

Spring 1-1-2011

# The Explanatory Value of the Cognitive Unconscious

Benjamin Thomas Pageler

University of Colorado at Boulder, pageler@colorado.edu

Follow this and additional works at: [https://scholar.colorado.edu/phil\\_gradetds](https://scholar.colorado.edu/phil_gradetds)



Part of the [Cognitive Psychology Commons](#), and the [Philosophy of Mind Commons](#)

---

## Recommended Citation

Pageler, Benjamin Thomas, "The Explanatory Value of the Cognitive Unconscious" (2011). *Philosophy Graduate Theses & Dissertations*. 11.

[https://scholar.colorado.edu/phil\\_gradetds/11](https://scholar.colorado.edu/phil_gradetds/11)

This Dissertation is brought to you for free and open access by Philosophy at CU Scholar. It has been accepted for inclusion in Philosophy Graduate Theses & Dissertations by an authorized administrator of CU Scholar. For more information, please contact [cuscholaradmin@colorado.edu](mailto:cuscholaradmin@colorado.edu).

THE EXPLANATORY VALUE OF THE COGNITIVE UNCONSCIOUS

By

Ben Pageler

B.A., Fort Lewis College, 1996

M.A., University of Colorado, Boulder, 2002

A thesis submitted to the  
Faculty of the Graduate School of the  
University of Colorado in partial fulfillment  
of the requirement for the degree of  
Doctor of Philosophy  
Department of Philosophy  
Institute of Cognitive Science

2011

This Dissertation Entitled:  
The Explanatory Value of the Cognitive Unconscious  
written by Ben Pageler  
has been approved for the Department of Philosophy  
and the Institute for Cognitive Science

---

Robert Rupert Committee Chair

---

Brad Monton Committee Member

Date \_\_\_\_\_

The final copy of this thesis has been examined by the signatories, and we  
Find that both the content and the form meet acceptable presentation standards  
Of scholarly work in the above mentioned discipline.

Pageler, Ben (Joint Ph.D., Philosophy and Institute of Cognitive Science)

The Explanatory Value of the Cognitive Unconscious

Thesis directed by Professor Robert Rupert

#### Abstract

This dissertation defends the explanatory value of unconscious mental states. Unconscious mental states are a subject's information-bearing states with certain properties. These states possess features that make them mental states but do not possess the properties that make them conscious mental states. The claim is that some human behaviors, driven by perceptual and cognitive mental states, are best explained by employing unconscious mental states of one kind or another. The general idea is that the explanatory value of unconscious mental states equals the number of a subject's behaviors valuably explained by their mental states minus the number of the subject's behaviors that are valuably explained by their conscious mental states.

Support for the claim that some human behaviors are best explained by employing unconscious mental states comes in three forms. First, there is a coherent information-processing model for mental states. This model can account for certain of a subject's perceptually and cognitively driven behaviors, in part or whole, with unconscious mental states. Second, independent criteria regarding human subject's conscious mental states support explanations that use unconscious mental states. Finally, unconscious mental states are supported by the success of models that employ such states found in modern psychological and neurological explanations of some human behaviors. This dissertation will also support the idea that the set of unconscious mental states in human subjects are moderately unified, made of various kinds, and are numerous.

Unconscious mental states in empirical psychology and neuroscience are often called a subject's "cognitive unconscious" (Kihlstrom 1987). The explanatory value of the "cognitive unconscious" has been recently attacked, for example by Searle (1990), as well as others. This dissertation defends the explanatory value of the cognitive unconscious against these attacks.

This dissertation is dedicated to my unconscious;  
without its help I could not have read, much less written, this work

## CONTENTS

CHAPTER		
	General Introduction.....	1
1	The Cognitive Unconscious: A Brief Sketch.....	4
1.1	The Human Mind and the Information-Processing Framework.....	4
1.2	Brief Sketch of the Cognitive Unconscious.....	8
1.2.1	Intuitive Illustration of the Cognitive Unconscious.....	13
1.3	Further Clarifications.....	15
1.4	Summary and Overview.....	16
2	The Three-Tiered Information-Processing Model of the Human Mind.....	17
2.1	Introduction to the Three-Tiered Model of Human Mind.....	17
2.1.1	Information and the Three-Tiered Information-Processing Model of Human Mind.....	20
2.2	Tier One: Transduction and Early Sensory Information-Processing.....	27
2.2.1	Information-Bearing States and Information Transmission.....	28
2.2.2	Transduction.....	30
2.2.3	Kinds of Information Processing and Coding Information.....	32
2.2.4	Low-Level Information Processing.....	35
2.2.5	The Importance of Coding for Perception and Cognition.....	40
2.3	Tier Two: Stages of Information Processing Leading to Percepts.....	44
2.3.1	Feature Detectors.....	44
2.3.2	Operators.....	46
2.3.3	The Percept and Perceptual ‘Inference’.....	49
2.4	Tier Three: Levels of Cognition.....	53
2.4.1	Low-Level Cognition: Object Recognition, Kinds of Memory, and Kinds of Learning.....	57
2.4.2	High-Level Cognition: LoT, Inference, and Abduction.....	60
2.5	A Brief Note on “Cognitive” Architecture.....	63
2.6	Summary and Conclusions.....	67
3	What and Where Are Mental States in the Information-Processing Model?....	68
3.1	The Differences between Representation and Information.....	68
3.1.1	Focus of Information.....	69
3.1.2	Expressive Role.....	75
3.1.3	Misrepresentation.....	78
3.2	Where Does Mental-State Explanation Begin: High-Roads, Middle-Roads, Low-Roads.....	82
3.2.1	High-Road: Tier-Three and Representation.....	84
3.2.1.1	Propositional Cognitive Systems.....	85
3.2.1.2	Non-Propositional Cognitive Systems.....	88
3.2.2	Middle-Road: Tier-Two and Representation.....	91
3.2.3	Low-Road: Tier-One and Representations. How Low Can Representation Go?.....	93
3.3	Application: Mental Representation in the Three-Tier Information-Processing Model.....	96
3.4	Summary and Conclusions.....	99
4	Demarcating Conscious States.....	101
4.1	Introduction.....	101

4.1.1	Phenomenal Judgments.....	104
4.2	Functional Demarcations for Conscious States in Humans.....	107
4.2.1	Conscious States as those Tied to Phenomenal Judgments and Types or Report.....	109
4.2.1.1	The Attention and The Working Memory Nexus: Posner, Dehaene and Naccache.....	110
4.2.1.2	Attention versus Awareness: Lamme.....	113
4.2.1.3	Multiple Drafts of Conscious States: Dennett.....	118
4.2.2	Conscious States as those without a Tie to Phenomenal Judgments or Report.....	122
4.2.2.1	Islands of Consciousness and Reverberation or High-Activation: Block and Zeki.....	123
4.3	Summary and Conclusions.....	127
5	The Explanatory Value of “The” Cognitive Unconscious.....	128
5.1	Unconscious Mental States: The Argument Outline.....	128
5.1.1	Unconscious Mental States: The Coding Architecture.....	130
5.1.1.1	Summary of the Overall Picture.....	134
5.1.1.2	Extending Conscious Mental States beyond Working-Memory, But Not Too Far.....	136
5.2	Unconscious Mental States: Three Criteria for Conscious Mental States.....	140
5.2.1	Criterion #1: Dual Subjective Judgment Criterion.....	142
5.2.2	Criterion #2: Justification Criterion.....	145
5.2.3	Criterion #3: Expressibility Criterion.....	149
5.2.4	Summary of the Role of Criteria for Conscious Mental States.....	152
5.3	Unconscious Mental States: Explanatory Success.....	153
5.4	Summary of Defense.....	156
5.5	Size, Kinds, and Unity: Varieties of Cognitive Unconscious.....	157
5.5.1	Merely Dis-Unified Unconscious Mental States.....	160
5.5.2	The Limited Cognitive Unconscious.....	161
5.5.3	The Medium-Sized Cognitive Unconscious.....	164
5.5.4	The Vast Cognitive Unconscious.....	168
5.6	What is the Right Picture?.....	170
5.7	Summary and Conclusions.....	176
6	Objections and Replies.....	178
6.1	Objection #1: Deflation.....	178
6.1.1	First reply to Deflation.....	180
6.1.2	Second reply to Deflation.....	183
6.2	Objection #2: Dependence.....	187
6.2.1	Searle’s Version of Dependence.....	190
6.2.1.1	Reply to Searlean Versions of Dependence.....	192
6.2.2	Ludwig’s Version of Dependence.....	195
6.2.2.1	Reply to Ludwig’s Version of Dependence.....	197
6.3	Objection #3: Identity of Mental States and Conscious States.....	201
6.3.1	Hanna’s Version of the Identity Thesis.....	203
6.3.2	Lind’s Version of the Identity Thesis.....	205
6.3.3	Reply to Identity of Mental States and Conscious States.....	207

6.4	Objection #4: All Mental States Can Become Conscious States.....	209
6.4.1	Reply to All Mental States Can Become Conscious States.....	212
6.5	Summary of Objections and Replies.....	214
6.6	Final Summary.....	215
	Bibliography.....	216

## FIGURES

### Figure

1. Overall Structure of the Human Mind.....	8
2. Trajectories of Information-Processing States.....	66
3. Perceptual and Cognitive Mental States.....	136
4. Summary of the Two Main Questions.....	158

### General Introduction

Unconscious mental states are of explanatory value in cognitive science, a multi-disciplinary inquiry into the human mind. The cognitive scientific approach combines philosophical, psychological, neural, and computational considerations when explaining intelligent human behavior. Often, cognitive scientists posit states internal to the human organism that control the organism's intelligent behavior. In doing so, cognitive scientists generally rely on an information-processing framework when accounting for certain human behaviors. For example, various theories in cognitive science suggest that the human mind is a certain type of information-processing system. Many of these theories converge on the idea that there is a certain information processing architecture of the human mind. The architecture of the human mind includes several distinct kinds of states that help to explain human behavior.

Importantly, many states in the architecture of the human mind are best described as mental states. As we will see, unconscious mental states comprise those information-processing states that can be considered mental but not 'conscious'. Conscious mental states are individuated by special criteria. These ideas are outlined, respectively, in chapters two through five.

This dissertation defends the idea that unconscious mental states are explanatorily valuable for a substantial number of human behaviors. Unconscious mental states likely comprise a fair amount of the mental states in the human information-processing architecture. For example, if conscious mental states are those that human subjects judge to be conscious, then there are a substantial number of mental states, which are not conscious, that valuably explain human behaviors. These states are a subject's unconscious mental states.

To argue for the explanatory value of unconscious mental states we need to establish that there are "mental states" which are "not conscious". This will require arguing for a large taxonomy of states and processes. For example, mental states can be treated as those states that

carry representational content. Within the domain of mental states, there is a distinction between perceptual states and cognitive states. Discussion below will show that part of a subject's mental architecture is comprised of both perceptual and cognitive states that are unconscious.

To defend this taxonomy requires first laying out the standard information-processing picture. This picture distinguishes between low-level information states, perceptual states, and cognitive states. Then what mental states are and where they are likely to begin in the information-processing picture must be shown. *Exactly* where mental-state explanation begins in the information-processing picture need not be established. What is important is that most information-processing explanations in cognitive science employ unconscious mental states. What unconscious mental states are and how they fit into the information-processing framework, in the form of both perceptual as well as cognitive states, will be defended in chapter five. Objections and replies are given in chapter six.

The notions of conscious and unconscious mental states are contentious. Adding the term "so-called" to the terms "conscious", or "unconscious", mental states would show this contention. However, including the term "so-called" throughout the thesis makes reading difficult. Take the contentious nature of so-called, conscious and unconscious, mental states as read throughout.

A brief note of clarification is also required. The term "cognitive unconscious" is an explanatory construct in cognitive science based in the information-processing framework. It likely includes both perceptual and cognitive states. So, the term "cognitive unconscious" can be slightly misleading. It is sometimes also referred to as the "psychological unconscious" (Kihlstrom 1990; Kihlstrom et. al. 1992).

The term "cognitive unconscious" is used here because it is used in cognitive science and helps

to differentiate the notion of unconscious mental states from how such states are used in psychiatry (see Kihlstrom 1987; Wilson 2002; Shevrin and Dickman 1980; Shevrin 1992; among others). For example, the term “unconscious” has often been associated with psychoanalytic theory, such as that of Freud or Jung (Freud 1915; Jung 1959). The cognitive unconscious differs from any psychoanalytic theory of unconscious mental states. For example, cognitive scientists endorse unconscious mental states without making assumptions about repression or universal psychological images (Freud 1915; Jung 1959). In this way, the *cognitive* unconscious should not be confused with other theories that employ unconscious mental states (Power and Brewin 1991; Shevrin and Dickman 1980).<sup>1</sup>

---

<sup>1</sup> This work was helped by discussions with Rob Rupert and Dugald Owen, among several others. Comments on an earlier draft from Rob Rupert and Brad Monton greatly aided this draft.

## Chapter 1: The Cognitive Unconscious: A Brief Sketch

This chapter introduces the cognitive unconscious as it is described in cognitive science. The present chapter also gives a brief illustration of some empirical work in support of the cognitive unconscious. The current chapter also gives a brief illustration of common views regarding unconscious mental states.

### 1.1 The Human Mind and the Information-Processing Framework

The human mind is often viewed as a control system, where internal states control various kinds of behavior. To explain certain human behaviors, such as the person moving their arm, psychologists often posit internal information-bearing states that control the behavior. A motion, such as a subject having their arm raised for them by another person, is not controlled by internal states. In this case, an external force raises the subject's arm. When an arm motion like this is caused by internal mental states then such a motion is sometimes explained as something the subject does (Dretske 1988). It may, for example, be an action rather than something that happened to the subject.

Of course, other systems, such as trees sprouting new leaves in spring also have behavior that is controlled by internal information-bearing states. Trees in spring sprout leaves based on internal states that are sensitive to the external temperature around the tree. To make internal information-bearing states that control a system into mental, or representational, states the control must, at least, be based on the content of the internal state (Dretske 1994). We will see more about the differences between mere information-bearing states and mental states in chapters two and three.

Empirical psychologists have found that the human mind is made up of various kinds of control systems. These systems are loosely integrated. Damage to one system often leaves other systems unaffected. Psychologists posit a difference between control systems that operate on

mere informational states and those that operate on representational, or mental, states. For example, the information processing states that occur in the human retina are not like the information processing states that underlie human decision making. The kinds of explanatory states in these two processes differ.

Representational states, as the term is used here, are states used in certain processes which, for example, track objects and properties through various changes in information. This suggests that the control of certain behaviors is best explained in terms of the content of the states being operated on. The information processing in the retina is not based on states with representational content whereas the information processing underlying human decision making is based on states with such content.

We can already see that psychologists often make several kinds of distinctions regarding the explanatory states they use when account for human behavior. The distinction that is of concern in this thesis is between two kinds of mental states. A further distinction, between mere informational states and conscious and unconscious mental states, is also of concern. All three of these distinct kinds of control states are of explanatory value to psychologists.

For example, not all of the internal explanatory states posited by psychologists are known by the subject whose behavior is being explained. When the states concerned are mental, this is a rough way to distinguish conscious and unconscious mental states. The term ‘conscious states’ is often thought to refer to control states with phenomenal properties. Phenomenal properties are notoriously hard to describe in language, especially non-circularly. Roughly speaking, phenomenal properties are supposed to be first-person properties of a human being’s mental system which are associated with a “subjective feeling” of having or being in that state (Block 1995). Phenomenal properties are supposed to ‘be like something’ for the subject who has them

(Nagel 1974).

Given that phenomenal properties are notoriously hard to understand, or even describe, this way of defining conscious states will be set aside. The separation of conscious and unconscious mental states employed here relies more closely on subject's various judgments about their own mental states. Not all of a subject's mental states are judged by them to be conscious. This is only a rough way to distinguish a subject's conscious and unconscious mental states. The distinction between conscious and unconscious mental states will be further explored in chapters four and five.

Further brief clarification is also required for the contrasting term "unconscious". Only information-bearing states that are best described as mental states will be considered unconscious states (Searle 1992). Trees, for example, have information-bearing control states that will not be considered unconscious. The control states in trees are not best described as mental states. So, the information-bearing states in trees will be considered non-conscious. This use of the term "unconscious" will save us from having to consider the information control states in a person's retina, for example, as part of their unconscious. Chapters three and five will give more argument in support of excluding the information states in the human retina in this way.

The picture of the human mind defended here breaks internal states that control various human behaviors into three general groups. There are explanations that use mere information states, and explanations using information-bearing states that can be further classified as conscious and unconscious mental states. It will also be necessary to break the kinds of mental states employed in psychological explanations into sub-groups as well. The three groups of states defended here and their subdivisions are illustrated in figure 1.

As we will see in chapter four, conscious mental-state explanations are likely made up of

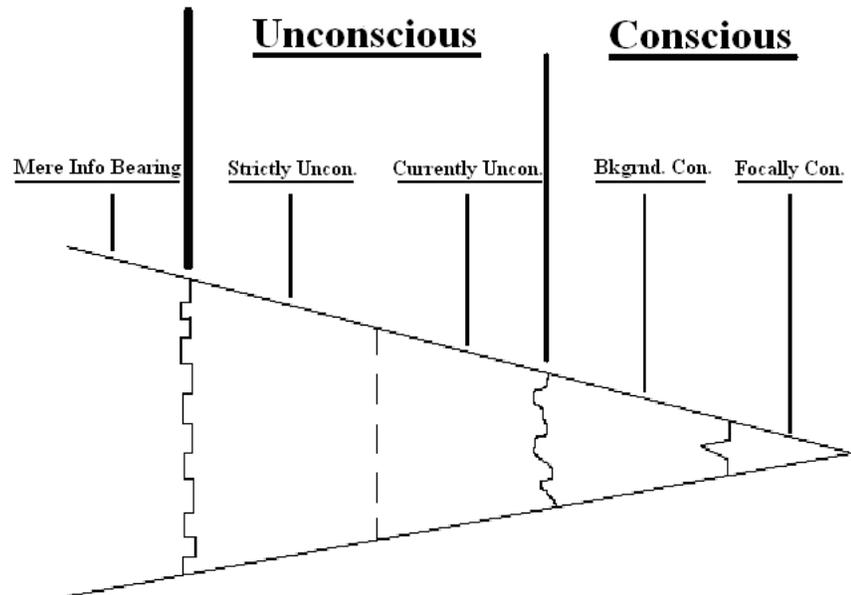
focally conscious states and background conscious states. These kinds of states are analogous to focal and peripheral vision. As we will see in chapter five and six, unconscious mental-state explanations are likely made up of two kinds of unconscious mental states as well. One kind of unconscious state can become a conscious state of the subject at some point or other. The other kind of unconscious mental state cannot become conscious to the subject who has them. These kinds of mental states are separate from non-mental information-bearing states, such as those occurring in the retina.

A few brief clarifications of figure 1 are necessary before we move on. Empirical researchers often refer to “the” unconscious, speaking as if the set of a subject’s unconscious mental states were unified into a single mental system (Kihlstrom 1990). However, the cognitive unconscious of a human subject, as the term is used by cognitive scientists, is unlikely to be a single mental system (see Lyons 2001; Vaina 1990). The set of a subject’s unconscious mental states could therefore be further sub-divided (Milner and Goodale 1995, p. 181).

Likewise, the general shape of the overall picture suggests the relative size of one set of states compared to others. This figure is meant to indicate that there are more states of some types than of others. This implies, for example, that the group of a subject’s strictly unconscious mental states, those mental states that cannot become a subject’s conscious mental states, is larger than the group of a subject’s mental states that are described as focally attended conscious states. These implications will be further explored in chapter five.

## Non-Mental

## Mental



**Overall structure of the human mind and its relation to non-cognitive information processing. From left to right, the sections indicate non-mental information-bearing states, strictly unconscious mental states, currently unconscious mental states, background conscious mental states, and focally conscious mental states.**

Figure 1

### 1.2 Brief Sketch of the Cognitive Unconscious

The empirical model of human behavioral control that employs conscious and unconscious mental states is quite popular. This can be seen in the frequent use of the terms “conscious” and “unconscious” in mainstream psychology and neuroscience literature (Wilson 2002; Kihlstrom 1984; 1987; Hassin, Uleman, and Bargh 2005). Modern empirical explanations of human behavior often assume that there are, at least, three main kinds of states that control human behavior. These states are mere information-bearing states, conscious mental states, and unconscious mental states (Milner and Goodale 1995; Fodor 1983; Pylyshyn 2003; Shevrin 1992).

The psychologist John Kihlstrom defines a subject’s cognitive unconscious as “mental structures and processes that, operating outside [the subject’s] awareness, nevertheless influence conscious experience, thought and action” (1987, p. 1445; 1990; 2004; see also Piaget 1973). A

subject's mental states, which function outside a particular set of operations, define the subject's verbal or mnemonic access to those states. This kind of division of access is often used to demarcate a subject's unconscious mental states.

What is important to note here is that, "the repository of latent [information] is one thing, but... it is the dynamic activity of unconscious [mental] contents that makes them really interesting" (Kihlstrom 2004, p. 97). Of special interest, then, for many modern empirical psychologists are the mental states of a subject that are active but are unavailable to certain operations, such as verbal report, explicit memory, or cognitive reflection (Evans 1982; Chomsky 2000; Marr 1982).

One example that illustrates the empirical account of unconscious mental states comes from social psychology (Wilson 2002). Consider a person's self knowledge. When subjects make consumer decisions not all of the explanatory mental states psychologists attribute to the subject are known to the subject (Nisbett and Wilson 1977). For example, consumers in a typical mall were asked to rate pairs of nylons in relation to other pairs. Unknown to the subjects, each pair of nylons was identical. However, subject's reported a variety of mental states which they thought explained their own behavior while rating the nylons. Few subjects realized the nylons were identical.

Oddly, consumers gave the last pair of nylons the highest rating, 4 to 1. In this way, psychologists found a left-to-right position effect to have the most influential control of consumer's choices (Nisbett and Wilson 1977, p. 243). A mental state representing position and a mental state of preference for this position, rather than the mental states reported by consumers, best explain each consumer's choice. No consumers, however, reported entertaining thoughts about position as an explanation for their decision.

These results illustrate mental state control because the criteria for representational content, which we will see in chapter three, were likely factors that explained the customers' behavior. Likewise, the mental states that psychologists employed for explaining this behavior were not ones that the consumers knew about or reported. This suggests unconscious mental states were involved. From such results it can be argued that psychologists sometimes find it explanatorily valuable to employ unconscious mental states (Wilson 2002).

Explanations that employ unconscious mental states can be found in many areas of cognitive science. The set of empirical explanations that employ unconscious mental states illustrate a subject's overall "cognitive unconscious" (Kihlstrom 1987). Again, the term "cognitive unconscious" is not meant to limit a subject's unconscious mental states to cognitive states alone. Unconscious perceptual states are also often included under the term "cognitive unconscious".

Various kinds of research illustrate a subject's cognitive unconscious. For example, research on human subjects' cognitive unconscious encompasses updated theory and evidence regarding subliminal perception. In subliminal perception research, psychologists have subjects report whether dim lights, soft tones, or other minimal stimulation is experienced or not. If a behavioral impact can be found, but the subjects report no experience of the stimuli, then subliminal perception is thought to be demonstrated (Merikle 2000). Such behavioral effects are often thought to involve a subject's unconscious mental states.

Research into the existence of human subject's unconscious mental states now includes much more than subliminal perception studies. Among other things, it includes studying effects of fully "*supra*"-liminal perceptual stimuli. That is, researchers now look at effects from stimuli that can be easily seen, for example, if the stimuli were attended or presented without influence from other stimuli. Illustrative cases can be found in current research projects that look at effects

from stimuli that are simply “unnoticed” or “masked” by other stimuli that are presented quickly after the target stimulus (Merikle 2000). Subjects without distraction or masking stimulus effects can report these perceptual stimuli whereas with the distraction or masking stimulus subjects cannot report these stimuli. Again, if behavioral impacts can be found but subjects report no experience of the stimulus then unconscious mental states are often thought to be demonstrated.

Research into the cognitive unconscious also includes studies of effects in the domain of normal perception. Human behavior controlled by normal perception, such as the perception of visual or auditory signals, is often explained by states that are considered mental. For example, certain effects in visual object constancy or filling-in effects during noisy auditory perception are often explained by mental states. Many of the states that help to explain behaviors controlled by normal perception are considered unconscious.

Of course, a subject’s cognitive unconscious also includes cognitive mental states as well. Cognition involves different operations from perception. However, both cognitive and perceptual operations are often described as operating on states with mental content. The distinction between perception and cognition will be further explored in chapter two.

For now, a simple illustration of a cognitive operation that likely involves unconscious mental states can be seen in an example given by the psychologist George Miller (1962; Halligan and Oakley 2000). In this example, psychologists ask someone to remember their mother’s maiden name and then ask the subject how they remembered the name. The first question will usually be accompanied by an answer whereas the second, most times, will not. This suggests the influence of unconscious mental states in the second case.

Another example of cognitive processes that operate on unconscious mental states is found in linguistics. Average speakers can accurately categorize a mixed set of grammatical and

ungrammatical sentences. Speakers are able to do so without the ability to report how they accomplish this task. For example, an average speaker can verbally report whether a sentence is grammatical or not but cannot report the states or processing that causes the intuitions which the subject reports. Many of the mental states required to explain the speaker's linguistic abilities are often thought to be unconscious mental states (Jackendoff 1992).

Cognitive operations, such as implicit memory and decision making, automatically learned behavior, or "rules" in reading or other linguistic processing are all thought to reveal unconscious mental states (Kihlstrom 1990). Behaviors that seem best explained by unconscious mental states can also be seen in studies involving prejudice and stereotyping. These studies look at very fast, un-reportable, race-classifications made by subjects. Such classifications may involve unconscious mental states that influence racial-stereotyping behavior (Ito and Cacioppo 2000; Wilson 2002).

Obviously, both perceptual and cognitive unconscious mental states seem quite valuable in empirical psychological and neuroscientific explanations of certain human behaviors. Subjects complete many psychological tasks even though they do not report, or know about, the mental states that explain their abilities. We will see in chapter four and five that merely being unable to report a state is not itself a demonstration of unconscious mental states. Alone, a subject's inability to report is not diagnostic of unconscious mental states (see Block 2007; 2008). What is important here is just that empirical scientists commonly endorse two kinds of mental-state explanations, those that employ conscious mental states and those that employ unconscious mental states.

The current range of explanations of behaviors controlled by cognitive and perceptual unconscious mental states suggests that these states are valuable for a substantial number of

psychological explanations. This result seems non-accidental given the success of the explanations that invoke unconscious mental states (Fodor 1990a; Kihlstrom 1987, 1990; Dretske 2006; Block 2007).

### 1.2.1 Intuitive Illustration of the Cognitive Unconscious

Common sense psychological explanations also grant the explanatory value of many unconscious mental states when giving day-to-day explanations of human behavior. These explanations also hold that states with representational content control many human behaviors (see chapter three). Explanations of perceptual or cognitive processes that control a person's behavior are illustrative cases. Not all such states are thought to be conscious mental states. This common practice adds intuitive motivation to the idea that unconscious mental states explain a substantial number of human behaviors.

In common sense psychological explanations, for example, many mental states are treated as unconscious simply because there is no need to treat them otherwise. Unconscious states, on common psychological explanations, are often cashed out in terms of internal states with representational content which are simply unknown to the subject or which the subject does not attend. For example, common psychological explanations will account for a person's behavior in terms of love that is unknown to the subject whose behavior is being explained. For many of the same reasons employed in empirical psychological explanations, common psychological explanations employ unconscious mental states to account for certain human behaviors.

