Title
Conceptual Role Semantics

Permalink
https://escholarship.org/uc/item/1sd8h2g5

Authors
Greenberg, Mark
Harman, Gilbert

Publication Date
2005-09-01

Peer reviewed

Conceptual Role Semantics

Mark Greenberg, UCLA and Gilbert Harman, Princeton
University
Draft as of 9/1/2005

1. Meanings Determined by Use

Conceptual role semantics (CRS) is the view that the meanings of expressions of a language (or other symbol system) or the contents of mental states are determined or explained by the role of the expressions or mental states in thinking. The theory can be taken to be applicable to language in the ordinary sense, to mental representations, conceived of either as symbols in a "language of thought" or as mental states such as beliefs, or to certain other sorts of symbol systems. CRS rejects the competing idea that thoughts have intrinsic content that is prior to the use of concepts in thought. According to CRS, meaning and content derive from use, not the other way round.

CRS is thus an attempt to answer the question of what determines or makes it the case that representations have particular meanings or contents. The significance of this question can be seen by considering, for example, theories of mind that postulate a language of thought. Such theories presuppose an account of what makes it the case that a symbol in the language of thought has a particular meaning. Some conceptual role theorists have not clearly distinguished this kind of question from questions about the nature of the meanings or contents of various kinds of representations. CRS, as we understand it, is consistent with many different kinds of positions on the latter question. For example, as we discuss below, CRS has no commitment to the view that the meaning of a symbol should be *identified* with its conceptual role.

Some discussions of CRS (e.g., Sellars, 1963; Harman, 1974, 1975, 1987) suppose that CRS must limit the relevant uses to those involved in inference, in reacting to perception, and in decisions leading to action. (In Section 5. below, we discuss versions that take an even more limited view of relevant factors.) But it is best to begin discussion by interpreting "conceptual role" in the widest possible way, considering a great variety of uses of symbols in thought, in order to be able to ask which uses if any might be relevant to meaning or content and how they might be relevant (see section 3.).

We propose to use the phrase *conceptual role semantics* or *CRS* in a very broad sense, according to which CRS includes any theory that holds that the content of mental states or symbols is determined by any part of their role or use in thought. There is a common use of the term that is more limited. In

this use, in order to count as a version of CRS, a theory must hold that the determinants of content include the role of the mental states or symbols *in inference* or in other purely internal mental processes. This restriction excludes information-based or indication theories of content (see section 5. below). By contrast, on our broader use of the term CRS, information-based or indication theories count as special versions of CRS.

In what follows, we will sometimes use the abbreviation "CCRS" for the sort of CRS that takes the recognition of internal inferential and implicational relations to be crucial to the meaning or content of some expressions or syntactic constructions. CCRS allows also other aspects of use, such as relations of symbols to perceptual input and to actions.[1] So, we will use the expression *CCRS* in the way that some theorists use the phrase *conceptual role semantics.*

Just how inclusive our broad understanding of CRS is depends on how broadly conceptual role or use is understood. For example, *teleological* theories of content give an important role to the evolutionarily determined "function" of symbols or symbol structures, where some such theories understand the notion of the function of a symbol or structure in a way that goes beyond the symbol's use or role as ordinarily understood (e.g., Dretske, 1988, 2000; Millikan, 1984, 1993; Neander, 1995; Papineau. 1987). We do not count such theories as versions of CRS. (We discuss these theories in section 6.3, below.)

One other point is that we understand conceptual role in such a way that it might be externally or non-individualistically individuated. Thus, if we consider myself and my twin on Twin Earth (Putnam, 1975), it is arguable that my symbol for water and his symbol for twater have different conceptual roles. For example, they have different relations to properties in the world, as I have often applied my symbol to H_2O, and he has often applied his to XYZ. If our uses are individuated externally, our uses are different since my uses are water-applications and my twin's are twater-applications.

When CRS is understood in our ecumenical way, much of the currently active debate is a debate between competing versions of CRS, such as between CCRS and information-based theories, rather than between CRS and other positions.

There are theorists, however, who reject CRS on even the most inclusive understanding of it. According to such theorists (Searle, 1980; Bonjour, 1998), the content of mental states is intrinsic to them, not explained by their use or the use of any sort of mental symbols, and the content or meaning of words and other symbols derives from the content of mental states. Such theorists reject CRS on any understanding of it.

---

[1] *Solipsistic* theories, according to which the *only* relevant conceptual role is inference (or other purely internal relations), are also special versions of CRS. We believe, however, that the only plausible versions of CRS do not restrict the relevant conceptual role to wholly external or wholly internal aspects of conceptual role.

It is important to emphasize something from the start. CRS supposes that in some sense meaning or content is determined by (and so *supervenes* on) conceptual role, but that does not imply that meaning and conceptual role are the same thing. Nor does it imply that any difference in conceptual role entails a difference in meaning. For example, to the extent that "giving the meaning" of an expression is providing a paraphrase or translation of the expression, CRS implies that the adequacy of such a translation or paraphrase is determined by the way expressions in the relevant languages are used in thought. CRS does not imply that any difference in relevant usage automatically calls for a difference in translation. (We return to this point in Section 5.2 below.)

There are three broadly different ways in which symbols can be used—in communication, in speech acts like promising that go beyond mere communication, and in thinking. CRS takes the last of these uses, the use of symbols in thought, to be the most basic and important use for determining the content of symbols, where that use includes (at least) perceptual representation, recognition of implications, modeling, inference, labeling, categorization, theorizing, planning, and control of action.

In one view (e.g., Katz, 1966), linguistic expressions are used mainly for purposes of communication and do not have a significant use in thought. In this view, the content of linguistic expressions derives from the content of the non-linguistic thoughts they express and CRS is relevant to language only to the extent that it provides the correct story about the contents of non-linguistic thoughts. In a contrasting view (e.g., Sellars, 1969), ordinary linguistic communication involves "thinking out loud," people sometimes think in language (but not only in language), and the use of language in thought determines meaning. In the latter view, CRS applies directly to expressions in natural language as well as to other symbols used in thinking.

CRS need not claim that the content of all expressions is determined by their use. (Indeed, CRS does not claim that all expressions of a language have functions or uses. Such a claim would be very implausible for very long expressions that never occur.) Many conceptual role or use theorists (e.g., Ryle, 1961; Peacocke, 1992) claim that the contents of simple expressions, such as words, are determined by their conceptual roles, and that the contents of complex expressions, such as sentences, are determined by the contents of their components and the way in which they are combined (see section 2.4).

More precisely, then, CRS holds that meaning and content (including the meanings of words and other symbols and the contents of mental representations) arise from and are explained by the role words, symbols, and other features of representation play in thinking of various sorts. CRS seeks to describe the relevant sorts of conceptual role and to explain how conceptual roles determine meaning and content.

In the first few sections, we examine CRS's treatment of a few fundamental issues (section 2), consider diverse examples of ways in which representations are used in thought, (section 3), and discuss how to investigate

the relevance of conceptual role to content (section 4). Next, in section 5, we turn to information-based versions of CRS and the challenge that they pose to versions that recognize other aspects of conceptual role. Finally, we consider a number of important objections to CRS.

2. Understanding Meaning

2.1 Understanding Oneself

According to one plausible version of CRS, the basic understanding one has of the meaning of one's own words and expressions consists in one's being at home with one's use of those words and expressions. It is a kind of know-how: one knows how to proceed. One can have that basic kind of knowledge of meaning without having any sort of theoretical understanding of meaning and without being able to say what is meant in any interesting way.

We believe that a correct account of this sort of knowledge must reject the popular but obscure metaphor in which basic understanding of meaning involves "grasping" something, as if such understanding consisted in getting one's mental hands around something (Frege, 1982; Dummett, 1991; Peacocke, 1999; Fodor, 2004). According to our understanding of CRS, although one's meaning is determined by and explained by the way one uses words and other basic symbols, one's understanding of one's own meaning need not consist in having an understanding of the way one uses these items. (Nor need it consist in having an understanding of truth conditions or anything else.) It might consist simply in having symbols with the relevant conceptual roles.

