that expresses the proposition *all nuts are fruit* refers to each individual member of the class of nuts: it says that each individual member of the nut class is also a member of the fruit class. Yet a sentence (or mental state) that expresses the proposition does not refer to each member of the class of all fruit.[[244]](#footnote-244) For it says that the class of fruit-things contains the entire class of nut-things, but it does not say that each individual member of the fruit class is a member of the nut class.

The class-oriented interpretation of the semantic content of a universal proposition has an important implication for the metaphysics of an instance of a belief that *all nuts are fruit*. For if the proposition says nothing about each individual piece of fruit, then the concept that presents FRUIT to the mind cannot represent each individual piece of fruit. Therefore, it must be metaphysically distinct from the type of concept that Bare Property Intentionality defines as a “class concept.” This implication restricts the options that a proponent of Bare Property Intentionality has for explaining of the truth-preserving causal tendencies of such a belief-instance. The content of a state cannot be interpreted in terms of class inclusion. This means that one cannot account for the truth-preserving causal patterns of such a mental state in terms of any intuitive awareness, whether innate or acquired, of an inference rule governing universal propositions concerning *classes*. In other words, one cannot account for mental truth-preservation over beliefs in universal propositions through an appeal to an immediate awareness of the conditional principle that says: *if class F is contained within class G, and class G is contained within class H, then class F is contained within class H*.

The truth-preserving tendencies of instances of PAs of the form *all Fs are Gs* cannot be interpreted in terms of the semantics of class inclusion or innate knowledge of deductive inference rules. Moreover, the assumptions of this essay prevent some a priori semantic connection between these mental states and objects from accounting for these tendencies. Therefore, causal relations between intuitive mental presentations of properties should be the explanation. For instance, it may be suggested that a person who has a belief that *all things that instantiate F also instantiate G,* anda belief that *all things that instantiate G instantiate H,* these two mental states cause him to believe that *all things that instantiate F also instantiate H*. The explanation for this causal tendency may be that a person's having instances of the first two beliefs creates two mental associations. The first obtains between his mental representations of the class of all Fs and his mental presentations of instantiations of property G. The second obtains between his mental representations of the class of all Gs and his mental presentations of instantiations of properties H. The combination of these mental associations beings about an instance of the belief that *all things that instantiate F also instantiate H,* because it brings about an association between he class of all Fs and his mental presentations of instantiations of H.

It is important to note that the view of mental truth-preservation being espoused is not that properties, *qua* universals, explain any causal relations between mental states (in the sense of physically bringing them about). Rather, the view depends upon the common-sense assumption that that in the actual world, the existence of some types of causal patterns between instances of certain types of mental states is inevitable. As a metaphysical theory, Bare Property Intentionality purports to explain why the states that, as an empirical matter, *do* tend to be caused by a given type of state have the particular semantic contents that they do.

A final objection to Bare Property Intentionality might focus on the metaphysical *asymmetry* between class concepts and instantiation-concepts. To recall, an earlier section argued extensively that in an instance of a belief such as *all nuts are fruit*, the constituent that presents the property NUT to the mind must, in turn, have a constituent that plays the logical-semantic role of a bare variable in a formal sentence. That is, the concept that intrinsically represents the class of all nuts must include a bare referential intention, which directs the mind toward some given entity (or class). Yet, the concept that presents the instantiation of FRUIT to the mind does not have a bare referential intention. From a metaphysical standpoint, one may wonder why the two types of concepts are not parallel.

An earlier section already explained some of the reasons that an instantiation-concept has to be a distinct type of concept from a class concept*.* Yet, from a more general philosophical perspective, it may help to point out that the central thesis of Bare Property Intentionality is that the mind intrinsically represents things through properties. Thus, the particular properties that the conceptual components of a PA-instance present to the mind, and not the class of objects (or individual object) that the state represents, determine the semantic content of that state. The features of PA-instances must function to allow that mental state to express, intrinsically, that a given property or other is instantiated by some class of entities. It follows, in turn, that the ability of a PA-instance to represent a class of objects (or individual object) intrinsically is secondary to and dependent upon the intrinsic ability of that state to present properties to the mind intrinsically.

Suppose that an instance of the belief that *all nuts are fruit* were composed of two (or more) metaphysically parallel concepts, each of which was composed of a bare referential intention and a mental presentation of a property. Then one would have to interpret that state to express the following proposition: *the class of all nuts is contained in the class of all fruit*. This would make mental presentations of objects, and not mental presentations of properties, the primary units of instances of propositional attitudes. In the Aristotelian interpretation of the logic of universal propositions, however, such a mental state represents the entire class of nuts but does not represent the entire class of pieces of fruit. This is because in order to represent the entire class of pieces of fruit, such a proposition would have to say something about the class membership of each individual piece of fruit.

Yet, such a proposition only “says something,” so to speak, about each individual nut. The Aristotelian interpretation of the semantics of universal propositions depends upon an assumption that *classes*, and not properties, are the units to which a sentence expressing such a proposition refers. Thus, if the components of instances of PAs are all class concepts, then each such concept must be interpreted to refer to a class of objects. The PA-instance would have to contain a concept of the class of all fruit in this case. However, if that were the case, the Aristotelian semantic interpretation would imply that the concept of the class of all fruit within this mental state does *not* refer to each individual piece of fruit. It is a basic assumption of Bare Property Intentionality, nonetheless, that an intrinsically representative class concept *does* refer to each individual member of a certain class. Thus, a proponent of Bare Property Intentionality cannot hold that an instance of a PA of the form *all Fs are Gs* is composed of two metaphysically parallel class concepts, which refer to the class of all Fs and the class of all Gs, respectively.

J. Conclusion

In summary, Bare Property Intentionality is a view of the metaphysical features of an instance of a propositional attitude of the form *all Fs are Gs*, such as an instance of the belief that *all nuts are fruit.* These features allow such a mental state to represent each individual nut intrinsically as instantiating the property FRUIT. According to Bare Property Intentionality, the ingredients in virtue of which such a mental state has this intrinsically semantic and representative content are:

* 1. An intrinsically semantic and representative class concept, which contains:

a. A bare referential intention (an intrinsically semantic and primitive intention of directing the mind toward some external entity or class thereof), and

1. A presentation to the mind of the primary property FRUIT
2. An intrinsically semantic basic instantiation-concept of the property FRUIT.

The fact that the concept of the class of all nuts contains a bare referential intention makes it the case that the mental state intrinsically represents some entity (or class) as opposed to representing no entity (or class) intrinsically. This concept also contains a presentation of the property NUT to the mind; and this possession allows the mental state to represent the class of all nuts intrinsically and exclusively, as opposed to any other class of objects. A basic instantiation-concept of FRUIT, as an additional ingredient of this state, gives the mind the ability to represent each individual nut as instantiating the property FRUIT. This instantiation-concept develops from innate logical, arithmetical, and metaphysical intuitions concerning the source of fruit-like perceptual experiences. For these reasons, the concept presents an abstract object to a person’s mind as being instantiated.

Bibliography

1) Block, Ned. “Advertisement for a Semantics of Psychology,” Midwest Studies in Philosophy X, Minneapolis: University of Minnesota Press, 1986, p. 615.

2) Block, Ned. “The Computer Model of the Mind,” in D. N. Osherson and E. E. Smith, (eds.), *Thinking: an invitation to cognitive science*, vol. 3 (1990). Cambridge, MA: MIT Press. Reprinted in *Philosophy and Cognitive Science*, Alvin I. Goldman (ed.), MIT Press, Cambridge, MA. 1993, p. 819.

3) BonJour, Laurence. *In Defense of Pure Reason****:*** *A Rationalist Account of A Priori**Justification*.Cambridge: Cambridge University Press, 1998.

4) BonJour, Laurence. “Is Thought a Symbolic Process?” Synthese 89: 1991, p. 331-352. Kluwer Academic Publishers, The Netherlands.

5) Churchland, Paul. “Eliminative Materialism and the Propositional Attitudes,” The Journal of Philosophy 78, 2 (February 1981): 67-90. Reprinted in *The Nature of Mind*, David M. Rosenthal (ed.), New York: Oxford University Press, 1991. p. 601.

6) Dretske, Fred*. Knowledge and the Flow of Information,* Cambridge, MA: MIT Press, 1981.

7) Fodor, Jerry. *Concepts: Where Cognitive Science Went Wrong*, New York, Oxford University Press, 1994.

8) Fodor, Jerry. *Psychosemantics,* Cambridge, MA: MIT Press, 1987.

9) Fodor, Jerry. “Semantics, Wisconsin Style,” in *Rerepresentation: Readings in the Philosophy of Mental Representation*. Stuart Silvers (ed.), Kluwer Academic Publishers, The Netherlands, 1989, p. 85.

10) Fodor, Jerry A. *The Elm and the Expert: Mentalese and its Semantics*, MIT Press, Cambridge, MA. 1994.

11) Fodor, Jerry A. “Fodor’s Guide to Mental Representation,” Mind Vol. 94, 6-100 (1985). Reprinted in *Philosophy and Cognitive Science*, Alvin I. Goldman (ed.), MIT Press, Cambridge, MA. 1993, p. 271.

12) Fodor, Jerry A. “Methodological Solipsism Considered as a Research Strategy in Cognitive Psychology,” The Behavioral and Brain Sciences III, 1 (March 1980), p. 63-72. Reprinted in *The Nature of Mind*, David M. Rosenthal (ed.) Oxford University Press, Cambridge, MA. 1993, p. 485.

13) Fodor, Jerry A. “Propositional Attitudes,” The Monist, LXI 4 (October 1978), reprinted in *The Nature of Mind*, David M. Rosenthal (ed.) Oxford University Press, Cambridge, MA. 1993, p. 325.

14) Fodor, Jerry. *Psychosemantics*, MIT Press, Cambridge, MA. 1987.

15) Frege, Gottlob. “On Sense and Reference,” P. Geach and M. Black, *Translations from the Philosophical Writings of Gottlob Frege*, second edition (Oxford: Blackwell, 1970); reprinted in Harnish, Robert M. (ed.), *Basic Topics in the Philosophy of Language* (Englewood Cliffs, NJ: Prentice Hall, 1994), p. 142-160.

16) Frege, Gottlob. “The Thought,” P. Strawson (ed.), *Philosophical Logic*, (Oxford: Oxford University Press, 1967), translated by A. M. and Marcelle Quinton; reprinted in Harnish, Robert M. (ed.), *Basic Topics in the Philosophy of Language* (Englewood Cliffs, NJ: Prentice Hall, 1994), p. 517-535.

17) Geach, P.T. “The Doctrine of Distribution," Mind 65 (1956) 67-74.

18) E. - H. Kluge. *The Metaphysics of Gottlob Frege*, Boston: Martinus Nijhoff Publishers, 1980.

19) Loar, Brian. “Can We Explain Intentionality?” In *Meaning in Mind: Fodor and His Critics*. Barry Lower and Georges Rey (eds.), Cambridge: Blackwell Press, 1991.

20) Lyons, William. *Approaches to Intentionality*. New York: Oxford University Press, 1995.

21) Putnam, Hillary. “The Meaning of ‘Meaning’,” in Gunderson (ed.), *Language, Mind, and Knowledge*, Minnesota Studies in the Philosophy of Science, 7. Minneapolis: University of Michigan Press, 1971.

22) Quine, W.V. “The Double Standard” in *Word and Object*, MIT Press: Cambridge, Massachusetts, 1960.

23) Russell, Bertrand. “On Denoting,” in Marsh R., (ed.), *B. Russell, Logic and Knowledge*, London: George Allen and Unwin, 1956; reprinted in Harnish, Robert M. (ed.), *Basic Topics in the Philosophy of Language* (Englewood Cliffs, NJ: Prentice Hall, 1994), p. 161-173.

24) Weller, Cass. “BonJour and Mentalese,” Synthese 113: Kluwer, The Netherlands, 1997, p. 251.

25) Wittgenstein, Ludwig. Tractatus Logico-Philosophicus, New York: Routledge, 1992.

**VITA**

Sudan Turner was born and raised in Pittsburgh, Pennsylvania. Currently, she lives in Redmond, Washington. At the University of Pittsburgh, she earned a Bachelor of Arts degree in Philosophy. At the University of Washington, she earned a Master of Arts in Philosophy in 1996 and a Doctor of Philosophy in Philosophy in 2004.