In day-to-day psychological explanations some mental contents are thought to have the potential to become, and then cease to be, conscious mental states. On this view, a state retains the same mental content while being either a conscious, or unconscious, mental state (Searle 1990, p. 586). For day-to-day psychological explanations this just means the explanatory states exist within a person's mental systems and these states can affect behavior in similar ways,

whether the states with that content are conscious or unconscious. As we will see in chapter six, forgetting and then remembering again is often treated as a single mental content that changes from being conscious to unconscious and then back to being a conscious mental state once again (see section 6.1.1).

Those who give day-to-day psychological explanations are not always careful about details. However, common-sense suggests that these accounts are intended to have unconscious mental states be those mental states that are explanatory but which the subject cannot report or remember, at least at the time their behavior is being controlled by such states (Finklestein 1999).

A few more examples will illustrate common psychological explanations that employ unconscious mental states. Think of a child at Christmas. The child in the wakeful state tears into new gifts. At the time, the child's attention is completely absorbed with the gift being opened. A moment later, the child may only be able to report those mental states that represented the gift they were attending. This suggests that the child had only conscious mental states of that present. However, memories of other presents, or the anticipation of more of them, remain intact explanatory states. These states are likely the unconscious mental states of the child at that time.

Common, day-to-day, psychological explanations that employ unconscious mental states are further illustrated in explanations of a subject's tone of voice. Subjects are often unaware of the mental states that control their tone of voice, for example, when the subject is frustrated or angry. Unconscious mental states are used to explain subject's behavior in these situations even when the subject does not know about or report such mental states.

The above examples show that common psychological explanations treat some explanatory states as unconscious mental states. This practice helps to both explain and unify cases such as those given above by treating them with a single class of explanatory state. The intuitive

illustrations of common psychological explanation given above add support to the idea that unconscious mental states have explanatory value.

### 1.3 Further Clarifications

This thesis not only defends the explanatory value of unconscious mental states but also the idea that some unconscious states cannot become conscious. Holding the laws of nature and the structure of the human organism fixed, this thesis defends the idea that the architecture of the human mind makes it impossible in the actual world for certain of a subject's mental states to become conscious. This splits the kinds of unconscious mental states attributed to subjects into two kinds. There are unconscious states that can become conscious for the subject who has them and those that cannot.

As we will see in chapter six, others theorists disagree with the idea that unconscious mental states have various kinds of explanatory value (Searle 1990; 1992; Horgan and Tienson 2002; Strawson 1994; Pitt 2004). For example, Searle argues that we need to invert many psychological explanations (1990). He thinks that mental-state explanations which do not rely on conscious mental states have problems. Searle thinks that psychological explanations that do not rely on conscious states fail to utilize *mental* states in their explanations. Searle holds, for example, that "no...predictive or explanatory power is added" by thinking of mental states as strictly unconscious. That is, nothing is gained, for Searle, by thinking that some of a subject's mental states cannot become conscious mental states of that subject (Searle 1990, p. 594).

This is wrong, as will be demonstrated by the explanatory value of strictly unconscious mental states. Along the way, it is prudent to also show unconscious mental states that can become conscious. If a theorist adds other properties to successful mental-state explanations, such as a links to consciousness, that theorist is required to show a necessary explanatory role for these additional properties. Some strictly unconscious mental states of a subject accomplish the

required explanatory work without needing to be conscious mental states, or having the potential to become conscious mental states of that subject.

The difference between unconscious states that cannot become conscious and unconscious states that can become conscious for the subject who has them complicates the overall taxonomy. Mental states are those states with representational content. Both perceptual and cognitive processes operate on states with representational content. Some of a subject's perceptual and cognitive mental states can be either currently or strictly unconscious mental states of the subject who has them (see figure 1).

The core of the argument will be that unconscious mental states, of whatever kind, are explanatorily valuable. The information-processing architecture of the human mind likely makes some mental content strictly unconscious. However, defending the idea of strictly unconscious mental states is not necessary for the main thesis. The structure of the human mind includes unconscious mental states and likely makes it impossible for some of a subject's mental states to become conscious mental states of that subject at any time.

#### 1.4 Summary and Overview

We can see from the above illustrations that both empirical psychological explanations and day-to-day psychological explanations accept that unconscious mental states have explanatory value. This suggests that a substantial number of human behaviors can be best explained by unconscious mental states. A subject's unconscious states are likely made up of various kinds. The explanations outlined above paint an initial picture of a subject's cognitive unconscious.

## CHAPTER 2: The Three-Tiered Information-Processing Model of the Human Mind

This chapter sets out, and defends, a ‘Three-Tiered’ information-processing model commonly used in cognitive science. The three-tiered model is not to be confused with figure 1. The three-tiered information-processing model divides low-level information states and processing from mid-level and higher-level states and processing. For example, this model divides the states and processes in the senses from perceptual states and processes and also divides both sensory and perceptual states and processes from cognitive states and processes.

The information-processing model is often not explicitly formulated in terms of mental states. In particular, it is often formulated without assumptions regarding conscious or unconscious mental states. This model will be used as a scaffold for later chapters. Later chapters will inquire, for example, about where mental-state, as well as conscious mental-state, explanations begin in the information-processing model.

### 2.1 Introduction to the Three-Tiered Model of Human Mind

Many modern psychological explanations rest on an information-processing approach when accounting for certain human behaviors. Information that is in the world is picked up by the senses. Information received by the senses eventually helps to form a subject’s perceptual and cognitive states. To understand the information-processing framework we need to understand information as well as various kinds of information processing. We will need to see how the information picked up by a subject’s senses, given certain information processing, leads to the subject’s perceptual and cognitive states. The notion of a ‘framework’ is meant to loosely circumscribe certain ideas that guide theory and practice in psychology and cognitive science that are based on information-bearing states and the processing of these states.

As an initial illustration of this framework, think of the informational states that occur in the human retina. A greater number of cells receive information than compose retinal output. There

are, for example, two-hundred and sixty million photoreceptors at the start of human retinal processing and only two million cells that send information out of the human eye for further processing (Gazzaniga et. al.1998, p. 124). In terms of information, it would seem like a lot of information has been lost. However, there is a fair amount of information processing that goes on even in the human retina. In this way, quite a bit of information is extracted in retinal processing as well.

One way to distinguish the three tiers of the information-processing model is illustrated by the states that explain subjects' reactions to the Müller-Lyer illusion. In this illusion, two main lines of the same length are perceived as having different lengths. The illusion occurs because one main line is tipped with short lines pointing inward while the other main line is tipped with short lines pointing outward. When viewing this illusion, subject's report that one of the main lines appears longer than the other.

Subjects' sensory states carry the information about the actual line length. When the main lines are measured, subjects can see that the lines are the same length. However, the perceptual states that subjects report when not measuring the main lines indicate that subjects see the main lines as having different lengths. Finally, subjects can be informed of the illusion. When subjects have measured the main lines and are informed the lines have the same length they often form conceptual states, such as beliefs, which carry the information that the lines are of the same length. Subjects form these cognitive states even in the face of perceptual information that the lines are of different length (Pacherie 2000).

The states involved in the above explanation set out three tiers of information processing and various kinds of states in each tier of processing. For example, sensory, perceptual, and cognitive states are shown when explaining human subjects' responses to the Müller-Lyer

illusion. Below we will set out criteria that distinguish these three-tiers of various kinds of states and the information processing involved with each kind of state.

Generally, the information-processing model is formulated in terms of stages where information is added to, or lost from, or extracted from the original signal. The information added, subtracted, or extracted does not just amount to noise but is relevant for preparing a state for later operations. For example, information processing that extracts certain aspects of the incoming signal do so relevant to later operations (Massaro and Cowan 1993). Processing that extracts information, as Marr notes “determine[s] what information is made explicit and hence what is pushed further into the background” (1982, p. 12).

Some information processing is linear and serial. This is the case in many modern computers. The information-processing approach to the human mind has often been thought to be limited to the ‘computer model of the mind’ (see Block 1990). This model suggests that the human mind processes information like a computer. Computer models are often linear and serial. For example, information processing occurs in one stage before another and, often, later information processing does not influence earlier information processing.

Neural models that are currently popular are generally not simple linear and serial models of information-processing of the human mind. Neural information-processing models often have stages of processing that can run in parallel to other stages of processing. Likewise, these models often have information that is processed in later stages influence how information is processed in earlier stages.

The computer-model of human information processing, as we will see more below, suggests that certain perceptual and cognitive systems operate on discrete units that are somewhat like geometrical units or the words and rules of language, respectively. These units are often called

“symbols”. Often information-processing systems, such as computers, employ symbolic units for certain kinds of systematic or even reason-like operations (Newell 1990).

The information processing in the human retina, for example, is unlike symbolic information processing. Retinal information processing does not operate on the kinds of symbolic units often thought to be involved in decision making or reasoning. This difference helps distinguish kinds of information processing which are important for explaining human behavior. We will see more about these two kinds of information-processes below. Needless to say, the information processing done in computers does not exhaust all types of information processing used in psychological explanations.

#### 2.1.1 Information and the Three-Tiered Information-Processing Model of Human Mind

There are several reasons information processing is applicable to explanations of some aspects of human behavior. One reason is that information can be a basis for constructing internal states that are about the external world. Certain natural relations, central to the notion of information, can allow one state to indicate conditions in another state. For example, tree rings indicate the age of the tree and ripples on the surface of a pond can indicate the impact of a stone. The rings of trees and ripples on ponds do this because they carry information. These states reduce the uncertainty regarding the age of the tree, or the location of the impact on the surface of a pond, to zero. Likewise, medical explanation suggests that certain bumps carry information that a patient has measles. The bumps are not the illness. The bumps are states that are directly correlated with the illness (Dretske 1988; Grice 1957). In this sense, informational states often *indicate* certain other states of affairs, such as the tree’s age, the location of impact, or that the measles virus is occurring in the patient.

In a somewhat similar way, thoughts are *about* things, perceptions are *of* things, reaches are *for* things. In each of these cases, the psychological states are success oriented and so, *prima*

facie, directed to things in an environment, for example. Naturalistically inclined psychologists and philosophers hold that natural meaning, indication, or information is a good basis for explaining, in purely natural terms, the directedness of the states used in psychological theories.

Naturalism more generally suggests that any explanation should account for events with natural entities and processes. Information-processing models of the human mind use only natural entities and processes to explain human behavior. For example, additions and subtractions of information in stages of information processing are accomplished by natural, physical, operations and result in only natural, physical, entities. In this way, information-processing accounts of mind are compatible with naturalism. This is one reason why psychologists use the information-processing framework.

As an illustration of an explanatory state that is directed to the world, think of perception. Perception is success oriented because, in general, if a subject does not get in contact with an actual object then they hallucinate, they do not perceive, the object. This success orientation of perceptual states requires contact, or a causal relation, with the object. Similar constraints hold for sensory and cognitive states as well. Information, or ‘natural meaning’, helps to understand how these various kinds of states are employed in psychological explanations.

More broadly, if a system’s cognitions about an object, perceptions of an object, or reaches for an object do not causally interact with the object in question, then the internal states controlling the subject’s behavior will not explain how the subject’s behavior is related to the object. Take reaching as an example. If a subject’s reaching behavior is not causally related to the object then the state that explains the subject’s behavior will not explain how they are reaching for that object. So, *prima facie*, psychological explanations of human behavior must include states directed to the world via causal relations. Psychologists do this in terms of

information-bearing states. Information about various aspects of the world, carried by the states controlling a subject's behavior, helps explain behavior directed to those various aspects of the world.

A plausible initial explanation for externally directed states is, therefore, found in the notion of natural meaning, natural signs, or information (Grice 1957). Natural signs are certainly not cognitive or perceptual states. Tree rings are not perceptions or cognitions about the age of the tree, for example. Perceptions and cognitions are very likely representational states. One reason tree rings are not perceptions or cognitions is the tree rings could never misrepresent the age of the tree. In this way, tree rings carry information about the age of the tree but never represent the age of the tree.

In chapter three we will see how perceptual and cognitive states can be described as representational states. For now we just need to see how to build perceptual and cognitive states from information states. One main difference between informational states and perceptual and cognitive states is that perceptual and cognitive states do not have to be veridical information carriers. As an illustration, think of perceptual illusions. The states involved in perceptual illusions are not veridical but still function in the system in a way linked to sensory information.

A standard way of characterizing information is in terms of conditional probability. Information is defined as perfect causal correlation between one state and, in most cases, another. For one state to carry information about another state, the probability of the occurrence of the second state, given the occurrence of the first state, must be equal to one (Dretske 1988).

One reason for defining information in this way is so that it will have certain intuitive properties. One of these properties is the condition that if one state carries information about B and it carries information about C then it carries information about B and C (Demir 2008).

Likewise, if one state carries information about A, this information can be transferred through various other states. That is, if state C carries information about state B, and state B carries information about state A, then state C carries information about state A.

As an illustration, think of the activation states in neural systems. When the activation of one state causally correlates in a reliable way with activation of another state, then the second state carries information about the first. On the standard view, the activation of a third state, based on the activation of the second state, can then also carry information about the activation of the first state. A good way to think about this kind of transfer of information is in terms of phones and phone lines (see section 2.2.1 below). If each state in a single information channel correlates with the others one-to-one, then later states can carry information about earlier states that do not directly cause the later states.

One-to-one causal correlation should not be read too narrowly. This notion is not meant to be symmetrical. Likewise, there are often multiple states in single causal channels, say of a phone line. Each of these states can all carry information about the original signal. This means there are many states correlated with the original signal. Likewise, multiple phone lines can carry information from one source. So, the notion of one-to-one causal correlation needs to be treated cautiously. The term just indicates the idea that the correlation of *each* state along a channel must be active when, and only when, a signal is active. This terminology will be clarified more below. In this way, many states that are caused by an original signal can carry information about the original signal.

For one state to carry information about another certain background conditions must be held fixed. For example, the states in the information channel must be reliable transmitters of the original information. In the case of phones, the phone line must be a reliable channel; otherwise

the information being transmitted will be lost or mixed with other information rather than simply transmitted.

Mere information-bearing states cannot be false or misrepresentative because they require perfect correlation in one direction. So, if information is to 'ground' what cognitive and perceptual states are about we will need to link information to other kinds of explanatory states, such as perceptual and cognitive states (Harnad 1990). This is because cognitive and perceptual states are not required to carry veridical information. Mere information-bearing states, such as those in the initial stages of retinal processing, must be veridical. The link between mere informational states and perceptual or cognitive states is in terms of information processing which prepares, say, retinal information for other kinds of operations.

As an example, think of the progression from information picked up at the senses to perception and cognition. The distinction between sensory states and perceptual and cognitive states relies on coding information for later operations. Coding is a physical process that physically shapes states for use by later operations. We can think of this in terms of neural activation states. Information processing sharpens, adds to, or removes activation so that the initial sensory state can function in later operations.

As a further illustration of shaping a physical state in terms of coding, think of key-coding. Coding a plastic card with a set of lines, such as a bar-code, prepares the state of the card for later operation in a scanner. We will see more detail about coding in section 2.2.3. For now, we can think of the difference between the three tiers of information processing in terms of coding sensory, perceptual, and cognitive states in different ways. The difference between how states are physically shaped, or coded, and therefore which operations they can enter into, helps determine whether the state is a sensory, perceptual, or cognitive state.

Sensory systems, when functioning properly, function as simple information-transmission channels. That is, sensory states co-vary one-to-one, along a particular information channel, with a particular kind of stimulation. Think of the response of auditory states to sound. A particular kind of wave length of sound activates a particular kind of auditory receptor. Later auditory states, such as perceptual auditory states, can co-vary one-to-many with information in the world. For example, two different auditory signals, which activate two different sensory channels, can activate a single perceptual state.

Many-to-one causal correlations should also not be read too narrowly. The term “many-to-one” just means that, for perceptual states and cognitive states, there are many ways to cause the single internal state considered a functional ‘end product’ of perceptual or cognitive processing. For example, think of categorical speech perception. In certain cases, subjects perceive the sounds *ba* and *pa*, for example. Sometimes, the original acoustic signal can have many *pa* features but still be heard as *ba* (Spivey and Dale 2006). Clearly, such perceptual states can be created by various sensory states. In this way, perceptual and cognitive states do not always carry the information that original sensory states do.

Different mappings of information, progressing from one-to-one on to many-to-one mappings, help to constitute the progression from information picked up at the senses to perceptual and cognitive states. This progression will allow psychological systems, which are sensitive to abstractions over certain informational changes, to function in ways that other information-processing systems, such as the senses, cannot. Such progression will likely take many stages of information processing to reach just the first stages of perceptual processing. The first stages of cognitive processing are even further along information-processing routes. Such processed states are costly since it takes a variety of resources to shape, or code, informational

states into states that can enter perceptual or cognitive operations.

Using terminology such as “one-to-one” and “many-to-one” can be confusing. For example, retinal cells sensitive to colored light fire for several different wavelengths of light. So, there is a one-to-many relation between color sensitive cells in the eye and wavelengths of light (see section 2.2.5 below). Even though a bit confusing, characterizing certain processes as many-to-one mappings is a common way to describe what is going on in perceptual and cognitive processing. For example, Pylyshyn notes “the fact that the mapping from the distal [three-dimensional] environment to a [perceptual state] is many-to-one entails that the visual system fails to distinguish certain [informational] differences” (Pylyshyn 2003, p. 134). Many other cognitive scientists also think this terminology is appropriate for clarifying the relation between the sensory signal and the perceptual, or cognitive, interpretations eventually given to that signal (Dawson 1991). Clearly, though, such terminology must be used cautiously.

What is important from the above is that sensory states are much more closely tied to particular stimulation than are either perceptual or cognitive states. Again, sensory states have a kind of one-to-one mapping with physical stimulation. Perceptual and cognitive states differ, for example, in that the mappings from stimulation to these states are a kind of many-to-one correlation. Perceptions are more closely tied to sensory states than are cognitive states. For example, perceptual states that occur without stimulation are hallucinations. Cognitive states that occur without stimulation are not always delusions. In this way, there is a looser tie between cognitive states and sensory stimulation than is the case for perceptual states.

How tight the tie is between information present, say, in an environment and an internal state helps to distinguish sensation, perception, and cognition. Again, the Müller-Lyer illusion is a good illustration. Perceptions are not always veridical sensations and perceptual states are not

always revisable in the face of cognitive states, such as beliefs, that may carry contradictory information. Thus, although intermediate perceptions can give rise to cognitive states, such as beliefs, they should not be confused with sensations or cognitive states. These differences are what psychologists are trying to indicate with terminology such as “one-to-one” and “many-to-one”.

From the above we can see why a “Three-Tier” information-processing model is applicable when explaining human behavior. These models divide information picked-up by the senses, for example, from perceptual and cognitive states. We can also see that the notion of “*three tiers*” of processing is a bit of a misnomer. The information processing between sensation and perception or that between perception and cognition requires many stages of processing within each tier. Clearly, though, functional differences do exist between sensation, perception and cognition (see Pylyshyn 1984; 2003; Fodor 1975; 1983). Hence, we get the three-tiered information-processing model of the human mind.

## 2.2 Tier One: Transduction and Early Sensory Information-Processing

It is interesting to note that some theorists characterize a subject’s cognitive unconscious in terms of all the information-bearing states that are received by the subject’s nervous system (Zimmerman 1986; 1989; Nørretrands 1998; Wilson 2002). For example, the number of bits of information that come into the human nervous system is extraordinarily large when considered in terms of the various receptors in each sense organ. Clearly, people do not have conscious experience of all these information-bearing states. Since subjects do not, or cannot, know about, report, or attend all the information-bearing states in their nervous system, some theorists characterize the subject’s cognitive unconscious in terms of all of the sensory information detected by a subject’s nervous system.

As we will see more clearly later, this is an overestimation of a subject’s unconscious mental

states. A subject's unconscious mental states are composed of perceptual and cognitive states, not all the information-bearing states that occur in, or even just after, the information that is picked-up at the senses. Again, the example of information pick-up and processing in the retina is a good illustration. The states, at least in the very initial stages, of retinal processing should not be included as part of a subject's cognitive unconscious. In this way we can distinguish the sensory stage of information processing from perceptual and cognitive stages of information processing. Later on, in section 2.4, we will see more about how to distinguish perceptual states from cognitive states. As a limited illustration, we can think about this distinction in terms of perceptual and cognitive states as information-bearing states that are available to different psychological operations. For now, all that is important is that these states are different from sensory states.

#### 2.2.1 Information-Bearing States and Information Transmission

To begin giving psychological explanations we must get information into the system. This information eventually helps explain how psychological states, such as perceptual and cognitive states, function. We need a connection to information so that perceptual states and cognitive states can in some sense be about, for example, external objects and properties. This connection, as we have seen, rests on the notion of naturally occurring signs, or information. Naturally occurring signs can be organized for artificial, or biological, purposes. Signals can be transmitted through states of a system designed to capture natural sign relations and to organize these natural sign relations for further information processing.

Think of a working telephone. If someone speaks into a phone when it is off the hook, and they have successfully contacted another phone, the states internal to the phones will reliably covary with the person's voice. If a channel of communication has been established with another working phone then the signal at one end of the connection will carry the information that is

present in the other phone. The fact that one phone can carry the information present in another is true because the signal at one end of the phone line is reliably correlated with what is occurring at the other end. This is also true of all the states along the phone line. These correlations come about because of the design of the phones and the connecting lines between the phones. Phones, and other artificial or biological devices, can capture sets of natural relations and put them to use, for example in information transmission.

For two physically distant states to be reliably correlated, the states between them, that constitute the information channel, will all need to co-vary with changes in the first state. For example, the states in one phone, and the wire, must all be reliably correlated with the states in the phone sending the signal. Reliable channel conditions are necessary for information *transmission* because we do not want any new information to be introduced into the signal that is being sent (Dretske 1988). So, one telephone connected to another, for example, has physical states that correspond to what is spoken into the phone and these physical states are then transmitted along the phone lines to the other working phone.

Information transmission is different from information processing. The channel conditions needed for information transmission require that no new information, or that only redundant information, is generated after the signal's source has caused the original signal (Dretske 1981, p. 115). In the case of a phone call, this just means that the voice on one end provides the only variable, or new, information, the only thing that needs to be listened to. Every natural sign 'downstream' from the variable source, in good channel conditions, causes only a redundant bit of information correlated with the original signal. In this way, the source of the signal is the only place variable new information is 'coming from' and not the phone or phone line itself. Subjects who are called on phones listen to the speaker of the message when receiving a phone call not

the transmission line or the phones. If variation occurred along the phone line, or in the phone itself, then this variation plus the voice would be the source of the signal received.

Information transmission does not alter the information. In information transmission the signal in the information is not processed. This is a crucial difference between information transmission and information processing. Neuroscientists, for example, distinguish between the view that various parts of the nervous system are “transmission cables” and the view that other parts are “information processors” (Debanne 2004). Processing alters the informational content, for example, by adding, removing, or extracting certain bits of information. The additions, subtractions, or extractions are relevant for later operations and so do not amount to noise. Some of the information processing in the nervous system helps to extract certain aspects of the information that is present in sensory states. To function, systems which transmit information, such as phones, must not process or ‘operate’ on information states - they must merely transmit the information.

### 2.2.2 Transduction

Transduction can be thought of as a kind of information transmission. Transduction is technically defined as the mechanistic process whereby information in one physical form is converted into another physical form (Fodor and Pylyshyn 1981, p. 157; Fodor 1983, p. 43). For example, a transduction process converts physical states of one kind into internal sensory states of the subject’s nervous system. Information transmission is illustrated in the human retina that transduces light into neural signals.

The notion of transduction is central to cognitive science and information-processing explanations of human behavior. How this notion is defined constrains how one is doing cognitive science (Pylyshyn 2003; Fodor and Pylyshyn 1981). Surprisingly, the difference between Gibsonian and traditional cognitive psychologists hinges on how transduction is

defined. If transduction, in vision for example, is defined as “directly detecting” high-level properties in the local environment, as Gibsonian psychologists hold, then perceptual processes will not have much to do when extracting the high-level properties of distal stimuli that are present in proximal stimulation.

If, however, transduction cannot ‘directly pick-up’ high-level properties, then some processing would be necessary to get from proximal stimulation to perception and cognition about the high-level properties of the distal stimulus. This is because the high-level properties that are perceived and cognized are not *directly* picked-up at the sensory level. There needs to be some information-processing between, say, the sensory signals which some stimulus provides and a perception of, or cognition about, the distal stimulus (Marr 1982).

Take the example of a subject seeing edible fruit. In the standard treatment of the information-processing framework, ‘edibility’ is not the appropriate level of description for a property entering the process of transduction. Transducers play a certain role in the information-processing framework in that they mediate contact between the perceptual and the *physical* world. So, transducers, in the traditional information-processing framework, must translate physically defined states, or states that figure in laws of physics, into states of the nervous system, for example (Fodor and Pylyshyn 1981).

This is why a particular definition of information is so important to cognitive science and why the notion of transduction, on the standard picture, is tightly linked to the notion of information. Transduced information is the link between internal states and the physical world. The link is explained in purely physical terms. The informative properties of natural states are what eventually allow perceptual or cognitive states directed to the world to be a part of a fully natural psychological explanation of human behavior.

The Gibsonian account of perception does not explain the connection between the world and the perceptual and cognitive system in non-semantic terms. This is precisely why the information-processing framework cannot define transduction as Gibsonians do. The information-processing framework must first account for perception and cognition in physical terms and so transduced states must be described in the vocabulary of, say, physics (for a detailed defense see Fodor and Pylyshyn 1981, Fodor 1983; 2009, Pylyshyn 1984; 2003; 2009).

There is some disagreement within the standard information-processing framework regarding how far traditionally defined transduction can reach. For example, Pylyshyn suggests that transduction can reach as far as feature detection (1984; see also section 2.3.1 below). Fodor, on the other hand, suggests that transduction is a kind of bare information transmission (1983). On Fodor's view, when information *processing* starts transduction ends (however, see Fodor 2008, p. 187, for a view more like Pylyshyn's). We will follow Fodor's early definition in the rest of this discussion.

### 2.2.3 Kinds of Information-Processing and Coding Information

If information transmission is not the same as information processing, what, then, is information processing? Strikingly, cognitive scientists use the phrase 'information processing' to refer to both symbolic and non-symbolic information-processing. As we have seen, a symbolic use of the term 'information processing' is where the operations in question use discrete units of information, such as symbols somewhat like interpretable words and rules for combining these symbols.

You can find cognitive scientists working under the banner of information processing regarding research into high-level decision-making. These theorists posit operations that use information-bearing states that are organized into discrete packets of information, such as words and rules for their combination (Newell 1982; Pinker 1999). Some of these operations mirror the

interactions between objects and object-properties involved in reasoning and decision making (Newell 1982). Symbolic operations can move symbolic units around, for example, in ways that mirror logical thought.

Chess is a good illustration of symbolic information processing. A person can ask what impact moving a bishop will have on their overall defensive position. Similarly, a symbolic computation can calculate the value of changing similar variables. If decision making, whether logical or not, requires symbolic information-processing then this use of the term 'information processing' represents just one kind of information processing.

Symbolic information-processing, or symbolic-computation, is familiar to computer scientists, especially early artificial intelligence researchers. For example, modeling the game of chess on a computer is fairly straightforward. In chess, there are well defined pieces, kinds of moves, and a space within which to move pieces. For each of these elements a data structure, or highly organized unit of information, can be made.