2.2 Understanding Someone Else

Some CRS theorists (Sellars, 1962; Quine, 1953, 1960; Davidson, 1973; Field, 2001) suggest that to understand the meaning of an expression built from resources that one does not use oneself, one seeks to find a paraphrase or translation into an expression built from resources one does use.[2]

---

[2] Some versions of CRS give a prominent place to the notion of translation (e.g., Quine, 1953; Harman, 1990). There are at least two distinct ways in which translation can figure in such theories. First, the notion of translation can be used to address questions about meaning statements (see the text below) or about the nature of symbols' meanings or contents. Second, as we discuss in section 4.2, considering translation can be a way of investigating the way in which conceptual role maps onto, or determines, content.

This might suggest treating "'Nichts' means *nothing*" to a first approximation as a variant of "'Nichts' is best translated into my system as 'nothing'." Let us call proposals that try to explain meaning statements (of the form '*e* means *m*') in terms of translation *translational accounts of meaning statements*. Various worries might be raised about such accounts. It might be objected that the suggested treatment can be shown to fail by comparing the translations into French of "'Nichts' means *nothing*" and "'Nichts' is best translated into my system as 'nothing'''. Sellars (1962) responds by rephrasing the proposal using "dot-quotation," where "·nothing·" is used to specify a type of expression that can appear in any language, categorized by its use in its language. Field (2001) notes that ordinary quotation often functions like Sellars' dot quotation (see also Recanati, 2001, p.641). We will not here try to decide whether translational accounts might provide an adequate treatment of meaning statements. CRS is compatible with such proposals even if not committed to them.

2.3 Meaningfulness

Barry Stroud has observed (personal communication) that there is an ambiguity in the remark pronounced "'Nichts' means nothing" between the claim that 'Nichts' has no meaning and the claim that 'Nichts' has a meaning and its meaning is *nothing*. In philosophical writing it is customary to use an italic font for the second interpretation, according to which the word is used to mention its meaning rather than to express that meaning, and a regular font for the first, as we have done above. In this chapter, we sometimes use italics in this way, to mention a meaning or content, and sometimes use it to mention an expression. Context will make it clear which role the italic font is playing.

We noted above that translation can be used to give an account of meaning statements. It also provides a sufficient condition for meaningfulness. If an expression *e* has an adequate translation into something meaningful in one's own system, *e* is a meaningful expression. Davidson (1974) appears to argue for the converse claim that an expression in another language is meaningful *only if* it has such a translation into one's own system. But even if the notion of translation provides the best account of meaning *statements* (of the form "*e* means *m*"), it does not follow that we must identify an expression's *being meaningful* with its having such a translation, and we should not do so, according to CRS. It is consistent with a translational account of meaning statements to hold that an expression in another system is meaningful in virtue of its conceptual role in that system even if nothing has a corresponding role in one's own system.

2.4 Compositionality

It is widely assumed that meaning is compositional in the sense that the meaning of a compound expression is composed of the meanings of its parts and the way they are put together. Fodor and Lepore (2002) argue that CRS cannot accept such compositionality, because the use of a complex expression is not composed from the uses of its parts and the way they are put together. (This is obvious, if Ryle is right in arguing that only simple expressions have uses.) As we have emphasized, however, CRS need not identify meaning with conceptual role, nor need it hold that the contents of complex expressions are determined by their uses.

Once this point is recognized, compositionality presents no obstacle. For example, CRS can certainly allow that the *translation* of a compound expression is composed from the translation of its parts. Versions of CRS that accept a translational theory of meaning statements can therefore allow a form of compositionality of meaning. So can any other version of CRS that supposes that meanings of simple expressions are determined by, but not identical to, conceptual role. On such a view, the meanings of complex expressions are neither determined by nor identical to their uses, but are derived from the meanings of the simple expressions of which they are composed.

3. Examples of Uses of Symbols

We now describe some noncommunicative uses of representations in maps, gauges, models and diagrams, mathematical calculations and other sorts of problem solving, lists, labels and naming, categorization of various sorts, inference, and planning. We also consider what features of these uses might be especially relevant to meaning or content, an issue we take up further in the following section.

3.1 Maps

Maps are used to communicate information about geographical areas but also perhaps more importantly in thinking about the geography of an area. A person might use a map in planning what route to take in order to get somewhere, perhaps drawing a line to sketch a possible route, maybe erasing it and trying another, in this way thinking by marking up the map. People use a map in order to get clear about relative locations or to estimate distances. Some people construct their own rough maps in order to get clearer about where things are, as a way of putting together various things they know. In this way people use printed maps on paper, maps on computer screens, and also internal "mental maps." Dropping breadcrumbs in order to indicate the way home is another way of using symbols to represent geographical features.

CRS might speculate that the representational content of maps is partly a function of ways in which maps are constructed on the basis of features of areas mapped and partly a function of the ways in which maps are used in planning routes and the like.

3.2 Gauges

A driver uses a fuel gauge in order to make sure there is enough gas in the gas tank. The driver uses a speedometer to tell how fast the car is going, perhaps to avoid a speeding ticket. People check thermometers in order to tell how hot an oven is or what it's like outside.

People also have internal gauges that indicate via hunger and thirst when they need food or drink. Sensations of pain function to indicate that parts of their bodies are suffering harm.

CRS might suggest here that what is indicated by the value of a certain feature of a gauge depends both on what the values of the feature normally depend on and how one reacts to various values of that feature. So, for example, hunger and thirst differ in content in that hunger normally arises from lack of food and normally produces the goal of eating, whereas thirst normally arises from dehydration and normally produces the goal of drinking.

3.3 Models and Diagrams

People use models and diagrams to help in planning marching band formations, football plays, battles, and seating arrangements. These are sometimes three-dimensional wooden constructions, sometimes sketches in pencil on paper, and sometimes internal "mental models." The spatial relations of marks can serve to represent other, non-spatial relations in a way that greatly aids thinking about those relations. Flow charts, pie charts, graphs, and Venn diagrams are examples.

CRS might suggest that the content of such models and diagrams derives in part from the role they play in planning. What makes a certain figure the representation of a band member, for example, might in part derive from the way the model in which it is a part is used to plan a marching band formation. What makes a certain rectangular piece of cardboard represent a desk might in part derive from its use in planning where a desk should go.

3.4 Mathematical Reasoning

People use representations of numbers to count and measure, calculate costs, balance checkbooks, and solve other problems. They do mathematics on paper and in their heads.

What makes certain figures stand for amounts of money in a bank account might depend partly on how the amounts relate to various transactions involving that account. What makes certain mathematical symbols stand for mathematical addition or exponentiation or integration might be in part what are taken to be good calculations involving those symbols. Learning the meanings of such symbols might in part depend on learning how to use them in mathematical reasoning. (Of course, representations of numbers can also be used in a wide variety of other ways, for example as memory aids and passwords.)

3.5 Lists

People make shopping lists and other "to do" lists. They keep diaries and schedules of appointments. They make lists of who to invite to parties. Their lists can be on paper and in the mind.

People solve problems or crimes by listing initial possibilities and ruling as many out as they can. Some puzzles can be solved in their heads, others require writing things down on paper. In trying to decide what to do, people make lists of considerations, trying to correlate those supporting one decision with others supporting another decision, so that the considerations can be crossed off, leaving easier problems.

CRS might suggest that what makes these things lists is at least in part that they are used in such ways. In support of this, notice that it is not the case that every sequence of representations is a list. For example, a sentence or a mathematical proof is not a list.

3.6 Envisioning Possibilities

In planning and related thinking people form representations of various possible scenarios, anticipating in their imagination or in some more external way what can happen and how others may react. It is possible that the conceptual role of certain terms includes the use of such terms in processes that model various possibilities and reasoning with these models. Relevant terms might include logical constants, modals (alethic, normative, and epistemic) like *may*, *might*, *can*, *must*, and *ought*, concepts of knowledge, possibility, and necessity, moral concepts, etc.

CRS might suggest that what makes representations representations of possibilities rather than representations of how one takes the world to be (beliefs) is the distinctive way in which the representations are formed and how they function in further thinking.

3.7 Reasoning and Implication

Reasoning often includes recognition of implications. How these deductive relations affect inference is not straightforward (Harman, 1995).

CRS might suggest that the meanings of certain terms is due in part to the role these terms play in recognizing implications. So, for example, CRS might suppose that a construction $C(\cdot,\cdot)$ functions as logical conjunction (*and*) for a person if and only if the person recognizes that $C(P,Q)$ immediately implies both P and Q and is immediately implied by them taken together. Similarly, CRS might suppose that the meanings of certain terms is connected to the recognition of certain immediate inconsistencies, so that one thing that makes a construction $N(\cdot)$ represent negation (*not*) is that P is treated as immediately inconsistent with $N(P)$. (Harman, 1986, appeals to notions of psychologically "immediate" implication and inconsistency. Peacocke, 1992, appeals to "primitively compelling transitions.")