1. W.V. Quine asserts that Brentano is responsible for reviving the use of this term by Scholastics. (See “The Double Standard” in *Word and Object* (MIT Press: Cambridge, Massachusetts), 1960, p. 219.) [↑](#footnote-ref-1)
2. It is plausible to believe that many intentional states, such as the belief that *Oedipus was an ancient king*, and the belief that *5 X 5 = 25*, *fail* to represent physical objects in the external world. Thus, the feature of intentionality must connect mental states semantically both with things that have physical features and with things that do not (such as mythological objects,abstract objects, and other mental phenomena). [↑](#footnote-ref-2)
3. Many propositional attitude-instances concern things that do not exist in the *actual* world, but which do have physical features in some *possible* world. For example, consider a case in which a person named Susan assumes that someone has stolen a group of items that she has actually misplaced: Susan may believe that *the person who stole my PDA is immoral*. While this mental state represents a particular thief in some possible world, it fails to represent any actual individual person, since, in the actual world, no one has stolen Susan’s PDA. Nonetheless, one may espouse a theory according to which Susan’s concept of *the person who stole my PDA* does successfully represent Susan’s *PDA*. One who endorses this viewwould consider Susan’s mental state to represent an existent object in the actual world. Yet, some intentional states definitely fail to represent *any* existent object at all. A person who does not own any cosmetics, for instance, may mistakenly believe that she does (suppose that she mistakenly checked off the term ‘cosmetics’ on her shopping list); and its absence may cause her to believe that *the person who stole my cosmetics package is immoral*. In such a circumstance, the relevant occurrence of this PA has a semantic content, but it has no actual referent or truth-value. (This issue is distinct from the issue of the semantic features of intentional states concerning mythological entities. Such beliefs *do* have referents; and therefore, one may evaluate them as being true or false. An instance of a belief that *all Greek goddesses are kind*, for example, successfully refers to a class of mythological entities as the relevant myth defines them. Thus, one may evaluate this belief as true or false in accordance with the features ascribed to Greek goddesses by the relevant socially accepted narrative, such as *The Illiad*. Moreover, one may consider this class of mythological objects to be made up of possible but non-actual entities.) [↑](#footnote-ref-3)
4. A propositional attitude-*instance* is a particular, concrete mental state that takes place during a discrete temporal interval. A person’s occurrent belief (at time t) that *all bachelors are sports fans*, for example, is an instance of a general PA-type. This PA-type can have many different instances. This is because there may be many different cases in which some given person is in a mental state with this content. In this discussion, all claims about the *constituents* of a PA refer to the concrete, particular ingredients of a PA-instance.

   While all instances of a given PA have *semantically* identical ingredients, moreover, different instances of the same PA-type may have different referents and truth-values in the actual world. For example, while each occurrence of the belief that that *all American school children are on summer vacation* has the same meaning and logical implications, an instance of this belief that occurs in July could be true, while any instance of this PA that occurred in December would have to be false. [↑](#footnote-ref-4)
5. A later chapter will define an intrinsically semantic state as one that has its meaning and its logical implications essentially, while defining an intrinsically *representative* state as an intrinsically semantic state that essentially has some truth-value or other and some referent (or class thereof). That is, what an intrinsically representative mental state represents may vary; and the truth-value of such a state may vary, as well. Yet, an intrinsically representative mental state necessarily represents *some* (possible or actual) entity, or set thereof; and such a state necessarily has *some* truth-value. While the meaning and logical implications of an intrinsically semantic or an intrinsically representative state remain the same through time, the actual object (or class of objects) to which an intrinsically representative state refers may change. This is because the properties that an object (or class thereof) has may change. For example, the set of all objects that belong to the class of all children in July of 1992 has a different membership than the set that belongs to this class in July of 2002. Therefore, an instance of the belief that *all children are well nourished* that occurs in July of 1992 refers to a different class of objects than the class to which an instance of this belief that occurs in July of 2002 refers.

   As a later chapter will argue, some mental states may be intrinsically semantic *without* being intrinsically representative. Thus, the claim that “intrinsically semantic” states have reference essentially does not mean that for every intrinsically semantic state, there is a particular object (or class of objects) to which it necessarily refers. Rather, it means that it is an essential feature of every intrinsically semantic and intrinsically representative instance of a PA that it refers to *some* possible or actual object (or class thereof). [↑](#footnote-ref-5)
6. One may wonder whether it makes sense to hold that intrinsically semantic entities “have” meanings, as opposed to *being* meanings themselves. However, the view put forth in this essay is that that intrinsically semantic entities have meanings as essential components, but are not identical to meanings. A propositional attitude-instance *cannot* be identical to the specific meaning that it has. This is because as a later chapter will explain, a PA-instance also contains semantically generic and non-semantic components. One reason to say that an occurrent concept that is an ingredient of a PA-instance is not identical to a meaning is that the meaning of such a concept is universal. For any given meaning may be had by more than one occurrent, concrete concept. [↑](#footnote-ref-6)
7. Not *all* intrinsic semantics views of mental representation endorse the metaphysical position that concrete, mind-internal concepts are even *a* source, let alone the only source, of semantic content for mental states. As a later chapter will explain, Frege held that mind-external, abstract objects termed ‘senses’, and not any mind-internal phenomena, were the only intrinsically semantic entities. Moreover, while the Fregean view is that senses are *identical* to meanings, others may hold that intrinsically semantic entities *have* meanings essentially, but are not identical to meanings. [↑](#footnote-ref-7)
8. The extrinsic semantics view of intentionality must also allow a person to have mental states that represent other phenomena inside her own mind. Therefore, one version of the extrinsic semantics perspective, i.e., the view according to which mental symbols have semantic content in virtue of contingent relations with *mind-external* objects, must acknowledge an exception. The caveat is that a person’s beliefs about any of her other mental states, events, or processes have semantic content in virtue of their contingent relations to these state-external, but mind-*internal* phenomena. (For the philosophical purposes of this essay, one person’s occurrent belief that concerns any existent mental phenomena of another person represents mind-*external* phenomena. This is because the PA-instance represents phenomena external to the mind in which *that* PA-instance occurs.) [↑](#footnote-ref-8)
9. In this introductory chapter, I make several general philosophical points about intentionality itself. I also make several general philosophical points about the basic metaphysical assumptions and implications of different *perspectives* on intentionality. As mentioned, in the final chapter, I will argue for a specific metaphysical version of the intrinsic semantics perspective. These ultimate arguments only *necessarily* apply to the specific types of intrinsically semantic concepts, which, according to my ultimate view, are constituents of PA-instances that concern *universal* propositions. The beliefs that *all cherries are red* and that *all fruit are spheres* are examples of this type of PA, while the beliefs that *some mothers are kind* and *that Joan is a kind mother* are not. The restricted application of my ultimate conclusion does not, however, limit the general points that I make in *this* chapter concerning the metaphysical assumptions and implications of various versions of the intrinsic and extrinsic semantics perspectives on intentionality. For this reason, it is acceptable that in this chapter, I use beliefs involving singular and existential propositions to exemplify general, theoretical points. [↑](#footnote-ref-9)
10. One can use both of the terms ‘cognitive content’ and ‘meaning’ to express connotation (iii) and connotation (iv). Yet, one can consider a proposition to be an abstract entity, while a mental state must be concrete. A proposition must, therefore, exist outside of any mind or particular occurrence of a mental state. The terms ‘cognitive content’ and ‘meaning’ also can refer to what is going on *inside the mind* of a person while he or she is in a given mental state. Hence, there is a *prima facie* philosophical distinction between the proposition and entailment relations that the content of a mental state involves, on the one hand, and the cognitive content or meaning of that state, on the other hand. Thus, one must distinguish connotation (i) of the term ‘semantics’ from connotation (iv). [↑](#footnote-ref-10)
11. As mentioned earlier, one can also interpret the notion of the “meaning” of a mental state in terms of the state’s being directed toward abstract but non-mental entities, such as universals or propositions. [↑](#footnote-ref-11)
12. While a proposition is an abstract, objective, and mind-external entity, an idea is a concrete, potentially subjective, and mind-internal entity. A common sense view of the mind would probably ignore the distinction between the semantic content of a state and its semantic constituents. Therefore, it would consider this content to involve “ideas.” Yet, metaphysical views of intentionality tend to explain semantic constituents in terms of concrete internal phenomena, while considering semantic *contents* to involve abstractions. As the final chapter will explain, Bare Property Intentionality says that the semantic *constituents* of mental states are internal concrete phenomena, which derive their semantic *content* (or meanings) from mind-external abstractions, such as universals. [↑](#footnote-ref-12)
13. An occurrent desire that *[x] be the case* is also a kind of propositional attitude-instance, since it has a semantic content. The desire that *no metropolitan area receive rain this evening* has a semantic content, i.e., the proposition that no metropolitan area receives rain on the evening of the day in which the subject has this occurrent belief. This proposition has certain logical implications, truth-conditions, and so on, which are metaphysically independent of the emotion that the thinker attaches to the contemplation of the (possible) state of affairs that the proposition concerns. While the essential emotional component of a desire is not relevant to its semantic content, the metaphysical assumptions of this essay do not imply that a desire intrinsically represents the world in a manner that one may evaluate as true or false. Therefore, with the exception of a discussion of the causal role of desires in logical thought processes that motivate actions, this essay will address the metaphysics of the semantic features of beliefs, and not that of desires, as the paradigmatic propositional attitudes. [↑](#footnote-ref-13)
14. One may wonder why it makes sense to believe that for a mental state to *be* intrinsically semantic it must have a content that is *also* intrinsically semantic. One may ask: why it is not enough for the state to have its semantic content intrinsically? This appears to be a quasi-linguistic question. Since the content of a propositional attitude-instance *is* its semantic content, the content of an intrinsically semantic PA-instance will be intrinsically semantic. An intrinsically semantic mental state *is* a state that has its semantic content intrinsically, which, in turn, just *is* a mental state with (at least some) constituents or ingredients that are intrinsically semantic. [↑](#footnote-ref-14)
15. It is important to note that this claim refers to concrete concepts as they occur in actual mental states, not to any type of abstract concept-*type*. This essay does not address the metaphysical nature of concept-types or their relation to the intrinsically semantic content of propositional attitudes. The view of intrinsically semantic concrete concepts that I will later put forth does imply that concept-types exist, since more than one occurrent concept can have the same intrinsic meaning. [↑](#footnote-ref-15)
16. Some metaphysical views, for example, endorse the existence of tropes, which are abstract but *particular* properties that attach uniquely to particular objects (at least in the actual world). These abstract entities are thus not universals, given that they may not be instantiated in more than one thing. [↑](#footnote-ref-16)
17. Note that more than one universal (or non-repeatable but abstract entity) may also instantiate one and the same higher-order universal. Thus, concrete entities are not the only types of entities that instantiate universals. [↑](#footnote-ref-17)
18. The symbolic position on mental representation belongs under the general category of empirical, cognitive- science orientations toward explaining the mind. From this perspective, thoughts can be modeled, analyzed, and explained as being analogous to computer states and processes. [↑](#footnote-ref-18)
19. This is Fodor’s Computational Theory of the Mind, which I will explain later. [↑](#footnote-ref-19)
20. One may think that it is incoherent to consider the Fregean view that mind-independent, abstract entities are the source of the semantic content of mental states to be a version of *internalism*. For if such entities exist outside the mind and provide mental states with their semantic features, then it would seem that Frege’s view would have to be externalist. Yet, there are reasons to consider some versions of it to be internalist. According to Frege, the relation whereby a mental state gains contact with such an abstract entity necessarily involves mind-dependent (and partially subjective) “representations.” (“The Thought,” P. Strawson (ed.), *Philosophical Logic*, (Oxford: Oxford University Press, 1967), translated by A. M. and Marcelle Quinton; reprinted in Harnish, Robert M. (ed.), *Basic Topics in the Philosophy of Language* (Englewood Cliffs, NJ: Prentice Hall, 1994), p. 530-531.) These internal representations provide a mental state with some subjective semantic content, in contrast to the objective content of the abstractions that they serve to grasp. The *relation* whereby such a representation grasps an abstract entity is a semantic *feature* of a mental state. Nonetheless, the metaphysical assumptions of the intrinsic semantics view require that any intrinsically semantic entity be semantically objective. Thus, the Fregean view of subjective mental representations *can* hold that these phenomena are an *internal* source of *semantic* content. Yet, it cannot hold that they are an internal source of *intrinsically* semantic content.