An algorithm for moving these elements can be constructed to reflect, say, the decision making processes that humans employ regarding one type of piece, its position, and the potential moves this piece has on a chess board. Likewise, something close to symbolic computation may be involved in processing leading up to object recognition (Pylyshyn 2003). Early vision, for example, may perform the function of segmenting 'parts' of the visual scene and then attempting to reassemble these 'parts' into whole objects that can be recognized by later processes (Biederman 1987).

To reach the stage of perceptual and cognitive information processing generally requires coded units of information. Codes are ways of organizing physical states so that these states can enter into particular kinds of operations. In this way, the notion of a code and the notion of an

operation are interrelated. Coded units of information are not always semantically interpretable. For example, coding of light and dark saliencies from retinal input, or grey-coding as it is often called, may be one example (Sterelny 1990, p. 67; also see below section 2.3.1).

One stage of visual processing is thought to operate on ‘vertex types’ (Pylyshyn 2003; Waltz 1975). These coded units are conjunctions of units that represent lines and can be combined to form representations of the outlines of objects. Although we do interpret the units as vertex types, they are just units that represent conjunctions of three ‘lines’ and their overall configurations represent how various vertexes might fit into object outlines.

Operations are sequences of states and the causal links between them. That is, the sequence of states and the constraints for their combination constitute an operation. Different causal links between the very same states can be used to differentiate operations. For example, we can distinguish operations that associatively link steps of a mathematical problem from those that link the states based on other principles. A subject who achieves the result of a mathematical problem based on mere associative memory is different from one who achieves this same result, using the same states, based on understanding.

Classically, computations in computers are a good illustration of operations and their relation to coded states. Computer states are coded, or physically shaped, for particular operations. For example, the electrical states of a computer are organized into states that can enter certain operations based on their shape. Single units can be used to represent, for example, vertex types. In this way the shaping of a coded state can sometimes be found in the wiring of the system that assigns status to a single node. Without the status of the node being designed into, or coded into, the rest of the system then the rest of the system will not recognize the unit as one available for certain operations. For example, something must separate states so that operations can

recognize, say, single units as representing one thing rather than another. Coded states, or system designs that allow single nodes to be treated as coded states, allow operations to perform particular functions.

Interestingly, codes can be distinguished from natural signs or other information, say, in a computer or a nervous system. Not all informational changes to an information-processing system are computationally relevant in terms of states that later processes can operate on. For example, a human's nervous system may change systematically with their body temperature, expanding and contracting, say. Without the information about the temperature of the nervous system being coded into computationally relevant forms the information about temperature cannot be operated on by later psychological processes. Hits with a hammer cause information changes in a computer system, as well, but this information is unlikely to be relevantly coded for later operations (Pylyshyn 2003, p. 5-11; Uttal 1967; Fodor 1983). In this way, coded information becomes the key to psychological explanations based on information processing. Only certain information-bearing states will be relevantly coded to function in operations we consider perceptual or cognitive.

#### 2.2.4 Low-Level Information Processing

Low-level information processing, such as that found in the retina, or in dendrites, is not like operations that work on perceptually or cognitively coded states. Dendrites, for example, are the parts of neurons that take in information. They are the 'receiving end' of the neuron. It was once thought that dendrites were passive channels of information transmission. However, it is now thought that dendrites process information as well (London 2005). Not only do dendrites sum incoming inputs they can also compute certain kinds of responses to input. So, low-level information processing can contribute to organizing informational states into coded states that can enter particular operations. However, the lowest levels of information processing do not

operate on coded states.

For example, structures of dendrites can allow low-level inputs at two sites to inhibit activation of a further part of the neuron. Input from either site is usually enough to activate the later part of the neuron. However, in certain dendrites input from *both* sites gives rise to strong inhibition thus canceling activation of the overall neuron. In this way, one part of a neuron computes something analogous to an exclusive-or, or XOR, function (Mel 1994). Such computation is not always based on symbols. However, information is being packaged in a way that can be handled by a particular later operation. Given the complex branching structure of most dendrites, quite a bit of low-level information processing can occur before information even reaches the neuronal body (Eliasmith 2003, p. 510).

Low-level information processing can also occur between whole neurons. For example, certain neural connections in the retina influence how other neurons in early visual processing fire, thus causing a difference in the original information received. One of these kinds of information-processes is called lateral inhibition (Ratliff 1965). Lateral inhibition is a way of wiring neurons together in such a way that contrasts in the original signal are enhanced. This kind of processing organizes incoming information in a way that transduction does not. Visual borders, for example, are sharpened (Wehr and Zador, 2003). This kind of early sensory processing is found in other modes of sensory processing besides vision, such as audition (Pantev et. al. 2004). In this way, even initial stages of information processing go beyond transduction since information is organized, or coded, in a way not present in the original signal.

Early information processing states are coded states but are not symbolically coded states. This low-level information processing involves the second sense of information processing, namely non-symbolic information processing. We must be careful to distinguish between uses of

the term ‘information processing’ when it is applied to processing that involves symbols and processing that does not involve symbolically coded states.

One further issue of clarification is needed when considering the term “information processing”. Only very early on in the nervous system are the states involved in the information processes directly transduced information-bearing states. Later states will be coded and so carry information about different aspects of the distal environment.

Take the example of information processing that can be described as an ‘XOR-gate’. These kinds of gates activate when they receive certain kinds of input, as we saw above. For example, they receive information from two sources and activate when one input, or the other, is present but not both. An information processor that can be described as an ‘AND-gate’ activates only if both inputs are active.

Now consider a system designed with two XOR-gates wired to an AND-gate. Each XOR-gate will fire if one or the other of its inputs is active. In this way, both of the XOR-gates outputs are directly correlated with information transmission from an initial signal. Whenever one of these XOR-gates is active it indicates that one or the other of its input states is active.

Consider further the output of the AND-gate with two XOR-gates as input. What information does the output state of such an AND-gate carry? Is it the same information as that carried by the original proximal input? No. It does not co-vary one-to-one with any of the original proximal information-bearing states. For example, it is no longer directly correlated with any of the four initial input states fed into the XOR-gates. It fires when one or the other of the two *pairs* of initial inputs is present but does not ‘care’ which exact input was originally present.

The output state of a system such as that just described is now correlated with a slightly more abstract property. For example, the output is now abstracting over the original input signals in a

way that correlates with the *situation* of one or the other of the two pairs of inputs firing but not with any of the particular input states that fire each of the XOR-gates. The output of the AND-gate ignores certain information present in the original signal.

Various ‘gate-stacking’ designs, as illustrated in this simple system, suggest the output of certain gates can be fed as input to others. In this way, different information-bearing states can be produced that are, for example, about different parts of the environment. The information in these states is not simply transduced or simply transmitted proximal input. Encoding information in this way allows later operations to have states that co-vary with high-level properties in the environment, say.

With enough information processing an internal state can co-vary with a high-level feature such as ‘edibility’. This kind of encoded information is beyond what mere sensory transduction, and simple information processing systems, can provide. This limitation of the information present in transduction and low-level information processing is also why standard cognitive scientific theories are not Gibsonian (see section 2.2.2 above).

As we have seen, only in situations of information transmission is the content of the proximal signal exactly preserved. When information processing occurs this is no longer the case. So, many of the states in post-transduced information processing do not carry the same information that was, or even can be, transduced by the sense organs, for example. This is because these later states are no longer in one-to-one correlation with information picked-up at the initial stage of sensation. What information processing does is to change the information that has been transduced by the senses.

Information-processing states in even the earliest stages of the nervous system, therefore, carry two kinds of information. They carry as much information about the, say, neural

interactions of the system they have gone through as they do about the distal stimulation itself. This is because, even at this early stage of information-processing, the information has been encoded, or physically organized, in a certain way. This is why direct perception does not occur for higher-level properties such as ‘edibility’ (Fodor and Pylyshyn 1981; Churchland, Ramachandran, Sejnowski 1994; Pylyshyn 2009). Coded states have information about both the system they are in *and* the world from which the information originated.

The above idea can be brought out further by looking at an example from Dretske (1981, p. 140). He gives the example of a system for coding an analog signal into digital form. For example, an analog speedometer can take any speed values from, say, 0 to 100. A digital converter attached to this analogue speedometer can be designed to emit a sound for four ranges of speed, say, 0-24, 25-49, 50-74, 75-100. So, when the analog signal is between 25 and 49 the digital converter emits a particular tone for that range. Dretske says that this digital process “necessarily involves the loss of information” (1981, p. 141). This is because the specific speed, say 43, is lost when the digital converter signals that the range of speed is between 25 and 49. The specific speed is lost but what is gained is a kind of classification or generalization. A system that can employ the coded states will be able to deal just with the coded information that covers a range of initial inputs.

So, to speak of this process as *just* a loss of information misses a crucial point for information-processing explanations (Hatfield 1992). Certain information processing systems add information during processing. The digital converter in Dretske’s example adds the information regarding the ranges that are likely significant or relevant to later operations. This “relevance-based information” is not simply in the original signal nor is it simply in the initial sensory processing. It is coded information. In this way, we can see the importance of salience

or relevance of coded information. Certain information processes both take away certain information, such as the specific speed, as well as add information, such as the relevant ranges of stimulation. This kind of coding, or physical organization, of information-bearing states prepares them for later operations, especially those we consider perceptual or cognitive.

#### 2.2.5 The Importance of Coding for Perception and Cognition

Processing information, such as adding relevant groupings or ‘classifications’ to incoming sensory states or extracting salient portions of information for later operations, is crucial for psychological explanations (Marr 1982, p. 12). These processes encode states, making them available for perceptual and cognitive operations.

Categorical speech perception, which crosses several stages of information processing, is an illustration. Human subjects can perceive the sounds *ba* and *pa*, for example. As noted earlier, sometimes the original acoustic signal can have many *pa* features but still be heard as *ba* (Spivey and Dale 2006). In this way, the internal state that controls the subject’s categorical perception carries information about the original acoustic signal as well as information about the categorical nature of processing relevant to the human acoustic system. This means the information in the perceptual state is as much about the original signal as it is about the categories coded into the original signal that are relevant to later operations.

The upshot is that coded information can function in later operations in ways that un-coded information cannot. For example, the categorical information encoded into an original sensory state can then help to cause behavior that the original signal cannot. Subjects report hearing ‘*ba*’, rather than ‘*pa*’, even when the sensory signal is noisy. Coding allows for the transition from transduced sensory states to perceptual states, for example. As a further illustration, coded states are what separate perceptual states about the Müller-Lyer illusion from sensory, or cognitive, states about the same set of lines.

Relevant coding of states for certain operations requires an information-processing system be designed in certain ways. In most nervous systems, for example, to achieve certain coded states several stages of information processing are required. Codes allow states to respond to equivalence-classes of stimuli. Certain equivalence-class responses are ways for states to capture the generalizations psychologists find useful in explaining some aspects of human behavior.

Categorical perceptual states, for example, respond to an equivalence-class of incoming stimuli. States that respond to equivalence-classes ‘classify’ a set of inputs as the same. That is, the response of these states to various bits of information in the stimuli is equivalent. Coding shows how states can ignore certain informational changes. This helps to clarify the relevant use of the term “many-to-one” psychological explanation and shows how psychologically useful states can emerge from information processing of incoming sensory states.

Many kinds of states respond to equivalence-classes. Not all such responses indicate states relevant for psychological explanation. As noted earlier, color receptors in the eye respond to equivalence-classes of wavelengths. These cells fire for various kinds of wavelengths. In this way, such states could be said to ‘classify’ input of various wavelengths as the same. However, the level at which this ‘classification’ occurs can be captured in physical terms. No psychological states need be invoked. The equivalence-classes that psychologists are interested in cannot be described in lower-level sciences, such as physics or biology (Pylyshyn 1984, p. 21).

Psychologists are interested in states that respond to equivalence-classes that ‘categorize’ various inputs in a way that does not coincide with the ‘classifications’ found in other sciences. Color cells in the retina, for example, ‘classify’ inputs in terms that can be captured in physics.

The generalizations that concern cognitive psychologists are not captured in physical or biological terms. Human movements can be caused by physical forces, such as gravity, or biological processes, such as reflexes. However, the particular kinds of explanatory states that psychologists employ respond to equivalence-classes of inputs at a particular level of description (Pylyshyn 1984).

For example, a state that responds to ‘edibility’ is an appropriate level of description for a psychological state. This is why psychological processes do not begin at transduction and why psychologically useful states require information processing. Only after processing can a state respond to an equivalence-class such as ‘edibility’ that is relevant to psychological explanations.

The explanatory success of psychology strongly suggests that states with particular kinds of equivalence-class responses are valuable for capturing certain generalizations about human behavior. States with these kinds of equivalence-class responses capture generalizations that are “lost in the diversity of [lower-level] causal connections” (Pylyshyn 1980, p. 112). Relevant psychological generalization, say in case of categorical perception, ignores multiple physical, biological, and behavioral differences that obtain. This is illustrated, for example, when subjects hear a noisy version of what they report as the acoustic signal ‘*ba*’.

The right level of generalization allows psychologists to explain particular behaviors. For example, although physical and neurological details explain how some of the dynamics of an incoming signal are processed, it is the characterization of the input as ‘*ba*’ that explains what subject’s report hearing. The important thing for psychological explanation, in the case of categorical perception, is that various kinds of information, in terms of physical or neurological details, can be ignored in favor of accounts that employ states with psychologically relevant categorical information.

As noted above, it is the particular kind of code of the initial sensory information that differentiates perceptual and cognitive states from sensory states. These differences are illustrated, for example, in the states that respond to the Müller-Lyer illusion. Explanations that employ states that respond to certain equivalence-classes of information illustrate psychological explanation (Pylyshyn 2003). Psychological investigation is thought to identify certain kinds of information-bearing states, such as those that respond to perceptual or cognitive ‘categories’, which are relevant to explaining certain human behaviors.

To briefly reconnect these issues to the broader discussion, the requirement of coding states is why it is wrong to characterize a subject’s cognitive unconscious in terms of all the information-bearing states transduced by the subject’s senses. Sensory states and other early information processing states do not count as coded psychological states, such as perceptual or cognitive states. The explanatory value of a subject’s cognitive unconscious drops off quickly if states beyond a subject’s perceptual and cognitive states are included.

Beyond the boundary of perceptual and cognitive states there is no reason to stop including other information-bearing states in the subject’s brain, nervous system, or the world. Without information being coded into computationally relevant forms the information cannot be operated on by later psychological processes. A computer’s computationally relevant states are not all the information-bearing states present in what we designate as ‘the computer’. Counting other information states as part of a subject’s unconscious loses explanatory value since there is no longer any explanatory distinction between the subject’s psychological states and all other informational states.

Coding is why, contrary to what some theorists hold, a subject’s cognitive unconscious is not composed of all the incoming sensory information present in a subject’s nervous system, or even

the states in the immediate information processing thereafter (Wilson 2002; Zimmerman 1986; 1989; Nørretrands 1998).

### 2.3 Tier Two: Stages of Information Processing Leading to Percepts

All the preceding detail regarding information, information processing, and coding establishes the connection between information and higher-level explanatory states such as perceptual or cognitive states. As already noted, perceptual and cognitive states do not always carry the same information as sensory states. The connection between perceptual and cognitive states and sensory states allows perceptual and cognitive states to be about the information states of the world.

The different ways perceptual and cognitive states carry information distinguishes them from sensory states. As we will see more in chapter three, perceptual and cognitive states have mental content. Simple information states, such as tree rings or sensory states, do not. So, distinguishing perceptual or cognitive states from sensory states is required to eventually show how perceptual or cognitive states can be representational.

The second tier of the information-processing model of human mind is perceptual processing. The way human brains get from sensory states and early information-processing to perceptual states is by stages of information processing. Stages of information processing in the human brain create coded information-bearing states that can be operated on by later processes. For example, only certain coded information-bearing states can enter perceptual operations (Uttal 1967). As we have seen, information not coded into computationally relevant forms cannot be operated on by later psychological processes. With coding, perceptual states can carry different information from that transduced by sensory systems.

#### 2.3.1 Feature Detectors

As we have seen, processing initial input is a way of repackaging information. For example,

repackaging information codes the information so that parts of the information are made more explicit. This coded information is then able to cause later operations to function in particular ways. The first stages of early information-processing are unlikely to count as perceptual or cognitive states. Some of the operations just beyond early information processing are called feature detectors (Pylyshyn 2003). The input states that enter initial feature detectors are sensory states or initial coded forms of sensory states. Some feature detectors are mechanisms that specialize in the detection of particular aspects of distal stimulation. Feature detectors respond only when information about a particular feature, or something closely resembling it, is carried by the state that is being processed.

As noted, not all coded states are representational. For example, the output from on-center/off-surround or off-center/on-surround cells forms the input to some early feature detectors. These states are an example of an early coding of information in vision. On-center/off-surround or off-center/on-surround cells respond to particular profiles of information from earlier stages of sensory transduction. For example, on-center/off-surround cells fire most rapidly when light falls on a central region of cells feeding input into one of these cells and decrease their firing if light falls on a peripheral region of cells feeding input into the cell.

Regions of input create receptive fields for, say, on-center/off-surround cells. The activation of these kinds of cells is a coded response to initial information. Receptive fields are a way of coding information for later processing. The response from an on-center/off-surround or a off-center/on-surround cell is not clearly representational. Such responses, however, are coded responses to initial input.

A feature detector can be built from initial coded information. One example is found in the classic view of detection of visual depth. Visual depth processing is thought to take the values of

light and dark variation from each eye as input. The states that are output from depth-detection cells are coded outputs of on-center/off-surround cells, for example. Depth processing takes the coded values from initial information processing, such as activations of on-center/off-surround or off-center/on-surround cells, and checks for mismatches between the intensities in the array of light reaching each eye. The output from depth detection is an initial perceptual state that can be said to be about the feature of depth.

An illustration of depth processing occurs when subjects view random dot stereograms in the right way (Marr 1982). Random-dot stereograms are two-dimensional arrays of dots where part of each array is slightly different from the other. Each eye is given a set of random dots that have most of the dots match locations relative to each eye. The mismatching area of dots is seen as if it were in depth. The output of such processing is then a coded version of the incoming information. The coded output combines information from each eye into a single perceptual state about depth.

There are a variety of different kinds of feature detectors. Another example of a feature detector comes from activations of on-center/off-surround or off-center/on-surround cells that ‘converge’ to form lines or edges. The values of variation of light and dark from each eye can form convergent projections. Such activation can be coded into outputs that are about lines or edges. In human vision, for example, the output of initial sensory processing can be input into line-detectors. Line-detectors activate when particular arrangements of light and dark, coded into on-center/off-surround or off-center/on-surround cells, can be considered lines or edges.

### 2.3.2 Operators

The output from different feature detectors can be related in further processing. For example, the outputs from several line detectors can be fed into later processing that detect ‘vertex types’. These processes operate based on points where, say, three ‘lines’ meet. The vertexes do not have

to be real since these kinds of detectors can fire for combinations of shadows and real edges.

The output from various vertex-detectors can then be combined by further operators. Often vertex types for two or three lines are interpreted as various possible ‘corners’ of an object. For example, vertex types are thought of as representing either real or drawn outlines of the corners of boxes (Pylyshyn 2003). The outputs from several ‘corner-detectors’ can be combined as well. With the right kinds of combinations, sets of ‘corners’ can be fed into ‘cube-detectors’ that fire only when a cube, or a cube-like visual outline, is present (Waltz 1975).

Operators combine coded inputs to form outputs that are further coded states. Pylyshyn notes that, “operators map...symbols onto symbols” (Pylyshyn 1984, p. 170). For example, certain operators take ‘lines’ as input and map them onto ‘vertex types’ and other operators take these latter states and map them onto states that are codes for ‘cubes’. Interconnecting feature detectors in this way is similar to the gate-stacking we saw earlier. The only difference here is that the outputs from feature detectors are more easily treated as interpretations of features, or descriptions, of stimuli. So, operators map the kinds of feature-coded states onto further coded states. As we will see more in chapter three, these kinds of symbolic operations are likely where representational content enters the information-processing explanation of the human mind.

For example, some operators combine features into states where the information fits descriptions for geometrical shape-classes (Pylyshyn 2003). Feature-detection and operators that combine the outputs of certain kinds of feature-detection are a kind of processing that goes beyond mere summarization of sensory input. Further kinds of computational constraints are required to coherently combine features to reach higher-level states, such as those that represent ‘cubes’.

Computational constraints for combining the various outputs of feature-detectors and

operators are necessary so that ‘legal’ combinations of features can result in states that are about coherent configurations of the real world, for example. Combinations of states by operators use different computational rules, or constraints, for processing than those present in early information processing. Certain operator constraints, for example, limit vertex types to those that are physically possible (Pylyshyn 2003, p. 102). Still further constraints then limit how operators can combine vertex types into ‘legal’ outlines of geometrical shape-classes.

In this way, operators can function as higher-level information processors than initial sensory information processors. The output combinations from operators can be seen as working on interpreted symbols. At this stage of processing, discrete packages of information exist that investigators can more easily label as ‘descriptions’ of distal stimulation. This is another level of processing since low-level feature detectors are not generally thought to take symbols as input. Depth extraction from random dot stereograms illustrates this point. Once depth has been extracted this coded unit can be interpreted. States input into a depth-detector are not valuably explained as representational states (Pylyshyn 2003; Waltz 1975; see chapter three).

Constraints employed for operators limit possible combinations of inputs. In this way, constraints are a bit like rules built into the visual system. Constraints, or even rules, are just limited ways operators can function. Constraints, for example, are the way operations are designed to function. They allow for ‘categorization’ of incoming stimuli that sensory systems cannot. Information-processing systems for visual perception are built in such a way that transitions between states can happen only in certain ways. This is part of the information that the visual system adds during its processing of initial sensory input.

The constraints in operators can be used to combine various interpretations of coded units of information so that ‘descriptions’ applicable to the world can be reached. For example,

constraints limit the overall interpretations for a set of ‘corners’ that can make certain shapes. The output from one ‘corner-detector’ may have several possible three-dimensional interpretations until it is combined with others. ‘Legal’ combinations of an overall set of possible ‘corners’ gives an interpretation that can ‘flow through’ and fix the set of local interpretations of each possible ‘corner’ to reach an interpretation for a single overall geometrical shape (Pylyshyn 2003).

With the right constraints and coded information the interpretations of the overall shapes of a simple scene can be fixed. As an illustration, the boundaries and surfaces of boxes in a simple line drawing can be found. This “constraint propagation” operation allows for fairly high-level analysis of the boundaries and surfaces of a visual scene and yields a perceptual ‘description’ of the world from initial sensory information.

Clearly, the initial parts of sensory systems themselves cannot provide such a ‘description’. Visual processing results in a set of states for each shape-class and their relation that is present, for example, in a line drawing. This kind of coded information processing will make explicit information about surfaces and boundaries that belong to one another. These kinds of higher-level interpretations are not available without several stages of information processing. Stages of processing from transduction through feature detection to operations, such as constraint propagation, are all required to reach the final states of perceptual processing.

### 2.3.3 The Percept and Perceptual ‘Inference’

The end point of perceptual processing can be called a full percept (Dretske 1978; Pylyshyn 2003). The operations in visual-perceptual processing, for example, result in a state that combines outputs from operations regarding surfaces and detection of depth. Visual-perceptual processing results in a state about “surfaces in depth” in a visual scene (Nakayama, He, and Shimojo 1995; Pylyshyn 2003).

A full percept is an explanatory construct which is involved in explaining both perception and hallucination. In hallucination, a percept can occur that is not currently being caused by stimulation. In such a case, the state still carries information as if it were caused by, say, certain surfaces in depth. Another example of a percept is illustrated in bi-stable visual perception, where a subject's conscious experience switches from one interpretation of a single stimulus to another. In these cases, states representing both interpretations remain active, say in the subject's brain, even when a subject cannot report one or the other percept.

A percept is the output of encapsulated early information-processing. Although there is influence from later parts of encapsulated processing to earlier parts, the processing is isolated, say, from cognitive processing (Pylyshyn 2003). We see such encapsulation from cognition when viewing visual illusions. For example, cognitive states that are informed about the illusory nature of a perception do not effect how the illusion appears to the subject.

The output from this kind of processing is thought to occur prior to object recognition or identification. Percepts can, for example, track 'objects' that are not yet identified by cognitive processes. This kind of tracking involves a state that points to, or indexes, the particular 'object' without identifying that object in a conceptual, or cognitive, way (Pylyshyn 2003). Such tracking characterizes just a part of what it means for a state to be a percept.

Although perceptual processing can be thought of as indicating the constancy of a stimulus, percept processing does not amount to object identification (Pylyshyn 2003; 2009). For example, people with intact auditory processing can track the movement of several acoustic signals in a dark room without being able to identify the sounds. People are also able to see a thing without recognizing the object. The ability to perceptually track an object without recognition can also be preserved in certain kinds of brain injury (Farah and Feinberg 2000).

People with certain kinds of brain injury complete a large portion of visual or auditory processing and can keep track of these elements over time. Such patients are not able to identify the objects causing stimulation. Hence, percepts occur prior to object recognition.

Generally, human subjects do perceive auditory and visual signals as certain kinds of things. However, this object-recognition ability is built on the basis of first forming a percept. For example, in vision, shape-classes can be tracked but do not identify the object that forms that shape class (Pylyshyn 2009). Adult humans can, say, see a table or chair as such only if they have first seen some thing or other which is later cognitively processed in terms of its identity. In this way, perceptual processing is isolated from cognitive processing and must occur prior to, say, object recognition (Fodor 2009). This treats the notions of percepts and perceptual processing in a technical way (Pylyshyn 2003).

Percepts are still highly stimulus dependent. Visual percepts, for example, do not remain active when the stimulus ‘object’ is not in sight (Pylyshyn 2003; but see Fodor 2009). Interestingly, these kinds of internal states do stay active when a distal stimulus is causing them but changing some, or all, of its properties (Pylyshyn 2009, p. 12; 2003). So, percepts track the sameness of a stimulus but not in terms of the identification that cognitive processes would form. For example, perceptual processes do not attach a label to an ‘object’ as a certain kind of conceptual thing. In vision, a percept can track an ‘object’ through changes in its location, shape, and color (Pylyshyn 2001). Tracking of this kind does not identify the object in conceptual terms.

Percepts of different sensory modalities are built fairly independently within each perceptual system. Full percepts are the highest ‘interpretation’ produced by perceptual processing of transduced information (Pylyshyn 2003). States of this kind are highly coded.

Visual percepts, for example, are the three-dimensional interpretations of, or ‘descriptions’ of, the two-dimensional retinal input picked up at the eyes. Visual percepts are employed, for example, in the case of viewing a line drawing. Because many two-dimensional projections can cause the same percept, visual percepts no longer co-vary one-to-one with proximal sensory information.

By this stage of processing the information is packaged in such a way as to make it ready for cognitive identification as well as other cognitive operations. However, percepts have not yet been delivered to cognitive operations and so are not categorized in conceptual terms.

Creating visual scene ‘descriptions’ requires complex information processing that is based on various kinds of constraints on how information can be combined. Sometimes, the processing that leads to a percept is thought of as a kind of inference, so-called ‘perceptual inference’ (Hatfield 2002). For example, the input states of feature detectors have been treated as premises, the processing as rules of inference, and the output as the conclusion of an argument (Fodor and Pylyshyn 1981).

However, the processing that creates a percept is not doing inferential reasoning in the cognitive, or rational, sense of the word (Pylyshyn 2003; Kanizsa 1985). Pylyshyn notes three key differences show the distinction between the operations that create percepts and true cognitive inference. Perceptual processes are more isolated than cognitive processes, they sometimes have non-rational results, and perceptual processes operate on different states than those in cognitive, or rational, inferences.