3.8 Mental Models

People use representations to think about implications. Given some information, they draw physical or mental pictures from which they read off further information. In determining what follows from assumptions people form mental models of possibilities, using the assumptions to eliminate possibilities and conclude that what is implied is what is true in the remaining possibilities (Johnson-Laird and Byrne, 1991).

People have mental models of how things work. Their models of how thermostats in refrigerators work influence what they do in order to adjust their temperatures (Kempton, 1987; Norman, 1988).

People reason to causes and other explanations by envisioning possible causes, perhaps using complex mental models of possible causes. People reason by analogy, using a model of one area, such as the flow of water through pipes, to form a model in another area, such as the "flow" of electricity "through" wires (Holyoak and Thagard, 1994; Lakoff and Johnson, 1980). Such uses of symbols in figuring out how systems work or what the cause of an event is may in part determine what the symbols represent.

3.9 Labels

People put marks or labels on things in order to recognize them later. At the gym, one puts a colorful label on a lock in order to distinguish it from the other locks in the locker room. Walking in the woods, one puts a mark on a tree to recognize it as the tree at which one turned left. Labeling an object allows a way to refer to it later: it's the one with the label. One might even use an actual feature of an object as a label: it's the tree with the distinctively

broken branch; its broken branch functions for one as a label. Numerals provide a common way of labeling many kinds of things: houses, contestants in sports events, automobiles, guns, complaints. Once items have been labeled, the labels can be used to manipulate objects for a variety of purposes. The letters and numerals on keyboards, telephones, and combination locks, and the icons on computer screens are some examples.

This sort of label – an identifying label – is used to mark a particular individual item. Proper names like *Peter*, *Chicago*, and *The Spirit of St. Louis* are also used as labels of that sort. Strawson (1974) discusses the functions such proper names or labels can have.

People also use labels to classify or categorize items. A bottle might be given a label with a scull and cross-bones on it to indicate that it contains a poison. Otherwise identical looking shakers might be labeled to indicate whether they contain salt, pepper, or sugar. Color-coding is a common way of labeling objects, such as files, in order to be able to classify them quickly. Items of clothing, such as uniforms, badges, priests' collars, or blind persons' canes, can function as labels to classify people. And the shape and color of road signs can serve as symbols to indicate the types of signs. Common nouns are sometimes used as labels of this second type. (We say more about classifying uses of labels in the next section.)

Although a label or name might be used to label several different items, identifying labels and proper names are not so used to classify the items as similar in some interesting way. There is no implication that the various people named *Peter* are similar except in having that name. The other sort of labels and common names are used to classify things as similar in certain respects. In this respect a proper name like *Sam* that is used as a name of several different people is multiply ambiguous in a way that a common name like *person*, which applies to any person, is not.

CRS might suggest that the two sorts of labels and names are semantically distinguished in part by these ways in which one uses them, how names or labels are assigned and how we use the labels in negotiating the environment.

3.10 Categorization

As we remarked in the previous section, common names and labels can be used to categorize things in various ways. Labels can be used as warnings: "poison," "flammable," "soft-shoulder." Or to indicate an "exit". Traffic signs indicate directions and distances to desired goals, gas stations, rest areas, diners. The content of such signs and labels is indexical, indicating that this is poison or flammable, that this road has a soft-shoulder, that this points the way to the exit, etc.

CRS might suppose that what gives content to a categorization of something as poisonous is in part that assigning this category to something enables one to treat it in an appropriate and safe way, and similarly for other danger categories. CRS might also suggest that what gives content to the

categorization of something as an exit is in part the use of such a categorization to enable a driver to use the exit as an exit by leaving the highway and similarly for other traffic signs.

People categorize certain geographical features of their environments as hills, mountains, rivers, lakes, fields, forests, plains, and so forth. CRS might suppose that these categorizations function in planning and practical reasoning in part by helping one get around in the world.

Symbols for categories of individual items have roles that are different in certain respects from symbols for categories of materials, substances, and stuff, as is indicated by different ways we use count nouns like *cat* and *mountain* and mass terms like *water* and *dirt* (Quine, 1960.)

People also categorize living things in various ways, as one or another type of plant or animal. CRS might suggest that this sort of categorization plays a role within a proto-biology of the natural world, according to which cats are animals, that are similar in their internal make-up, with similar organs arranged similarly, with similar more or less specified behaviors, this proto-biology helping to guide behavior in interactions with cats and other living things.

Sometimes people seem to categorize things in terms of function, artifacts like knives, watches, and pencils, for example. Parts of artifacts are also often categorized functionally, for example, the steering wheel and brakes of a car. CRS might suppose that the content of such categorizations depends in part on the way they facilitate the appropriate use of such artifacts.

Parts of living things are often categorized functionally, for example, eyes, hearts, and lungs. People are classified functionally as having certain occupations, as doctors, farmers, soldiers, teachers, and burglars. Such functional categorizations facilitate understanding of what things do and how they work.

Functional categorizations connect with evaluation and CRS might treat such connections as important to the meanings of the functional categories and the evaluational concepts. A good *X* is an *X* that functions well. Good eyes are good for seeing. A good knife cuts well. A car's brakes are good if they enable the car to stop quickly. A good safe-cracker is quick and quiet at getting a safe open. There is something wrong with an *X* that does not function well. An *X* ought to function in a certain way. There is something wrong with a teacher whose students do not learn. A bad farmer does not do well at farming. These same "conceptual connections" apply also to evaluations of people as people: a coward is not a full or good person, for example. Of course, it is less clear in the moral case how to treat being a person as a functional role. (We note some complications about functional classifications in the next section, below.)

4. Investigating Conceptual Role

We have now described some of the ways we use representations to think with. (We will mention others as we go along.) CRS is concerned with the various roles that aspects of our representations play in such thinking, and maintains that the content of those representations is determined by these roles.

4.1 Possibility Test

One way to investigate the contribution of use to meaning is to consider how a thinker describes certain imaginary possibilities. For example, one aspect of Mabel's use of concepts is her firm belief that all cats are animals. Other aspects include her firm beliefs that there are cats now, there have been cats in the past, and there will be cats in the future. Another aspect is the way she applies the concept *cat* to particular things. In order to assess the relative importance of these aspects of Mabel's use we might ask her how she would describe the imaginary discovery that all the things that people like Mabel have ever called *cats* are really radio-controlled robots from Mars (Putnam 1962, Unger 1984). Her saying, "That would be to discover that cats are not really animals," would be evidence that her firm belief that cats are animals may not be as important to the content of her concept of a cat as other aspects of her use. In this kind of example, the way that the thinker describes certain imaginary possibilities is itself an aspect of the thinker's use of the concept. Rather than putting the point in terms of evidence, we could say that the way in which Mabel describes certain imagined cases plausibly makes a certain contribution to the content of her concept. We do not consider here the different question of the possible relevance to content of what Mabel would do if the imagined cases became actual.

A similar issue arises about Mabel's concept of a witch. Mabel applies this concept to various people and also accepts some general views about witchcraft, including the view that witches have magical powers of certain specified sorts. We can ask Mabel how she would describe the possible discovery that no one has the relevant magical powers. Would she describe this as showing that there are no witches or as showing that witches do not after all have magical powers? If Mabel says that this sort of discovery would show that there are no witches, that is some evidence that her acceptance of the general views is more important to the content of her concept of a witch than her judgments that various people are witches.

In sum, her characterization of such imagined cases might show that Mabel's acceptance of certain theoretical assumptions is more central to the content of her concept of a witch than it is to the content of her concept of a cat.

4.2 Translation Test

To know the meaning of someone else's words is often to know how to translate them into your language, and to understand what an experience is like for another person or what it is like to be that person is often to know how to translate that person's outlook into yours. So, another way in which CRS might study how conceptual role determines meaning is to see how it might determine good translation. This too is a useful heuristic.