    Another reason to consider such a view to be internalist is that the intuitions behind externalism imply that a mental state has semantic content in virtue of some kind of *causal* relation to mind-external and *concrete* entities. However, an abstract entity cannot have a causal relationship with anything. Thus, it would appear that the Fregean view of abstract entities as the only source of intrinsically semantic content would have to be internalist, by default. [↑](#footnote-ref-20)
21. Hillary Putnam, “The Meaning of ‘Meaning’,” in Gunderson (ed.), *Language, Mind, and Knowledge*, Minnesota Studies in the Philosophy of Science, 7. (Minneapolis: University of Michigan Press, 1971.) While this example has many implications concerning the construal of semantic content as computational, as internal, and as being subsumed under psychological generalizations that predict behavior, etc., here, we are only concerned with the way that Twin Earth cases illustrate the externalist notion of content. [↑](#footnote-ref-21)
22. Some have thought that it was either intuitive or philosophically desirable to hold that both Eric and TwinEric’s thoughts represent the disjunctive class of H2O or XYZ. [↑](#footnote-ref-22)
23. One can only guarantee this identity of internal cognitive value before the discovery of the chemical constitution of the substance referred to with ‘water’ (in English). Given the difference between the chemical constitution of water on Earth and Twin Earth, the relevant knowledge acquired on Earth at that time will be that this substance is made of H2O, while Twin Earth residents will learn that this substance is made of XYZ. Under the reasonable assumption that Eric and TwinEric will learn these respective facts, the internal cognitive value of at least some of their beliefs concerning this substance will explicitly represent it *as having* certain molecular features. Since the substance is H2O on Earth and XYZ on Twin Earth, at some point Eric will explicitly think about H2O, while TwinEric explicitly thinks about XYZ. Hence, the validity of the thought-experiment depends upon restricting the comparison between Earth and Twin Earth mental states to the time before the discovery of the molecular features of the substance that English-speaking people denote with the word ‘water’. [↑](#footnote-ref-23)
24. As an earlier section explained, a semantic feature is the property that an entity, e.g., a mental state or a sentence, has in virtue of which it has semantic content. A semantic content, in contrast, is some specific meaning, extension, or set of truth-conditions belonging to the belief, sentence, etc. I described externalism as the view that a mental state’s relations to things outside the mind at least partially determine its semantic features and content. Since no perceptual experience actually is a *semantic* constituent of a mental state, therefore, it may seem strange to suggest that such experiences determine the semantic content of a state (and thus constitute its feature of having such content). Yet, a perceptual experience is a relation to an entity outside the mind, and such a relation forms a link in a definite and unique causal chain between, for instance, the water-meaning concept instantiated in Eric’s occurrent belief that *water is wet* and some actual aggregation of H2O molecules. Under a more sophisticated set of philosophical specifications, the externalist considers this kind of external relation to determine the fact that Eric’s occurrent water-meaning belief is about H2 O (as opposed to XYZ or NaCl, for instance). [↑](#footnote-ref-24)
25. This assumption is required at least in cases concerning perceptual concepts, which the concept of water should be in this thought-experiment. [↑](#footnote-ref-25)
26. This claim also depends upon the above-mentioned assumption of empiricism on concept acquisition. [↑](#footnote-ref-26)
27. A person may have learned such a fact from a high-school textbook; and in the process of acquiring the information, she may have an indirect, or once-removed perceptual experience of the war ⎯ from a textbook photograph of a musket in use during one of its battles, for example. Nonetheless, she acquires the information itself from the writers of the relevant text, who in turn, acquired the information from secondary research into the reports of other researchers, and so on. The causal connection that makes her belief *about* the Revolutionary War, according to the externalist view, is some non-perceptual but historical connection between the student’s belief (specifically, the concept of the Revolutionary War contained in the belief) and the mind-external event that constituted the war itself. [↑](#footnote-ref-27)
28. Moreover, it would seem that if PA-instances have such intrinsically semantic entities as their ingredients, then the networks of causal and inferential relations that PAs have with other PAs will be determined *by* the semantic contents of these constituents, and not vice versa. [↑](#footnote-ref-28)
29. I will explain a version of this specific view in a later chapter. The fact that cherries are the only things that normally cause an instance of the mental symbol for cherries, for instance, would provide such a one-to-one relationship. [↑](#footnote-ref-29)
30. Most actual concepts are not simple but complex; however, these concepts should be analytically reducible to concatenations of primitive concepts. [↑](#footnote-ref-30)
31. If some of these concepts turn out to be complex, i.e., not primitive, then the internalist atomist would say that the set of all the primitive concepts making up each of these three concepts is what determines the semantic content of the belief. [↑](#footnote-ref-31)
32. This is the Causal Theory of Content, which Fodor, Kripke, and others have proposed in various versions. Causal theories say that a person’s mental symbol represents a thing when his (or her) interaction with that thing normally and appropriately causes him (or her) to have a mental state involving that symbol. [↑](#footnote-ref-32)
33. In the next two chapters, I will base my evaluations and criticisms of two versions of the extrinsic semantics perspective upon their ability to meet these explanatory criteria. The ability of a symbolic theory of the metaphysics of propositional attitudes to meet these criteria is independent of whether a given belief-instance has the form *all Fs are Gs, a is G,* *the F is G*, *or some Fs are Gs*. Therefore, despite the fact that my ultimate metaphysical theory only concerns PA-instances of the form *all Fs are Gs,* such as an occurrent belief that *all fruits are spheres,* it is acceptable that the examples of PA-instances that the second and third chapters use do not have this form. [↑](#footnote-ref-33)
34. Fodor suggests that the predictive ability of IH is its most obvious justification, and that any empirical psychological science will have true generalizations about behavior that involve intentional explanations. (See “Fodor’s Guide,” p. 279, and *The Elm and the Expert*,p. 3-4.) Churchland also discusses this basic criterion in “Eliminative Materialism and the Propositional Attitudes,” *The Journal of Philosophy* 78, 2 (February 1981): 67-90. Reprinted in *The Nature of Mind,* David M. Rosenthal (ed.), (New York: Oxford University Press, 1991), p. 601. [↑](#footnote-ref-34)
35. BonJour in discussion. [↑](#footnote-ref-35)
36. This theory may seem to be so deeply rooted in common sense as to be indubitable. Others, however, have asserted that the very idea that we possess such discrete, individual states as beliefs and desires is a merely a highly questionable although practically useful piece of “folk psychology.” For it is hard to reconcile the claim that definable semantic entities such as propositional attitudes exist in the mind with the view that the mind is the brain. (See William Lyons, *Can We Explain Intentionality*? (1991), and Paul M. Churchland, “Eliminative Materialism and the Propositional Attitudes” (1981).) [↑](#footnote-ref-36)
37. Gottlob Frege in “On Sense and Denotation” first established this criterion. (Reprinted in *Basic Topics in the Philosophy of Language*, Robert M. Harnish (ed.), (Englewood Cliffs, NJ: Prentice-Hall, 1994), p. 148.) [↑](#footnote-ref-37)
38. BonJour, Laurence. “Is Thought a Symbolic Process?” Synthese 89: 1991, Kluwer Academic Publishers, The Netherlands, p. 336. [↑](#footnote-ref-38)
39. “Methodological Solipsism as a Research Strategy,” reprinted in *The Nature of Mind, David* M. Rosenthal (ed.) Oxford University Press, and Cambridge, MA. 1993, p. 486, and “Fodor’s Guide to Mental Representation,” in *Philosophy and Cognitive Science,* Alvin I. Goldman (ed.), MIT Press, Cambridge, MA, 1993, p. 288-289. [↑](#footnote-ref-39)
40. “Propositional Attitudes,” in The Monist, LXI 4 (October 1978), reprinted in *The Nature of Mind*, David M. Rosenthal (ed.) Oxford University Press, Cambridge, MA, 1993, p. 328-329. (Fodor describes logical form as being required to explain what I have termed the practical syllogism assumption.) [↑](#footnote-ref-40)
41. The notion of syntax is related intricately to the notion of a language: a system of symbols interpreted to represent objects, their properties, and connections between objects and properties. The syntax of a language involves a set of symbols broken into kinds, and a sub-system of rules: imperatives for and prohibitions against combining certain kinds of symbols with others to make sentences. For example, the content of the belief that *Bus 11 goes downtown* is, according to the symbolic view, a sentence-like entity constructed from different kinds of symbols: those standing for objects, like buses and the area termed ‘Downtown’, etc., and those standing for properties or relations, such as the action of *going to x*. The idea that PAs have syntactic properties thus means that they are constituted by strings of symbols that belong to different kinds, which kinds, in turn, are organized and combined in accordance with rules. One syntactical rule may be that to construct a PA, an object-symbol must be combined with a property symbol. [↑](#footnote-ref-41)
42. This is the formal structure of the syntactical relations between a given PA, the states that it causes, and the states that cause it. [↑](#footnote-ref-42)
43. As explained, in virtue of being arranged according to syntactical rules, the structure of the complex symbols that constitute PA-instances mimics the logical structure of propositions. This structure, moreover, determines which syntactical rules will govern the *causal* relations of the mental symbols. These relationships thus tend to be truth preserving, because sets of syntactico-causally related intentional states mimic the form of inferentially related propositions. For example, consider a set of PA-instances such as: 1) a belief that *John is a bachelor*, 2) a belief that *a bachelor is an unmarried adult male*, and 3) a belief that *John is unmarried*. The constituents of these tokens have a deductive logical relationship that is identical to the logical relationship between members of a set of propositions {P, Q, R}, where P is the proposition associated with the sentence which expresses the content of (1), Q is the proposition associated with (2), etc. Thus, if a subject’s having beliefs (1) and (2) cause her to have (3), the pattern of causal relationships between the constituents of PAs 1, 2, and 3 is formally analogous with the inferential relations between members of the set {P, Q, R}. (For various stages of this (ultimately rejected) view, see “Fodor’s Guide”, p. 273, Psychosemantics, (MIT Press, Cambridge, MA. 1987), p. 78-79; and The Elm and the Expert, p. 8-9.) [↑](#footnote-ref-43)
44. One may think that this belief also represents the class of all male objects, and thus that it does not *exclusively* represent the class of bachelors. Yet, the truth-conditions of the belief are that each object that belongs to the class of bachelors also belongs to the class of males. Thus, this belief does not represent the entire class of males, because it does not represent each individual male. Rather, the only entire class that the belief represents is the class of all bachelors. Moreover, as Aristotle’s class-membership theory of logic shows, the predication of the belief, i.e., the assertion of class membership that it makes, is applied to each and only member(s) of the class of bachelors. For it only asserts that each member of the class of bachelors belongs to the class of males; and it says nothing exclusive about each member of the class of males. The only assertion or predication that the belief that makes about males is that *some* males belong to the class of bachelors. [↑](#footnote-ref-44)
45. “Is Thought a Symbolic Process?” Synthese 89: Kluwer: The Netherlands, 1991, p. 334. The implications of BonJour’s argument have profound importance for the general symbolic view of the ontology of mental representation, and specifically for the *internalist* symbolic perspective. If, as his objection entails, there is no coherent sense in which one can have an internally based understanding of a symbolically constituted thought, and if, under some acceptable definition of ‘understanding’, a mental state’s having semantic content requires that its subject be able, in principle, to understand that content from the inside, then the intentional features of PA-instances cannot be constituted solely by symbols. [↑](#footnote-ref-45)
46. “Complex semantic constituent” means the entire set of individual constituent symbols which, when combined in a structured form, make up one complex constituent. In the *verbal* language sentence ‘Bus 11 goes downtown’, for instance, the individual constituents ‘Bus 11’, ‘goes’, etc., combine to make one complex sentence. Presumably, the semantic constituency of a symbolically constituted mental state is analogous to the constituency of such a sentence in the relevant way. [↑](#footnote-ref-46)
47. This analogy is owed to BonJour, in discussion. [↑](#footnote-ref-47)
48. *Ibid*., p. 336. [↑](#footnote-ref-48)
49. While the exact sentences ((2) through (6)) don’t appear in his text, BonJour has agreed (in conversation) that this is a faithful paraphrase of his reasoning. [↑](#footnote-ref-49)
50. Note that this inference is cogent only on the assumption that there is a conceptual distinction between the (mental analogue of) *inscribing* a symbol with semantic content and *understanding* that content. I will discuss this assumption shortly. [↑](#footnote-ref-50)
51. *Ibid*., p. 336. [↑](#footnote-ref-51)
52. Cass Weller, in discussion. [↑](#footnote-ref-52)
53. Weller, “BonJour and Mentalese,” Synthese 113: Kluwer, The Netherlands, 1997, p. 261-262. [↑](#footnote-ref-53)
54. BonJour, in discussion. [↑](#footnote-ref-54)
55. Some arguments may be formally circular without committing the *fallacy* of begging the question. That is, there are some cases in which it may be legitimate for one of the premises of an argument to include the truth of the conclusion, i.e., in situations where the relevant assumption is a conceptual or logical truth. [↑](#footnote-ref-55)
56. *Ibid*., p. 252 (quote of BonJour, p. 332). [↑](#footnote-ref-56)
57. *Ibid*., p. 253. [↑](#footnote-ref-57)
58. *Ibid*., p. 253. [↑](#footnote-ref-58)
59. BonJour (in discussion) notes the difference in the justificatory demands of these assumptions. [↑](#footnote-ref-59)
60. *Ibid*., p. 336. [↑](#footnote-ref-60)
61. *Ibid*., p. 334. [↑](#footnote-ref-61)
62. *In Defense of Pure Reason: A Rationalist Account of A Priori Justification*, Cambridge: Cambridge University Press, 1998. [↑](#footnote-ref-62)
63. BonJour, in discussion. [↑](#footnote-ref-63)
64. “On a confusion about a function of consciousness,” Behavioral and Brain Sciences 18 (2): 227-287.