For example, operations that create percepts have access to only a limited kind of information, they operate on this information in only fixed ways that rational inference would sometimes treat in other ways, and perceptual processes do not operate on beliefs and desires, the

typical states of rational inference (Pylyshyn 2003, p. 68; Stich 1978). It is better to think of perceptual information processing as applying ‘natural principles’, or constraints, to inputs rather than applying rules of inference to premises (Pylyshyn 2003).

Percepts provide the kinds of states required for cognitive operations. For example, perceptual descriptions result in states that represent perceptual constancies, perceptual categories, and perceptual saliencies. Without this kind of stabilization of incoming sensory input cognitive operations would be awash in varying details of information that impact sensory systems (Pylyshyn 1980). So, perceptual processing forms the stable ground upon which cognitive operations perform (Fodor 2009). Likewise, the tie to sensory information helps to ground the information in perceptual states, and thereby cognitive states, so that these states do not completely uncouple from the real world (Harnad 1990). In this way, perception helps to ground the meanings of cognitive symbols that would otherwise be interpretable in multiple ways.

#### 2.4 Tier Three: Levels of Cognition

Tier-three states and processes are what are considered cognitive. Tier-one processing operates on sensory states, or mere informational states. A fair amount of information-processing results in tier-two states and processes that can be considered perceptual states. Perceptual states are significantly processed compared to sensory states but are not considered cognitive states. The point at which cognitive processing is thought to occur is when a percept is made available to certain operations, such as recognition, recall, learning, or kinds of rational inference. When this occurs the states treated in these ways have cognitive function and can be considered cognitive states.

There are many forms of cognitive operations. Memory for, or recognition of, a percept, for example, are two forms of cognition. Other states, such as beliefs and desires are also cognitive

level states. As Sutherland notes, one of the defining features of cognitive states is “control over behavior by an event or a relation between events that is not now present” (Sutherland 2005, p. 422). Specific kinds of behavioral control, often when the stimulus is not present, form cognitive states and operations. That cognitive states control behavior often when information is not present distances such states from the information carried by sensory, and veridical perceptual states.

Pylyshyn, among other theorists, endorses an encapsulation view of the processes that create percepts and cognitive states and operations. Encapsulation of perceptual processing has been called modular processing (Fodor 1983). Perceptual processing, for example, is supposed to end prior to object recognition, which is a kind of cognitive operation (Pylyshyn 2003; Kanizsa 1985; Nakayama et. al. 1995). Dretske also notes that “the field of perception can be said to lie between the field of sensory processes on the one side and of cognitive processes on the other” (Dretske 1978, p. 107). Dretske goes on to state that cognitive processes “begin with the perceived object as a given and tend to concentrate on such processes as recognition, recall, association, attention, understanding, problem solving, and thinking” (Dretske 1978, p. 107). This is a standard way to formulate the three-tiered information-processing model.

Perceptual tracking, for example, can be done without applying cognitive individuation criteria. This kind of tracking state is required for cognitive processing. One way to think of the difference between the operations that build percepts and cognitive operations is in terms of how these two types of processes might ‘classify’ a single external event. Perceptual processing in the visual system, for example, might treat a particular kind of input as the same “object” when cognitive operations would classify this input as several different objects. For example, a plane, an ejecting pilot, and finally an open parachute might be ‘classified’ as a single percept

(Pylyshyn 2003, p. 214). Again, perceptual states that respond to a set of stimuli as an equivalence-class are not necessarily classifying the stimuli in cognitive terms. There are clearly equivalence-class responses in perception. These perceptual ‘classifications’ can be treated differently by cognitive processes.

For example, perceptual and cognitive processes can sometimes come to different ‘conclusions’ regarding the same stimuli. So, these two kinds of processes have different “individuation criteria” (Fodor 2008, p. 176). “Individuation criteria” are often thought to apply only to cognitive states, such as seeing a stimulus as falling under a particular conceptual classification. However, perceptual states do refer to entities as “the same” without those ‘classifications’ being based on conceptual categories (Pylyshyn 2001; Harnad 2005; Pacherie 2000). In this way, perceptual states seem to be individuating external world stimuli as some kind of ‘perceptual individuals’ (Pylyshyn 2003).

Conceptual categories can disagree with perceptual ‘conclusions’ about certain stimuli being “the same”. Conceptual categorization is one way cognitive states respond to equivalence-classes of stimuli. The equivalence-class responses of cognitive states do not always match those of perceptual states. For example, perceptual constancies, perceptual categories, and perceptual saliencies are not always treated as cognitive ways of classifying stimuli. So, it seems that perception and cognition are applying different kinds of individuation criteria.

In whatever way we characterize the difference between perceptual tracking and cognitive categorization, there is clearly a difference that importantly separates cognitive states from perceptual states. For example, one object that is cognitively known to be singular can be perceptually seen as several “objects”, or some stimuli that are seen as a single “object” can be cognitively known to be several different objects. This distinction shows how perceptual and

cognitive states can differ. If we use the term “individuation criteria” for this difference, then different individuation criteria seem to be applied, for example, in the case of the parachuting pilot and his plane.

Percepts can tell us that a thing is present but not what it is (Fodor 2009). Cognitive states tell us what a thing is. In this sense, object identification is where percept processing leaves off and cognitive processing begins. Likewise, as we saw in the case of the Müller-Lyer illusion, subjects can have perceptual states that suggest the lines are of different lengths while cognitively knowing the lines are of the same length. The percept is not changed by the contrasting cognitive states. Subjects can cognitively re-interpret the status of a visual percept, for example, but cannot change whether the visual system gives one kind of percept or another. So, cognitive states are not percepts. Cognitive states, generally, have limited impact of the content delivered by the perceptual system.

There are several levels of cognitive processing. The first levels of cognitive processing are memory and object recognition. Cognitive operations of these kinds occur just after percept processing. Seeing a stimulus as falling under a certain category, for example, can begin only after a stable percept has been built and passed on to later operations. Learning is another simple form of cognition. Being able to learn certain kinds of conditioned responses to stimuli requires that they are first perceived in stable ways. Learning and conceptual classification are considered cognitive particularly because the stimulus does not need to be present for these cognitive operations to function properly. Higher-level forms of cognition are also possible. These higher forms of cognition include operations such as cognitive inference.

Examples of each of these kinds of cognitive processes will be treated in turn. Given the amount known about various cognitive operations, sections below will only provide the briefest

sketch of a few forms of cognition.

#### 2.4.1 Low-Level Cognition: Object Recognition, Kinds of Memory, and Kinds of Learning

The output of percept processing forms the input to cognitive processing. One of the first stages, or lowest-levels, of cognition is object recognition. The first thing required for object recognition, in the normal sense, is that the state referring to the object remains active even when the object is no longer present. Although it is logically possible that a creature only recognizes an object when the object is present this is not a general kind of cognitive state. The point is that object recognition is a further process than building a percept for the object. This is true even in the logically possible case where object recognition occurs only when the stimulus is present. When internal tracking states function even in the absence of the stimulus such states go beyond percepts (Pylyshyn 2003; Sutherland 2005).

Object recognition can occur, for example, when a visual stimulus goes behind a rock but a tracking state remains active and also tracks the stimulus as the same when it returns a short time later. If new states are tokened within the system each time the object disappears and reappears then the states needed for normal object recognition are not present. This simple kind of memory and object recognition requires comparison of a current and a past percept. In this way, object recognition is beyond simple percept processing.

Think of this kind of case in detail. In each case of perceiving an object, say before and after it has disappeared behind a rock, a new perceptual state is created. We do not perceive the object when it is out of view. However, we can remember the object and we can classify it as the same object when it returns to view. This kind of memory and classification require that a single state treats various perceptual inputs as the same.

In the case of seeing an object and then classifying it as the same when we see it again, a single state treats two percepts as the same. In the case where mnemonic operations are

involved, a single state treats a percept and, say, an internally caused mnemonic state as the same. States that have equivalence-class responses such as this are generally considered cognitive since the classification or mnemonic operations do not require the object to be present. This means that the information provided by the object does not have to be caused by the object at the time of cognitive processing that is relevant to the object. Such processing differentiates cognitive from sensory and veridical perceptual processing. Cognitive states are, generally, independent from the need for information to be currently caused by the stimulus the cognitive states indicate.

In the case where two percepts are treated as tracking the same object we can see several cognitive operations in play. For example, object classification, recognition, and short-term memory (STM) are all demonstrated. An internal state that tracks a distal stimulus when it is not currently, but has recently, caused stimulation, allows a subject crude forms of classification and memory. Reverberation of activation of a state in connectionist neural nets, for example, could maintain a state in active memory. This state could then be compared to another incoming percept. Finally, if a mnemonic and a perceptual state are treated as relevantly similar, a crude form of object recognition, or cognitive classification, is demonstrated since the cognitive state responds equivalently to the mnemonic state and the new percept. States that treat two percepts, or a mnemonic state and a percept, as equivalent require a kind of coding beyond that needed for building a percept. States with equivalence-class responses to percepts, for example, go beyond the perceptual stage of information processing.

When a state is reused to classify a stimulus as the same thing over longer periods of time long term memory (LTM) is demonstrated. LTM is a cognitive capacity to store states beyond those that have recently been active. So, mnemonic states of either STM or LTM are cognitive

since they carry information about an object that is often no longer present. Memories are not like percepts in that they do not indicate the present occurrence of the object or property they carry information about. Likewise, they often generalize the encoded information so that only the gist of, but not the exact information regarding, the original signal is retained.

Other forms of low-level cognitive states are illustrated by certain forms of learning. Think of the elements in classical conditioning, for example. When a rat is trained on a certain object, or feature, and when this training is associated with reward, the combination of these two elements creates a discrete unit that indicates reward in the presence of the object or feature (Rescorla 1988).

Different associations can be shown to control behavior in differing ways. The important thing is that learned states can control behavior even when the stimulus they are about is not present. For example, it takes longer for certain systems to learn a combination of states that does not give reward, AB-, when they have already been trained that each element does give reward, A+ and B+ (O'Reilly and Munakata 2000). Independent training on the rewarded elements requires more training when previous, non-rewarded, training states are present. This extra training time suggests the previously learned states are interfering with later learning even when the non-rewarded situation is not present.

The above example shows that the states of the past learning are controlling behavior even when what they are about is not present. This shows a kind of cognitive processing because the presence of the non-rewarded states is no longer needed. If the state remains active for information that is not present it is clearly being operated on by processes that are not sensory. Likewise, learning is not generally considered a kind of hallucinatory perceptual state. In this way, learning states can be considered cognitive and distinct from sensation and perception.

In certain kinds of learning, percepts from different sensory modalities, such as visual and auditory percepts, can be associated. Cross modal learning associates percepts from two modalities only after they have been constructed by each perceptual system. When a system links the percept of a light and that of a sound, say, in stormy weather, this can lead to adaptive behavior. The behavior of the system can be controlled by a state even when what the state is about is not present. For example, a creature can hide in anticipation of thunder simply from seeing lightning. In this way, one of the associated stimuli can come to stand in for, or predict, the other before what the second state is about becomes present. Adaptive behavior that occurs before the on-set of a second stimulus reveals a cognitive level operation since the information from the second state does not have to be present in the environment for it to control behavior appropriately.

Other forms of simple learning are also possible. For example, learning can link a current percept with a distant memory. In other cases, learning is not dependent on current perceptions at all. For example, two mnemonic states can be associatively linked. There are, therefore, many kinds of simple learning that can be considered operations on cognitive states since they go beyond percept construction and do not require current information present in the environment to function properly. These are just a few of the kinds of low-level cognitive operations available to various information-processing systems.

#### 2.4.2 Higher-Level Cognition: LoT, Inference, and Abduction

Higher forms of cognition are not limited to object-recognition, or simple forms of memory and learning. Higher-level forms of cognition, such as decision making, are also of interest to cognitive scientists. Several of these higher-level forms of cognition seem amenable to the symbolic information-processing approach. For example, some higher forms of cognition, such as problem solving in chess, are often computed by purely syntactic formula that can mirror

human logical inference (Newell 1982). These models often use semantically interpretable symbolic units. These symbolic units relate to one another in a way that is abstractly similar to that of chess pieces on a board.

Symbolic cognition requires particular kinds of information processing. Language of Thought (LoT), for example, is a computational means for preserving certain relationships between internal states. LoT is not only a theory about how parts of the human mind work. It is a viable way to implement certain kinds of information processing. A LoT system can be thought of in terms of logically structured internal states. For example, a state can be coded with three parts, XYZ, and the parts of this unit can then be operated on in certain ways. Transitions that mirror semantic relations, such as truth-preservation, can therefore be carried out. For example, if we know that one thing is true and that another thing is true, a formula representing both can be created that preserves what we know about each element.

Other relations can be preserved in LoT systems, as well. For example, there is a relationship between a subject thinking that the sky is falling and fearing that it is falling. Some internal state must preserve the relation that what is being thought and what is being feared are the same thing. The concept FALLING, for example, is linked to the concept SKY, and this whole unit can then move between contexts where the content is thought about and then feared.

Certain forms of cognition require this kind of explicit preservation of the relations between the parts of internal states. Although LoT is not considered a natural language, the metaphor to natural language is apt. LoT states have parts in a sense somewhat like written or spoken sentences. Like language, there are rules for the combination of the parts of LoT states. For example, written sentences can have the same parts and the parts can be interchanged. Such systems can have states such as 'man bites dog' or 'dog bites man' and these states will mean

different things based on rules for how the parts are combined.

Based on the structure internal to LoT states, a particular kind of information processing can be accomplished. LoT based information processing allows for an explanation of several kinds of cognition that other kinds of information processing cannot explain. For example, LoT systems are compositional, systematic and productive (Fodor and Pylyshyn 1988).

Roughly speaking, the idea of syntactic compositionality is that complex states are built up from parts and rules for their combination (Fodor 1987, p.138). It is highly unlikely that there would be a normal LoT system that could have an internal state made up of several parts, for example, without being able to also operate on all of the parts of this state individually.

Likewise, a LoT system can have states with the parts organized in various ways. As we saw above, such systems can have states such as ‘man bites dog’ or ‘dog bites man’ and these states can mean different things even though the parts are the same. This kind of systematicity in operations of states is explained by LoT systems.

Finally, LoT systems are often thought to be productive. That is, LoT systems have, at least in principle, unbounded capacities in regard to compositionality. What this means is that LoT systems, in principle, have the ability to compose and decompose arbitrarily complex internal states from various parts and rules for their combination. Such rules will make relevant computational differences for conjuncts such as “AB” and “BA”.

Logical inference is another form of higher-level cognition. LoT states, with certain additional rules for combining states, can help to explain rational transitions between internal states. LoT systems with the features of productivity, systematicity, and compositionality do not guarantee rationality. Although LoT systems need to acquire, or be specially built with, rules for rational transitions, structured LoT states are still valuable parts of explanations for rational

behaviors.

Deductively valid transitions, for example, are the simplest kinds of rational transitions. Transitions in chains of deductive reasoning are based on certain rules and the comparison of internal parts of LoT states. Deductive transitions are well explained by LoT states with the right rules for combining parts of internal states. Likewise, more complex rational transitions, such as those involved in induction, may be accounted for by LoT systems. To the degree that inductive transitions can be explicitly formalized by rule-governed interactions between states and their parts, then induction can be explained by LoT systems, with the right additional rules, as well.

Not all forms of higher-level cognition may be amenable to a LoT approach. Abduction is one example. Abduction requires finding relevant relations among a large variety of cognitive states. Finding relevant relations, for example, is likely based on the ability to draw information from cognitive states throughout the entire cognitive system (Fodor 1983, p. 105; 2000). Systems able to find relevant relations among a large number of cognitive states seem to accomplish this feat based on sensitivity to properties of the entire set of cognitive states possessed by that system (Fodor 1983, p. 107).

Abductive cognitive operations do not fit LoT models since LoT models are often based on 'local computational profiles'. This means that LoT computation works on states and their parts and are not sensitive to the properties of the entire set of cognitive states present in the system. Abductive cognitive phenomena require sensitivity to the information present in a large part, if not all, of the cognitive states in the information processing system. Although, non-local cognitive phenomena, such as abduction, present problems for LoT information processing, they are thought to be explainable by other kinds of information processing (Fodor 2008). For example, certain systems that combine symbolic and connectionist processes may help to

account for abduction (d'Avila Garcez, Gabbay, Ray, and Woods 2007).

The important point from the above is that stages of cognitive processing go well beyond sensation and the forming of veridical percepts. Although the focus of much cognitive science research, such cognitive states and processes likely form only a part of a subject's overall mental architecture.

### 2.5 A Brief Note on "Cognitive" Architecture

An important distinction in the three-tiered information-processing model of mind is the distinction between processes and architecture. Processes, as we have seen, are various operations on particular kinds of states. An information-processing architecture is simply the interrelation of various information-processing operations, such as perceptual and cognitive processes (Vaina 1990; Aizawa 2003; Ebdon 1993). There is no generally agreed definition for "cognitive architecture". So, setting out the process-architecture distinction in this way may not conform to other ways of setting out such a distinction (see Pylyshyn 1980; 1984). However, this is a simple enough way to formulate the distinction for our purposes.

In the case of the three-tiered model of human information processing, the interrelation of psychological processes make up the "cognitive architecture". However, the architecture of interest to psychologists is not limited to cognitive processes. The interrelation of perceptual processes and cognitive processes each form important parts of the information-processing architecture of human subjects (Fodor 2009).

We can imagine a "cognitive architecture" in terms of neurological circuits. Adult human brains are divided into areas of processing and pathways that connect these areas in specific ways. One possible architectural relation may be that certain areas of processing are more peripheral than others. For example, perceptual operations, motor control, and language processing may be peripheral operations linked to more central problem-solving operations

(Ebdon 1993; Fodor 1983). Many theorists endorse, for example, a cognitive architecture built of specialized mechanisms in interaction with more general purpose mechanisms (Goodale and Arbib 1998; Fodor 1983; 2008; however, see Carruthers 2006).

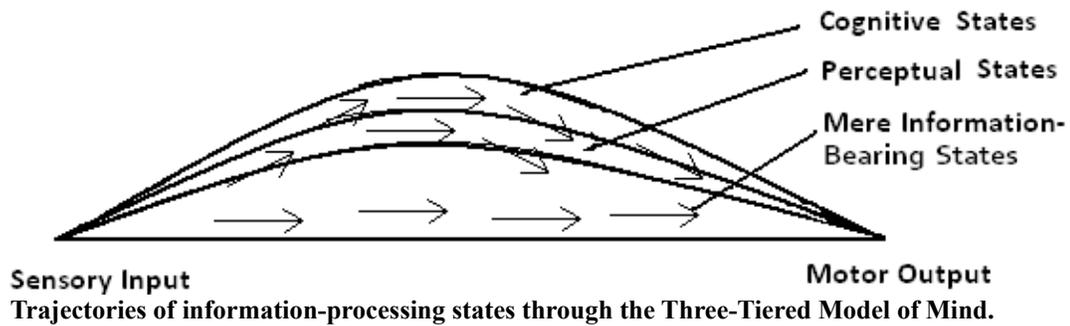
What does matter is that after maturation, the organization of information-processes in adult human brains is thought to be fixed and universal (Pylyshyn 1980). That is, the relation of memory and attention, say, is fixed in that there is a particular relation between these processes that does not alter significantly over time. Likewise, this organization is supposed to be the same in all average human beings. So, for example, human subjects can track only a few moving items in a large set of moving items and can remember only a certain amount about these items.

The architectural organization, or relation of processes, is not altered when subjects gain knowledge, for example. Even gaining a new capacity, such as reading, leaves the architecture organization, or relation of processes, generally fixed. Empirical studies suggest that there is both a limited amount of resource and limited number of pathways that define the human information-processing architecture. This kind of architectural limit constrains the kinds of states that have access to various processes. An information-processing architecture constrains how, and even whether, certain states are accessed by certain processes or coded in certain ways. Not all the information in an information-processing architecture is available to all processes.

We can think of the architecture of the three tiered model as defining trajectories information will take through the system. For example, we can imagine information-processing trajectories in terms of something like figure 2. This is a very rough way of thinking of information-processing trajectories. However, figure two will serve the purpose of illustrating the point about information-processing architecture.

In this figure there are divisions of processing between sensory information processes,

perceptual processes, and cognitive processes. Very roughly, sensory information can be thought to start at the left-hand side of the figure and motor output can be thought to end at the right-hand side of the figure. Within these divisions, we can imagine certain trajectories sensory information might take on the way to controlling motor output.



Given the information-processing architecture of the human mind, the trajectories of

information will have various paths through figure 2. Some sensory information will directly control behavior, undergoing minimal information processing before it gains control over certain motor output. Such information would control behavior directly without being operated on by perceptual or cognitive processes. Other trajectories of information-bearing states that control behavior pass through perceptual processing, going on to control behavior without any cognitive processing. Finally, some behavioral control would require information to pass through all the tiers to reach cognitive processing. The architectural design of the information processing system would determine such trajectories.

The kinds of limits architectural constraints impose on trajectories information can take are important in defining the three-tier information-processing model. For example, defining the input and output of perceptual processing has direct bearing on the thesis that perceptual processes are independent of cognition (Pylyshyn 2003, p. 125). Likewise, architectural divisions limit the access cognitive processes have to certain kinds of information. Architectural divisions regarding which states can be coded in certain ways will also be important later on

when trying to define conscious and unconscious mental states.

## 2.6 Summary and Conclusions

There is much more to say about the information-processing framework. For example, motor output of various kinds has not been addressed. Likewise, many of the architectural relations, for example between perception and motor control that bypass cognition, have also barely been addressed (Milner and Goodale 1995; Pylyshyn 2009). Suffice it to say here that there is much more to the information-processing framework. However, we have enough of a picture of this framework in place to scaffold later discussion. This discussion will be important, for example, when we discuss where mental-state, as well as conscious mental-state, explanations begin in the information-processing framework.

The present chapter has provided just a sketch of the information-processing framework which can be used for later discussion regarding mental states and the division between conscious and unconscious mental states. The information processing approach can be roughly divided into three “tiers” of information processing. As we have seen, these tiers include sensory transduction, perception, and cognition.

### CHAPTER 3: What and Where Are Mental States in the Information-Processing Model?

This chapter addresses two questions. First, what are mental representations? Second, where is mental-state explanation best applied in the three-tiered information-processing model?

General information-carrying states are not considered mental representations. Likewise, early information processing is not generally considered an operation on mental representations. So, we need to see just how early information-processing states differ from mental representations and where mental-state explanations best enter the information-processing model. A majority of cognitive scientists hold that mental states have a particular explanatory value and first appear somewhere after the first-tier of information processing.

#### 3.1 The Differences between Representation and Information

States that carry information about a particular thing or property are not like representational states about that thing or property. Representational states about a thing or property carry information in the way that all states do. However, representational states differ from mere information-bearing states in a number of ways. For example, representational states are explicitly focused on a particular aspect of the information they carry. Likewise, representational states can correctly occur in a system even when they are not caused by the thing or property they are about. So, although representational states must be grounded in information they function in different ways from mere sensory information states. Finally, representations can misrepresent.

To have these differences, while still being linked to information, representational states must be both coded for an explicit aspect of information and play different causal roles from those played by simple information-bearing states. The first instances of mental states can best be thought of as instances of non-derived representational states (Crane 1998). Non-derived representational states are those that do not depend on other representational states for their

content. Standard psychological theories, when using mental-state explanations, rely on the notion of information, especially the notion of coded information discussed above. So, here we just need to see the unique aspects of mental states used in psychological explanation that differentiates them from other coded information-bearing states (see Fodor 2009; Pylyshyn 2003; 2009).

### 3.1.1 Focus of Information

A standard form of the information-processing model of mind identifies the mental content of a state with the regular cause of that state (Fodor 2009; Pylyshyn 2009). This links mental states to information. Generally, however, various kinds of states carry an enormous amount of information. For example, a single picture will carry information about several shaped surfaces, several animals, several giraffes, one family of mammals, as well as channel conditions along which the information passed (Fodor 2008, p. 177). Singular representational states are focused on just one aspect of the information available. To focus on another aspect of information, say of the same scene, requires another representational state. That is, of all the information that a state carries, some part of that information must be made explicit for a single representation to have determinate content.

To be focused on one aspect of information, a state must be coded so that this one aspect is made available to later operations. Coding, as we have seen, is a process which physically shapes states. Shaping states allows for a focused aspect of the information to be made explicit for later operations. For example, specific kinds of coding will allow an explanation employing this state to go forward in terms of the content of the state, rather than all the information the state carries.

An analogy helps to see how this works. Information states can be thought to have many hooks or barbs, each hook or barb linking the state to various aspects of the information causally

related to the state. Coding can be thought to “pull out” one of these hooks so that later operations can work on just this aspect of the information present in the state. Fodor calls the focused information coded into a state the information *recovered* from a state carrying massive amounts of information (Fodor 2008, p. 180; see also Marr 1982).

Certain ways of coding information will result in focused information in a state being made physically explicit so that later operations can work on only that aspect of information. We can imagine physically shaping, or coding, a state in terms of neural activation. Sharpening a sensory activation pattern in certain ways can make some of the information in the state explicit (Marr 1982, p. 12).

For instance, to make a state that is about the shape of the giraffes in a picture, rather than the mammals, the family of animals in the same picture, or any of the channel conditions along which the information traveled, an informational state must have this one aspect of its information physically coded, or made explicit, in a way that later operations can use. One way this can work is in the case of extracting form from visual sensory information. Some of the information processing in the visual system filters input in a way that causes aspects of the information in the state to be made salient. Light and dark saliencies, or grey-coding, for example, that is extracted in this way can be used to extract lines and eventually visual forms (Sterelny 1990, p. 67). When this kind of coding is done a state can be said to be about visual form rather than any other information in the state.

As noted, simple coding does not amount to making a state a mental representation. We can see this in the case of coded states of early visual processing or even the states in computer programs. The states of computers, for example, are coded for various operations and so have focused information. Although these states are described as having focused information they are

not generally considered mental representations. Although coded, any content computer states carry is only derived from the intentions of the computer programmer. So, a theory of representation does not underlie all forms of coding. Since not all states that are coded forms of information are *mental* representations, particular kinds of coding are required for mental representations to exist.

For a state with focused information to become a mental representation it must not be dependent on an outside semantic interpretation. To have original content, a state must be initially coded by processes that are not themselves mental. This will result in the first instance of mental representations.

For example, original content likely exists in states that mediate certain kinds of animal perception. This fits a fairly widespread version of the information-processing picture of the human mind. Percepts are coded states about particular aspects of the world that do not require other intentional states for their coding (Pylyshyn 2003; 2009).

So, the first instance of mental content, or non-derived mental representations, is an atomic, or primitive state, such as a percept. Atomic representations are never composed of elements that are themselves representational. A representational state about a giraffe, for example, is a complex state constructed from atomic representations of shape, color, and distance. Likewise, the first instance of mental representations cannot be about non-existent entities since representations need to be, at least initially, grounded in information (Fodor 1990a, p. 101). Veridical atomic percepts, therefore, are the most likely ground upon which all other representations are built.

Since no state can, in the first instance, carry information about something that was never present we can never perceive non-existent things. However, we do imagine and think about

non-existent things. When we imagine non-existent things the information in these states is dependent on previous experience. For example, imagination combines previously experienced content. Such combinations allow us to form mental representation of non-existent things. So, one way to fill out the information processing view in regard to representing non-existent things is to say that imagined contents are complexes whose parts have actually been experienced (Rupert 1999, p. 351, n. 34).