4.2.1 Color Concepts

If Mabel applies certain words to objects on the basis of perception in ways that match your applications of your color terminology, that may be a reason to translate Mabel's words into your corresponding color words. If a bear's color perception works similarly to that of humans, allowing bears to make discriminations of color of the sort that humans make, that may be a reason to "translate" their color experience into ours—that is, to understand them as seeing colors much as we do. To the extent that a rabbit's color perception works in some other way, perhaps enabling the creature to make different sorts of discriminations between objects from the ones we make, it may be hard to translate rabbits' experience into ours and hard to gain understanding of how things look to them. Since there are such differences even among people, who may have one or another form of color blindness, or may be totally blind, a similar point holds there also. A congenitally blind person may have at best a very impoverished understanding of what perception of color is like for someone with normal human color perception.

What about the color words used by a congenitally blind person who relies on others for information about color? One kind of CRS might interpret the blind person's use of 'red' as meaning something like *having the perceptual feature that sighted members of my community call 'red'*. But CRS need not take this position. A different version of CRS might hold that the blind person's conceptual role for 'red', though different from a sighted person's, nevertheless manages (through reliance on sighted members of the community) to determine the same content had by sighted persons' 'red'.

What about someone who has normal color perception and terminology at one time but then loses color vision? CRS may be able to allow that the person still remembers how red looks. Perhaps CRS would understand this as a case in which the conceptual roles are still there but are blocked, as in a sighted person wearing a blindfold.

Some versions of CRS assume that there is a non-conceptual content of mental states that is not determined by considerations of conceptual (or functional role) (Block, 1998; Peacocke, 1983). Other versions of CRS claim to apply to all aspects of the phenomenal content of mental states. Consider a possible interpretation of Mabel's visual experience that attributes an inverted spectrum to her. This interprets the experience Mabel has looking at something red as like the experience you have when you are looking at something green, and similarly for other colors. Without special reasons for

13

such an interpretation, a CRS that aims to explain phenomenal content would speak against it, holding that, if color concepts and words are functioning in the same way for both Mabel and you with respect to the external colors of objects, that contributes to making it the case that the non-inverted interpretation is the correct one.

There might be a consideration on the other side if Mabel's internal mechanisms were somehow inverted, so that what happens internally when Mabel sees red is like what happens internally when you see green, where the differences in internal mechanisms constitute differences in the internal use of symbols. Or this might not be relevant. If you accept this sort of CRS, you might approach this issue by trying to determine what would make for the best translation between Mabel's mental life and your own.

4.2.2 Moral Concepts

Consider a different sort of case, the interpretation of moral thinking and terminology of people in a different culture, call them *the Amarras*. Imagine (Dreier, 1990) that the Amarras make two contrasts, using the words *ret* and *wreng* for one contrast and *rit* and *wrig* for the other. The things the Amarras take to be *ret* are of the sort that you and other people in your society tend to consider morally right and the things the Amarras take to be *wreng* are of the sort you and yours tend to consider morally wrong. However, the Amarras do not take themselves to have reasons to be motivated toward what they take to be *ret* and do not take themselves to have reasons to be motivated to avoid what they take to be *wreng*. On the other hand, the Amarras do take themselves to have reasons to be motivated toward what they call *rit* and to avoid what they call *wrig*, although what they consider *rit* and *wrig* are quite different from what you and yours consider right and wrong, respectively.

How should you interpret their words *ret*, *wreng*, *rit*, and *wrig*? Which best correspond to your 'right' and 'wrong'? Should you translate them as agreeing with you about what is right and wrong while lacking your interest in doing what is right? Or should you translate them as thinking that different things are right or wrong from you? Suppose the latter option is better, so that *rit* and *wrig* are better translated as *right* and *wrong* than are *ret* and *wreng*. CRS can use that as an indication that the connection with motivational reasons is an important aspect of the meaning of moral terms like *right* and *wrong*.

Would this conclusion imply that people cannot believe certain things are right and wrong without being motivated to do what is right? What about someone who uses moral concepts and terminology in your way for years but eventually decides that morality is bunk and loses the motivations? And what about psychopaths who lack the sort of human sympathy that seems important for moral motivation (Blair, 1995)?

For CRS, such issues are similar to those that arise about the color concepts of non-normal perceivers and similar methods might deal with them.

For example, Hare (1952) suggests that a moral sceptic's use of moral terminology might be such that the sceptic's 'good' is best interpreted as *the sort of thing you call 'good'*. On the other hand, a normative conceptual role theory (Greenberg 2005) could hold that the fact that a sceptic *ought* to have the relevant motivations makes it the case that the sceptic's 'good' has the same content as others' 'good'.

4.2.3 More or Less Functional Concepts

Suppose you are trying to determine the meaning of a symbol T in Zeke's thought. Zeke tends not to apply T to something unless it has the function of collecting dust, crumbs, or other relatively small particles or objects from floors or other surfaces. This observation may suggest the hypothesis that T should be translated as 'broom'. However, Zeke uses T for anything that has that function, regardless of its construction or composition. For example, Zeke uses T for vacuum cleaners and sticky sheets of paper that are used to pick up dust. If you are inclined to conclude that T does not mean *broom*, that would indicate that your concept of a broom is not a purely functional concept.

By contrast, suppose Zeke tends not to apply U to an object unless the object has the function of slowing or stopping the system of which it is a part. This observation raises the hypothesis that U means *brake*. Also, Zeke uses U for anything that has that function regardless of its construction or composition. For example, Zeke uses U for tennis shoes when they are given the function of slowing bicycles and for electromagnetic fields when they are given the function of slowing space ships. If you are inclined to think that this aspect of U's conceptual role does not undermine the hypothesis that U means *brake*, that would suggest that your concept of a brake is more of a functional concept than your concept of a broom.

There seems to be a spectrum of artifact concepts from predominantly functional ones, of which *brake* or *clock* may be examples, to concepts that are not only functional but have additional aspects. Although something must have a certain purpose in order to count as a typewriter, a drill, or a stapler, not just anything with that purpose is a typewriter, a drill, or a stapler. For some concepts, composition or construction seems to matter. For others, history is important. For example, arguably a musical instrument that is very like an oboe doesn't count as an oboe if it was independently developed by Australian aborigines. It is not part of the historical family of oboes. Thus, if Oscar uses a term for all oboe-like musical instruments, that term does not mean *oboe*.

In this section, we have illustrated how one can investigate conceptual role by considering imaginary possibilities and by asking how to translate expressions. A remaining question for CRS is whether it is possible (and if so, how) to give a systematic account of what determines which aspects of conceptual role are relevant to content and what their precise relevance is.

5. Limited Versions of CRS: Indication

In Section 3., above, we discussed a variety of ways in which symbols are used. We have also mentioned that some CRS theorists propose to limit the relevant factors to perceptual input, inner mental processes, and output in the form of action. The first and third of these are concerned with relations between symbols and the world, the middle is concerned with relations of symbols to each other. This might cover all the cases we discussed.

In this section, we turn to special versions of CRS that restrict the relevant conceptual role to the first of the three factors, namely perceptual input. Verification theories of meaning (e.g., Ayer, 1936; Quine, 1960) are an historical example of such a restricted CRS. We will be concerned with the more recent information-based or indication theories (Dretske, 1986, 2000; Fodor, 1987; 1990; Stampe, 1977). There is an active debate between such theories and CCRS (the kind of CRS that holds that inner uses are essential to determining content). We will suggest that information-based theories encounter a range of difficulties that push them to include inferential relations and actions in the relevant conceptual role.

5.1 Information-Based Theories

Information-based theories hold that the content of a symbol depends only on the information about the environment carried by an internal tokening of the symbol. So, an internal occurrence of a token of 'red' indicates or carries the information that there is something red in the environment, where such indication might be analyzed as a kind of counterfactual, causal, or nomic dependence.

One problem for such views is that it is difficult for them to do without intentional notions such as the *application* of a symbol to an object, i.e., using a symbol with the intention to characterize an object as falling under it, as in "That's Bill" or "That's a cow" (see Greenberg, 2001). The straightforward way to give an information-based account is to say, roughly speaking, that a symbol has the content of the property whose instantiations normally or optimally covary with the symbol's *application* (e.g., Boghossian, 1989). Many things other than water – deserts, *thoughts* of deserts – may covary with the *occurrence* of my mental symbol for water. But, leaving aside mistakes, only water covaries with the *application* of the relevant mental symbol.[3] The problem is that the notion of an application of a

---

[3] We here ignore the different problem for information-based theories of what makes it the case that a symbol means *water* rather than, for example, certain patterns of nerve cell stimulations, or some other more proximal or distal correlate of the

symbol is plainly an intentional notion that, at least on the face of it, needs to be explained in terms of internal aspects of the use of symbols such as the intentions or other mental states that cause the occurrence of the symbol.