    (One may object that there are entities, such as linguistic tokens, and maybe individual concept- types themselves, which have representational or semantic content without having any consciousness. Yet, ‘access-consciousness’ here refers to a property of a state of a mental system itself, and not to individual semantic constituents of such states, or of their verbal or written expressions.) [↑](#footnote-ref-64)
65. *In Defense of Pure Reason*, p. 163. [↑](#footnote-ref-65)
66. *Ibid.,* p. 162. [↑](#footnote-ref-66)
67. “Is Thought a Symbolic Process?” p. 334. [↑](#footnote-ref-67)
68. *In Defense of Pure Reason*, p. 163. [↑](#footnote-ref-68)
69. See, for example, Gottlob Frege, “The Thought: a Logical Enquiry,” P. Strawson (ed.), *Philosophical Logic*, (Oxford: Oxford University Press, 1967), translated by A. M. and Marcelle Quinton; reprinted in Harnish, Robert M. (ed.), *Basic Topics in the Philosophy of Language* (Englewood Cliffs, NJ: Prentice Hall, 1994), p. 531. [↑](#footnote-ref-69)
70. Some would say that in addition to entertaining a semantic content, i.e., mentally tokening a set of symbols (or entertaining a set of abstractions), one must also engage in an action of affirming or asserting this content in order to qualify as having a belief. Yet, this issue is immaterial to the present concern. [↑](#footnote-ref-70)
71. Shortly, I will discuss Functional Role Semantics in the context of the Computational Theory of the Mind and its ability to explain mental truth-preservation. Then, I will explain how FRS is the only alternative for CTM to explain the semantic content of symbolically constituted mental states. [↑](#footnote-ref-71)
72. Obviously, the conceptual position on intentionality avoids the regress by holding that the constituents of representative states are intrinsically semantic. Hence, this view can explain the truth-preserving causal powers of mental states through a person’s internal understanding of their concepts — without encountering a regress of higher-order conceptual states. Therefore, the folk wisdom that mental states such as beliefs and desires are about things in virtue of involving concepts *can* meet the first explanatory condition that successful theories of intentionality must satisfy. [↑](#footnote-ref-72)
73. “Methodological Solipsism as a Research Strategy,” p. 486, and “Fodor’s Guide,” p. 288-289. CTM can be described in terms of six tenets, the first five of which actually constitute his Representational Theory of the Mind (RTM), and which when combined with the sixth tenet, become CTM. I refer to the entire set as CTM, since this discussion concerns Fodor’s computational view of the truth-preserving character of series of PAs. The first three tenets are, for all intents and purposes, consistent with what I refer to as the common sense view of PA ontology: (i) PA ascriptions should be construed *realistically* (because this is the best way to explain behavior); (ii) PAs are relational, e.g., Sharon’s believing that *Bus 11 goes downtown* is a relationship between her belief-act and *whatever* entity is properly expressed by the linguistic object ‘Bus 11 goes downtown’; and (iii) PAs are relations to mental representations. (The schema of these tenets is owed to Charles Marks, in discussion.) The current discussion begins with the fourth tenet, according to which the content of PAs is linguistic and symbolic. [↑](#footnote-ref-73)
74. It is questionable whether a purely syntactical, and thus ultimately, mechanical, account of truth-preservation among mental states can be considered a logical account. This is because it is reasonable to interpret the adjective ‘logical’, as applied to a *mental* process, to refer to a person’s being aware of the entailment relationship between a set of propositions. Yet, if a purely syntactical process of transforming a system from one state to another does not necessarily involve internal understanding of any propositional entailment relations, then the process is not a logical process. Rather, it is a rule-driven causal process. Of course, the question of how one can consider these causal processes to be logical is the topic of this section of the chapter. [↑](#footnote-ref-74)
75. “Methodological Solipsism,” p.486, and “Fodor’s Guide,” p. 288-289. (The Practical Syllogism assumption concerns the truth-preserving character of PA-instances in a causal sequence, and although the functionalism of CTM is not yet relevant to this discussion, we must note that its fifth tenet attempts to account for these causal properties of PA-instances by defining PAs as functional states. ( “Fodor’s Guide,” p. 279).) [↑](#footnote-ref-75)
76. “Propositional Attitudes,” in The Monist, LXI 4 (October 1978), reprinted in The Nature of Mind, p. 328-329. [↑](#footnote-ref-76)
77. As opposed to being made up of these elements, *atomic* mental symbols are identical to them. [↑](#footnote-ref-77)
78. “The Elm and the Expert,” p. 8-9. [↑](#footnote-ref-78)
79. ## “Fodor’s Guide”, p. 281, and p. 288; and Psychosemantics, MIT Press, Cambridge, MA. 1987, p. 78-79.

    [↑](#footnote-ref-79)
80. Later, this essay will discuss the kinds of systemic mechanisms through which *formal* rules for categorization implement causal rules for state transformation. [↑](#footnote-ref-80)
81. “Methodological Solipsism” p. 486, and “Fodor’s Guide,” p. 289. [↑](#footnote-ref-81)
82. For crude illustrative purposes, let us imagine some quite simple examples of the syntactical categories and rules governing the symbols of M; and let us model these categories and rules using standard quantificational notation. The first is a logical set, that is, a set of categories for symbols representing syntactical relationships between symbol strings. The abstract structure of these relationships is formally analogous to truth-conditional relationships between propositions. Symbols of this category, therefore, play roles analogous to the formal language quantificational symbols for conjunction, negation, inclusive disjunction, and so on.

    Next is a metaphysical set, that is, a set of categories for symbols whose syntactical roles are formally analogous to those of formal language symbols for properties and individuals. In first-order quantificational notation, a property or relation symbol, e.g., ‘F’ for ‘is French’, is only properly attached to a symbol for an individual, e.g., ‘a’ for ‘Anne;’ and ‘Bxy’ (for ‘x boards y’) is properly attached to ‘a’ and ‘b’ (for ‘Bus 11’). Similarly, in the hypothetical mental language *M* there is a category of symbols (like property symbols) that can only properly attach to symbols belonging to another category (like object symbols). Moreover, such an attachment result in a complete symbol string that plays a syntactical role analogous to the truth-conditional role of a proposition, e.g., ‘Fa’ plays a role like the proposition *Anne is French*, and ‘Gab’ plays a role like the proposition *Anne boards Bus 11*.