The need for focused information necessary for representational content is shown by the role focused informational states play in psychological explanations. To explain general navigational behavior, for example, high-level causes are invoked to explain regularities in a system's behavior. When explaining navigational behavior we pick out, for instance, a rat's seeing a route as explanatorily relevant to the rat's general navigation ability. Theorists think the equivalence-class responses called 'seeing', and not, say, the exact brain state, are more 'proportional' to explaining navigational behavior. Furthermore, seeing was enough to explain the general navigation abilities since nothing more than this, and some motor behavior, was required to explain the navigation (Yablo 1992, pg. 277). Seeing creates a kind of focused informational state relevant for explaining certain behaviors.

As a further example, think of explanations of particular learned behavior. Once trained in certain ways, an internal state can guide a rat's behavior in a T-maze. Learned states can guide a rat's navigational behavior in ways simple seeing cannot. For example, seeing can guide the animal down one long hall to a T-junction and control the animal's behavior at this choice point. A rat in a T-maze can then be trained to focus on information about how a rewarding situation comes about. Say the animal was rewarded when it ran down the long arm of the maze and made a right turn into a white arm. In such cases the rat receives information that right turns are

rewarded, that going into the white arm of the maze is rewarded, and that going to this *place in the room* is rewarded. This information about place is independent of any cues within the maze, such as right turns or the color of the maze arms (Sutherland 2005).

Training rats in certain ways creates focused informational states that guide the rat's behavior. Certain kinds of training result in a state about the location in the room, rather than other cues inside the maze. By moving the entire maze experimenters have found that rats with certain training choose inter-maze routes which have never resulted in reward, such as making certain turns or going down certain colored hallways (Sutherland 2005).

When certain kinds of behavior needs explained, such as direct navigation through a new part of an environment, more detailed causes need to be invoked than, for example, simply seeing the route. More detail allows the explanatory causes to remain proportional to the more detailed behavioral effects being explained. Theorists posit a particular aspect of information in the learned state as the cause for direct navigation through a new part of the maze.

For example, to explain navigation in new parts of an environment, theorists often account for a rat's navigation by invoking a state with focused information, such as a cognitive map. A state with this kind of focused information helps explain the rat's ability to directly navigate through parts of an environment it has never been trained on. In these cases, the information about place, rather than inter-maze cues, directs the rat's behavior. Rat navigation is just one example where focused information, or mental content, is explanatorily valuable.

Think of a symbolic information-processing explanation for certain human behaviors. Such cases are another example where focused informational states are used. Imagine a person who has perceived a situation as an emergency. A rule such as "if in an emergency then call for help" allows us to predict and explain what a person will do in this situation. Imagine the physical

situation where the person who has noticed the emergency has also noticed a phone nearby and dialed 9 and 1. What is going to happen next (Pylyshyn 1984)?

Physics would be able to say, at a micro-structural level perhaps, exactly what will occur in terms of the movement of micro-particles. However, other predictive accounts can be pitched at the mental-level. Such predictions only go forward based on states with a particular focus of information. The person will perform the action “dial 1” because this action completes the overall desire of calling for help. The state with this explicit information captures what will happen in the face of a variety of different physical ways this action can occur.

When we say that a system perceived, recalled, intended, or acted we ignore the unlimited number of physical ways in which information can be acquired and the unlimited number of physical ways this information can control behavior (Pylyshyn 1984, p. 11). A person can learn that dialing 911 is a way to call for help from a friend, a sign, or a book and can dial the number with their finger, toe, or tongue. In situations that people perceive as emergencies they signal for help rather than flex their arms and legs. In this way, information states focused on just the information regarding signaling for help account for the behavior. The information in such states is focused well above states that carry neurological or micro-particle information. Such focused informational states capture the generalizations psychologists employ when accounting for human behaviors.

In this way, cognitive psychological explanations are, generally, autonomous from explanations in other kinds of disciplines, such as physics or neuroscience. Psychological explanation is autonomous because the equivalence-class responses of states with content are not those found in other kinds of explanations. Explanations that involve states with content, for example, need only assume some mechanism or other controls the behavior in question.

Psychological explanations of this kind do not have to worry about the details of how this task is carried out physically or neurologically. The person's behavior in our example of an emergency is perfectly well explained, for example, by a state focused on the information "dial 1".

Explanations of behavior controlled by mental representations need the external relational properties that are essential parts of natural information. Equivalence-class responses linked in the right way to sensory input allows psychological states to be about the world. Such a link is needed to explain how a set of formal transitions between internal states constitutes part of a psychological capacity. Mental-state explanations can proceed only when the internal information-bearing tokens are linked to a regular cause in the right way. Representational states can then be said to help control coherent behavior in a particular environment (Egan 1995).

Although a necessary part of psychological explanation, formal transitions between internal states do not provide a complete explanation of the behavior in question. Only when we have focused information, or content, is the psychologically driven behavior fully explained (Peacock 1994a; Egan 1995). The focused information, or content, is being used to explain particular behaviors that would otherwise remain unexplained. For example, in the case of trained rats, content is playing an explanatory role, rather than environmental cues, since the cues may never have been experienced before. With the right kinds of equivalence-class responses a level of explanation employing mental content can be applied when explaining what we take to be psychological behavior.

### 3.1.2 Expressive Role

Another typical function of representational states is to control behaviors that "express" the focused information, or content. Generally, expressing content occurs when behavior is controlled by the content in a state without the information the state is about being present. That is, when a state expresses content it need not be currently caused by what the state is about.

Expressive uses of content are correct uses even when the representational state is not currently caused by what the content of the state is about. Expressing content is partly a functional notion that helps to explain various kinds of behavior. For example, behavior directed toward information, when that information is not presently causing the state controlling the behavior, is valuably explained by mental representations (Sutherland 2005; Tolman 1948).

Expressing content is not the same as having content. Expressing content is a functional notion tied to, at least potential, behavioral control. Content is dependent on a state being caused by original information (Fodor 1990b). To have content the state must be connected to veridical information in some way or other at some time or other. However, behavior controlled by informational states is not always best explained by what is currently causing the state (Fodor 1990b). For example, we are reminded of certain contents by the most diverse stimuli. Behavior controlled by mnemonic states expresses the content of the mnemonic state even when the mnemonic state is caused by something other than what its content is about.

Often, when no sensory information about a thing is causing occurrences of perceptual or cognitive states about that thing we can see the content is still expressed in certain behavior. Take expressions of perceptual content, such as hallucinations. A perceptual state not currently caused by the information it is about will still be operated on by later processes as if it carried information about what is present. This is because behavior can still be controlled by hallucinatory states. When Hamlet hallucinates that a dagger is before him, and reaches for it, the hallucination controls where, and how, he reaches. When operations work on states, say in terms of their coded shape, the actual cause no longer matters to these operations. So, reaching can be guided by hallucinations. Later operations are isolated from how the state is currently caused. However, later operations can still control behavior based on the coded shape a state has

rather than its actual cause. Such behavior expresses the content of originally perceived information.

Later operations, internal to the information-processing system, merely operate on active states based on their coded shape regardless of how the coding came about. Explanations of behavior controlled by psychological states still require content but the expressive function is explanatorily valuable too. The behavioral control system operates formally, for example, on the shape of states alone and so behavioral-control states are isolated from their causes. A state that is not currently caused by what it is about can control much of the same behavior that states with currently veridical information can. The state in question is treated as having the same focused information, or content, as a state which is caused by what it is currently about. Such treatment is due to the form, or coded shape, of the state and how operations treat that state.

The case of mnemonic states controlling behavior is another example of expressing content. Again, mnemonic states must initially gain content from veridical information at some time or other but this information does not have to be causing the mnemonic state for the state to control behavior based on that information. Of course, remembering does not amount to hallucination, so the information in a memory state needs to control other kinds of behavior. When Hamlet hallucinates that a dagger is before him and reaches for it, his grip size is not that for grasping a pencil. His grip size is likely based on past veridical information with daggers. This is true even though information from an actual dagger is not currently causing the state controlling grip size. To express the content ‘as if’ he were reaching for a dagger, mnemonic information about average dagger size and weight must control his behavior in the appropriate way. Hallucinations and memories cause different dispositions to reason and act in various ways. However, behaviors controlled by either hallucinations or memories can still express content, for example,

related to daggers when no dagger is present.

Behaviors controlled by states about non-existent things express content that is not currently, nor may it ever have, or will be, experienced. Behaviors controlled by, for example imagining, planning, or thinking, can express content of non-existent, or merely imagined, things. For instance, thinking about the content ‘pink platypai’ employs a conjunctive state that combines veridical information about the color pink and veridical information about platypai. The conjunction of these two contents is not veridical content since it has, likely, never been experienced. Using a state with the content ‘pink platypai’ to control actions or inferences does not amount to hallucinating or remembering. The state with this content is employed in operations that do not treat the content as perceptual or mnemonic. For behavior to express the content ‘pink platypai’ the content would need to play certain kinds of roles in guiding inference or action. Generally, imagining, planning, or thinking, are operations that control behaviors which are distinct from those controlled by perceptual or mnemonic states.

Expressive use of content is central to the explanatory value of mental representations and further clarifies the notion “many-to-one”. Expressive uses of content are one of the features that help to separate representational uses of informational states from those dependent solely on information causing a state. For instance, rats with certain kinds of training can be guided by content about the end of a maze. These rats are guided by such content at the very beginning of their run through the maze (Tolman 1948). So, their behavior expresses information that is not currently causing the state guiding their behavior. There is still a dependence on original information to allow such states to have content (Fodor 1990b). However, expressive uses of content are still partly a functional notion that is explanatorily valuable to psychologists.

### 3.1.3 Misrepresentation

Finally, misrepresentation must also be taken into account when considering the differences

between simple information-bearing states and representational states. Fodor calls misrepresentation the flip side of expressive uses of focused informational states (Fodor 1990b). Both expressive uses of a representational content and misrepresentational uses of that content are caused by things other than what the content of the state is about. However, misrepresentations are not a correct use of the expressed content.

Misrepresentations are the expressive function regarding a content “gone wrong”. For example, when a subject thinks, rather than merely imagines or entertains the idea, that platypai really are pink then misrepresentation occurs. We can tell this is occurring when the conjoined content ‘platypai’ and ‘pink’ controls certain behaviors. Certain uses of the conjoined content misrepresent the color of platypai. Another kind of misrepresentation occurs in hallucinations, for example, when a subject inappropriately combines the content ‘platypai’ with the content ‘present’. Certain uses of this content misrepresent the presence of platypai.

Clearly, expressive uses of content and misrepresentational uses of content are closely tied together. Both are uses of content to control behavior when the state is caused by something other than what it is about. Of course, not all expressions of content are misrepresentational uses of that content. Both correct and incorrect uses of content express content by controlling certain behaviors. Misrepresentational use of the content just controls behavior we deem to be incorrect.

The explanatory requirement, then, is to find the contexts that determine inappropriate expressions of content. Causes cannot be relied on here since both expressive uses and misrepresentational uses of contents are caused by things other than what the content of the state is about. One difference between correct and incorrect expressions of content is that representational systems are generally disposed to avoid one but not the other (Fodor 1990b, p. 187). Representational systems avoid misrepresentational uses of content because misusing the

content often leads to abortive actions (Fodor 1990b, p. 188).

Based on behavior controlled by the state, projects for naturalizing representation may have a way to distinguish correct expressive uses of a representation from those where the state misrepresents (Fodor 1990b; 1987; Antony and Levine 1991). For example, a state's role in behavior control could show that the state is being used in a misrepresentational way. If, when a creature is jumping a gap, it misses, then this may show a misrepresentational use of the content. Likewise, if a navigating system is maneuvering a corner and runs into the corner, then the organism's behavior might show that it was guided by a misrepresentation. If a person is deluded about the presence of platypai, without hallucinating the presence of platypai, then the person will likely look for platypai without finding any.

To have a state with content be used in error, then, it must have a particular function that is being frustrated in a certain context. If the world does not match the expressed content then certain behaviors fail. However, as Fodor notes, evaluations of error only get a grip when representational states have functions that lead to a system not getting what it wants (1990b, p. 188). For example, mental states, such as desires, often influence the evaluation regarding whether a state has misrepresented or is playing the right role. In the case of navigation, we can only explain the behavior as being guided by a misrepresentation if we also assume that the creature wanted to control its behavior *correctly* but did not do so.

Although misnavigation may be a way to evaluate the correctness conditions of some representational state, these evaluative contexts assume other content in the creature, such as desires. For example, such explanations assume that the representational system did not want to miss the jump or mis-manuever the corner. This means that determining focused content must come before determining misrepresentation, for naturalistic accounts of error, because

evaluations of error are impossible to determine without bringing in the notion of other mental contents (Fodor 1990b).

Relying on other mental content to determine contexts of misrepresentation means that evaluation of misrepresentation is tied up with other mental representations besides the state under consideration. Representational content must be fixed prior to evaluation of misrepresentation (Fodor 1990b). So, to account for misrepresentation we need a naturalistic and functional account of the representational system not getting what it wants (Fodor 1990b, p. 188).

One possible way to account for misrepresentation is with a state with content having a particular role in behavioral control. The naturalistic account of misrepresentation may require a behavioral component, as well as other mental states. These components help to establish contexts we can analyze as inappropriate uses of content. However, not all content directly controls behavior. This is the case, for example, with representational states which are isolated from behavioral control, such as some perceptual states. Some perceptual states can also misrepresent and so would need a functional context in which to evaluate their misuse *internal* to the system. One possible function for internal content may be to feed information into the rest of the system in particular ways. When this function goes wrong misrepresentation occurs internal to the system. For example, in the case of perception, percepts need to be fed to the rest of the system as indicating the presence of the stimuli the perceptual state is about. Later systems, operating on misrepresented perceptual content, will not necessarily malfunction if given a misrepresentation. These operations will, however, be operating on misrepresented content. If there is no way for a naturalistic theory to determine misrepresentation internal to an information-processing system, then a naturalistic account of misrepresentation may necessarily

require a behavioral component.

3.2 Where Does Mental-State Explanation Begin: High-Roads, Middle-Roads, Low-Roads  
Clearly mental-state explanations have some value to cognitive psychologists. States with focused informational aspects, and contexts of expression and misrepresentation, help psychologists to account for certain behaviors. These kinds of information-bearing states are usefully described as representational mental states.

The question now is where mental states, and so mental-state explanations, begin in the three-tier information-processing model. Information processing is ubiquitous in the natural world. In particular, information processing is ubiquitous throughout subjects amenable to mental-state explanation, even in those parts of such subjects not amenable to mental-state explanations. For example, many automatic reflexes in humans involve a kind of information processing that is unlikely to involve mental states. Many cognitive scientists apply mental-state explanations in the second tier of the information-processing model of the human mind. However, it is not universally agreed that this is the best place to begin mental-state explanations.

A natural demarcation between simple information-bearing states and mental states in the three-tier information-processing model is in terms of the collection of capacities and mechanisms that create coded states that meet the above criteria for mental status. Very few collections of natural information processing capacities create internal states with these properties.

In one very plausible account, such collections are a “persisting set of integrated capacities... [that] contribute, distinctively and non-trivially” to explaining certain kinds of flexible behavior (Rupert 2009, p. 42). Generally, such integrated capacities are thought to occur in certain perceptual systems. Such systems often allow for the occurrence of more than single mental states. Furthermore, the standard picture holds that mental states generally occur within

persisting organisms with a fair amount of structural complexity (Rupert 2009; Fodor and Pylyshyn 1981; Lyons 2001).

These assumptions help theorists in psychology and cognitive science explain various behaviors. This set of explanatory practices forms a fruitful and robust research program. Applying mental-state explanations to states that do not fit these assumptions leads to a quick decline in the explanatory value of mental states. Without these features information-bearing states all around the creature could be included as ‘psychological’ states. However, the equivalence-class responses important for psychological explanations are limited to the kind of information-bearing states with representational features.

Applying the notion of mental states to states in information-processing systems with different equivalence-class responses, such as the retina, is logically possible. However, alternate ways of explaining retinal function, which do not employ psychologically describable equivalence-class responses, are also available. We can explain the equivalence-class responses of photoreceptive cells in the retina in physical terms. These alternate explanations are more parsimonious in that they do not posit special states, those with content, to account for retinal function. Only when the special features of psychological states are needed should they be used. Psychologists find these features of informational states apt for capturing the particular generalizations they are after when accounting for certain kinds of behavior.

States with content are special states that capture generalizations of interest to psychologists (Pylyshyn 1980). The equivalence-class responses of retinal cells do not capture such generalizations. States with equivalence-class responses we think of as content do not occur in other explanatory domains (see section 3.1.1 above). In this way, the explanatory value of mental states remains high in psychological domains and declines outside of such explanatory

domains.

The fact that mental states are explanatorily valuable does not tell us where the successful explanatory application of such states begins within the information-processing model of the human mind. A philosophical debate exists, for example, over whether early visual processing uses discriminations that should be described in terms of mental content. One example of this debate centers on whether such processing discriminates cracks from shadows (see Shapiro 1993). Psychologists must characterize the content of the discrimination task correctly so as to account for what the visual system is actually doing. Even if the visual system does not discriminate cracks from shadows it would seem to be involved in detecting *something* external, for example ‘crackdowns’, or just certain kinds of boundaries that are “thin, dark, marks that could be either shadows or cracks” (Segal 1989, p. 208). However, the debate over whether such explanations are best put forth in terms of mental content remains unresolved.

Below we will look at several levels of processing where mental states may fruitfully be applied. The states in cognitive processes are generally agreed to be mental states. How far below this kind of processing mental-state explanations can be fruitfully applied is a further question. Below we will see that there is explanatory value in applying mental states at perceptual levels of processing which occur at levels below that of cognitive operations. This is a fairly standard picture in cognitive science.

### 3.2.1 High-Road: Tier-Three and Representation

This section covers one general approach to applying mental states in the information-processing picture. This picture suggests that representational states occur only within cognitive processing. The first version of this picture holds that creatures exhibiting linguistic behavior are the best, and only, place to apply mental-state explanations. The second version of this picture holds that non-propositional states, still at the cognitive level, are also a good place to apply

mental-state explanations. As noted, mental-state explanations are standardly applied in a lower-tier of the information-processing model than cognitive processes.

### 3.2.1.1 Propositional Cognitive Systems

One traditional picture in philosophy regarding mental states treats only very particular cognitive systems, those that can use propositional states within a natural language, as having mental states. This picture bases mental state attribution on linguistic behavior (Sellars 1964; Davidson 1982). Davidson, for example, denies that mental-state explanations apply to non-linguistic creatures because, he holds, non-linguistic systems cannot be legitimately interpreted as having propositional states. He argues that if a creature exhibits linguistic behavior we can interpret them rationally through their language use and that this behavior is the only evidence for attribution of mental states.

Briefly, propositional states are token mental states within individual cognitive systems that are structured in ways analogous to sentences in a natural language. Propositional states are the contents of various attitudes, such as hopes or fears. As we saw in chapter two, contents about the sky falling can figure in a subject's hopes or fears (see section 2.4.2). The content that figures in a subject's hopes can be moved to a context where it figures in a subject's fears and will still remain the same content. Parts of propositional states, then, are like that-clauses in sentences.

Propositional states have exchangeable parts that, when combined in the right ways, can control, for example, linguistic behavior. Rules are ways of combining the parts of propositional states so that content can be preserved in various ways. Such rules are sensitive to the parts of propositional states. As an illustration, rules for forming propositional states can result in representations for a pencil being in a bottle and the bottle being in a box. Further rules, about the transitivity of the 'in' relation, can operate on these propositional states so that operations

based on these states can reach a state with the conclusion that the pencil is also in the box (Johnson-Laird 1980). Some theorists hold that linguistic behavior is the only behavior that reveals the need for propositional mental states and that such behavior is the only kind that requires mental-state explanations.

Other philosophers and cognitive scientists have had success applying propositional states when accounting for certain kinds of adult human behaviors not related to language (Pylyshyn 1984). Non-linguistic behavior can also show that a system has combined propositional states in rational ways. Based on where a system looks for the pencil in the above example we can also surmise that the system has reached a correct conclusion. This is true even for systems without linguistic abilities. Attempts to model this kind of behavior based on explicit sentence-like mental states have met with some success (Newell 1982; Anderson and Lebiere 1998). So, linguistic cognitive systems are not the only systems well explained by propositional mental states (Field 1978).

Adult humans are thought to have cognitive systems which can flexibly learn and remember, and which can operate on very specific kinds of states in vary specific kinds of ways. For example, adult human cognitive systems can operate on structured states in ways that mirror rational thinking. The states in such systems function much as sentences in a language do and can easily perform certain kinds of logical calculations (Fodor 1975). These capacities are based on operations that are sensitive to propositional states and their sub-parts (Fodor 1975; Marcus 2001). As already noted, such systems can support variations of states such as ‘man bit dog’ and ‘dog bit man’ and can perform simple deductions. Although these cognitive operations are language-like they do not necessarily require linguistic behavior or the ability to use language. In this way, many theorists think propositional mental states can help account for cognitive, but

not necessarily linguistic, behavior.

Theorists explaining rule-based behavior often invoke sentence-like, or propositional, states (Field 1978). Such explanations are tied to the existence of whole sets of other propositional states. For example, when we say that an adult human's cognitive system has one propositional state, such as a belief that something is the case or a desire that something is the case, we assume a background of many other propositional states also exist in that system.

Take the motivation to get a beer from the refrigerator as a case in point. This motivation is explanatory only in the context of other propositional mental states, such as the belief that the beer is located in the refrigerator. The motivation explains a behavior only given the presence or absence of other propositional states. If the person knows a guard dog is watching the refrigerator, thinks provoking the dog will have bad consequences, and desires to avoid these consequences, then the belief and desire about the beer may not motivate behavior of going to get the beer.

Each of the other required propositional states then requires further propositional states as well. In this way, propositional mental-state explanations require a network of other propositional states. Given other capacities, such as productivity, compositionality, and systematicity, this profile of adult human cognition requires individually structured states, a network of their interrelations, and capacities to operate on such states in certain ways. Each of these features of the cognitive system work together to form a part of the picture of adult human cognition based on propositional mental states. Such systems are very sophisticated.

Other kinds of human cognitive systems also have extremely robust capabilities in the face of variation in the world. They can accomplish the same task in a variety of ways in different environments. Likewise, they can treat a variety of different physical stimuli as the same. For

example, they can treat the characters in a puppet show, at least in certain ways, like real people (Frith and Frith 2005; Rupert 2009; Fodor 1987). As cases in point, pretense and deception also show the sophisticated abilities of these systems. Very few information processing systems can engage in such behavior. Likewise, scientific and medical achievements are obviously not possible with most other kinds of information-processing. Other information-processing systems cannot accomplish these feats, specifically because they do not have the propositional states and rule-based capacities that are thought to constitute much human cognition.

Applying mental-state explanations only to cognitive systems with propositional states is a highly limited use of mental-state explanations. Other theorists find mental-state explanations valuable when accounting for behavior controlled by non-propositional states. Some of these theorists hold that mental-state explanations should still be applied only to cognitive capacities but extend mental-state explanation to non-propositional states.

#### 3.2.1.2 Non-Propositional Cognitive Systems

There is an analogy between some of a subject's internal mental states and parts of public language. Although this analogy motivates some philosophers to limit mental-state explanations to systems with propositional states, this motivation is not a standard one in all of cognitive science or even certain parts of philosophy (Marras 1973; Fodor 1983; Dretske 1988). For example, many mental-state explanations treat mental states as states that do not need propositional structure (Pylyshyn 2003, 2009; Jackendoff 1992; Freedman, Riesenhuber, Poggio, and Miller 2001; Dietrich and Markman 2003).

Certainly the explanatory value of mental-state explanations seems to go beyond linguistic creatures or even the linguistic systems in language using creatures. Many theorists often hold that certain kinds of non-propositional states are also mental states. Even so, many such theorists still think mental-state explanations should be limited to cognitive processes.

Many accounts of behaviors driven by internal information-bearing states, for example, treat systems that can learn and remember as systems with mental states. Such pictures still limit mental-state explanations to cognitive operations but extend mental-state explanations beyond propositional states. Some of the behavior exhibited by preverbal human infants and higher primates, as well as non-linguistic cognitive systems within language using creatures, can be successfully explained with non-propositional mental states (Jackendoff 1992; 1994; Godfrey-Smith 1992; Sterelny 1990). Such behavior is thought to be based on cognitive processes operating on non-propositional mental states.

For example, many cognitive scientists believe there are non-propositional, but mental, states in creatures that can learn by way of operant conditioning. Theorists such as Dretske, for example, suggest that informational states that are incorporated into the flexible control of operant conditioned behavior count as mental states (Dretske 1988). Operant conditioning is a cognitive stage of processing. On Dretske's theory none of the states in processing that lead up to behavior controlled by operant conditioning are considered mental states. However, according to Dretske, behaviors controlled by states involved in operant conditioning are mental states. Cognitive systems capable of operant conditioning do not seem to require propositional states. Even without propositional states, such systems are still valuably explained as using mental states.

Other theorists extend mental-state explanation to other kinds of learning as well. As we saw in the case of rat navigation, information about location in a room can be learned (Sutherland 2005; Tolman 1948). This kind of learning seems to employ a focus of information as well as the ability to express such states at the beginning of a maze, for example, during direct navigation. It is likely that such states can also misrepresent. The states involved in such

learning seem best described as mental states that do not have propositional structure.

Other examples of non-propositional mental-state explanations of behavior controlled by cognitive operations are easy to find. The flexible behavioral control of adult ground squirrels in response to certain environmental stimuli is a case in point. Two different predators offer various degrees of threat to these squirrels. Sterelny notes, “raptors pose a very urgent threat, and hence [the squirrels’] predominant response is simply to take cover. Snakes pose a more enduring but less urgent danger. Adult squirrels need to assess the level of threat [snakes present]... [T]hese decision problems are not simple, for the threat depends on the size, temperature and species of the snake as well as the terrain. No simple response – like that of taking cover against raptors – will do” (2006, p. 293; Owings 2002). Adult squirrels, for example, will cautiously approach, jump back, and re-approach a snake several times exploring the level of threat the snake presents (Owings 2002, p. 19).

Explanations of ground squirrels’ behavior regarding raptors likely do not require representational states. The kinds of information-processing states that control the behaviors squirrels display in response to raptors do not have the features necessary for mental status. However, the states that control squirrels’ behavior in response to snakes are helpfully described as mental states. The right focus of information, ability to express this information, and to use it as a misrepresentation allow us to treat the states that control the squirrels’ behavior in response to snakes, but not to raptors, as mental states (Godfrey-Smith 1992). The states that help to explain some of a ground squirrel’s behavior are mental states but are not considered propositional mental states. So, non-propositional mental states do seem to be valuably employed when explaining some kinds of behavior.

Points of clarification are required when employing non-propositional mental-state

explanations. One clarification regards the natural demarcation between information-bearing states that generally underlie stimulus-response systems and those states best described as cognitive but non-propositional mental states. One way to show a system is employing mental states is to analyze its behavioral control. If an internal state, activated by a certain cause, has the right focus of information, and can play certain roles in an information processing system then the system controlled by this state will behave flexibly in response to a variety of input (Godfrey-Smith 1992). These features allow a system to be valuably described as using mental states.

When behavioral control is done in the right way it will require explanation in terms of whether the system “notices” a stimulus, or how it “classifies” that stimulation. These are mental-state explanations. With certain kinds of behavioral control, the explanation of the system’s behavior will no longer be able to effectively go forward merely in terms of information-bearing states that generally enable stimulus-response behavior. The explanation will require mental states and cognitive capacities (Bermudez 1998; Peacocke 1994a; Gold and Stoljar 1999; Rescorla 1988). However, the mental states in such explanations are unlikely to be propositional mental states.