A different problem for standard informational theories is that they have fewer resources than other versions of CRS for dealing with such problems as necessarily co-referring expressions and necessarily co-instantiated properties. For example, an informational theory cannot appeal to inferential or implicational considerations to distinguish the concept of a unicorn from the concept of a gremlin (assuming unicornhood and gremlinhood are both necessarily empty). And, similarly, an informational theory cannot appeal to a concept's role in reasoning to solve the problem (Quine, 1960) of whether the concept refers to rabbits, undetached rabbit parts, or temporal stages of rabbits.

It is also natural to appeal to internal aspects of conceptual role to address the problem that not everything that carries information has meaning or content. For example, for a creature to have a concept of, say, red, it is not enough that there be some state or condition of the creature whose instances or tokens carry the information that there is something red in the environment. The relevant tokens must figure appropriately in the creature's psychology. In response to this kind of problem, some theorists move away from pure information-based accounts by taking into account how the internal tokening of a symbol relates to other internal states in a way that might affect how the creature acts to satisfy its needs (Stalnaker, 1984, pp. 18-19; Dretske, 1986, 2000; Fodor, 1990, p. 130).

There are other issues on which even the purest information-based theories tend to appeal to internal aspects of conceptual role. For example, Fodor (1998, p. 35, 163-165; Fodor and Lepore, 2002, pp. 18-22) holds that what makes it the case that a complex symbol – one that is composed of other symbols arranged in a certain way – expresses a particular concept is the symbol's relations to the simple symbols of which it is composed. Another example is that it is difficult to see how to give an account of the content of logical constants without appeal to internal relations (Fodor, 1990, pp. 110-111).

Fodor's (1990) asymmetric-dependence theory, perhaps the best-known informational theory, attempts to deal with some of the problems discussed in this section,[4] but it has generated a battery of objections (Loewer and Rey 1991) and few if any adherents, and we think it is fatally flawed (Greenberg 2001).

Fodor and Lepore have argued that CRS must give up the extra resources available to versions of CRS that are not purely informational; we criticize

---

symbol's occurrence.

[4] Fodor's asymmetric dependence theory is designed to do without the notion of an application of a concept (see Fodor, 1990, 89-131). For Fodor on co-extensive and co-instantiated symbols, see his 1994, pp. 39-79; also his 1990, pp. 100-101; 1998, pp. 163-165.

this argument in the next section. (A terminological caution: Fodor and Lepore use the term "conceptual role semantics" or "inferential role semantics" for (roughly) the views that we are calling "CCRS"; thus, in their terminology, information-based theories of contents are rivals to conceptual role theories, rather than as in our terminology special versions of them.)

5.2 Fodor and Lepore's Dilemma

We have so far argued that it is not easy to see how meaning or content could be explained entirely in terms of information or indication without appeal to internal uses of terms. In other words, it is hard to see how CRS can avoid being CCRS.

We now consider an argument by Fodor and Lepore that is supposed to provide a threshold objection to any form of CCRS (Fodor 2000; Fodor, 1998, chap. 4; 1990, pp. ix-xi; Fodor and Lepore, 1992). (See also the discussion of this argument in sections 3-4 of the Meaning Holism chapter in this volume.) Fodor and Lepore begin by assuming plausibly that that no two people accept exactly the same internal inferences and implications. Given that assumption, they argue that CCRS faces the following dilemma. Either

    (a) every such internal aspect of the way one uses one's terms is relevant to the terms' content, or

    (b) only some such internal aspects are relevant.

If (a), according to Fodor and Lepore it follows that no two people ever mean the same thing by any of their terms (or ever have thoughts with the same contents). This conclusion, they maintain, has the following implications, which they take to be absurd:

    (c1) that no two people can ever agree or disagree with each other about anything

    (c2) that intentional explanation collapses since no two people ever fall under the same intentional laws.

    If (b), according to Fodor and Lepore it follows that CRS is committed to the analytic-synthetic distinction, a distinction that (according to them) has been decisively undermined by Quine.

    However, Fodor and Lepore's presentation of their alleged dilemma is flawed. Consider their argument if horn (a) of the dilemma is chosen. That argument rests on, among other things, the following assumption:

    (aa) that, if all aspects of internal use are relevant to meaning and the aspects of one person's internal use are not exactly the same as those of another person's, then the two people do not mean the same thing by their terms, and

    Assumption (aa) is indefensible because, as we emphasized at the beginning of this chapter, even if all aspects of internal use are relevant to meaning, there can be differences in such use without a corresponding difference in meaning. To say that a given aspect of internal use is relevant to meaning is to say that there is a possible case in which a difference in that

aspect makes for a difference in meaning, not to say that a difference in that aspect always makes for a difference in meaning. (Similarly, whether the number of students in a class is odd or even depends on the number of students in the class, but that does not imply that two classes with different numbers of students cannot both have an even number of students.)

In response, Fodor and Lepore might try to argue that no plausible version of CCRS has the consequence that differences in the determinants of content do not imply differences in content. But such a response would require consideration of the merits of different possible versions of CCRS; the point we have made here is that Fodor and Lepore have failed in their attempt to offer an in-principle threshold objection to all versions of CCRS.

It is also worth noting that two people who mean different things by their terms can still use those terms to agree or disagree with each other. Mary can disagree with John by saying something that they both know is true only if what John said is false. Mary can agree with John by saying something that they both know is true only if what John said is true. To take a very simple example, suppose that Mary and John do not mean exactly the same thing by their color terms in that the boundaries between what counts for them as *red* and *orange* are slightly different and the boundaries between what counts for them as *green* and *blue* are slightly different. Still, they disagree about a color when John calls it *red* and Mary calls it *green*.

The claim that intentional explanation collapses if no two people have the same contents can also be disputed. It may be that intentional explanation requires only a notion of similarity of content (Harman, 1973; 1993; Block, 1986). Fodor (1998, 30-34) has objected that, according to CRS, to have *similar* content is to be related to at least many of the *same* contents, which presupposes sameness of content. But CRS is not in fact committed to any such account of similarity of content.

Thus, horn (a) of the alleged dilemma for CCRS is harmless.

According to horn (b) of the alleged dilemma, the claim that only some aspects of internal conceptual role are relevant to meaning commits the CCRS theorist to an analytic-synthetic distinction of a sort that Quine is supposed to have shown to be untenable.

We have three things to say about this horn. First, there are coherent versions of CCRS that do not accept an analytic-synthetic distinction yet take some but not all aspects of internal conceptual role to be relevant to meaning. As we have observed in discussing (aa), from the claim that a given aspect of conceptual role, a certain belief for example, is part of what determines that a symbol has a given meaning, it does not follow that someone without the belief cannot have a symbol with the same meaning. Thus, the belief's relevance to the meaning of the symbol does not imply that the belief is analytic. (See also section 4 of the Meaning Holism chapter.)

Second, various distinctions may qualify as some kind of analytic-synthetic distinction. Whether Quine's (or others') arguments undermine the particular distinction to which a given CCRS is committed depends on the

details of each case. (See Rey, 1993; 1995 for discussion.) For example, Peacocke (2002) has made out a strong case that Quine's arguments do not apply to the particular kind of analytic-synthetic distinction to which Peacocke's (1992) version of CCRS is committed. Similarly, Fodor's own informational theory of content is committed to conceptual truth, though arguably not to an objectionable version of the analytic-synthetic distinction (e.g., Fodor, 1998, p. 14 & fn. 8).

Third, Quine's attack is aimed at a traditional notion of analyticity according to which analytic truths are *a priori*. But CCRS need not accept that knowledge of conceptual role is *a priori*. As we noted above (section 2.1), a thinker can have a symbol with a certain conceptual role without having a theoretical understanding of how she uses the symbol.

We conclude that Fodor and Lepore have not yet refuted CCRS.

6. Further Objections to CRS

According to CRS, conceptual role determines and explains content. Searle, (1980, 1992) vigorously argues for the opposite view. Searle argues that mental states have intrinsic content that explains and is not explained by the conceptual roles such states have in thinking. Other symbols have derivative content by virtue of having some relation to the intrinsic content of mental states. Linguistic representations are used to express people's thoughts. States of a computer program have derived content through people interpreting them as having content. A translation of a term into another language is good to the extent that the translation expresses an idea with the same intrinsic content as the idea expressed by the term being translated. Although we can appeal to linguistic use in assessing translations, that is not because use determines content but because content determines use, in Searle's view.