    Finally, one can imagine a quantification set, that is, a set of quantification categories for symbols whose syntactical roles are formally analogous to those of formal language symbols for the numbers of individuals that fall under a certain description. As ‘(x)’ represents the idea that all objects fall under a certain description, e.g., ‘(x)(Fx)’ expresses that all individuals are French in first-order quantificational notation, in the mental language *M* there are symbols that play this kind of syntactical role. [↑](#footnote-ref-82)
83. Of course, the actual symbols that constitute a mental language of thought are not necessarily related, in a physical or structural way, to the symbols of any natural written or spoken language. For the present philosophical purposes, however, the assumption that these symbols constitute intentional states must explain why such states cause each other in a truth-preserving manner. This means that the syntax of this mental language must categorize its symbols into discrete categories. These categories play certain unique logical roles and are uniquely associated with certain metaphysical categories. The reason for this necessity is that symbolically constituted PA-instances could only cause each other in a truth-preserving manner because of this kind of syntactical organization. [↑](#footnote-ref-83)
84. It may seem to be implausible that the syntactic system that (presumably) governs mental symbols is structured like the syntax of first-order logic. However, the deductive principles governing propositions of the logical form *P or Q* are tautological and thus universal. Therefore, it is reasonable to think that these principles, including the principle expressed in rule *M1*, can be codified in the causal rules of any syntactical system that is responsible for truth-preservation among its formulas. [↑](#footnote-ref-84)
85. Inputs to an organic system must be of a perceptual nature, even if these inputs are ultimately processed into symbolic or cognitive content. Outputs to an organic system are of a behavioral nature, even if the behavior is an utterance or an inscription of a symbol-token for the purposes of communicating some cognitive content. Obviously, inputs to a purely mechanical or electrical system involve either purely macro-level physical forces (and their effects upon states of the system), or micro-level mechanical (or electrical) impulses. [↑](#footnote-ref-85)
86. State S2, in turn, is also defined in terms of conditional relationships between inputs, outputs, and state-state transformations. [↑](#footnote-ref-86)
87. This principle must be considered a necessary or metaphysical truth, because it could only be an analytic or conceptual truth if it concerned some relationship between the concepts of red and green. Yet, according to the theoretical framework of this essay, concepts are intrinsically semantic entities, while the purpose of this discussion is to evaluate the view that the semantic constituents of thoughts are symbolic. In that case, there should probably not be such a thing as a “conceptual” truth. [↑](#footnote-ref-87)
88. Ned Block, “The Computer Model of the Mind,” in D. N. Osherson and E. E. Smith, eds., Thinking: An invitation to cognitive science, vol. 3 (1990), Cambridge, MA: MIT Press, reprinted in *Readings in Philosophy*, p. 819, 830. (Block notes that this vision was first expressed by Fodor, 1968.) [↑](#footnote-ref-88)
89. reference? [↑](#footnote-ref-89)
90. A machine that didn’t give change would stay in S2 (having received an extra nickel), and “wait” for the next nickel or dime input. [↑](#footnote-ref-90)
91. Jeri Hanly and Elliot Koffman, (eds.), *Program Design in* ***C***, (Addison-Wesley Longman, Inc.: Reading, MA, 1996), p. 7. [↑](#footnote-ref-91)
92. Ned Block describes it (more realistically) it in terms of the *degree* (as opposed to the presence) of macro-level electrical charge, e.g., in terms of a 7-volt and a 4-volt charge. [↑](#footnote-ref-92)
93. The binary numerical system is a symbolic system, i.e., a system for representing numerical quantities. All numbers can be represented in this system through some combination of just these two symbols. In contrast, the Arabic numeral system requires ten symbols: i.e., ‘1’ through ‘9’, plus ‘0’. [↑](#footnote-ref-93)
94. “The Computer Model of the Mind,” p. 821. [↑](#footnote-ref-94)
95. *Ibid*., p. 828. [↑](#footnote-ref-95)
96. *Ibid*., p. 830 (Block notes that he was first exposed to this idea in Fodor (1975)). [↑](#footnote-ref-96)
97. On the surface, Functional Role Semantics appears to be quite similar to the philosophical position termed Conceptual Role Semantics (CRS). Yet, it is important to recognize the distinction between them. The purpose of CRS is to reductively define the semantic content of a concept (and thus the semantic content of a mental state) in terms of the role that such a concept plays in a network of deductive and inductive inferences. (CRS can apply to a view according to which the semantic constituents are non-symbolic, as well.) [↑](#footnote-ref-97)
98. “Fodor’s Guide,” pp. 281-282. (The aspect of the view that accounts for meaning in terms of the relations between mental states is all that concerns our discussion. Nonetheless, there are “two-factor” versions of this view, in which the functional definition of a PA also involves its causal relations with entities in the external world.) [↑](#footnote-ref-98)
99. On the other hand, CRS would identify their semantic contents with the *inferential* roles of these propositions, and not with the individual propositions themselves. [↑](#footnote-ref-99)
100. As mentioned above, CRS would say that this content is only reductively analyzable in terms of its deductive and inductive *inferential* relationships with the other relevant conceptual states, that is, PAs. A problem with CRS on this point concerns how to individuate conceptual roles such that they map one-to-one with our pre-theoretic intuitions concerning the identity of concepts. If we assume that the concept of a bird must have a distinct semantic content, and thus a distinct conceptual role, from the concept of a fish, then how do we make sure such their conceptual roles are individuated in a manner consistent with this intuition? For as Block tells us, we can’t appeal to any analytic notion of the definition of these concepts. This is because the distinction between analytic and synthetic is one of the grounds upon which meaning holism is adopted. More importantly, how do we establish that different people (who are willing to make radically different sets of inferences from the concept of a bird) both have the same concept? See “Advertisement,” p. 628-629. [↑](#footnote-ref-100)
101. “Fodor’s Guide,” p. 291 [↑](#footnote-ref-101)
102. This empirical point does not challenge the background Practical Syllogism assumption of the Intentional Hypothesis. This is because this assumption does not claim that most causal relationships between propositional attitudes are truth preserving. Rather, the Practical Syllogism assumption is that given observations of a person’s behavior and environment, we can coherently explain and reliably predict further behavior by assuming that some of the causal relationships between her beliefs and desires are truth-preserving. [↑](#footnote-ref-102)
103. It is clear that for any PA-type, there is more than one possible “ideal” model of truth-preserving causal patterns, since each proposition entails (and is entailed by) more than one other proposition. [↑](#footnote-ref-103)
104. I owe this general suggestion to BonJour, in discussion. [↑](#footnote-ref-104)
105. “Fodor’s Guide,” p. 291 [↑](#footnote-ref-105)
106. For the purposes of this discussion, I will not assume that the symbolic analogue of if-then clauses have the same syntactic properties as the material conditional of first-order logic. This is because the truth-functional definition of the material conditional in first-order logic includes deductive patterns that are not part of the reasoning patterns of the ordinary person. For example, contrary to the truth-functional definition, people generally think that that in the absence of a causal relation between x and y, propositions of the form if x then y are undetermined when x is false. [↑](#footnote-ref-106)
107. That is, the principle that *-P V Q <-> P->Q*. [↑](#footnote-ref-107)
108. As I suggested earlier, there are also deductive processes governed by rules for inferences grounded in conceptual truths that are non-syntactical (and non-formal). An example of this type of truth is the principle that *nothing is both red and green all over*. [↑](#footnote-ref-108)
109. In other words, I can think that *John is a bachelor* without having any given tautology in mind. Yet, the set of entailment relations of the semantic content of that thought is identical to the set of entailment relations belonging to an infinite number of complex propositions. This is the set of propositions that combine the simple proposition *John is a bachelor* with some tautological proposition, such as *if A then A*, (This set has an infinite number of members). [↑](#footnote-ref-109)
110. Moreover, not only are there an infinite number of propositions that attach to a given set of inferential relations. In addition, the causal networks of each PA-type support a very large number of different truth-preserving models. As explained above, the combination of any proposition and a tautology, e.g., J3, has the same set of entailment relations as that proposition, e.g., J1. Thus, the causal network of the belief that *John is a bachelor*, for instance, has not only *at least* the number of models of ideal truth-preservation as the number of beliefs, such as the belief that *John is male*, whose semantic contents its own semantic content entails. Rather, the causal network of J1 has at leastthe number of models of truth-preservation as the number of tautologies, i.e., *infinitely* many. Distinct abstract structures will often attach to distinct model networks, however, and distinct abstract structures will often belong to the entailment networks of distinct propositions. Hence, since the choice of *which* ideal network to associate with the PA-type would be arbitrary, the choice of which proposition to identify with the semantic content of a given propositional attitude would be arbitrary. In addition, the syntactic idealization proposal cannot assume that the functional patterns of each PA *qua* symbolic state are even partially parallel to the inferential patterns of any one *proposition*. This is because the proposal has to allow there to be more than one theoretical or schematic representation of the causal networks of each PA-type. [↑](#footnote-ref-110)
111. As mentioned, since there are infinitely many propositions and only finitely many possible PA-types, the point was to match the latter one-on-one with members of a subset of propositions. [↑](#footnote-ref-111)
112. See “Is Thought a Symbolic Process?”, p. 342, and “Fodor’s Guide,” p. 291. [↑](#footnote-ref-112)
113. One escape for CTM would be to deny that there is such a thing as semantic content. Rather, CTM could endorse the position that only mental symbols exist (along with the functional and syntactically-driven relations between them). However, that position would contradict the metaphysical assumptions of this essay, and it could not meet the explanatory requirements put forth in the introductory chapter. [↑](#footnote-ref-113)
114. Fodor acknowledges the apparent inconsistencies between the externalist conception of semantics and the computational (symbolic) conception of semantics. In fact, he devotes several chapters of The Elm and the Expert to resolving the inconsistencies between the claim that intentional content is informational and the claim that intentional processes, such as truth-preservation, are computationally implemented. Since the truth-preservation of intentional states is only an explanatory generalization of empirical psychology, and not a conceptual or necessary truth, he proposes to resolve the problem by saying that there is some mechanism which makes it the case that the relevant intentional properties and computational mechanisms are reliably correlated, or co-instantiated. This co-instantiation is a metaphysical contingency, which only has to be true most of the time (in our possible world) in order for the empirical laws of psychology to explain behavior (p. 24-25). [↑](#footnote-ref-114)
115. Citation. [↑](#footnote-ref-115)
116. To ask what makes it the case that a mental state may be false is to ask a question about a feature of the *mental state*. It is not to ask for a definition of truth or falsity, in the sense of asking what it is for a proposition to be true, or in the sense of asking about the nature or existence of mind-external, objective reality. In this essay, however, the discussion of intentional features of mental states does assume the existence of an objective, mind-external reality, which intentional mental states *represent* truly or falsely. [↑](#footnote-ref-116)
117. Cambridge, MA: MIT Press, (1987). [↑](#footnote-ref-117)
118. New York, Oxford University Press, (1994). [↑](#footnote-ref-118)
119. In *Psychosemantics*, Fodor explains that he considers the complex casual criterion (or set of criteria) to be jointly *sufficient* for instances of a mental symbol to represent a particular kind of object exclusively. In his discussion of “The Slightly Less Crude Causal Theory,” he says: “The Crude Causal Theory, together with psychophysics, provides a plausible sufficient condition for certain symbols to express certain properties: viz., that tokenings of those symbols are connected to instantiations of the properties they express by *psychophysical* law” (p. 113). He also says: “The condition I’m imagining is supposed to be *sufficient* but not *necessary* for ‘proton’ meaning *proton*. For all I care, there may be other sorts of routes that would connect concepts to their instances in a semantically relevant way; for example, ones that don’t depend on the existence of psychophysical laws” (p. 124). [↑](#footnote-ref-119)
120. While Fodor often refers to these mental representations using the term ‘concept’, connoting an idea or something intrinsically semantic, his theory actually involves extrinsically semantic mental *symbols*. [↑](#footnote-ref-120)
121. This essay does not concern the semantics of the entire proposition *this is a sheep*. Rather, it uses instances of the belief that *this is a sheep* as examples of mental states that contain occurrences of the mental symbol SHEEP. [↑](#footnote-ref-121)
122. Fodor bases these criteria, at least in part, upon Fred Dretske’s ideas on misrepresentation *in Knowledge and the Flow of Information*, Cambridge, MA: MIT Press, (1981). [↑](#footnote-ref-122)
123. In his discussion of the “Crude Causal Theory,” Fodor explains that what it means for this type of causal relation to take place generally is for there to be a “nomological — lawful — relation between … the property of being an instance of the property *horse* and the property of being a tokening of the symbol horse.” (Psychosemantics, p. 99.) Fodor later amends the Crude Causal Theory and constructs the “Slightly Less Crude Causal Theory.” As he explains in the amended version, the “lawful” conception of what it is for such a causal relation to take place generally does *not* mean that every horse causes the mental symbol to occur in someone’s mind, or that horses always cause such symbols to occur. (Ibid., p. 100-101.) [↑](#footnote-ref-123)
124. See “Semantics, Wisconsin Style,” in *Rerepresentation: Readings in the Philosophy of Mental Representation*. Stuart Silvers (ed.), Kluwer Academic Publishers, The Netherlands, 1989, p. 8. [↑](#footnote-ref-124)
125. *Psychosemantics*, p. 107. [↑](#footnote-ref-125)
126. *Ibid.*, p. 106. [↑](#footnote-ref-126)
127. Fodor assumes only that the asymmetrical causal condition is a necessary condition for falsehood. Yet, he expresses skepticism that it could be insufficient. “I don’t think there could be … *adventitious* asymmetric dependence: worlds in which ‘A’ means B v C, but it *just happens* that there’s a law that C’s don’t cause ‘A’s unless B’s do (and no law that B’s don’t cause ‘A’s unless C’s do) …. At a minimum, to suppose that there could be adventitious asymmetric dependence begs the question against causal theories of content” (pp. 109-110). [↑](#footnote-ref-127)
128. *Ibid.,* p. 108. [↑](#footnote-ref-128)
129. *Psychosemantics*, pp. 99-100 and 107-108. [↑](#footnote-ref-129)
130. In *Meaning in Mind: Fodor and His Critics*. Barry Lower and Georges Rey (eds.), Cambridge: Blackwell Press, 1991, p. 120. [↑](#footnote-ref-130)
131. Loar uses the term ‘concept’ to refer to entities that Fodor consider to be mental symbols. [↑](#footnote-ref-131)
132. This class of objects contains all the past and present occurrences of the SHEEP symbol in that person’s mind. [↑](#footnote-ref-132)
133. *Ibid*., p. 123. [↑](#footnote-ref-133)
134. *Ibid*., p. 120. [↑](#footnote-ref-134)
135. Loar’s notion of a demonstrative concept is similar to the more intuitive philosophical notion of an atomic perceptual concept. It is easy to see the parallel between this notion and the notion of an atomic concept, although not all types of atomic concepts are perceptual in content. This is because a concept of a certain basic type of perceptual experience cannot be reductively analyzed into discrete components; and many common concepts of physical properties appear to be complex combinations of concepts of perceptual experiences. [↑](#footnote-ref-135)
136. *Ibid*., p. 120-122. [↑](#footnote-ref-136)
137. *Ibid*., p. 122. [↑](#footnote-ref-137)
138. Fodor denies the existence of “socially deferential” and “demonstrative concepts.” Only words, he says, and not mental symbols or concepts, could even in principle have their intentional content defined by a community. The notion of a concept whose referent is determined by the beliefs of a community is incoherent, for it only makes sense that whatever a concept represents has to do with something inside the mind or with its connections to the properties of actual objects in the world, and not with what someone else arbitrarily pronounces it to mean. ("Replies [to critics]” in *Meaning in Mind*, p. 285-286.) Fodor also argues that there is no such thing as a “recognitional” concept, whose content is defined in terms of an individual’s perceptions as induced by objects in the external world. For every alleged recognitional concept is implicitly intended to refer to an object in the mind-external world, not just to a cause a kind of perception. ("There Are No Recognitional Concepts, Not Even RED,” http://www.nyu.edu/gsas/dept/philo/courses/concepts/. ) [↑](#footnote-ref-138)
139. “Can We Explain Intentionality?,” p*.*120-121. [↑](#footnote-ref-139)
140. *Ibid*., p. 120. [↑](#footnote-ref-140)
141. Such concepts can also refer to a natural kind, the general kind of object that causes any perceptual experience, and other types of objects. [↑](#footnote-ref-141)
142. *Ibid*., p. 123. [↑](#footnote-ref-142)
143. *Ibid*., p. 121. [↑](#footnote-ref-143)
144. *Concepts*, p. 75. [↑](#footnote-ref-144)
145. *Ibid.*, p. 73. [↑](#footnote-ref-145)
146. *Ibid*., p. 78. [↑](#footnote-ref-146)
147. *Ibid*., p. 78-79. [↑](#footnote-ref-147)
148. Larry BonJour made this suggestion in conversation. [↑](#footnote-ref-148)
149. See Dretske’s *Knowledge and the Flow of Information* for the philosophical notion of a learning period in establishing the semantic and epistemic relation between a cognitive template, a concept and an object in the world. [↑](#footnote-ref-149)
150. This assumption is not that the property of being a cause of SHEEP occurrences is instantiated 100% of the time that the property of being either a sheep or a goat is instantiated. Rather, it is that the property of being either a sheep or a goat is instantiated 100% of the time that the property of being a cause of SHEEP occurrences is instantiated. [↑](#footnote-ref-150)
151. This does not imply that the normal perceptual conditions and statistical predominance qualifications are not reasons for asymmetrical dependence. Rather, the suggestion is that the unique and superior connection between sheep and the relevant perceptual template is the ultimate *cause* of the fact that non-sheep only trigger the template rarely and under abnormal perceptual (or cognitive) conditions. [↑](#footnote-ref-151)
152. One cannot construct a hierarchy of degrees of “fitting the template” in terms of the phenomenal criteria that it encodes: an object either arouses these perceptual memories or does not. [↑](#footnote-ref-152)