### 3.2.2 Middle-Road: Tier-Two and Representation

Further pictures regarding where mental-state explanations begin in the information-processing model extend mental states beyond cognitive states and operations. For example, certain non-cognitive states, such as those in perceptual processing, are often treated as mental states. This is because the features of mental states are applicable to certain kinds of non-cognitive states. The standard picture holds that certain non-cognitive states can have a focus of information, and can be used to both express and misrepresent their focused information, or content.

Applying mental-state explanations to non-cognitive states extends such explanations to the second-tier of the three-tier information-processing model. The second tier of this model is the level of perceptual processing. Extending mental-state explanation in this way is popular in cognitive science (Pylyshyn 2003; 2009; Fodor 2009). For example, when percepts occur in a human subject's visual system, the person may have an experience as-of surfaces spread out in depth, whether or not the stimulation represented by such content is present in the environment. This is the case when subjects appropriately view random-dot stereograms. In such cases, percepts function like mental states because they have a focus of information, or content, and this content can be expressed, sometimes in contexts where we think it is misrepresenting. So, percepts seem to be amenable to mental-state explanation.

Middle-road positions attribute mental content to states in perceptual processing, earlier than cognitive information processing (Fodor 1983, 2009; Pylyshyn 2003). Theorists holding this position think that the explanatory value of mental representations does not begin too early, such as in transduction or the first stages of information processing. So, they do hold that mental states take a fair amount of information processing to be created. However, they hold that mental states also occur well before cognitive processing.

For example, in categorical perception, such as perceiving 'ba' and 'pa', information is added to incoming sensory states to achieve a distinct, or categorical, mid-level perceptual state. This additional information makes percepts, as Pylyshyn notes, "correspond to meaningful categories in terms of which objects are [later] identified" (Pylyshyn 2003, p. 134). These categories are not yet cognitive classifications of stimuli. Early visual states, as an illustration, respond to shape-classes that are like geometrical figures. These states represent something like the outlines of shapes in line drawings. Individual representations for lines are bound together in such a way

that each line belongs to some ‘object’. These kinds of equivalence-class responses can later be used by cognitive operations to identify objects. However, these states are not yet identified in cognitive ways (Fodor 2009).

The basic idea for representation in perceptual processing is that information-bearing states have been altered so that a focus of information is made explicit. This gives perceptual states at least one of the features required for mental status. This allows such states to function in later operations in the absence of the information that triggered the state. In the context of perceptual processing, when a percept occurs without being caused by what it is about it is often a hallucinatory misrepresentation. However, the amount of processing required for percepts gives these states a focus of information and these states can function as expressions of this content with contexts for misrepresentation.

Fodor and Pylyshyn, among others, have shown that these aspects of perceptual states allow us to valuably describe these states as mental representations (Fodor 1983; Pylyshyn 2003). In fact, making sense of reference without cognitive labels, and therefore representation before cognition, may be required for cognitive science (Fodor 2009, p. xiv). What this means is that there are internal states that mediate between information pick-up, for example, at the senses, and later cognitive processing. These mediating states have representational features that help scientists explain the behavior of the systems in question. So, there seems to be a psychological level of processing going on well before cognition (Fodor 1983; 2009; Pylyshyn 2003). Other theorists also agree that a legitimately psychological level of explanation can occur after stages of transduction but before fully cognitive processing occurs (Dretske 1978; Hatfield 1988; Rey 1997; Pinker 1999).

3.2.3 Low-Road: Tier-One and Representation. How Low Can Representation Go?  
Full percepts, such as visual states about shapes in depth are created from combinations of

other states. Some states, earlier than the construction of full percepts, are also sometimes valuably treated as mental states. For example, visual states prior to the construction of full percepts differentiate, say, depth from surfaces and color. These states are distinguished from transduction and information processing in the sensory periphery but are sometimes treated as representational states. For example, representations of lines are combined to form representations of vertex-types and these latter are combined to form representations of surfaces. These kinds of explanations extend the notion of representational content to some of the states in the earliest parts of second-tier information-processing.

Some theorists posit mental representations even lower than the second-tier of information processing. We must be cautious in our use of mental-state explanation at too low a level in the information-processing picture of mind. Very little information processing has occurred by the time early information-processing states are activated.

For example, David Marr calls the image impinging on a subject's retina a representation (Marr 1982, p. 31). He adopts this language from common talk that treats graphs as representations. However, the image at the retina is a direct physical projection from the environment. No information processing has yet occurred, so extending mental-state explanation this far loses much of its value. For example, such states do not have very focused information or an original way to misrepresent. Graphs are not mental representations, and only derive a focus of information or an ability to misrepresent from our use of them. It is our mental representational *use* that gives them a focus of information or makes them express or misrepresent content. This is also the case with states at the retina and those immediately transduced by the retina, for example.

Even though transduced images from the retina are not generally considered representations,

still, very low-level information-processing states are sometimes considered mental representations. Consider the output of the retinal-thalamic track that carries minimally processed information to the amygdala. When explaining what was processed in the amygdala, mental-state explanations are sometimes employed. This route of information processing and transmission has been termed the “quick and dirty route” to fear activation (LeDoux 1996). That is, the visual stimuli that enter the eye are transmitted within a very short time span and with limited processing to the subject’s thalamus and then on to the subject’s amygdala. However, fear processing at the amygdala is often characterized in terms of representational content.

For example, the states that can activate a fear response seem based, to some degree, on focused information, or content, in the signal and so the states controlling this response are sometimes explained in terms of mental features (LeDoux 1996, p. 67 and n. 47 p. 311). A potential threat is ‘recognized’ and the amount of threat is computed in relation to the distance from the individual. In humans, for example, “snake-looking” stimuli will illicit a large fear response, more so when they are close to the individual. As an illustration, this occurs to anyone who has mistaken a garden-hose for a snake (Cunningham 1997, p. 458; Lyons 1980, p. 59; Ledoux 1996).

Amygdalar fear states are not very focused information states. The processing of these information states is often characterized in terms of template-matching. Template-matching may be a way to explain these states without invoking mental representations. However, the fact that this processing computes threat in relation to how close the individual is to the threat suggests a certain amount of focused information exists in these states. Likewise, these states have a clear potential for misrepresentational use. Amygdalar fear states can occur in the absence of the stimulation they are about and so seem able to express any focused information they may have.

For these reasons representational states might be involved at the amygdalar stage of processing, even if the processing is a simple form of template-matching.

If representational states are explanatorily valuable at low-levels of information processing then some states before perceptual processing begins may be representational. These states may not even be processed to the level of having features such as perceptual constancies, perceptual categories, or perceptual saliencies.

There is a risk, however, in attributing mental representations at such a low-level of information-processing. If the states in this processing are representational we may also have to attribute mental representations to very simple information-processing states. Attributing mental states to the information processing in frogs, human knees, certain thermostats, as well as other simple information-processing systems is a contentious position. Certainly, though, some very early information processing states such as those in the amygdala seem amenable to a degree of mental-state explanation.

### 3.3 Application: Mental Representation in the Three-Tier Information-Processing Model

Explanations for a substantial number of behaviors require non-propositional representational states at the perceptual level. Furthermore, views that endorse perceptual representation are common to many cognitive scientific explanations. For example, the criteria for mental-state status fit one case from the philosophical literature, namely Swampman's first sensory guided movement. Explanation of this "creature's" first movement requires non-propositional perceptual mental states.

Swampman is a molecule for molecule duplicate of a famous philosopher created in a swamp from a lightning bolt (Davidson 1987). Before any causal interaction with the local environment Swampman seems to possess all the dispositions to behave that were possessed by the original philosopher (Fodor 1994). For example, it will say "Hello" to the philosopher's friends and walk

to the philosopher's home from the swamp. Many theorists are skeptical, however, that this "creature" would possess any mental states at all. That is, at least before it has causal interaction with the local environment (Papineau 1996).

The criteria given above for treating non-propositional perceptual states as mental do not apply to Swampman before *any* causal interaction with the local environment. The above criteria require the "creature's" internal states to be based on information. That is, the internal states guiding behavior need to be grounded in sensory states which do actually co-vary to a high degree with information from the local environment (Pylyshyn 2001; 2003; Harnad 1990). For example, tracking states, activated by objects around this "creature", must be available for its behavioral control before non-propositional perceptual representations will be useful in explaining some of its behaviors. Simple dispositions alone will not suffice since an actual connection to information is needed.

However, upon Swampman's first *sensory* guided movement, a non-propositional perceptual representation will best explain its movement. For example, the "creature" will reach out and grasp local environmental objects, it will ward off bugs, and it will navigate around plants in the swamp. It will do so based on representational states with focused perceptual contents that have a natural context for expression and misrepresentation. To explain these behaviors we must attribute representations in its perceptual processes of the distance between the eyes or ears as well as the speed of sound (Evans 1982). These representational states are needed for correct movement in relation to the environment.

The actual and coherent behavior of this "creature" in relation to the local environment will be best explained by representational content immediately upon causal interaction with the environment. The "creature" will not need to acquire any real history, or learning. These

representations are not cognitive, since they are still very stimulus dependent and not yet involved with operations such as learning, memory, or object recognition. However, these states can still count as mental states. Such immediate representations are non-cognitive because object recognition, for example, requires treating two or more of these lower-level representations as being about the same thing and the initial interaction between this “creature” and the environment does not yet cause this level of processing.

Does Swampman *initially* navigate the swamp with a representation that has the content ‘water’, or the propositional content ‘there is a tree there’? No. Non-propositional perceptual representations do not require cognitively individuated or propositional content of this kind. For example, they do not require content in terms of cognitive classifications, nor are they propositionally structured internal states (Pylyshyn 2003, p. 134). Such representations are about perceptual categories, such as shape classes.

These representations occur before cognitive processes and so are not classified in cognitive ways (Pylyshyn 2003; 2009). They are something like the perceptual version of “bare demonstratives” (Fodor 2008). They can track ‘perceptually individuated’ entities. As Fodor notes, “under certain circumstances...things in the world ‘grab’ mental indexes” without these indexes being cognitively classified (Fodor 2009 p. xv).

Explanations employing non-propositional perceptual mental states do not need to make a ‘natural-kinds only’ assumption about the content guiding this “creature’s” navigational behavior, say (Rupert 1999, p. 344). This is because, as Pylyshyn notes, non-propositional perceptual representations are based on a kind of reference that is “nonconceptual because it does not refer to things that have certain properties or that fall under certain conceptual categories” (Pylyshyn 2009, p. 5).

Non-propositional representational explanation, at the perceptual level, can proceed even if the representational states do not track natural kinds. The tracking states do not individuate kinds, properties, or individuals in the ways that cognitive states do. They merely need to segregate, or individuate, incoming information in a way that keeps track of them perceptually (Pylyshyn 2003, p. 99). Clearly, though, a ‘natural-kinds only’ constraint does begin to enter the explanatory picture when states enter certain kinds of cognitive processing, such as object-recognition.

The coherent behavior of Swampman in relation to its local environment will best be explained by some kind of representational content (Evans 1982; Egan 1995; Peacocke 1994b; Bermudez 1998; Rey 1997; Hurley 2003; Sterelny 1990; 1995; 2003). This is true even though it will not require cognitive representational states, such as those used in object classification. In this way, the explanation of Swampman’s first sensory guided movement fits the perceptual representation, or middle-road, account of where to apply mental-state explanations.

### 3.4 Summary and Conclusions

This chapter has shown what mental states are as well as various kinds of useful application of these states in psychological accounts. This helps determine where the explanatory value of mental states first appears in the information-processing framework. Percepts are a level of processing where mental-state explanation is safely employed.

We can also see from the above further reason to deny the claim that a subject’s cognitive unconscious is made up of all the bits of information picked-up at the subject’s senses. As noted earlier, this is likely an overestimation. Many theorists hold that a subject’s cognitive unconscious is made up of mental states but not all the information picked-up by their senses. Three features are required for a state to have mental status. First, the state must be coded in such a way as to have a focus of information. Second, the state must have an expressive

function. Third, theorists must have a way to determine when the state misrepresents. So, the best explanatory value of unconscious mental states begins after a certain amount of information processing (Kihlstrom 1990).

This chapter also provided reason to believe that not all of a subject's mental representations are cognitive states. Generally, percepts are as low as representations go. So, a subject's unconscious mental states can plausibly be composed of both perceptual and cognitive states. We will see more on this issue in chapter five.

## Chapter 4: Demarcating Conscious Mental States

This chapter first explores an empirical definition of conscious states. The definition of conscious states explored here, based on a subject's phenomenal judgments, is an attempt not to beg the question for, or against, subjective qualities. This chapter then canvases several theorists' demarcations for where conscious mental states begin in the human information-processing architecture.

### 4.1 Introduction

In cognitive science and philosophy, basic mental states are often distinguished from conscious mental states. From the point of view of psychological explanation, mere mental states are the norm. Generally, mental states without special features, such as those that make a mental state conscious, are the states employed in mental-state explanations. So, we must make sense of what normal mental states are supposed to be divided *from* when considering the difference between a subject's conscious mental states and the rest of their mental states. The nature of conscious mental states is a highly contentious issue. However, many theorists still think that conscious mental states are of explanatory value and this provides good reason for making sense of conscious mental states (Manson 2000, p. 149).

Often conscious states are thought to be distinctive kinds of mental states. Conscious states are often thought to be distinct because these states have properties called subjective qualities. However, the nature of subjective qualities is notoriously hard to describe in public language, especially non-circularly (Block 1995). So, it is hard to say, objectively, just how conscious mental states are distinctive based on these supposed subjective qualities. The core scientific methodology by which researchers distinguish conscious mental states from other mental states is a subject's judgments. For example, subjects often judge some, but not all, of the mental states attributed to them as having subjective qualities. Initially, then, the best naturalistic method for

demarcating a subject's conscious mental states relies on the subject's judgments.

Because subjective qualities are private, they are supposed to be directly known only by the subject who has states with these qualities. Subjective qualities are supposed to make the subject the best authority regarding their conscious mental states. For an outsider to know about the supposed subjective qualities of a subject's conscious mental states, or even that such qualities exist, they must ask the subject about the qualities and what the qualities are like (Nagel 1974). Some theorists hold that without the subject's judgments regarding subjective qualities there is no way to know about these qualities (Jackson 1982).

The existence of subjective qualities, and what they are like, is, therefore, ascertained from a scientific perspective by certain kinds of behavior. For example, researchers often rely on a subject's verbal reports when ascertaining which of a subject's mental states have qualitative features. Other times, non-verbal probes can be used to ascertain whether subjects judge the mental states ascribed to the subject to have subjective qualities. For example, psychologists have a subject press a button when the subject judges that a subjective experience of a light occurs. Button presses, or other behavior, are supposed to indicate the subject's conscious mental states.

Views about subjective qualities are often formulated in terms of what is called state-, rather than creature-, "consciousness" (Manson 2000b; McBride 1999; Antony 2008; Bayne 2007; Rosenthal 1993; Block 2007; Dehaene et. al. 2006). State-conscious mental states are those that subjects judge to have subjective qualities. Creature-consciousness is a notion that pertains to the whole creature which marks when the creature is asleep or awake. State-consciousness does not assume anything about whole subjects coming into, or going out of, wakeful states. As Manson notes, such a distinction is necessary for the debate regarding unconscious mental states.

As he says, “in so far as we believe in the distinction between conscious and unconscious *mentality* and believe that the distinction is not the same as the distinction between conscious and unconscious *subjects*” we need a distinction between state- and creature-consciousness (2000b, p. 410, emphasis added).

For example, some theorists believe that sleeping subjects can have subjective experiences even when they currently have no creature consciousness. This occurs when the creature is dreaming (Flanagan 1995). If this is true, then the creature will have state-conscious mental states even when it lacks creature-consciousness.

The distinction between state-, and creature-, consciousness is also important because state-conscious mental states are of ethical value. If it could be shown that a human subject had state-conscious mental content without being awake then we would likely alter our treatment toward such a patient. If a patient, for example, displayed some type of internal indicator of state-consciousness while also being behaviorally unresponsive, or vegetative, we would count the internal indicators of state-consciousness as significant to an ethically valuable mental life (Naccache 2006; Block 2007, p. 484). This is true even if the patient displayed no signs of a sleep/wake cycle and so lacked creature-consciousness.

Many cognitive scientists, who are dubious of intrinsic subjective properties such as private experiences, search for an objective demarcation for a subject’s state-conscious mental states. For example, they look for behavioral, or neural, correlates that may identify the states underlying a subject’s judgments about certain states being conscious. These objective features can then be used to demarcate the conscious states within human subjects.

Several kinds of theorist, who otherwise differ in their approach to conscious states, still agree on the explanatory value of conscious states. For example, many theorist hold that only certain

mental states pass objective tests thought to demarcate them as conscious states. Various theorists suggest different tests that are supposed to be diagnostic of conscious states. On such models, no intrinsic subjective qualities are invoked.

#### 4.1.1 Phenomenal Judgments

The best objective methodology for getting, at least an initial, understanding of which states in a subject are conscious states is to ask the subject which states they judge to be conscious. A subject's reports are the primary objective data regarding conscious states. However, this method is only an initial approach, as we will see later, because simply asking subjects to report their conscious states is not always the best method for demarcating the states theorists think are the conscious states of a subject. So, neural, rather than behavioral, criteria may be a better objective method for ascertaining which of a subject's mental states are conscious. Finding a neural correlate of conscious states, however, would initially rely on a subject's judgments. Generally, we can tell such judgments occur only through the subject's behaviors.

Phenomenal judgments are often considered to be propositional states that humans form *regarding* their conscious states. The corpus of a subject's phenomenal judgments seems to reveal both states with supposed subjective qualities as well as a relation among such states. For example, phenomenal judgments in the domain of vision reveal the difference between visual subjective qualities, such as experiences of red and green, as well as the relation between subjective qualities, such as focused visual experience and peripheral visual experience.

To better understand phenomenal judgments, think of the case where a person is shown a bright red balloon. When circumstances for viewing the color are good, and the observer is appropriately oriented, the subject might form the judgment "I see red now". We think this because subjects in these cases sometimes say things such as "I see red now". In the right kinds of circumstances, we think that the subject's verbal report expresses a judgment about an internal

mental state. Such judgments need not have caused public behavior. Subjects report conditions where they have had unexpressed judgments regarding their conscious states (Dennett and Akins 2008; Khilstrom 1990; Ericsson and Simon 1980). The kinds of behaviors that indicate a subject's phenomenal judgments are often used as the basis for demarcating the subject's conscious states.

When humans express judgments about the mental states they take to be conscious, such as seeing red, we may think they express more than just the representational content about the external color. Subjects may also express something about the subjective qualities of certain mental states. Chalmers, for example, holds that the content of phenomenal judgments is in part constituted by an 'acquaintance' with certain kinds of subjective qualities (Chalmers 1996, p. 183).

Chalmers notes a paradox for such a position, however. There will very likely be a physical, or functional, explanation of why humans are disposed to make the phenomenal judgments they do. Likewise, there will be a physical, or functional, account for why subjects judge these states to have the particular subjective qualities they are judged to have (Chalmers 1996, p. 164; Dennett 2007). However, subjective qualities are not part of physical, or functional, accounts. So it seems that much, if not all, of the picture about phenomenal judgments can be cashed out without relying on subjective qualities. This is true whether or not subjective qualities actually exist (Dennett 2007).

Empirical studies of phenomenal judgments suggest there are a limited number of states subjects judge to be conscious states. Subjects often do not judge all of their mental states to be conscious mental states. Many theorists try to explain these particular limits of a subject's judgments regarding which of their states are conscious by finding out which of a subject's

mental states are accessed by particular systems, such as the subject's verbal report or working-memory systems.

Due to the information-processing architecture, some of a subject's subsystems only have contact with certain mental states. The limited access certain sub-systems have in regard to the overall mental states in the subject is often used to explain the limited judgments subject's have regarding which of their states are conscious. Limited access is just one way to demarcate a subject's conscious mental states. As we will see a bit more below and in chapter five, coding is another way to account for why subjects judge some, but not all, of their mental states to be conscious.

Phenomenal judgments are issued from parts of the human mental system that do not have direct access to the objects or properties causing such judgments (Chalmers 1996). The sub-systems issuing a subject's phenomenal judgments are simply given bare information differences and can therefore only operate on the processed information in terms of the states they receive. On either access or coding accounts of limited phenomenal judgments, the design of the human mental system results in a subject giving qualitative judgments regarding only some of their mental states. Chalmers states, for example, that

as far as central processing is concerned, it simply *finds itself* able to make distinctions, and it *knows* it is able to make distinctions, but it has no idea how it does it... If one asks it, 'What is the difference between these states?' it has no answer to give beyond 'They're just different,' or 'This is one of *those*,' or 'This one is *red*, and this one is *green*.' When pressed as to what that means, the [centralized] system has nothing left to say but 'They're just different, qualitatively'... We would expect these states to be quite 'ineffable': the [centralized] system lacks... any further relevant information, so there is nothing it can say about the states beyond pointing to their similarities and differences... it is natural to expect the system to use the language of 'experience' and 'quality' to describe its own cognitive point of view on [some of its own states] (Chalmers 1996, p. 272).

Generally, only some of a subject's mental states are judged by the subject to be conscious. So, only some of a subject's mental states are coded or accessed in a way that results in phenomenal judgment. For example, there may be a special system that issues phenomenal

judgments only for states that enter this system. We will see reasons why a special kind of coding is a better account for limited phenomenal judgments later, and in the next chapter.

For now, we can think of the cause of phenomenal judgments in terms of either specially coded states or states accessed by a special system. Coded or accessed information states ‘veil’ some of the information they carry from certain operations (Chalmers 1996). For example, states about color carry information about nanometer and neural spiking-frequency that is not coded for, or accessed by, the operations of central systems. Given the lack of information that central systems have about certain states, these systems will issue phenomenal judgments for just these states. The descriptions central systems offer of these states will be in terms of subjective experience, or different ‘*qualitative*’ features.

Both supporters and skeptics of subjective qualities agree that humans do make fairly precise judgments as to which of their mental states are thought to have subjective qualities. Human subjects judge quite a variety of their mental states to have various kinds of subjective qualities. For example, bodily sensations, emotions, musical anticipation, occurrent memories, thoughts, and other perceptions are all processed in vastly different ways. Each of these kinds of states is, at times, judged to have various subjective qualities (but see Pitt 2004).

A variety of different kinds of theory find some demarcation for a subject’s conscious mental states explanatorily valuable, even when they disagree about which of a subject’s mental states are conscious. The rest of the chapter will address a variety of these theories regarding just which of a subject’s mental states are conscious.

#### 4.2 Functional Demarcations for Conscious States in Humans

Diverse theorists hold that conscious states are explanatorily valuable. Most theorists base this idea on the fact that subjects make phenomenal judgments about certain of their own mental states. Not all theorists base the demarcation for a subject’s conscious mental states on a

subject's phenomenal judgments. So, there are a variety of theories regarding which of a subject's overall mental states are conscious. These theories assume what will later be defended, that *unconscious* mental states are also of explanatory value. The next few sections canvas models that treat conscious mental states as a subset of a subject's overall mental states.

An immediate question arises regarding the method of demarcating a subject's conscious mental states used by various theorists. Not all theorists rely, in the end, on a subject's judgments to demarcate a subject's conscious states. Empirical demarcations of conscious states form two classes. In the first class, a tie to phenomenal judgments forms one empirical way to demarcate a subject's conscious states. Another class of empirical demarcations for a subject's conscious states does not, in the end, rely on the subject's phenomenal judgments. Initially, virtually all positions regarding a subject's conscious states start with a subject's phenomenal judgments.

In the next few sections, we will consider four general demarcations for a subject's conscious states. First, we'll consider Posner's Attention Model and Dehaene and Naccache's Working-Memory Model. These models suggest that conscious mental states are identical to those which subjects judge to be conscious without experimental cues. Another model, recently advocated by Lamme, also suggests that conscious states are tied to a subject's phenomenal judgments. On this model, however, the link between conscious states and a subject's phenomenal judgments is looser than Posner's model or Dehaene's model. Lamme identifies a subject's conscious states with those they judge to be conscious with the right experimental cue. A third model, given by Dennett, suggests that phenomenal judgments may actually engender conscious states. This will present a very different kind of demarcation for a subject's conscious mental states. Finally, we will briefly look at models that do not, in the end, tie a subject's conscious states to the subject's

phenomenal judgments in any way.

Each of these methods of demarcating a subject's conscious mental states presents a functional divide of one kind or another. Each of these different functions is the basis for demarcating a subject's conscious states that avoids using the notion of intrinsic subjective qualities.

#### 4.2.1 Conscious States as Those Tied to Phenomenal Judgments and Types of Report

It is fairly common to see theoretical demarcations of a subject's conscious states that are tied to the subject's phenomenal judgments. There are several types of model of this kind in the literature. One intuition motivating these positions is that conscious states are somehow linked to the "subject". For example, Clark notes such a link is what it means for a subject, rather than some of its parts, to have a conscious experience. However, all mental states in a subject are linked to the subject who has these states. So, demarcating certain states as those linked to the subject does not demarcate a sub-set of mental states that can plausibly be considered a subject's conscious states.

Often, phenomenal judgments are thought to arise from certain kinds of systems within subjects. Special access pictures are one way to account for the fact that subject's judge some, but not all, of their mental states to be conscious. Phenomenal judgments may arise, for example, when "currently transduced information [is put] into contact with stored knowledge and expectations...in ways that support...reasoned, self-intelligible, and flexibly goal-responsive behaviors" (Clark 2007, p. 589; Block 2007, p. 484). This position risks endorsing Cartesian Materialism where all and only conscious states of a subject exist in a sub-system of the subject specially connected to "the subject".

Treating conscious mental states as those accessed by certain systems avoids using an internal "subject" as the criterion for demarcating conscious mental states. However, such pictures still

equate conscious mental states with states accessed by certain of a subject's sub-systems.

Empirically, these models are quite risky in that no single consciousness-system has been found in human subjects. Later, arguments that show an alternate picture will be given.

One way to treat the notion of access, for now, suggests that conscious a subject's conscious states should be identified with states in the subject's focal attention or working memory systems. A second way to treat the notion of access suggests that a subject's conscious states should be identified with a larger capacity but more fragile buffer of mental states than those in a subject's focal attention and working memory. These will be the first two models covered below.

A final picture still links a subject's conscious states to the subject's phenomenal judgments does so in a way that is not fixed prior to, for example, asking the subject about their conscious states. This final picture does not rely on a particular system in a subject as a demarcation for the subject's conscious states. We will look at each of these models in turn.

#### 4.2.1.1 The Attention and Working Memory-Nexus: Posner, Dehaene and Naccache

The first positions that link a subject's phenomenal judgment and the subject's conscious mental states are Posner's Attention Model and Dehaene and Naccache's Working Memory Model (Posner 1994; Dehaene, Naccache, Sackur, and Sergent 2006; see also Baars 1988; Sligte et. al. 2008).

Historically, the paradigmatic test for conscious states has been a subject's verbal report. In early tests of subliminal perception, for example, researchers had subjects attempt to identify very low tones or quickly flashed visual stimuli. This research tries to identify a threshold of conscious perception. The verbal report approach has been, and still is, a central methodology in subliminal perception research that attempts to demarcate a subject's conscious mental states (Dixon 1971; Merikle 2000).

In this and other ways, a subject's verbal report has been a standard test criterion for the

subject's conscious states. This method of demarcation assumes a strong tie between a subject's conscious states and the subject's phenomenal judgments. Subjects report that they cannot focally attend to all the stimuli present in their environment. Likewise, subjects report they cannot focally attend all the mental states within their own mental systems. Subjects often report that the states they cannot attend are not conscious states. So, a possible way to think about which of a subject's conscious mental states is in terms of those states in a subject's focal attention (Posner 1994).