CRS denies that an explanation of conceptual role by appeal to intrinsic content has any force unless it reduces to some version of CRS. Perhaps explanations of particular occasions of the use of a mental symbol E will invoke m, the content of the symbol. But what explains E's having content m? In order to explain why E has the role it has, Searle would have to explain why it has content m, but his appeal to intrinsic intentionality has no resources to explain it (though he thinks that biology may ultimately be able to explain intrinsic intentionality). In particular, what is wanted is an explanation of why something has a particular content that also accounts for why something with that content has a given role. CRS has an explanation of E's having content m that explains this (though, as we discuss in section 6.3, there are difficult issues about, for example, whether and how actual use can explain a term's having a certain normative role). We now consider some worries about this explanation.

6.1 Circularity Objection

One worry about the explanation provided by CRS is that it might be circular. Consider the suggestion that the meaning of logical conjunction (*and*) is determined in part by the fact that one immediately recognizes that a conjunction implies its conjuncts. Fodor (2004) objects that any such account is circular because to recognize an implication presupposes thinking of the items in the implication relation as having content.

A defender of the suggestion might respond that the relevant recognition of implications does not involve such thoughts about symbols. It is enough that one is at home in using the symbols in the relevant way. One simply and directly treats a conjunction as implying its conjuncts. (See also the discussion in section 2 of the Meaning Holism chapter.)

But in order to make such a response work it is necessary to show that the relevant conceptual roles can be specified without reference to the content of the symbols.

Peacocke (2002) offers a version of CRS that is explicitly circular in explaining aspects of conceptual role in terms of what a person is "entitled" to accept, where entitlement is a normative epistemological notion that is itself to be explained in terms of intentional content. More precisely, according to Peacocke, "there is a large circle of interrelated notions, including entitlement, knowledge, and even intentional content itself, each of whose elucidations ultimately involves the others."

A related objection is that conceptual roles are interrelated and cannot be specified in isolation from one another. A structuralist like Saussure (1916) says that one's concept of red is partly defined in terms of colors like green that are in a certain respect excluded by something's being red. Sellars (1956, section 19) writes that "one can have the concept of green only by having a whole battery of concepts of which it is one element." Similarly, Wittgenstein (1969) says, "When we first begin to *believe* anything, what we believe is not a single proposition but a whole system of propositions. (Light dawns gradually over the whole.)" How can the conceptual roles of concepts be specified if they are interdependent in this way?

One response to this problem (e.g., Peacocke, 1992, pp. 9-12) is to suppose that, where there is such interdependence, there is a system of connected conceptual roles. (Of course, in the case of color concepts, there are connections through perception to items in the environment in addition to the interconnections among those concepts.) Two people can be said to have the same color concepts to the extent that they both have systems of concepts that satisfy certain conditions. (Compare our discussion above about when people might count as having the same color concepts.)

This idea fits with Ramsey's (1931) suggestion that references to theoretical states and processes be replaced with existentially quantified variables in an overall theory. It also fits with the idea that conceptual roles

are analogous to roles played by symbols in the running of computer programs.[5]

Such analogies open CRS up to objections on various fronts. One is that, if conceptual roles can be specified in the manner suggested, then it should actually be possible to build a robot directed by a computer program in which symbols have the relevant conceptual roles. While some defenders of CRS welcome that conclusion, Searle argues that it reduces CRS to absurdity.

6.2 Chinese Room Objection

Searle (1992) summarizes his basic argument against any computationally friendly version of CRS in the slogan that syntax is not enough for semantics. However, that slogan is misleading as an objection to CRS. The idea that syntax is not enough for semantics is obviously correct if what is meant is simply that expressions with different meanings might have exactly the same syntactic form. The sentences "Jack loves Mary" and "Sue hates Allen" mean different things but have exactly the same syntactic form, say, "(N (V N))". However, that obvious point by itself is no objection to computationally friendly CRS. CRS does not make the false claim that syntax in the ordinary sense is sufficient for semantics.

Searle takes conceptual role to be a purely syntactical matter in the following sense: conceptual role is to be defined entirely in terms of operations on certain symbols without any appeal to meaning or content. Of course, as emphasized above, conceptual role can also involve using symbols in relation to non-linguistic things in the world, as in perceptual responses or in practical reasoning leading to action. So Searle must understand "purely syntactic operation" to include these cases also.

Searle's famous "Chinese Room" argument tries to show that syntax in this second sense is not sufficient for understanding. The argument has a number of different targets. For our purposes, we can treat the argument as seeking to show by example that a person can know how to use symbols and be at ease with their use without having any understanding of what they mean.

The argument begins by supposing, for the sake of *reductio*, that a given person who speaks and understands only a dialect of Chinese thinks using a system whose elements have specifiable conceptual roles. According to CRS, this speaker's understanding of Chinese consists in his or her being disposed to use and using the symbols in the right way. So, CRS is committed to thinking that any other person would have the same understanding of Chinese if the other person used those elements in the same way.

---

[5] It should be noted, however, that, as Peacocke (1992) recognizes, his account makes use of contentful notions in a way that cannot be eliminated through Ramsey's suggestion.

The argument continues as follows. We are assuming that the relevant conceptual roles are specifiable, so consider a specification of those roles. Given that specification, it would be possible in theory to construct a robot that would have a central processor running a program that would allow the robot to follow those rules. And, if that is possible, it is in theory possible to replace the central processor in the robot with a room containing a person knowing only English and so not knowing any Chinese, who nevertheless could blindly follow the rules. Although the person doing the processing might use the symbols in accordance with the rules, he or she would not understand the symbols. So, it seems that, contrary to CRS, the use of symbols in the relevant way is not sufficient for understanding the meaning of those symbols.

Searle's Chinese Room Argument has generated an enormous response (beginning with the responses to Searle, 1980, in the same issue of the journal). We will not try to summarize this response.

Instead we mention only the following possible response. It might be suggested on behalf of CRS that the role of symbols being used to simulate a person who has certain concepts (the Chinese speaker in the example) is not the same as the role of the symbols in the Chinese speaker that express the relevant concepts. The original Chinese speaker is using the symbols to think with. The person processing a simulation – *the simulator*, for short – is using the symbols to simulate someone who uses the symbols to think with. One sign of this is that the Chinese speaker does not normally think about the symbols whereas the simulator must think about them.

But can this response be developed without circularity? As formulated, the response is circular because it explains conceptual role in part by mentioning what the subject is thinking about, which is to explain conceptual role in terms of intentional content, whereas CRS seeks to do things the other way round, explaining intentional content in terms of conceptual role.

We think that the response to Searle is not circular: it is not the case that the differences in conceptual role between the Chinese speaker and the simulator show up only at the level of contentful descriptions (though that level offers a easy way of describing the differences). The problem is that Searle has given us a strong reason for thinking that the simulator's symbols, either on paper or in his mind/brain (if we assume that he fully internalizes the process) do not have the same conceptual roles as the symbols of the Chinese speaker. In particular, he tells us that the simulator has not been taught to speak and understand Chinese but has been taught to follow rules that capture the conceptual role of the Chinese speaker's symbols. The result is that the simulator's symbols should not have the same conceptual role of those of the Chinese speaker, but those of someone who is simulating the speaking of Chinese. The point is most obvious if we take the original case where the simulator is in a room and the input and output are slips of paper with marks on them. The symbols in the actual Chinese speaker's mind/brain are connected to certain perceptual states and actions. The candidate symbols

in the simulator are connected to very different perceptual states and actions (perceptions of certain slips of paper with certain figures on them coming into the room and actions of making certain marks and passing slips of paper back out).

Even if we suppose that the simulator internalizes the whole process, including the room, and simply responds to utterances in Chinese with utterances apparently in Chinese, the problem remains.[6] Whether the simulator has the same conceptual role as the Chinese speaker depends on how the connections are organized, not just on whether the inputs and outputs are the same. So Searle faces a dilemma. If, on the one hand, he stipulates that the overall conceptual role of the simulator's symbols, including their internal organization, is now identical to that of the Chinese speaker, then he no longer will be able to rely on the strong intuition that the simulator does not understand Chinese. CRS theorists can plausibly maintain that a "simulator" who can interact with Chinese speakers and the world in just the way that Chinese speakers do – and whose internal symbolic organization is the same as that of Chinese speakers – understands Chinese.[7] (Theorists who believe that conceptual role cannot explain understanding may not be convinced, but the present point is only that the Chinese Room does not provide such theorists with a refutation of CRS.)