153. P. Geach and M. Black, *Translations from the Philosophical Writings of Gottlob Frege*, second edition (Oxford: Blackwell, 1970); reprinted in Harnish, Robert M. (ed.), *Basic Topics in the Philosophy of Language* (Englewood Cliffs, NJ: Prentice Hall, 1994), p. 142-160. (All page references to this essay are to the reprint in Harnish.) [↑](#footnote-ref-153)
154. P. Strawson (ed.), *Philosophical Logic*, (Oxford: Oxford University Press, 1967), translated by A. M. and Marcelle Quinton; reprinted in Harnish, Robert M. (ed.), *Basic Topics in the Philosophy of Language* (Englewood Cliffs, NJ: Prentice Hall, 1994), p. 517-535. (All page references to this essay are to the reprint in Harnish.) [↑](#footnote-ref-154)
155. It may seem that the term ‘the Queen of Thebes’ in this sentence functions solely as a complex part of the predicate ‘is the Queen of Thebes’, so that it only contributes to the ascription of the property of being the Queen of Thebes. However, in this essay, I use this term (and others of its kind, such as ‘the mother of Oedipus’) as a distinct linguistic unit that refers to a unique possible individual. Thus, the term ‘is’ within sentences such as ‘Jocasta is the Queen of Thebes’ does not function merely as a part of a predicate, but also as a symbol for an identity relation that holds between the referents of the two terms, e.g., ‘Jocasta’ and ‘the Queen of Thebes’. [↑](#footnote-ref-155)
156. *Ibid*., p. 143. [↑](#footnote-ref-156)
157. Nonetheless, the term ‘Thebes’ within the denoting phrase ‘The Queen of Thebes’ *is* a proper name. However, this fact does not prevent the relevant complex term from exclusively denoting a unique possible individual. (See fn. 11 below). [↑](#footnote-ref-157)
158. While truth-functional compound sentences, such as ‘Jocasta is the Queen of Thebes and Jocasta is the mother of Oedipus’, are complex and contain simpler sentences, these are not intensional sentences. This is because truth-functional compounds do not contain ‘that’ clauses that refer to sentences or propositions in themselves. In other words, the from an intuitive perspective, the sentence ‘Laius believes that Jocasta is the Queen of Thebes’ refers either to the sentence ‘Jocasta is the Queen of Thebes’ or to the *proposition* that *Jocasta is Queen of Thebes*. For it says, among other things, that either this sentence or this proposition has the extrinsic property of being believed by Laius. Yet, the truth-functional compound above does not ascribe any property to either the sentence or the *proposition* that Jocasta is Queen of Thebes. Rather, it expresses the truth of the proposition itself through containing the sentence itself. Thus, truth-functional compound sentences cannot be intensional sentences, because they have truth-functional contexts only and not intensional contexts. [↑](#footnote-ref-158)
159. *Ibid*., p. 146. [↑](#footnote-ref-159)
160. *Ibid*., p. 144. [↑](#footnote-ref-160)
161. In Marsh R., (ed.), *B. Russell, Logic and Knowledge*, London: George Allen and Unwin, 1956 (reprinted in Harnish, Robert M. (ed.), *Basic Topics in the Philosophy of Language* (Englewood Cliffs, NJ: Prentice Hall, 1994), p. 161-173.). All page references to this essay are to the reprint in Harnish. [↑](#footnote-ref-161)
162. *Ibid*., p. 162. [↑](#footnote-ref-162)
163. In this passage, Russell appears to exclude proper names, such as ‘Bertrand’, from the category of ‘denoting phrases’. This exclusion may seem to be justified because proper names do not denote “in virtue of their form,” but rather, solely in virtue of conventional practice. Moreover, the wide majority of proper names are used to denote more than one individual. Yet, in another work, Russell asserts that a proper name *can* be semantically reduced to an implicit definite description because of the meaning that the user of such a name associates with her utterances (or inscriptions) of it. Thus, according to this version, a proper name such as ‘Bertrand’ could be reduced to a closet definite description, such as ‘the author of “On Denoting” ’, etc. In this case, a proper name *would* be a “denoting phrase.” (See *The Problems of Philosophy*, New York: Oxford University Press, 1997, p. 54.) [↑](#footnote-ref-163)
164. Obviously, the singular term ‘the wife of Oedipus’ contains a distinct singular term: ‘Oedipus’. For this reason, it may seem to be inappropriately complex. For a thorough analysis of the logical and semantic import of ‘the wife of Oedipus’ as a definite description would certainly have to express the import of ‘Oedipus’ as a definite description. Yet, Russell himself uses examples of definite descriptions on the same level of complexity, such as ‘the King of France’, to illustrate his own theory of the semantic content of such terms (see p. 165). [↑](#footnote-ref-164)
165. Salmon in *Basic Topics*, p. 107, and S. Marc Cohen in lecture and discussion guide my explanation of the three forms of propositions. [↑](#footnote-ref-165)
166. “On Denoting,” p. 163-164, 169. [↑](#footnote-ref-166)
167. *Ibid*., p. 166. [↑](#footnote-ref-167)
168. *Ibid*., p. 169. [↑](#footnote-ref-168)
169. *Ibid*., p. 169. [↑](#footnote-ref-169)
170. Some intensional sentences have intensional contexts that contain no definite descriptions. Examples of these are sentences that ascribe beliefs in general propositions, like ‘Oedipus believes that all Thebans are noble’, and those that ascribe beliefs in non-unique existential propositions, such as ‘Oedipus believes that some Thebans are rich’. [↑](#footnote-ref-170)
171. Russell’s argument can be interpreted to reject the Fregean semantic view that a term has an ordinary conceptual meaning, in addition to the Fregean metaphysical view that this ordinary conceptual meaning derives its content from a sense, which is an abstract, mind-independent, *and* intrinsically semantic object. [↑](#footnote-ref-171)
172. It seems to me that since such terms express unique sets of properties, e.g., wife of Oedipus, and since the sentences that they occur in express that such properties are instantiated, it follows that such terms indirectly refer to the individuals that exist and instantiate these unique properties. [↑](#footnote-ref-172)
173. “On Denoting,” p. 169. [↑](#footnote-ref-173)
174. *Ibid*., p. 169. [↑](#footnote-ref-174)
175. Recall that I have acknowledged that the name ‘Oedipus’ occurs in the definite descriptions ‘the wife of Oedipus’ and ‘the mother of Oedipus’. This does not distract from the analysis because Russell’s ultimate position was that such names could be reduced to definite descriptions, as well. [↑](#footnote-ref-175)
176. Of course, Frege did not use the term ‘concept’ in this way. According to him, a concept is the denotation of a monadic predicate, that is, a function from objects to truth-values. For example, the denotation of ‘is a horse’ is a function that maps each horse onto The True and each non-horse onto The False. (He also held that The True and The False are abstract objects (See “Sense and Reference,” p.147).) I will not use the term ‘concept’ in this technical, Fregean sense. [↑](#footnote-ref-176)
177. Since senses are intrinsically semantic, while some abstract objects, such as properties, are not semantic at all, some abstract objects are not senses. [↑](#footnote-ref-177)
178. In contrast, Frege uses the term ‘concept’ to refer to a mind-independent, abstract object that plays the non-semantic role of a *function*. Frege did not appear to endorse the notion of mind-internal *and* objective semantic entities; rather, he appeared to believe that the “representations” that are constituents of mental states were essentially *subjective* in character. [↑](#footnote-ref-178)
179. In *Concepts: Where Cognitive Science Went Wrong* (New York: Oxford University Press, 1998), Jerry Fodor objects to Frege’s metaphysical arguments that senses must be non-mental abstract objects. Fodor’s argument centers on the first semantic function that Frege thought senses had to accomplish: to be the mode of presentation (MOP) of the referent of a term to a speaker. (The mode of presentation of the referent of the term is what the speaker (or hearer) of the term interprets the term to express, i.e., what he or she associates with the term.) Fodor’s main objection is roughly that there must be more than one way to grasp an MOP. He considers this to imply that MOPs must be mental entities, i.e., concepts. Fodor also argues that that MOPs must be mental entities because they must be able to cause other thoughts and behavior. [↑](#footnote-ref-179)
180. p. 145. [↑](#footnote-ref-180)
181. I owe this schema to S. Marc Cohen in lecture and discussion. [↑](#footnote-ref-181)
182. “The Thought,” p. 519. (Concerning the qualification that it is not true that the sense of every sentence is a thought, Frege probably means that some non-sensical sentences don’t express logically coherent thoughts. Moreover, there are clearly senses, i.e., senses of individual terms, which are not senses of sentences. Yet, these senses often become components of the senses that express complete thoughts, according to Frege’s principle of the compositionality of senses. Thus, the immateriality of the latter implies the immateriality of the former.) [↑](#footnote-ref-182)
183. *Ibid*., pp. 525, 527. [↑](#footnote-ref-183)
184. Nonetheless, a materialist about the mind may insist that neural processes are physical relations that depend for their existence upon the existence of particular minds, although not upon particular mental states. [↑](#footnote-ref-184)
185. One may believe that the only immaterial things are those such as concepts, emotions, sensations, and so on, all of which depend for their existence upon the existence of particular minds. [↑](#footnote-ref-185)
186. “The Thought,” p. 531. [↑](#footnote-ref-186)
187. E. - H. Kluge, *The Metaphysics of Gottlob Frege*, (Boston: Martinus Nijhoff Publishers, 1980), p. 198. [↑](#footnote-ref-187)
188. *Ibid*., p. 198-199. [↑](#footnote-ref-188)
189. “The Thought,” p. 525 [↑](#footnote-ref-189)
190. *Ibid*., p. 525. [↑](#footnote-ref-190)
191. *Ibid*. p. 525. Frege continues to offer what appears to be an epistemological argument for the conclusion that if ideas such as the his color-sensation of the leaf could belong to both his consciousness and that of his companion, then it would be possible to compare them or to know if they were the same. Since this is presumably impossible in principle, Frege concludes that ideas must be unique. Of course, one could consider this reductio fallacious because it confuses the truth of an identity relation with the possibility of knowing it. [↑](#footnote-ref-191)
192. There is a philosophical notion of a subjective sense of a term. This conception of a subjective sense, however, only differs from the conception of an objective sense in that the objectively defined properties that a term expresses may vary from person to person. Yet, the content of such a particular sense is still objective in that it is not constituted by any unique, experiential features. [↑](#footnote-ref-192)
193. One may think that these are different versions of one and the same argument. Yet, this is not the case, as the former line of reasoning concerns the unique possession of mental entities with subjective content, while the latter line of reasoning concerns the absurd implications of denying the objectivity of the truth-value and semantic content of senses. [↑](#footnote-ref-193)
194. In this essay, I will assume that this conception of ideas applies to ideas of any kind of thing, and not just to ideas of perceivable objects. For one can have sensory images, subjective associations and experiences of abstractions as well as physical objects. [↑](#footnote-ref-194)
195. Grossmann (p. 29) tells us that in earlier works, Frege held that ideas were either sense-impressions or images. [↑](#footnote-ref-195)
196. “On Sense and Reference,” p. 144. [↑](#footnote-ref-196)
197. One must take note of the fact that one and the same term may have more than one sense. In other words, Oedipus may associate a sense s with the term ‘Jocasta’ that does not include the feature of being Oedipus’s mother and does include the feature of being Oedipus’s wife, while Laertes may associate a sense s1 with ‘Jocasta’ that does include the former and excludes the latter. In addition, the sense that an *individual* associates with a term may presumably vary over time*.* Perhaps more precise renderings of M1, M2 and M3 would substitute ‘the set of senses {s1, s2, ..., sn} of a term t’ for ‘the sense of a term t’. They might also interpret the assumption that the sense of a term determines its reference differently. That is, this assumption could mean that the particular sense associated with a given utterance or inscription of a term secures the entity that t denotes. Thus, M1 would say that the particular sense sn associated with t during a given utterance or inscription by an individual is *one and the same entity as, and not merely a qualitatively identical copy of,* the sense that any other given individual gets in contact with when he associates *that* sense sn with t. The first part of M2 would thus be a conceptual truth, namely, that the particular sense sn associated with a particular token of t is qualitatively identical to that of any other individual who gets in contact with that sense. The second part of M2 would deny the existence of a mind-independent universal sense sn that the relevant particular thoughts instantiate. M3, in contrast, would affirm the existence of this mind-independent sense. [↑](#footnote-ref-197)
198. One may attempt to resolve this latter question through an appeal to Frege’s semantic definition of a sense: that which secures the referent of a term. For since more than one person uses the same term to refer to the same kind of object, it would make sense, prima facie, to think that this semantic definition requires that terms have the same sense for different users. Yet, Frege himself says that one and the same term can have more than one sense, and that more than one sense can denote the same object. Thus, the mere fact that more than one person can refer to the same object with the same term does not entail that the sense of a term must be held in common. For each person could, theoretically, have his own sense of the term and still refer to one and the same object with that term. [↑](#footnote-ref-198)
199. As explained above, we can consider his metaphysical arguments concerning the nature of senses of sentences to apply equally to the question of the nature of the senses of terms. [↑](#footnote-ref-199)
200. “The Thought,” p. 526. (I assume that ‘recognizable’ here expresses what Frege means by ‘bearable’ or ‘graspable’.) [↑](#footnote-ref-200)
201. One may wonder whether the ideas from which Frege distinguishes senses are general types of ideas (albeit types belonging uniquely to particular individuals) or mere particular manifestations of these types. It seems that an idea uniquely associated with a sentence S by an individual would have to be a *type* of idea only manifested in the mind of that one individual. In other words, the unique association must be the subjective and unique qualities that define the idea as a type, and not the trivial metaphysical fact that a particular manifestation of this idea at a certain time can only belong to her consciousness. For this is the only relevant and interesting interpretation according to which the idea could belong uniquely to the content of her consciousness in a way which *prevents* it from being identical to the sense of the sentence (which is held in common by many). It could be that the idea that an individual associates at a given time with a particular utterance of S is distinct from that which anyone else associated with that particular utterance. Nonetheless, this idea could be a member of a general pool of ideas associated with utterances of S. All people could draw the ideas that they associate with these utterances from this general pool. In that case, there could be a different occasion in which the sentence was uttered (or inscribed); and the same individual could associate a different idea with this token of S, and a different individual could associate an idea with this token which was qualitatively identical to the idea which the first individual had on the first occasion. This interpretation would clearly not answer the question how the uniqueness of the idea to that individual prevents it from being identical to the sense of the sentence, which is common to many individuals. [↑](#footnote-ref-201)
202. One could object that all Frege has established is that senses cannot be dependent for their existence upon being tokened by any specific individual, not that they can’t be dependent for their existence, *qua* types, upon being grasped or instantiated in the mental state of *some* individual. This challenge centers on the question whether universals must be instantiated in order to exist, which is a metaphysical issue that is immaterial to the topic of this essay. Yet, I will later argue that abstract entities, namely properties, are a source of semantic content for propositional attitudes. [↑](#footnote-ref-202)
203. “The Thought,” p. 526. [↑](#footnote-ref-203)
204. It may be thought that Frege is discussing what it would mean to describe the sense or thought of the sentence indexically,and not the sentence itself. However, it is clear that he is discussing a sense as the content of the Pythagorean theorem, for example, and as a thing that “belongs necessarily to the sentence.” Therefore, even if the thing that he thinks would be describable only indexically *is* the sense of the sentence, and not the sentence itself, the necessity of describing this *sense* indexically would mean that such a sense could be true for one person and false for another. As such, the truth of a sentence used to express such a sense would be relative to the individual who utters it. For a sentence refers to its truth-value, and the sense of a term determines its referent. If there were only one sentence that was the PT, then if it had different senses for different individuals, one and the same sentence could be true and false. Thus, it would be necessary to describe both the sentence *and* its sense indexically*.* [↑](#footnote-ref-204)
205. *Ibid*., p. 526. (There is an epistemological conclusion in this passage as well: “it would remain doubtful whether anything at all comparable occurred in the consciousness of others.” This is clearly an inference that the indexicality of the sense of a mathematical sense implies that we can never know whether others understand or entertain this sense. This inference is immaterial to the metaphysical issues of this essay. For the fact that the reductio assumption implies that we can’t know that we are communicating with others, or that others have the same thoughts that we do, implies nothing about whether the metaphysical view that senses are mind-dependent entities is true.) [↑](#footnote-ref-205)
206. Most philosophers would attach several qualifications to this crude definition-form, e.g., that the causal relationship must be based in some non-dispositional property, that the causal relation must take place under normal conditions (however that notion is defined), and that the sense-experience, as a type, must occur in a group of people, as opposed to one individual (See p. 525). [↑](#footnote-ref-206)
207. In “Sense and Reference” (p. 147-148), and in “The Thought” (p. 518), Frege himself challenges the supposition that the truth-*value* of a sentence may be ascribed to that sentence as a *property* of it. One may resolve the inconsistency by considering a distinction between a truth-value, i.e., the property of x being true or x being false, and the actual abstract *objects* that Frege terms “ ‘the True’ or ‘the False’.” As mentioned, Frege holds that sentences refer to *these* objects. [↑](#footnote-ref-207)
208. “The Thought,” p. 527. [↑](#footnote-ref-208)
209. Clearly, a propositional-attitude instance may contain more than one intrinsically representative concept. In this case, each such concept must be applied accurately to some given object (or class thereof) in order for the entire mental state to represent those objects (or classes) accurately. Hence, each intrinsically representative concept that is an ingredient of a PA-instance must accurately represent the objects to which it is applied in order for the state to be true or false. [↑](#footnote-ref-209)
210. One may think that this theory of a distinction between the feature of being intrinsically representative and the feature of being intrinsically semantic is analogous to (or the same as) the distinction in logic between a distributed term and a non-distributed term. Geach notes that while Aristotle was not aware of any formal theory of term distribution, the theory “is integral to the ‘Aristotelian’ doctrine of categoricals and syllogisms.” ( See P.T. Geach,"The Doctrine of Distribution" Mind 65 (1956) 67-74.)