Other theorists characterize a subject's conscious mental states in terms of states in a subject's working memory. Working memory models suggest that when a subject's mental states reach their working memory system they are broadcast to several other systems, such as reasoning and voluntary action control (Dehaene et. al. 2006; Prinz 2000). This makes states in a subject's working memory the states that the subject can declaratively remember or verbally report. Subject's often judge the states in their working memory to have subjective qualities (Dehaene et. al. 2006). Therefore, states in a subject's working memory are thought to constitute the subject's conscious mental states.

Focal attention and working memory models can be combined. Data that support focal attention or working memory models for the demarcation of a subject's conscious states suggest the radical hypothesis of "sparse" conscious representations. Sparse models of a subject's conscious states suggest there are a very limited number of conscious mental states in a person's mental system. For example, there may be only three to four mental states in a subject's focal attention or working memory systems at any one time.

Some theorists think data supporting focal attention or working memory models show only a few of a subject's mental states are conscious at a time (O'Regan and Nöe 2001). Most models

of conscious mental states suggest that a subject's conscious states are a subset of the subject's overall mental states. The focal attention and working memory models take this idea to an extreme.

The case of binocular rivalry is one example that suggests a limited notion of conscious states in a single subject. In these experiments subjects are presented with two different visible stimuli, one to each individual eye. Subject's report that in this set up, after a brief period of time, only one of the stimuli is represented in a way that is judged to be conscious at a time. For example, subjects report that one of the stimuli becomes visible and then the other (Blake 2001). This data clearly suggests, based on the subject's reports, a demarcation for conscious states in a subject's mental system. Such data can also suggest a very limited number of conscious states in a subject at a time. Neurologically, both states representing each stimulus remain active. That one state becomes visible at a time shows clear limitation regarding the states subjects can focally attend, remember, or seemingly judge to be conscious.

Further data can also seem to support the conclusion that there are only a sparse number of conscious states in human subjects at a time. Experimental results in change blindness studies also seem to suggest that people only have limited a number of conscious states. Change blindness occurs when stimuli presented to a subject are switched in such a way that the change in the stimulus is not detected by the subject, at least as far as the verbal reports of the subject indicate (see Mitroff, Simons, and Franconeri 2002). For example, if two static photographs of very similar visual scenes are alternated in the right way, then salient parts of one photograph can be eliminated from one picture and subjects will not report a difference in the photos. This data can suggest that subjects have conscious states of only a few parts of the stimuli with which they are presented.

Results other than those from binocular rivalry and change blindness can also suggest that subjects have only a sparse number of conscious states at a time. This next kind of data comes from reading research. When English speaking subjects read, for example, the text outside of the focus of their attention and working memory can be altered in certain ways. In a striking illustration, the entire page can appear to subjects to be covered in English writing but is in fact mostly made up of non-English symbols. Only the area of focused attention or that involved in working memory required for reading has real English text (McConkie and Rayner 1973; Rayner and Bertera 1979). The window needed for coherent reading ability may be as small as a few characters or, perhaps, as long as eighteen characters (see Hunziker 2006). Subjects do not report anything out of the ordinary when reading a ‘page of text’ composed of English characters only where they focus their attention. This can suggest that subjects have only a sparse number of conscious mental states at a time.

Several theorists endorse the idea that mental states in a subject’s focal attention or their working memory are the conscious mental states of that subject (Lahav 1993; Leman 1981; Hirst 1995; Place 2000; Dehaene et. al. 2006). The limited states in focal attention or working memory form the core set of states subjects usually judge to have subjective qualities. States in focal attention, or working memory, form a sparse number of the overall mental states in a subject’s mental system. Likewise, these states are often judged by subjects to be conscious, or to have ‘qualitative’ features. Based on such data, several theorists support an identity between a subject’s conscious states and the states in a subject’s focal attention or working memory. This would mean that a subject’s conscious mental states are very limited in number.

#### 4.2.1.2 Attention versus Awareness: Lamme

Other theorists think that conscious states are not as severely limited as the attention and working memory models suggest. It is true that subjects’ focal attention or working memory

systems have access to only a few items at a time. However, conscious states may not be as limited as reports of attended and remembered stimuli suggest. Attended or remembered items may be selected from an overall set of mental states poised for selection by a subject's focal attention or working memory.

If a subject's conscious mental states are selected by focal attention or working memory then another demarcation of a subject's conscious states can be seen. Lamme (2006) and others, for example, have endorsed a broader demarcation for a subject's conscious mental states than those in the subject's focal attention or working memory. On this picture, a subject's conscious states would be more numerous than those states that enter the subject's focal attention or working memory.

Much of the support for the sparse model of a subject's conscious mental states comes from certain interpretations of empirical results, such as those from change blindness studies. Alternate interpretations suggest a different demarcation for conscious states. For example, the label "change *blindness*" suggests visual changes are invisible to the subject who cannot report such changes. However, an alternate interpretation suggests that some of the changes that are unreported by subjects may actually be detected, and count as conscious states, but are simply quickly forgotten and so unreported. Such an interpretation would mean that change blindness is simply a kind of amnesia rather than a kind of blindness (Wolfe 1999; Rensink 2000).

Interpreting change blindness results in this way has implications regarding a subject's conscious states. Blindness, but not quickly forgetting states, is related to subjective qualities. For example, in perceptual situations subjects often glance at stimuli that they will later report being unable to remember. However, many of the states in these perceptual situations are generally judged to be conscious. Unable to remember changes, subjects simply do not report

such changes, but may not be *blind* to such stimuli. The changes in change blindness studies may not all be invisible to the subject and so change blindness results may not be a good measure of a subject's conscious states.

Clarification is required for interpreting other results as well. Only on particular interpretations do certain results seem to support the focal attention and working memory model of conscious states in human subjects. For example, clarification regarding foveal reading and binocular rivalry results can suggest a broader picture of a subject's conscious states.

In the case of foveal reading, although subjects need only focally see, and have short-term memory for, a small portion of the visual page, this does not suggest the rest of the display is invisible. In this respect then, the results from reading research do not suggest a very limited number of conscious states in subjects. Peripheral visual states for the rest of the page, for example, can still be considered conscious based on subject's reports. This suggests a broader number of conscious states outside the region of states in a subject's focal attention and working memory.

The same is true for binocular rivalry results. Although we should believe subjects regarding their reports that the visual stimuli alternate between being visible and being invisible, this result is not enough to show that subjects have only a sparse number of conscious states. Other states, such as a subject's peripheral visual states, some of their tactile states, or some of their auditory states are reported to remain conscious even when the focal visual stimuli are switching. Subjects do not become blind, numb, or deaf to many of their mental states in circumstances where binocular rivalry occurs.

These alternate interpretations and clarifications of experimental results extend the number of mental states that count as a subject's conscious states. This view endorses that some but not all

of a subject's mental states are conscious but extends the number of conscious states in a subject beyond just a sparse few. This position is especially strengthened by the fact that subjects judge and report many more states as conscious states than sparse consciousness pictures suggest.

Lamme's alternate theory, likewise, demarcates a larger set of a subject's states as conscious than focal attention or working memory models (Lamme 2006). On this alternate view, a high-capacity, but fragile, buffer of states are the conscious mental states in a human subject's mental system. That is, more states than those present in focal attention or working memory are judged by subjects to be qualitative states. Such states may form a 'buffer' of states available for phenomenal judgment if subjects are given appropriate cues to report these states.

Lamme's model extends the number of a subject's conscious states, but does so in a constrained way that does not count overly many of the subject's mental states as conscious mental states. One way he does this is by relying on subject's *cued* judgments regarding which of their mental states they take to be conscious. Experimental results suggest both that subjects judge a wide variety of states to count as conscious states and, when appropriately cued, judge more states to be conscious than those present in their focal attention or working memory systems at any one time (Sperling 1960; Lamme 2003).

For example, when subjects are presented with a brief visual array of twelve letters and numbers arranged in rows and columns, they report that all of the items in the display were judged by them to be represented by conscious states. When simply asked to report all of the items they saw in the array, subjects generally report only three or four of the items. At first, this result seems to support a sparse model of conscious states. However, Sperling and Lamme, among others, have shown that if subjects are appropriately cued they can report *any* three or four items from large displays (Sperling 1960; Lamme 2003).

This experimental design suggests a broader number of conscious states than sparse pictures do. Here *cued*-report or cued-judgments estimate that the number of conscious states in a particular subject's mental system is greater than the three to four items in the subject's focal attention or working memory. Three to four items are just the number of states subjects can focally attend or remember at once. This is not the number of states they judge to be conscious.

There are two reasons to believe the broader view of a subject's conscious states is correct. First, subjects judge and report that they consciously experience representations of more items than they can report. For example, they report that items slip away as they try to simply report or remember states representing these items. Second, subjects can report any part of the array if they are cued appropriately. This data supports subject's judgments and reports of being conscious of more than three or four states at once. It seems that a subject's conscious states are selected by their focal attention, working memory, or other systems for report rather than conscious states being just the states in a subject's focal attention or working memory.

On this picture, many full percepts for vision, audition, or tactile sensation, for example, create states that are judged to be conscious states by subjects. This alternate model extends the number of a subject's conscious states beyond just those in focal attention or working memory. The above clarifications and reinterpretations also suggest subjects have a large, but fragile, buffer of states that can be judged to be conscious. Such states are fragile because the process of judging and reporting them to be conscious disrupts the number of states available for such judgments and reports. Later sections endorse a picture something like Lamme's, with a bit of clarification regarding the notion of a "buffer" of states (see chapter five). Clearly, though, a subject's conscious mental states are not limited to mental states in the subject's focal attention or working memory systems alone.

#### 4.2.1.3 Multiple Drafts of Conscious States: Dennett

Dennett also links a subject's conscious mental states to the states the subject can report (Dennett 1991; but see also Tononi and Edelman 1998; Kinsbourne 1988; O'Brien and Opie 1998). Dennett's position regarding the demarcation of conscious states can be difficult to understand. One reason this is true is because he does not posit a fixed demarcation for a subject's conscious mental states. This makes his position both sophisticated and subtle.

Also, as Akins notes, Dennett does not exactly have a positive theory for demarcating a subject's conscious mental states. Instead he has a clear set of conceptual issues to avoid and a few empirical points to include in his model. These criteria leave a lot of room for more determinate theories in line with Dennett's ideas (Akins 2002; 1996). The above issues are just a few of the difficulties involved in understanding Dennett's theory. Below, a brief sketch of the relevant parts of Dennett's view is provided.

One of the conceptual issues that Dennett is adamant about avoiding is Cartesian Materialism. Cartesian Materialism suggests there is a special place where, and time when, all and only conscious mental states of a subject exist. Dennett does not believe there is a single place and time, say in neural processing, where all and only conscious mental states occur in a subject. Cartesian Materialism is a picture that endorses special access to a particular mental system in a subject as the way to demarcate the subject's conscious mental states. Dennett avoids this access picture of a subject's conscious mental states (see also Baars et. al. 1997a).

Further, one of the empirical points Dennett is concerned to include is an empirically based skepticism regarding the notion of a unified domain of a subject's conscious states (Marcel 1993; Dennett 1991). For example, when subjects respond to a light they judge to be consciously experienced, they can indicate this judgment by a verbal response, a button press, or by blinking. Each of these responses results in different reaction-times when done in isolation.

Interestingly, when researchers ask subjects to report in these three ways on a single trial the various ways of responding can be in conflict. The methods of report can disagree about the conscious status of a representation regarding the same stimuli on the same trial in the same subject (Marcel 1993). Marcel argues that “different ways of reporting have differential access to an experience... [and this difference between responses] calls into question the existence of a unitary reflexive consciousness... responsible for report” (1993, p. 174; see also Barlow 1982).

The criteria that Dennett wishes to avoid and those he wishes to include, at times, seem to make him endorse an attention or working memory model of the demarcation for a subject’s conscious states (Dennett 1978). For example, he states that drawing attention “thereby promote[s] the influence...of whatever is occurring [where the subject’s attention is drawn]... rendering it reportable and recollectable” (Dennett and Akins 2008, p. 4325). Presumably such states are then rendered conscious to the subject. Other times, however, he does not wholeheartedly endorse the motivation that makes some neuroscientists, computer scientists, and psychologists think the states in a subject’s attention or working memory are the only conscious states in a subject (Dennett 2001).

One reason these difficulties exist regarding Dennett’s position is that Dennett endorses a variable, or non-fixed, demarcation for a subject’s conscious states. In this way, Dennett’s demarcation of a subject’s conscious states can be seen as much more complex than simple demarcations. The complexity in Dennett’s position comes from the fact that his demarcation is also not continuous over time. Likewise, Dennett’s demarcation for a subject’s conscious states is only rough on any occasion that it does exist. Although his model is more dynamic and non-continuous he still believes there is a difference between a subject’s conscious and unconscious mental states. In Dennett’s own terms, this difference is “huge” and “real” just as there is a huge

and real, but only rough, transition between night and day (Dennett and Akins 2008, p. 4322; Dennett 2005, p. 170).

One of the main points in Dennett's positive model comes from Marcel's data. Variance in modes of a subject's response, even on the same trial, regarding their conscious states makes Dennett endorse a demarcation for conscious states that is 'probe dependent'. That is, Dennett endorses a demarcation that varies from trial to trial, or even within a single trial, depending on the probe used to ascertain which states a subject judges to be their conscious mental states (Dennett 1991, p. 247; Dennett and Akins 2008, p. 4325). For example, asking subjects to report verbally or to blink when a stimulus is judged to be represented consciously are two ways to probe the subject regarding their conscious states. These two probes can demarcate different states as conscious even on the same trial. In this way, the demarcation of conscious states, for Dennett, is probe dependent.

On Dennett's view, a probe is a cue, like a question, or a mode of response, like pressing a button, that causes an internal state to become fixed in certain ways (Dennett 1991, p. 315). Probes do not have to be external, or verbal, they can be caused by a subject's own covert attention, for example (Dennett and Akins 2008, p. 4325). So, isolated people can still have conscious states on Dennett's view (Dennett 2000, p. 379; Dennett and Akins 2008; Dennett 1991). Probes are functions that settle states into a determinate content which can then be judged by the subject to be conscious or not (Dennett 2001, p. 222; Dennett and Kinsbourne 1992, p. 253; Dennett 2000, p. 379).

For Dennett, the probe dependence of a subject's conscious mental states means that the transition some mental states undergo in becoming conscious mental states is made determinate in a subject only retrospectively (Dennett and Akins 2008, p. 4323; Dennett 1995). As an

analogy, Dennett believes that we cannot tell that a “speciation event happening now” *is occurring*, since it is a transition that is not determinate until it has *later* effects (Dennett and Akins 2008, p. 4325; Dennett 1991, p. 247). He holds a similar view regarding a subject’s mental states becoming conscious. For Dennett, there is no such event as a mental state “becoming conscious” in the subject prior to the probe that makes it conscious (1991, p. 241).

Probes can be perceptual or conceptual. So, on Dennett’s view a variety of kinds of content in the subject’s mental system can be considered conscious on some occasion or other. For Dennett, probe dependence also means that there is no absolutely timed demarcation for when a subject’s conscious states occur. On Dennett’s view, to ask about either the place or the time of conscious state “on-set”, at least within a certain narrow temporal window, is a question based on conceptual confusions (Dennett and Kinsbourne 1992). Although there is a real distinction between conscious and unconscious mental states, on Dennett’s view, at least on every occasion that conscious content is determinate, there is only a “dusk-like” blur, or rough demarcation, for where or when such mental content occurs (Dennett 2005, p. 171; Dennett and Akins 2008, p. 4322). That is, there is no exact time or place where conscious mental states exist in the information processing stream. On Dennett’s view, only after a probe or after a rough window of time can conscious states be said to exist in a subject.

Despite all this complexity, Dennett is not against the explanatory value of unconscious mental states. As Dennett notes, “the [processing] that elevates a content to the clout of conscious recallability is largely not conscious” (2005, p. 170; Dennett 1991, p. 275). Elsewhere he quotes Thompson as saying, “whatever intermediate entities there may be in the *process* of grasping an object, it is not these representations that we are conscious *of*” (2000, p. 363; Dennett and Akins 2008, p. 4322). So, clearly there is an explanatory role for conscious and

unconscious mental states on Dennett's view. At least some of a subject's mental contents, on some occasions, count as conscious states. Even with the complexity of Dennett's view, not all of a subject's mental states will count as conscious states at a time.

Dennett ties his demarcation between conscious and unconscious mental states to subject's reports. Processing underlying report is a central kind of probe that creates conscious mental states in a subject. Dennett suggests that humans employ a social way of thinking that helps determine when the effects of some mental state indicate a conscious state (Dennett 1991, p. 238-9; Dennett and Akins 2008, p. 4326). However, as Dennett explains, even verbal report is only a merely "normally sufficient" effect for retroactively determining conscious content (Dennett 1991, p. 140). This makes pinning down a subject's conscious mental states quite complex.

The above description is a very simplified version of Dennett's overall model. Likewise, it is only a small part of his complete theory. Still, even this sketch paints an extraordinarily flexible picture for demarcating conscious from unconscious mental states in human subjects.

#### 4.2.2 Conscious States as those without a Tie to Phenomenal Judgments or Report

A further class of models does not endorse any tie between subject's phenomenal judgments or their ability to report certain states and the subject's conscious mental states. These theories suggest that a subject's ability to phenomenally judge or report certain states as conscious is extraneous to demarcating the subject's conscious states.

Ned Block, for example, suggests that a subject's conscious states can exist in the subject divorced even from the *possibility* of that subject being able to report the conscious states (Block 2007, p. 498). In this way, Block's picture completely divorces the subject's systems of report and phenomenal judgment from the subject's conscious states. Zeki suggests a picture very like Block's, with a few core differences (Zeki 2007, p. 585; Zeki 2003, p. 217).

The reason for covering models that do not tie conscious states to a subject's phenomenal

judgments or their ability to report these states is a methodological problem (Block 2007). A problem exists because there is no clear way to simply drop or retain methods tied to phenomenal judgment and report without making certain contentious assumptions (Block 2007; Clark 2007, p. 587). We can, just by stipulation, drop or retain methods based on phenomenal judgment, or we can, with more sophisticated argument, attempt to find a “mesh” among various data and arguments that construct reason for or against dropping this method.

Either way, there is an overall problem in dropping, or relying on, methodologies involving subject’s judgments or reports. Highly contentious theoretical assumptions must be made on either theoretical approach. The following models drop the reliance on phenomenal judgments and report as a way to demarcate a subject’s conscious mental states.

4.2.2.1 Islands of Consciousness and Reverberation or High-Activation: Block and Zeki  
Block’s view is that a subject’s conscious states can exist even when the subject is unable, even in principle, to know about, judge, or report that these states are conscious. This view divorces a subject’s phenomenal judgments or ability to report from demarcating the subject’s conscious states.

Block’s approach requires finding a core (neural) realizer of conscious states that does not include the subject’s ability to form phenomenal judgments regarding, or to report, these states. This is a difficult empirical task, as Block notes, since report and judgment are crucially bound up with how we, at least initially, objectively determine that conscious states are occurring in a subject (Block 2007, p. 483; Cowey and Stoerig 1995). Block’s method for demarcating a subject’s conscious states therefore, initially, relies on the subject reporting the conscious states. However, he eventually drops this method.

Block gives a set of considerations for denying that the method for demarcating a subject’s conscious states essentially involves a subject’s phenomenal judgments regarding, or ability to

report, these states. Block believes that a core (neural) realizer for conscious states is isolated from, and so does not involve, the systems underlying a subject's ability to issue reports or phenomenal judgments regarding their conscious states. This means that, although judgment and report can be initially used, the systems underlying these abilities in a subject are not essential to demarcating a subject's conscious mental states.

According to Block, a neural realizer stripped of systems underlying report and judgment forms a core demarcation for a subject's conscious mental states. For Block, only the initial way of finding conscious states relies on neural systems involving report or phenomenal judgment. Once these initial methods are dropped, a 'core' neural realizer can be identified that is stripped of these inessential operations.

Block begins isolating what he calls the core (neural) realizer of a subject's conscious states from the subject's systems of report and judgment based on Sperling's and Lamme's data. Like Lamme, Block notes that the number of a subject's conscious states seems to overflow the subject's ability to simply report or judge that conscious states are occurring, at least without appropriate cueing. Cued-reports or judgments are key to showing that more states are conscious than subjects initially report or judge to be conscious. For example, Block believes that subjects can consciously see all of a quickly flashed visual display and that this information slips away as the subject tries to simply report or judge that it is conscious.

Sperling's data and that of Lamme are not enough for Block to find a core (neural) realizer that is truly divorced from phenomenal judgment and reportability. To make this next move, he needs to get away from the notion of possible report or phenomenal judgment as part of the demarcation for a subject's conscious states (Block 2008, p. 311; Naccache 2006). To take this step, Block relies on data where researchers disrupt the neural signal in a subject at various

times. By doing so, experimenters can manipulate whether subjects issue reports, or phenomenal judgments, regarding whether stimulation causes conscious states (Block 2007, p. 310-11).

For example, the feed-back loop of activation between V1, primary visual cortex, and MT(V5), a motion processing area, seems to be involved when subjects issue reports or phenomenal judgments about experiencing motion. If the activation in the loop, what is sometimes called ‘reverberation’, is disrupted on the way to V5 from V1, or on the way back from V5 to V1, no phenomenal judgment or report of conscious states is issued by the subject.

However, if the signal is allowed to complete its loop then phenomenal judgments ensue, at least based on subject’s reports (Block 2007, p. 496). For Block, this data indicates that completion of the feed-back loop is the core (neural) realizer of a subject’s conscious states as of motion. He believes conscious states can be realized in this way even when phenomenal judgments or reports are not made, and cannot be made, for such states by the subject who has these states.

Block concludes that subjects can have conscious states even in cases where they cannot, even in principle, form phenomenal judgments or reports regarding such states. Block simply argues that there is no more reason to believe systems for report or phenomenal judgment are central to conscious states than to deny such claims. He, therefore, concludes that subjects can have conscious states outside their ability to report or form phenomenal judgments. Block’s view, therefore, allows for multiple isolated “islands of conscious experience” to exist distributed within a single individual even when the subject cannot form judgments about, or report, these states. He forms this conclusion in terms of conscious states occurring in isolated modular processing (2007, p. 481). We will see more on Block’s view in section 5.1.1.2.

Zeki’s picture is very similar to Block’s. Zeki differs from Block in two important ways.

First, he thinks the core (neural) realizer of conscious states is based on the level of activation in an area rather than reverberation between two neural areas. Second, Zeki emphasizes various notions of ‘micro-conscious’, ‘macro-conscious’, and ‘unified-conscious’ states. Generally, these are not major differences, so Zeki’s picture is very similar to Block’s.

‘Micro-conscious’ states, for Zeki, are not tied to a subject’s ability to report or judge those states to be conscious (Zeki 2003; Zeki 2007, p. 584 and 586). This part of Zeki’s picture is very like Block’s. However, level of activation in a specialized processing area is enough, on Zeki’s view, for this kind of conscious state (2003, p. 125). So, Zeki disagrees with Block about the core (neural) realizer of a subject’s conscious states.

‘Macro-conscious’ states, for Zeki, then bind sets of two or more ‘micro-conscious’ states. Such states are the combination, say, of isolated representational states. These states also count as conscious states on Zeki’s model, even if a subject cannot report or judge such states to be conscious.

Finally, ‘unified-conscious’ states bind all the ‘micro-‘ and ‘macro-conscious’ states into something that can be judged and reported by the subject as conscious (2007, p. 585). Only these last states are tied to a subject’s ability to report or form phenomenal judgments. For Zeki, damaging systems underlying report or phenomenal judgment systems, for example, leaves a number of conscious states intact in a subject at the micro-, or macro-, conscious levels.

Several forms of data suggest the main difference between Block’s and Zeki’s pictures. Patients with lesions to their primary visual cortex, V1, sometimes report having conscious experience as of motion. The motion processing center, V5(/MT), is intact in these subjects and does show the requisite level of activity for processing motion. This data suggests Block is wrong about reverberation being the core neural realizer of conscious states (see Block’s own

statements roughly to this effect in his 2007, p. 497-8). If this activity alone is sufficient for conscious states, then, reverberation is not necessary for demarcating a subject's conscious states (Zeki 2007, p. 586).

Important for Zeki, the right level of activation in a processing area can occur without involvement of more frontal areas. These areas of a subject's brain are involved in a subject's ability to report and judge certain states to be conscious. So, Zeki too isolates a subject's conscious states from the subject's ability to judge or report these states to be conscious.

The data Zeki cites lead him to conclude that isolated occurrences of high activity are enough for what he calls 'micro-conscious' and 'macro-conscious' states. Given the possible isolation of high levels of activity from other processing in a subject's mental system, Zeki also holds that islands of conscious states can exist in a subject without the subject's knowledge, or the subject's ability to report or judge that such states are occurring. This 'islands of consciousness' view is very much like Block's picture.

#### 4.3 Summary and Conclusion

This chapter has briefly covered several possible demarcations for where the explanatory value of conscious mental state explanations could begin in the information-processing picture of the human mind. From the above we can see that a variety of theories hold that conscious states have an explanatory value. We can now turn to a more detailed defense of the explanatory value of unconscious mental states.

## CHAPTER 5: The Explanatory Value of “The” Cognitive Unconscious

This chapter provides a defense of the explanatory value of unconscious mental states. The claim is that some human behaviors, driven by perceptual and cognitive states, are best explained by employing unconscious mental states of one kind or another.

The support for this claim comes in three forms. First, there is a coherent information-processing model for mental states which account for certain of a subject’s perceptually and cognitively driven behaviors, in part or whole, with unconscious mental states. Second, independent criteria support explanations that use unconscious mental states. Finally, accounts that employ unconscious mental states are supported by the success of such models in modern psychological and neurological literature. Later sections of this chapter cover various possibilities for the kind, number, and unity of unconscious mental states in human subjects.

The term “cognitive unconscious” is used to capture the above claim regarding a subject’s perceptual and cognitive unconscious mental states. As noted earlier, the term “cognitive unconscious” does not restrict a subject’s unconscious states to merely cognitive states. Perceptual states are also included under the label “cognitive unconscious”. The term “cognitive unconscious” is generally employed to distinguish unconscious mental states as they are employed in empirical psychology and neuroscience from unconscious states as they are employed in psychoanalytic theories (Shevrin 1992; Shevrin and Dickman 1980).

### 5.1 Unconscious Mental States: The Argument Outline

Establishing the explanatory value of unconscious mental states requires arguing that two classes of a subject’s mental states explain different behaviors of the subject and by further arguing that one class of mental states is plausibly construed as the subject’s conscious mental states. The best defense of the explanatory value of unconscious mental states simply shows that some of a subject’s mental states explain behaviors that the subject’s conscious mental states do

not explain.

If a subject's conscious mental states do not explain all of a subject's behaviors which are explained by the subject's mental states then unconscious mental states explain some of the subject's behaviors. So, the explanatory value of unconscious mental states equals the number of a subject's behaviors valuably explained by their mental states minus the number of the subject's behaviors valuably explained by their conscious mental states.

1)  $\exists x \exists y \exists z$

2) X is the class of a subject's behaviors explained by the subject's mental states

3) Y is the class of a subject's behaviors explained by the subject's conscious mental states

4) The non-overlap of X, Y = Z.