If, on the other hand, Searle stipulates that the Chinese thinker continues to manipulate symbols according to the now-internalized rules for simulating the conceptual role of the Chinese speaker, we lack good reason to think that the overall conceptual role of the simulator's symbols is the same as the overall conceptual role of the Chinese speaker's symbols. To put it in the intuitive way again, the relations of some of the simulator's symbols (the ones that are supposed to correspond to those of the Chinese speaker) will be controlled by other symbols (the ones that specify the rules for manipulating the former symbols). We have not been given reason to think that that is precisely how the Chinese speaker's symbols are organized. (And, once again, if it is stipulated that the stipulator's internal states are organized in the same way as those of the Chinese speaker, it is plausible to claim that the simulator is no longer simulating, but speaking and understanding, Chinese.)

In sum, a set of instructions for taking symbols and manipulating them in a way that gives the simulator's symbols the same conceptual role as a Chinese speaker's mental symbols may be self-defeating. For part of the conceptual role of a Chinese speaker's mental symbols may be that they are not manipulated in accordance with that set of instructions.

---

[6] The following response applies, *mutatis mutandis*, if we suppose instead that the Chinese speaker is in the room manipulating slips of paper.

[7] As noted above, same conceptual role is not necessary for same content, but it is sufficient. Searle's argument depends on claiming that the simulator has the same conceptual roles as the Chinese speaker (and therefore that CRS entails that he has the same contents as the Chinese speaker).

6.3 Objections that Conceptual Role Is Nonfactual

A related issue about CRS is whether it is a purely factual matter what the conceptual roles of a given person's symbols are. Some symbols, for example those in a computer program that is running on a particular computer, may have their conceptual roles in virtue of facts about design. Assume that it can be a completely factual matter whether someone has designed a system so as to instantiate a particular computer program.[8] In that case, to the extent that a symbol's conceptual role is determined by design facts, its conceptual role can be a purely factual matter. But CRS is supposed to apply to the content of concepts of someone who has not been designed or programmed by anyone.[9] Can it be in the same way a matter of fact whether such a person's concepts have the relevant conceptual roles?

Suppose CRS says that a person's concepts have the relevant conceptual roles as long as the system *can be interpreted* as instantiating the relevant conceptual roles. Kalke (1969) and Searle (1992) object that there will always be a way to interpret *anything* as running any given computer program. If they are right *even taking into account relevant external relations*, this version of CRS reduces to behaviorism. But once external relations are taken into account, it is far from obvious that they are right.

Apart from that worry, an actual system may break down or wear out or not have enough capacity to carry out certain tasks it is programmed to do. What distinguishes those aspects of the system that are defects or limitations from those that are part of the program, as it were? CRS needs to distinguish those aspects of a system that reflect conceptual roles of components and those aspects that reflect processing limitations, noise, damage, and mistakes. But how does such a distinction reflect facts about the system itself (Wittgenstein, 1953; Kripke, 1982)? Such issues have spawned a large literature (e.g., Boghossian, 1989; Horwich, 1990, 1998; Lewis, 1983; Millikan, 1990; Pettit, 1990; Pietroski and Rey, 1995; Soames, 1998).[10]

A different but related issue is whether actual dispositions to use symbols in thought are the right sort of thing to determine content. Some theorists have thought that conceptual role must have a normative element (Kripke 1982). For example, we mentioned above that Peacocke's (2002) version of CRS explains some aspects of conceptual role partly in terms of conditions that

---

[8]This is to assume that it can be a factual matter what the content of a designer's intentions are.

[9]We have noted above that some theorists (e.g., Dennett, 1995; Millikan, 1984, 1993; Neander, 1995) appeal to evolution as a source of something that takes the place of design. A certain sort of learning might function similarly (Dretske, 1986, 1988). The worries in the following paragraphs may still apply.

[10] Greenberg (2001) shows that the so-called "disjunction problem" familiar from information-based theories of content (e.g., Fodor, 1990) is another way of presenting the same group of issues.

"entitle" someone to accept something, where entitlement is a normative notion. Apart from the circularity worry already discussed, one might also worry whether it could be a purely factual matter whether a certain normative condition obtains. Greenberg (2001; 2005) discusses a view that can be understood as a normative version of CRS – the view that a thought's having a certain content is in part explained not by the thinker's being disposed to use symbols in certain ways but by the thinker's being subject to standards requiring her to do so.

We will not try to answer the questions raised in this section, although we do not think they pose insuperable difficulties for CRS.

7. Summary

CRS says that the meanings of expressions of a language or other symbol system or the contents of mental states are determined and explained by the way symbols are used in thinking. According to CRS one's understanding of aspects of one's own concepts consists in knowing how to use one's symbols and being at ease with that use. Understanding expressions in other systems may involve interpreting or translating those expressions into corresponding symbols of one's own system.

Many different aspects of the way symbols are used are relevant to their meaning or content. There seem to be three main categories of uses, having to do with perceptual input, internal thinking, and output in action. Information-based or indication theories that attempt to rely only on perceptual input face difficulties that put pressure on them to rely on other aspects of conceptual role. Worries about CRS include possible circularity, how to respond to Searle's Chinese Room Argument, and whether there are facts about conceptual role. Whether these worries can be satisfactorily addressed is a matter of current debate.[11]

8. References

1. Ayer, A.J., (1936). *Language, Truth and Logic*. London: Gollancz.
2. Blair, R.J.R., (1995). "A cognitive developmental approach to morality: investigating the psychopath." *Cognition*. 57, pp. 1-29.
3. Block, N., (1986). "Advertisement for a semantics for psychology," in *Midwest Studies in Philosophy* 10: 615-678. Minneapolis, MN: University of Minnesota Press.
4. Block, N., (1998). "Is experiencing just representing?" *Philosophy and Phenomenological Research* 58.

---

[11] We are indebted to Ram Neta and James Pryor for useful comments on a prior draft.

5. Boghossian, P., (1989). "The Rule-Following Considerations". *Mind* 98: 507-549.

6. Bonjour, L., (1998). *In Defense of Pure Reason*. London: Cambridge University Press.

7. Davidson, D., (1973). "Radical interpretation." *Dialectica* 27: 313-28.

8. Davidson, D., (1974). "On the very idea of a conceptual scheme," *Proceedings and Addresses of the American Philosophical Association* 47: 5-20.

9. Dennett, D. (1995). *Darwin's Dangerous Idea*. New York: Simon and Schuster.

10. Dilthey, W., (1883). *Einleitung in die Geisteswissenschaften*. English translation by R. A. Makkreel and F. Rodi, *Introduction to the Human Sciences*. Princeton, NJ: Princeton University Press, 1989.

11. Dreier, J., (1990). "Internalism and Speaker Relativism," *Ethics* 101, pp. 6-26.

12. Dretske, F., (1986). "Misrepresentation." In Bogdan, R., editor, *Belief*. Oxford: Oxford University Press.

13. Dretske, F., (1988). *Explaining Behavior*, Cambridge, Mass: MIT Press.

14. Dretske, F., (2000). *Perception, Knowledge, and Belief: Selected Essays*. Cambridge: Cambridge University Press.

15. Dummett, M., (1991). *The Logical Basis of Metaphysics*. London: Duckworth.

16. Field, H., (2001). "Attributions of meaning and content," in H. Field, *Truth and the Absence of Fact*. Oxford: Clarendon Press. Pp. 157-174.

17. Fodor, J., (1987). *Psychosemantics*. Cambridge, MA: MIT Press.

18. Fodor, J., (1990). *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press.

19. Fodor, J., (1994). *The Elm and the Expert: Mentalese and Its Semantics*. Cambridge, MA: MIT Press.

20. Fodor, J., (1998). *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.

21. Fodor, J., (2000). *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. Cambridge, MA: MIT Press.

22. Fodor, J., (2004). "Having Concepts: A Brief Refutation of the 20th century." *Mind and Language*, 19, pp. 29-47.

23. Fodor, J. and Lepore, E., (2002). *The Compositionality Papers*. Oxford: Clarendon Press.

24. Frege, G., (1982). "Über Sinn und Bedeutung" in Zeitschrift für Philosophie und philosophische Kritik.

25. Greenberg, M., (2001). *Thoughts Without Masters: Incomplete Understanding and the Content of Mind*. (University of Oxford, D.Phil. Dissertation).