     Aristotle lays out a schema of four categories of propositions, namely, those of types A, E, I, and O. Only the terms that are preceded by ‘every’ in the following schemata of these propositions are *distributed* terms:

     A: Every F is identical to some G (All Fs are Gs).

     E: Every F is distinct from every G (No Fs are Gs).

     I: Some F is identical to some G (Some Fs are Gs).

     O: Some F is distinct from some G (Some Fs are not Gs).

     Despite the analogy with the distinction being put forth between intrinsically representative and merely intrinsically semantic concepts, however, the current discussion of Bare Property Intentionality’s view of concepts only applies to propositions of the form *All Fs are Gs*. This is because this essay only puts forth a view of the semantic content of propositional-attitude instances whose semantic contents take the form of universal propositions. These are occurrences of beliefs that *All Fs are Gs*.

     In addition, Aristotle puts forth his view of class inclusion, categorical syllogisms, and the corresponding schema of proposition-forms within the context of a whole theory of logic. In contrast, Bare Property Intentionality’s distinction between intrinsically representative and merely intrinsically semantic concepts only concerns the semantic features of mind-internal concepts. That is, the theory uses this distinction to help answer the question of which mind-internal concepts represent things in an objective manner. For according to the theory, the manner in which an intrinsically representative concept represents a thing (or each member of a class of things) allows that concept to be a semantic ingredient of a truth-evaluable mental state. [↑](#footnote-ref-210)
211. In “The Thought,” Frege assumes that mental representations, or “ideas,” are unique to an individual’s consciousness. He argues that if senses were not intersubjectively accessible, i.e., if senses were unique to the consciousnesses of individuals, then certain senses, such as the sense of the mathematical sentence ‘a2 + b2 = c2’, could be true for some and false for others. By reductio, he concludes that a sense cannot be identical to any idea. (P. Strawson (ed.), *Philosophical Logic*, (Oxford: Oxford University Press, 1967), translated by A. M. and Marcelle Quinton; reprinted in Harnish, Robert M. (ed.), *Basic Topics in the Philosophy of Language* (Englewood Cliffs, NJ: Prentice Hall, 1994), p. 526.) (All page references are to the reprint in Harnish.) Another implication of this reductio is that a subjective mental phenomenon, for Frege, may be defined as one that lacks an objective truth-value. [↑](#footnote-ref-211)
212. I owe this example (of a contingently false universal proposition) to Derek Dawson. [↑](#footnote-ref-212)
213. This chapter does not specifically concern the metaphysics of PA-instances whose contents are existential or singular propositions, such as a belief that *some nuts are fruit* or a belief that an individual person, such as Bill Gates, is rich. However, *one* of this chapter's metaphysical conclusions concerning intrinsically representative concepts of classes of objects also applies to concepts of partial classes of objects and concepts of individual objects. This conclusion is that all intrinsically representative concepts must include a bare referential intention. (In contrast, concepts that merely present properties to the mind without a bare referential intention are only intrinsically semantic.) [↑](#footnote-ref-213)
214. There are intrinsically semantic concepts of the instantiation of *non*-physical properties, such as the concept that presents the property VALID to the mind in an instance of the belief that *all syllogisms are valid*. Bare Property Intentionality accounts for the development of instantiation-concepts in terms of a person’s mental process of abstraction over his sensory experiences of *perceptible physical* objects. Instantiation-concepts of non-physical properties are either innate, or they develop from those of physical properties. [↑](#footnote-ref-214)
215. A later section will explain the selection of the term ‘intention’ to describe these mental events. It will also explain the primitive character that prevents them from being propositional or volitional. [↑](#footnote-ref-215)
216. Naturally, there are other philosophical views of what it is for more than one object to have something in common, all of which attempt to avoid the conclusion that universals actually exist. One position is that more than one object may belong to the same “primitive natural class,” the notion of which, as befitting its name, cannot be further analyzed or reduced. Others attempt to *explain* the notion of a natural class, by defining it in terms of quasi-objective similarity (or resemblance) relations among objects. [↑](#footnote-ref-216)
217. A proponent of Bare Property Intentionality says that *this* principle applies to the content of *all* intrinsically semantic mental states, i.e., those about physical entities and those about non-physical entities. [↑](#footnote-ref-217)
218. Intrinsically semantic concepts that are not intrinsically representative, such as the instantiation-concept of FRUIT, must also present properties to the mind essentially. Yet, such concepts do not intrinsically and exclusively represent any entity in a manner that may be an ingredient of a truth-evaluable proposition. The instantiation-concept of FRUIT does not represent the class of all fruit, or some sub-set of this class, or any individual piece of fruit. Therefore, it is not necessary that this concept represents any thing or class that exclusively instantiates this property. [↑](#footnote-ref-218)
219. A later section will discuss how Bare Property Intentionality accounts for mental truth-preservation. [↑](#footnote-ref-219)
220. Frege held that a sentence refers to its truth-value (“On Sense and Reference,” in *Basic Topics,* p. 148). [↑](#footnote-ref-220)
221. The same philosophical point could be made with an example of two sentences that refer to propositional attitude-instances of the form *all Fs are Gs.* [↑](#footnote-ref-221)
222. In other words, the terms in the intensional context of the sentence combine to refer to the intrinsically semantic constituents of Eric’s belief, namely, its concepts. One could also hold, however, that these terms combine to refer to the semantic *content* of the belief (as opposed to referring to the concepts themselves). In either case, the truth conditions of the sentence will involve the occurrence of the relevant concepts within the mind of Eric; and it is reasonable to assume that the sentence, in some sense, refers to these entities. Given that an intensional sentence of the form Eric believes that *x* may not be true, i.e., there may actually be no such concepts in Eric’s mind; the terms in the intensional context of the sentence may refer to possible but non-actual entities. That is, it can be the case that Eric has such a belief only in a possible world, but not in the actual world. [↑](#footnote-ref-222)
223. The question of how *instantiation*-concepts connect the mind with properties is the focus of a later section of the chapter. [↑](#footnote-ref-223)
224. Even if it were not a mere component of the concept, there would still be a regress. [↑](#footnote-ref-224)
225. Both types of presentations are *intrinsic* to the concept, however. For the metaphysical assumptions of this essay require that: (1) in either case, the concept intrinsically represents a class of objects; and (2) Property Principle governs this relationship. Therefore, the concept must bear some essential relationship to some property instantiated by all and only members of the relevant class. This relationship is the concept’s presentation of the property to the mind. Thus, a concept intrinsically presents some property to the mind, independently of the question whether this presentation is an intrinsically semantic representation in-itself. (My definition of the notion of an intrinsic relation is as follows: x bears an intrinsic relation R to y just in case x exists and bearing R to y is part of the identity conditions of x, that is, just in case *-Rzy* entails that *z ≠ x*. The number 5, for example, bears an intrinsic relation to the number 4: the relation EQUAL TO 1 + y; and an individual water molecule bears the intrinsic relation IS CONSTITUTED BY y to some group consisting of two particular hydrogen atoms and one particular oxygen atom.) [↑](#footnote-ref-225)
226. One may object that this conclusion does not follow, because an act of merely entertaining a proposition without affirming or denying it is an act of considering the possibility that the proposition is true. Considering the possibility that a universal proposition is true, in turn, is an act of representing some class of objects in at least one possible world, although it is not an act of representing any class of objects in the actual world. Therefore, one may conclude that except in the case of contradictory universal propositions, for a universal proposition to be presented to one’s mind implies that an intrinsically representative state occurs in one’s mind.