26. Greenberg, M., (2005). "A New Map of Mental Content: Putting Normativity in Its Place." Forthcoming in E. Sosa and E. Villanueva (eds.), *Philosophical Issues*, 15.

27. Hare, R. M., (1952). *The Language of Morals*. Oxford: Oxford University Press.

28. Harman, G., (1973). *Thought*. Princeton, NJ: Princeton University Press.

29. Harman, G., (1974). "Meaning and semantics," in M. K. Munitz and P. K. Unger (eds.), *Semantics and Philosophy*. New York: NYU Press. Pp. 1-16.

30. Harman, G., (1975). "Language, thought, and communication," in K. Gunderson (ed.), *Language, Mind, and Knowledge: Minnesota Studies in the Philosophy of Science* VII. Minneapolis, Minn.: University of Minnesota Press. Pp. 279-98.

31. Harman, G., (1986). *Change in View*. Cambridge, MA: MIT Press.

32. Harman, G., (1987). "(Nonsolipsistic) conceptual role semantics," in E. Lepore (ed.), *New Directions in Semantics*. London: Academic Press. Pp. 55-81.

33. Harman, G., (1990). "Immanent and transcendent approaches to the theory of meaning," in R. Gibson and R. B. Barrett, eds., *Perspectives on Quine*. Oxford: Blackwell.

34. Harman, G., (1995). "Rationality," in E. E. Smith and D. N. Osherson, eds., *Thinking: Invitation to Cognitive Science, Volume 3*. Cambridge, MA: MIT Press. Pp. 175-211.

35. Holyoak, K. J. and Thagard, P. R., (1994). *Mental Leaps: Analogy in Creative Thought*. Cambridge, MA: MIT Press.

36. Horwich, P. (1990). "Wittgenstein and Kripke on the Nature of Meaning". *Mind & Language*, 5, pp. 105-121.

37. Horwich, P. (1998). *Meaning*. Oxford: Oxford University Press.

38. Johnson-Laird, P. N. and Byrne, R. M. .J, (1991). *Deduction*. Hillsdale, NJ: Erlbaum.

39. Kalke, W., (1969). "What's wrong with Fodor and Putnam's functionalism," *Nous* 3: 83-93.

40. Katz, J. A., (1966). *The Philosophy of Language*. New York: Harper and Row.

41. Kempton, W., (1987). "Two theories of home heat control," in N. Quinn and D. Holland, eds. *Cultural Models in Language and Thought*. New York: Cambridge University Press.

42. Kripke, S. A., (1982). *Wittgenstein on Rules and Private Language : An Elementary Exposition*. Cambridge, MA: Harvard University Press.

43. Lakoff, G. and Johnson, M., (1980). *Metaphors We Live By*. Chicago: University of Chicago Press.

44. Lewis, D., (1983). "New Work for a Theory of Universals*," Australasian Journal of Philosophy*, 61: 343-377.   Reprinted in

Lewis, D., (1999). *Papers in Metaphysics and Epistemology*. Cambridge, UK: Cambridge University Press, pp. 8-55.

45. Loewer, B. and Rey, G., eds., (1991). *Meaning in Mind: Fodor and His Critics.* Cambridge, MA: Blackwell.

46. Millikan, R.G., (1984). *Language, Thought, and Other Biological Categories: New Foundations for Realism*. Cambridge, MA/London: Bradford Books, MIT Press.

47. Millikan, R.G., (1990). "Truth Rules, Hoverflies, and the Kripke-Wittgenstein Paradox," *Philosophical Review* 99: 323-353. Reprinted in her 1993, pp. 211-239.

48. Millikan, R.G., (1993). *White Queen Psychology and Other Essays for Alice*. Cambridge, MA: MIT Press.

49. Nagel, T., (1974). "What is it like to be a bat?" *Philosophical Review*, 83: 435-50.

50. Neander, K. (1995). Misrepresenting and Malfunctioning," *Philosophical Studies* 79: 109-141.

51. Norman, D., (1988). *The Psychology of Everyday Things*. New York: Basic Books. Reprinted as (1998). *The Design of Everyday Things*. London: MIT Press.

52. Papineau, D., (1987). *Reality and Representation*. Oxford: Blackwell.

53. Peacocke, C., (1983). *Sense and Content*. Oxford: Oxford University Press.

54. Peacocke, C., (1992). *A Study of Concepts*. Cambridge, MA: MIT Press.

55. Peacocke, C., (1999). *Being Known*. Oxford. Clarendon Press.

56. Peacocke, C., (2002). "Three principles of rationalism," *European Journal of Philosophy* 10: 375-97.

57. Pettit, P., (1990). "The Reality of Rule-Following," *Mind* 99: 1-22. Reprinted in (2002). *Rules, Reasons, and Norms*. Oxford: Oxford University Press.

58. Pietroski, P. and Rey, G. (1995). "When Other Things Aren't Equal: *Saving Ceteris Paribus* Laws from Vacuity," *British Journal for the Philosophy of Science* 46: 81-110.

59. Putnam, H., (1962). "It ain't necessarily so," *Journal of Philosophy* 59: 658-71.

60. Putnam, H., (1975). "The Meaning of 'Meaning,'" in Putnam, H., *Mind, Language, and Reality* (Philosophical Papers, volume II). Cambridge, U.K.: Cambridge University Press.)

61. Quine, W. V., (1953). "The problem of meaning in linguistics," in W. V. Quine, *From a Logical Point of View*. New York: Harper and Row. Pp. 47-64.

62. Quine, W. V., (1960). *Word and Object*. Cambridge, MA: MIT Press.

63. Ramsey, F. P., (1931). *The Foundations of Mathematics*. London:

Kegan Paul, Trench, Trubner.

64. Recanati, F., (2001). "Open Quotation," *Mind*, 110: 637-687.

65. Rey, Georges, (1993). The unavailability of what we mean," *Grazer Philosophische Studien* 46: 61-101.

66. Rey, Georges, (1995). "Keeping meaning more in mind," *Intellectica* 21: 65-80.

67. Ryle, G., (1961). "Use, usage, and meaning," *Proceedings of the Aristotelian Supplementary Volume* 35: 223-30.

68. de Saussure, F., (1916). *Course de linguistique générale*. Paris: Payot.

69. Searle, J., (1980). "Minds, brains, and programs," *Behavioral and Brain Sciences* 3.

70. Searle, J., (1992). *The Rediscovery of the Mind*. Cambridge, MA: MIT Press.

71. Sellars, W., (1956). "Empiricism and the philosophy of mind," in *Minnesota Studies in The Philosophy of Science, Vol. I: The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, edited by Herbert Feigl and Michael Scriven (Minneapolis: University of Minnesota Press: 253-329).

72. Sellars, W., (1962). "Truth and correspondence," *Journal of Philosophy* 59, 29-56.

73. Sellars, W., (1963). "Some reflections on language games." In his *Science, Perception, and Reality*. London: Routledge & Kegan Paul.

74. Sellars, W., (1969), "Language as thought and as communication," *Philosophy and Phenomenological Research* 29: 506-527.

75. Soames, S., (1998). "Skepticism about Meaning: Indeterminacy, Normativity, and the Rule-Following Paradox," *The Canadian Journal of Philosophy*, Supplementary Volume 23, *Meaning and Reference*, Kazmi, A.A., ed., pp. 1-39.

76. Stalnaker, R., (1984). *Inquiry*. Cambridge, MA: Bradford Books, MIT Press.

77. Stampe, D., (1977). "Towards a causal theory of linguistic representation," in P. French, T. Uehling, and H. Wettstein (eds.), *Midwest Studies in Philosophy*, 2. University of Minnesota Press, Minneapolis, MN, pp. 42-63.

78. Strawson, P. F., (1974). *Subject and Predicate in Logic and Grammar*. London: Methuen.

79. Wittgenstein, L., (1953). *Philosophical Investigations*. Oxford: Blackwell.

80. Wittgenstein, L. (1969). *On Certainty*. G.E.M.Anscombe and G.H. von Wright (eds.). D. Paul & G.E.M. Anscombe (trans). Oxford: Blackwell.