     However, while merely considering the possibility that a universal proposition is true is an act of representing some class of objects in at least one possible world, under the metaphysical assumptions of this essay, it is not an act of representing any thing in a manner that *may be evaluated as true or false*. For an entity is truth-evaluable because (1) it represents a class of objects (or individual) as having a given property, (2) the class (or individual) either does or does not have the property, and (3) it is not the case that the class (or individual) both has and does not have the property. Thus, for a semantically contentful entity to be evaluable as true or false is for it to be evaluable as true or false in the actual world. For if a theory of intentionality does not define the truth-evaluability of a semantically contentful entity to be modally restricted, its account of truth-evaluability fails to meet condition (3). [↑](#footnote-ref-226)
227. One might object that the fact that a mental state is assertorically neutral does not mean it is not a PA-instance. For example, fearing that x is a PA (schema), and desiring that x is a PA (schema). However, this essay only concerns the features and ingredients of instances of intrinsically representative PA-instances, namely, those that can be evaluated as being true or false. Suppose, for example, that Susan fears, but does not have enough reason to believe, that all of her credit cards have been stolen. Moreover, assume that Susan is not aware of the fact that the wallet containing all of her credit cards merely is hidden under her couch pillow. Although the proposition *all of Susan’s credit cards have been stolen* is false, it is not the case that her occurrent mental *state* of fearing that *all of (my) credit cards have been stolen* is false. For despite its propositional content, this mental state is not evaluable as *being* true or false. [↑](#footnote-ref-227)
228. According to the semantics of first-order predicate logic, the appropriate way to interpret this sentence is: *for all objects x, if the propositional function Bx is satisfied by x, then the propositional function Sx is satisfied by x*. While this shows that it may not be appropriate, from some perspectives, to interpret the semantics of this sentence in terms of class membership and class inclusion, concepts of classes of objects are a basic unit in the metaphysics of Bare Property Intentionality. To address the possibility that a class represented by a mental state may not exist, I appeal to the notion that some concepts have possible classes of objects as their extensions. One who has merely misplaced a wallet, but who believes it to have been stolen, may come to have beliefs about a possible but non-actual class of objects: the class of all people who participated in the theft. For instance, the owner may come to believe that *the people who stole my wallet must have entered through though a window*. Such a PA-instance has intrinsically semantic and logically structured content, even though it does not refer to an actual object or class. A proponent of Bare Property Intentionality does so by considering the owner’s concept of *the people who stole my wallet* to refer to a possible but non-actual class. [↑](#footnote-ref-228)
229. This does not mean, of course, that an entity may not be semantic without involving a bare referential intention. A word or sentence contains only symbols, and yet, it may represent something extrinsically without including a BRI. Thus, BRIs distinguish instances of an intrinsically representative PA from three types of entities: (1) *extrinsically* semantic entities, such as words and sentences, (2) metaphysically objective and *non*-semantic entities, such as individual objects, properties, events, substances or stuffs, and so on, and (3) metaphysically subjective entities, i.e., mind-internal sensations, emotions, associative phenomena, perceptual states, etc. [↑](#footnote-ref-229)
230. Being semantically generic, the extension of a bare referential intention *in-itself* is numerically indefinite. Thus, in principle, a concept that represents any number of objects, including an individual object, can include a bare referential intention. [↑](#footnote-ref-230)
231. Nonetheless, the property that combines with a bare referential intention cannot be a contradictory property. Otherwise, the relevant state will not succeed in representing any possible object. [↑](#footnote-ref-231)
232. A BRI must be more than that through which a person’s mind is directed toward the mind-external world, since there can be intrinsically representative mental states about mental phenomena. [↑](#footnote-ref-232)
233. An instantiation-concept is not a philosophically or theoretically *sophisticated* concept of the relation whereby an object instantiates a property. An instance of a belief that *all abstract objects are instantiated properties*, for example, contains such a sophisticated concept. This type of concept is not relevant to the present discussion, because it belongs to the most abstract category of mental representations, while Bare Property Intentionality only concerns the most simple and basic class concepts and instantiation-concepts. [↑](#footnote-ref-233)
234. If an instantiation-concept of the property FRUIT intrinsically represented this property, then each PA-instance that includes this concept would have to intrinsically present this property to the mind in a manner that one may evaluate as true or false. As mentioned earlier, the Property Principle implies that if x represents an individual abstract object a in an intrinsically semantic manner, then x presents to the mind some property F that is exclusively instantiated by a. Since properties are abstract objects, then in an instance of the belief that *all nuts are fruit*, the instantiation-concept would have to present the property FRUIT to the mind through presenting to the mind some second-order property exclusively instantiated by FRUIT. This would imply a vicious infinite regress. In addition, it seems to be highly implausible that within an instance of a common-sense level belief such as *all nuts are fruit*, there is a concept that represents the abstract object FRUIT in the same sense in which there is a concept that represents the class of concrete objects that are nuts. [↑](#footnote-ref-234)
235. The sentence ‘all cherries are red things’ is more appropriate than the sentence ‘all cherries are red’ to describe the semantic content of this belief. This is because the current discussion concerns instance of propositional attitudes of the form *all Fs are Gs* only. Yet, the semantic content of this mental state is still a universal proposition, independently of which sentence is used to describe this content. That is, the mental state presents the property RED to the mind as being instantiated by each individual cherry. While for all philosophical intents and purposes, the property term ‘RED’ refers to the same abstract object as the property-term ‘RED THING’, the former term appears to be more appropriate theoretically. [↑](#footnote-ref-235)
236. Most types of fruit, and thus most pieces of fruit, tend to taste tangy-sweet, juicy, and fleshy to humans. However, humans commonly eat some types of fruit, such as peanuts, that do not taste these ways, and which many people do not consider to be fruit. Science tells us that a piece of fruit is: "The ripened ovary or ovaries of a seed-bearing plant, together with accessory parts, containing the seeds and occurring in a wide variety of forms.” Nonetheless, there is a secondary definition of ‘fruit’: “An edible, usually sweet and fleshy form of such a structure. A part or an amount of such a plant product, served as food….” This definition appears to reflect common usage of the term ‘fruit’ to refer to the type of edible object that does cause tangy-sweet, juicy, and fleshy perceptual experiences in humans. [↑](#footnote-ref-236)
237. While some types of fruit, such as peanuts, almonds, hazelnuts, etc., do not tend to cause these perceptual experiences in humans, it is nonetheless true that the *only* objects that do tend to cause these experiences simultaneously *are* pieces of fruit. Moreover, this discussion concerns the actual intuitive semantic content of a BIC of the property FRUIT. Most people do think of fruit as objects that cause tangy-sweet, etc., sensations; and a secondary definition of ‘fruit’ does include the tangy-sweet and fleshy features. Thus, an account of the average person’s acquisition of a basic and commonplace concept of the property FRUIT should include this type of perceptual association. For the issue here is the accuracy of the account of the content of actual BICs given by a proponent of Bare Property Intentionality ⎯ not the accuracy of these concepts themselves. [↑](#footnote-ref-237)
238. A proponent of Bare Property Intentionality does not hold that the property FRUIT is instantiated by *all* and only objects that tend to taste tangy-sweet, juicy, and fleshy at the same time. [↑](#footnote-ref-238)
239. For the purposes of this discussion, it is acceptable to consider the above type of intuition to concern the primary property FRUIT in itself, even though the properties that actually define fruit as a class are more technical and biological. Bare Property Intentionality’s account of instantiation-concepts is not concerned with the features of sophisticated scientific beliefs. Moreover, the average person is not thinking of precise biological definitions when he develops the concept of fruit or has an instance of a belief that *all nuts are* fruit. [↑](#footnote-ref-239)
240. A proponent of Bare Property Intentionality assumes that the human mind is *innately* disposed to have the relevant predecessor intuitions of axiomatic logical, arithmetical, and metaphysical principles, when triggered by perceptual experiences of the external environment. The intuitions that cause the development and occurrence of a BIC must be innate, because the assumption that these intuitions are *learned* appears to imply a vicious regress. The underlying reason for all facets of this implication is that the view that such intuitions are acquired through *learning* is the only alternative to the view that they are innate. Yet, whatever type of process an individual would undergo to learn these intuitions would have to *involve* intuitions that presented the same types of axiomatic principles to the mind. It is clear, at least, that the logical and arithmetical relations are embedded in axiomatic principles. This implies that a person cannot *infer* such principles without assuming equally axiomatic principles. Moreover, there are no mental presentations of these relations whose contents could be acquired solely through perceptual experience of the external environment. For a person would have to possess mental presentations of these relations already, in order for perceptual interaction with physical objects alone to bring them about. [↑](#footnote-ref-240)
241. Other mental phenomena (including PA-instances) and mind-external stimuli trigger all subsequent occurrences of the concept. [↑](#footnote-ref-241)
242. Intuition (3) concerns something (or things) outside the mind as having tendencies to cause certain perceptual experiences. [↑](#footnote-ref-242)
243. In this discussion, a mental state that concerns the causal agent of a phenomenon concern instantiations of properties. These instantiations are concrete, and they can have effects upon other concrete events in the spatio-temporal world, such as occurrences of qualitative mental states. In contrast, other mental states concern the ultimate *source* of the secondary properties that are dispositions to cause these experiences. These types of intuitions concern properties as abstract objects, and not just the concrete instantiations of properties. Of course, a basic instantiation-concept is both about an instantiation of a property and about the property itself. [↑](#footnote-ref-243)
244. It does *imply*, under the assumption of existential import, that at least one member of the fruit class is a member of the nut class. [↑](#footnote-ref-244)