So, 5) Z is the class of a subject's behaviors explained by the subject's unconscious mental states

6)  $Z \neq 0$  because

6a) There is a coherent information-processing model that includes unconscious mental states

6b) Naturalistic criteria for subject's conscious states allow for unconscious mental states

6c) There are numerous successful empirical models that employ unconscious mental states

To support this set of claims, especially the conclusion regarding unconscious mental states, requires arguing for two sets of mental states, that the two sets of mental states do not overlap in terms of the behaviors they explain, and that one set is plausibly construed as a subject's conscious mental states. This leaves a set of states plausibly described as the subject's unconscious mental states.

Chapters two and three argued for a set of states plausibly considered a subject's mental states. Chapter four covered some popular accounts of conscious mental states. The current chapter covers a further account for mental states that can plausibly be considered a subject's conscious mental states. Finally, the current chapter also argues that the two sets of states do not overlap by illustrating how they explain different behaviors of a subject. Given that one set of mental states is plausibly considered the subject's conscious mental states, an explanatorily valuable set of mental states is left that can be described as the subject's unconscious mental

states.

### 5.1.1 Unconscious Mental States: The Coding Architecture

One kind of support for the claim that unconscious mental states are explanatorily valuable is a coding architecture. The most likely coding architecture for human behavior, given the current evidence in psychological and neurological literature, shows that the coding processes needed for creating mental states are plausibly different from the coding processes needed for creating conscious mental states. For conscious and unconscious mental states to explain different behaviors these states must have different causal roles. The different causal roles of these two classes of mental states arise from the different ways a subject's conscious and unconscious mental states are coded. Mental states that are coded in different ways show how a coding architecture can support employing both conscious and unconscious mental states to explain certain human behaviors.

As we saw in chapters two and three, creation of information-processing states that have mental status requires various kinds of coding. Coding gives mental states special properties beyond those of mere information-bearing states. Forming information-bearing states into mental states has a cost in terms of extra information processing. This cost is why not all of a subject's information-bearing states are coded as mental states.

Conscious mental states, on most or all naturalistic theories, have properties beyond mere mental states. Objections to this assumption will be addressed in chapter six. For now, on naturalistic theories, it seems that the creation of information-bearing states which can be considered conscious mental states requires further information processing beyond that required for creating mental states. As noted, coding information-bearing states has a cost in terms of information-processing resources. So, relative to creating mental states, the special construction processes needed for creating conscious mental states cost extra in terms of further information

processing. Further information processing is why not all of a subject's mental states are coded as conscious mental states.

Limited information-processing resources in a coding architecture constrain the number of states that can be coded in certain ways. Limiting resources in a way that constrains the mental states that are, or can be, coded as conscious mental states is plausible, at least given the naturalistic assumption that different coding processes are likely needed for these two kinds of states. Employing both conscious and unconscious mental states allows researchers to explain more of a subject's behaviors in various ways. Various kinds of mental states help to both explain and unify certain classes of behavior by accounting for each class with a single kind of explanatory state. With the right information-processing architecture, various kinds of conscious and unconscious mental states can plausibly explain different behaviors of a subject.

An analogy shows how this picture can work. Making keys and limiting the operations in which certain keys function is analogous to making mental states and conscious mental states that each control different behaviors of a subject. The explanatory states employed in psychological theory start off as mere information-bearing states. Bits of metal of the right size and strength appropriate for making keys are analogous to information-bearing states transduced by the appropriate sensory processing systems of a subject. Some of these information-bearing states can directly control certain of the subject's behaviors, such as reflexes.

Other processing allows information-bearing states to acquire an explicit focus of information. By analogy, bits of metal can be shaped into templates for various kinds of keys. Shaping gives the bits of metal more explicitly focused information relative to later operations, such as machines that can cut grooves into the key-templates. As we saw in chapter three, amygdalar states, analogous to shaped bits of metal, are information-bearing states that can be

employed in template matching operations. We also saw in chapter three that the level of focused information relative to later operations in amygdalar states may not be enough to create states with mental status.

Mental states more clearly arise at a level of processing that gives states a more determinate focus of information relative to certain later operations. By analogy, cutting grooves into key-templates gives the metal more focused information relative to particular operations, such as opening doors. Likewise, when information-bearing states have some of their information made explicit in the right way it gives these information-bearing states a determinate focus of information (see chapter two; Fodor 2008; Marr 1982). Along with other features this can give such states mental content (see chapter three).

The process of cutting keys for opening doors determines what kinds of keys are made. By analogy, information-bearing states can be considered perceptual and cognitive mental states based on the different information in these states. The various kinds of information made explicit in information-bearing states determine whether perceptual or cognitive states are created and thereby which operations they function within. This analogy shows how mental states can be made from information-bearing states.

Finally, for the analogy to conscious mental states to go through, some of the keys need to be coded in a further way. For example, some of the keys can be color-coded. Regulating processes which allow only color-coded keys to enter certain operations limits the doors these keys can operate within. Conscious states, by analogy, are also limited in the operations within which they can function. That is, at least on the naturalistic assumption that not all of a subject's mental states are conscious states. Conscious states are mental states with a special code that allows them to have unique causal properties and so to control a unique set of behaviors.

A further analogy for conscious states can be cast in the form of DNA-coding. Only one kind of coded state allows for certain decoding effects. Other forms of DNA have focused information, analogous to mental content. However, different codes in certain of these states would allow them to have unique causal effects.

DNA and keys are not considered mental codes because the type of information in the states and the operations in which they function are not psychological states or operations. Not every information-processing system is valuably described as a mental system operating on mental states. However, the analogies are still apt. Furthermore, no single center of processing is needed to code states in this particular manner nor is it necessary to posit a single system within which these specially coded states operate. By analogy, there need be no single “consciousness system” in human subjects.

The appropriate information-processing architecture allows for coding differences in various mental states. These differences make a subject’s mental states have unique causal effects different from mere information-bearing states and other of the subject’s mental states, such as the subject’s conscious mental states. So, employing unconscious mental states, as contrasted with a subject’s conscious mental states, is coherent, given an appropriate information-processing architecture controlling some of the subject’s behavior.

The coding architecture just outlined does not prevent some of the *content* of a subject’s unconscious mental states, at least at certain times, to become the content of the subject’s conscious mental states. However, it does mean that, based on coding differences, a subject’s unconscious mental states can influence the subject’s behavior in ways the subject’s conscious mental states cannot.

Radically, even if all of the content of a subject’s unconscious mental states were mirrored in

the subject's conscious states this would not argue against the explanatory value of unconscious mental states controlling some of the subject's behaviors. Although it is implausible that a subject's mental content is doubled in two sets of states, a subject's unconscious mental states could, even on this radical picture, still have unique control over some of the subject's behaviors.

None of the previous considerations conclusively show that the information-processing architecture of the human mind does employ unconscious mental states. It simply shows a coherent model for dividing the human mind in this way. We will see more reasons below why we should consider the information-processing architecture of the human mind to be divided in a way that allows for conscious and unconscious mental states.

#### 5.1.1.1 Summary of the Overall Picture

The information-processing architecture of the human mind just outlined is complex, especially when perceptual and cognitive states are included for consideration. The various distinctions between kinds of mental states in human subjects are summarized in figure 3. This figure is created by superimposing something like figure 1 onto something like figure 2. Figure 3 is not a model for the causal interrelation of the states pictured. Nor is it a picture of how content flows in the human information-processing system.

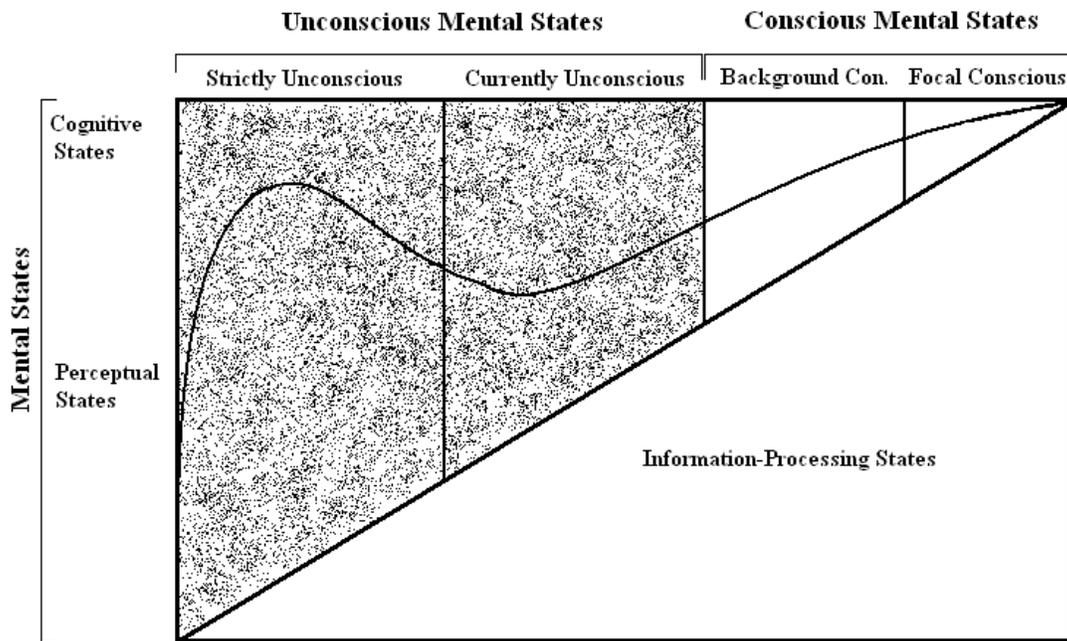
However, figure 3 does show that a subject's mental states are information-processing states that are coded in various ways. Of the subject's mental states, perceptually and cognitively coded mental states can be distinguished. Further dividing perceptual and cognitive mental states into states coded as mere mental states and those coded as conscious mental states can be done. This picture makes for two main classes of a subject's perceptual and cognitive mental states, namely the subject's conscious and unconscious mental states.

As we will see more below, allowing for further divisions is also explanatorily valuable. It is likely that two kinds of conscious mental states, and two kinds of unconscious mental states,

explain various behaviors of the subject. Strictly unconscious mental states explain some of a subject's behaviors that the subject's currently unconscious mental states do not explain. A subject's strictly unconscious mental states are mental states that cannot become conscious for that subject. A subject's currently unconscious mental states can become conscious if the subject is given the right cue. Finally, as we saw in chapter four, a subject's focally attended conscious states are distinct from the rest of the subject's conscious mental states, thereby arguing for two kinds of conscious mental states (see section 4.2.1.2 above).

There is also reason to think the number of a subject's strictly unconscious perceptual states outnumber the subject's strictly unconscious cognitive states. This pattern is reversed for a subject's currently unconscious mental states. This picture of the number of a subject's various kinds of mental states is briefly defended in section 5.6. Other pictures regarding the number of a subject's conscious and unconscious mental states are covered in sections 5.5.1-5.5.4.

Importantly, different pictures regarding the number of various kinds of a subject's unconscious mental states are not required for the general thesis to go through. All that needs to be shown for the main thesis is that unconscious mental states, of whatever kind, help psychologists to account for various human behaviors. The coding architecture just outlined simply shows the possibility of employing unconscious mental states in an information-processing model of the human mind.



Psychologically Explanatory Information-Bearing States  
**Perceptual and cognitive mental states as seen within  
conscious and unconscious classes of mental states**  
**Figure 3**

5.1.1.2 Extending Conscious Mental States beyond Working-Memory, But Not Too Far  
Given the current state of psychological and neurological literature, researchers seem to find

it useful to employ conscious and unconscious mental states when explaining certain human behaviors. To show the explanatory value of unconscious mental states in human subjects is not only possible but plausible is most easily done within a particular picture that contrasts human subjects' conscious and unconscious mental states.

For example, recent debate regarding the isolation of a subject's conscious states from the subject's systems for report gives one picture of the contrast between conscious and unconscious mental states within human subjects (see Block 2007). Subjects, at a time, can voluntarily report, for example, the three to four states active in their working memory. One version of this picture, as we have seen, holds that a subject's conscious mental states are greater in number than the three to four states active in the subject's working memory. The most plausible accounts of the data extend a subject's conscious mental states beyond the states in the subject's working

memory but not to all of the subject's mental states.

Block, for example, gives reasons for extending conscious mental states beyond the special processing system of a subject's working memory (Block 2007; see also Lamme 2006). As we saw in chapter four, some researchers limit conscious states to the three or four items active in a subject's working memory (see section 4.2.1.1). Block uses data like Sperling's and Lamme's to show that the states which subjects judge to be conscious are greater in number than those in the subject's working memory.

As we have already covered, subjects say they see more than three to four items when large visual arrays of letters and numbers are briefly displayed (Sperling 1960; Lamme 2003). Even though subjects only report three to four items from such displays, subjects can be cued to report phenomenal judgments for *any* three to four items in fairly large visual arrays. The fact that subjects can report any three to four items when appropriately cued suggests that subjects are correct when they report that they consciously represent more than three to four items at once. It also suggests a subject's conscious states extend beyond those active in the subject's working memory.

Block goes on to use data from frontal lobe lesions, which damage working memory circuits, to show that damage to working memory does not limit the number of states subjects judge to be conscious (Heath et. al. 1949; Pollen 2008). From these data Block rightly concludes that a subject's conscious states are not isolated to the three to four states active in the subject's working memory.

Block goes on to suggest that a subject's conscious states can occur in the subject's encapsulated information-processing systems, such as modular processing (Block 2007, p. 497-8). For example, Block suggests that states which are strongly activated, but were not accessed

by the subject's working memory, could form the neurological basis for a subject's conscious mental states. Block believes these states are not only outside a subject's working memory but also occur within isolated processing of the subject.

To illustrate, Block cites data showing that disrupting low-level reverberation in a subject's brain also disrupts the subject's report of conscious states. Low-level reverberation underlies the strength of activation of a state and is isolated from a subject's systems of report. For example, such activation can occur in encapsulated processing. So, Block thinks a subject's conscious states are isolated from the subject's systems of report and can occur within the subject's encapsulated processing. As we have seen, Block believes a subject's conscious mental states can occur without the subject's knowledge or their ability to report these mental states.

When subjects report their conscious states the subject's systems for report must be involved. These states are generally thought to occur in a subject's working memory. This fact is what leads several researchers to limit a subject's conscious mental states to those active in the subject's working memory. Block wants to extend a subject's conscious mental states well beyond those the subject can report.

However, Block notes a problem. Disrupting activity in isolated processing, such as low-level reverberation in modular processing, disrupts a subject's reports of conscious states. Unfortunately for Block, this does not show that activity in isolated processing is the sole basis for a subject's conscious states. Disrupting low-level reverberation also disrupts other processes as well.

Block attempts to give reasons for extending a subject's conscious states to include those active in a subject's encapsulated processing. However, there is little data that would support such an extension. There are states between those in a subject's working memory, or states in the

subject's systems for report and phenomenal judgment, and those in the subject's encapsulated processing. Since Block holds that states actually in a subject's working memory are not the only mental states a subject judges to be conscious, he argues that it is "hard to see why" states between a subject's working memory and those in encapsulated information processing would count as a subject's conscious mental states (Block 2007, p. 498). So, he thinks states *potentially* in a subject's working memory are not good candidates for being the subject's conscious mental states.

However, the states coded as potentially available to a subject's working memory, or systems for report and phenomenal judgment, are good candidates for a subject's conscious states. Coding is a way to allow researchers to count states outside working memory as conscious without extending a subject's conscious states into the subject's isolated processing, such as modular processing. This would mean that none of a subject's conscious mental states are un-reportable by that subject, at least when the subject is appropriately cued.

When subjects are appropriately cued they judge states potentially in their working memory to be conscious, and often judge such states to have been conscious before the cue occurred. So, a set of states between those in the subject's working memory and those in the subject's encapsulated processing, which are coded in the right way, are plausible candidates for the states causing a subject's cued phenomenal judgments (Pollen 2008; see section 5.2.1 below).

States in a subject's focal attention or working memory, which cause a subject's un-cued phenomenal judgments, are almost universally accepted as conscious states of the particular subject. These ideas lead to the conclusion that states in a subject's working memory or focal attention and states currently coded to be available to the subject's working memory and focal attention are the states most plausibly considered the subject's conscious mental states. These

states do not include all of a subject's mental states. For example, these states do not include the mental states in a subject's encapsulated processing.

As noted, a subject's conscious mental states should not be treated as the states that are available *in* any particular mental system of the subject, such as the subject's working memory. Characterizing the occurrence of conscious mental states as just those states processed by a particular mental system is a form of Cartesian Materialism (Dennett 1991). No such 'consciousness-system' has been found. Coding approaches to conscious mental states can avoid positing such a system.

Instead, we can characterize a subject's conscious mental states as states that are available *as* certain kinds of coded states. The mental states of a subject that are available in a certain format are potentially accessed by special processing systems, such as the subject's working memory. In this model, no particular system need be invoked for the formatting processes that create a subject's conscious mental states.

In this model, a subject's conscious mental states are still linked to the subject's knowledge of, and ability to report, these states, at least when the subject is appropriately cued. Given the right information-processing architecture and given the above data, this coding model is a possible, as well as plausible, treatment of human subjects' mental states that allows for both conscious and unconscious mental states.

## 5.2 Unconscious Mental States: Three Criteria for Conscious Mental States

Including conscious and unconscious mental states in the information-processing model of the human mind helps researchers explain certain data. For example, explaining data regarding a subject's phenomenal judgments, and related criteria, also supports the use of unconscious mental states.

It is often thought that behaviors controlled by a subject's conscious mental states do not

exhaust all of a subject's behaviors controlled by the subject's mental states. So, besides just being a plausible and coherent model of mind, other criteria also support the use of unconscious mental states when explaining some human behaviors.

Extending a subject's conscious mental states beyond those states active in the subject's working memory, but not as far as mental states in the subject's encapsulated processing, leaves a set of mental states that are plausibly considered unconscious mental states. The idea has been to differentially limit the coding-architecture outlined above in regard to a subject's conscious and unconscious mental states.

This position is supported by independent criteria linked to a subject's phenomenal judgments. Criteria linked to a subject's phenomenal judgments do not exhaust mental-state explanations for that particular subject. So, these criteria leave a significant number of a subject's behaviors explained by mental states that are not conscious. Conscious mental states are judged by subjects to be mental states that have subjective qualities. A subject's conscious mental states occur in a way that makes the subject believe just those states have subjective qualities. Whether or not subjective qualities exist, subjects judge certain mental states to occur in this special way (Dennett 2007).

A subject's phenomenal judgments are thought to pick out a subclass of the subject's mental states that both researchers and the subject consider to be conscious. However, subjects can sometimes be mistaken in their initial judgments about which of their own mental states are conscious states. So, a further criterion required for a subject's conscious mental states is also needed. This criterion is related to the justification subjects have regarding their phenomenal judgments.

Finally, many researchers, as well as most subjects, believe that human subjects can express

the content of their conscious states only in particular ways (Rey 1988). So, a third criterion related to expression can also be included when considering which of a subject's mental states are conscious.

Each of these criteria is used individually to demarcate a subject's conscious mental states (Dennett 2007; Dretske 2006; Rey 1988). However, these criteria can be seen to pick out the same states. So, the combination of these criteria helps to show which of a subject's mental states are best treated as conscious. The right kind of coding of a subject's mental states shows how certain of these states can cause subjects to form justified phenomenal judgments that can be expressed only in particular ways.

Spelling out a plausible 'consciousness code', or the form of states that have the right causal effects to be considered conscious, will be the task of the next few sections. Some of the nuances for each criterion need to be left undeveloped since any actual demarcation between conscious and unconscious states should be left open for the time being (see sections 5.5-5.6 for potential demarcations). The important point for each of the present criteria is that some mental states are not currently, or cannot be, coded as states that fulfill these criteria. So, based on these criteria, various kinds of unconscious mental states are still valuable explanations for a variety of human behaviors.

#### 5.2.1 *Criterion #1: Dual Subjective Judgment Criterion*

The first criterion that helps to support the claim that unconscious mental states have explanatory value relies on a subject's phenomenal judgments. The discussion here follows much current literature in psychology and cognitive science regarding a subject's judgments about their conscious mental states (Lamme 2006; Sperling 1960; see also sections 4.2.1.2 and 5.1.1.2 above). As we have seen, subjects judge, with or without cues, that some of their mental states have subjective qualities. So, a subject's cued and un-cued phenomenal judgments help to

distinguish classes of mental states.

Two of the classes of a subject's mental states picked out by the subject's phenomenal judgments are plausibly considered the subject's conscious and unconscious mental states. The states that subject's judge to be conscious in various ways can plausibly be considered the subject's conscious mental states. Subjects do not seem able to judge all of their mental states as phenomenal states at a time. Furthermore, subjects do not seem able to judge some of their mental states as phenomenal states at all. So, subject's phenomenal judgments help to show which of their mental states are conscious as well as which are unconscious.

The criterion concerning a subject's phenomenal judgments has two parts. First, subject's can judge some of their mental states to be conscious states with or without cue. Second, many of the states subjects judge to be conscious are also judged by the subject to *have been* conscious states prior to the time when the subjects judged these states to be conscious. If a subject judges certain mental states to be conscious, with or without cue, and to have been conscious, especially prior to receiving a cue, it is good reason to believe that such mental states are, and were, conscious. This is true whether or not subjective qualities exist.

Data regarding a subject's phenomenal judgments suggest such judgments are not exhaustive of all of a subject's mental states. So, unconscious mental states are supported by this criterion. Current evidence suggests that some of a subject's mental states cannot cause phenomenal judgments at all. Other of the subject's mental states cause phenomenal judgments only when the subject is appropriately cued.

Of the states that cause phenomenal judgments only when the subject is cued, some are also judged by the subject to not have had subjective qualities before the cue occurred. The mental states of a subject that cannot cause phenomenal judgments are the subject's strictly unconscious

mental states. The states that subjects judge to have subjective qualities at a time but also judge to not have had subjective qualities before this time are the subject's currently unconscious mental states. Therefore, various kinds of phenomenal judgments pick out a subject's two kinds of conscious as well as the subject's two kinds of unconscious mental states.

To illustrate some of these different classes of mental states think of two examples. First, think of someone who has just been reminded of something. The subject will often report that the content of the mnemonic state is now conscious. However, they will also often report that the content of the mnemonic state was not conscious before they were cued to report that state. This is different from the phenomenal judgments subjects issue when they are cued to attend their diplopic visual content. Think of double vision in human subjects. Diplopic content is illustrated when subjects are asked to look at a distant object but attend a close object. In such cases, subjects often report that the close object is represented by a double image (Tye 2008). Once cued to this double image, subjects almost always report that the two images are conscious and were conscious before they were cued to report the double images. This contrasts with subjects phenomenal judgments for mnemonic content that became conscious when the subject was appropriately cued.

So, the combination of two of a subject's phenomenal judgments picks out the subject's conscious, and unconscious, mental states. A subject's phenomenal judgments distinguish the subject's conscious states from the subject's currently unconscious mental states. A subject's currently unconscious mental states can cause phenomenal judgments with the right cue. However, these states will not cause the subject to form judgments that the states were conscious before the cue occurred. Whether subjects are cued or not, the subject's strictly unconscious mental states never cause phenomenal judgments. So, current evidence regarding human

subjects' phenomenal judgments supports several classes of mental states, in particular two kinds of conscious, and two kinds of unconscious, mental states.

On this picture, a subject's conscious states cause the subject's various phenomenal judgments. So, phenomenal judgments simply help to demarcate a subject's conscious mental states from the subject's unconscious mental states. Phenomenal judgments do not create a subject's conscious mental states. The various ways mental states give rise to phenomenal judgments, or do not give rise to such judgments, helps show different classes of mental states.

A subject's unconscious mental states either do not cause phenomenal judgments at all or do not cause the kinds of phenomenal judgments that indicate a subject's conscious mental states. Data regarding the different profiles in human subjects' ability to form phenomenal judgments shows the explanatory value of attributing two kinds of unconscious mental states to human subjects. This is true when assuming, as data suggests we should, that a subject's phenomenal judgments, of whatever kind, do not occur for all of a subject's mental states.

#### 5.2.2 *Criterion #2: Justification Criterion*

Subject's phenomenal judgments are not the final authority establishing that certain of their mental states are, or are not, conscious states. Subjects are sometimes mistaken regarding their judgments about their conscious mental states. For example, sometimes subjects simply do not notice certain mental states that researchers and cued subjects count as conscious (Lamme 2006; Sperling 1960). Diplopic visual content is one clear example (Tye 2008).

Phenomenal judgments must be caused in the right way to indicate conscious mental states. When subjects are not in the know and do not have to guess about which mental states are causing their phenomenal judgments then such judgments indicate conscious mental states. This shows a second criterion that helps demarcate a subject's conscious mental states.

The discussion of this criterion closely follows Dretske (2006). Dretske holds that an

unconscious mental state is the “absence of a justifying reason” for a subject’s own thoughts or actions (2006, p. 170). For example, as Dretske notes, in cases where subjects behave based on unconscious perceptual content, “there exist, in what S perceives, [explanatory] reasons why she [, say,] pushes the right button without there being, in what she perceives, [justifying] reasons for pushing the right button” (2006, p. 170). He notes that this distinguishes “explanatory reasons, the reasons *why* S does A or believes P, from justifying reasons, S’s reasons *for* doing A or believing P, the reasons S (if able) might give to justify doing A or believing P” (2006, p. 168).

The idea is just that there are reasons for a subject’s thoughts or actions and then there are reasons the subject can judge for themselves, or report to others, as reasons for their thoughts and actions. The latter states, if they cause phenomenal judgments, are a subject’s conscious mental states, since a subject can cite these states as reasons for their phenomenal judgments. The reasons for thought or action that subjects can judge for themselves or report to others as reasons for thoughts or actions are not thought to be exhaustive of all of the subject’s mental states (Nisbett and Wilson 1977; Wilson 2002; Ericsson and Simon 1980). So, this criterion leaves an explanatory role for unconscious mental states.

Doris summarizes this criterion well. He notes that unconscious mental states figure in “cases where actors would reject the [mental state] *causes* of their behavior as *reasons* for their behavior, or, to put it another way, where the *motivations* of behavior would not be judged fit to figure in *justifications* of behavior “ (Doris 2009, p. 66). Such cases illustrate the kinds of behaviors where the subject’s unconscious mental states explain some of the subject’s behavior.

The notion of mental states that justify a subject’s phenomenal judgments needs to be treated carefully. Mental states that justify a subject’s behaviors generally consist of both conscious and unconscious mental states. Likewise, a subject in the know about unconscious mental states

causing their behavior can report that such states are causing their behavior. However, subjects in the know about their unconscious mental states can only report indirect knowledge of the states. If subjects in the know report phenomenal judgments for their unconscious mental states the subject does not report the states “from their point of view”. States that make up a subject’s point of view are states the subject can report as their reasons for belief or action when they are not in the know and when they are not guessing.

An example helps to illustrate the notion of mental states in a subject’s point of view. There are cases where a subject’s childhood memories are mental states that the subject cannot report as a reason for their further thoughts or actions, at a particular time. If a subject cannot, for example, declaratively remember the color of their childhood house at some time they will often say that they do not know, and are not conscious of, the color of the house. This does not mean the mnemonic mental state about the color of the house is not present in the subject. The state may be able to cause certain behaviors of the subject without the subject’s knowledge. However, the subject cannot, at this particular time, attend to any mental state that they would report as justification for forming a thought or action based on the color of the house (Feldman 1988). So, subjects cannot form justified phenomenal judgments from their point of view regarding these kinds of states at this time.

If a subject in this situation goes on to form a correct belief or an appropriate action based on a state about the color of the house, the occurrence of such a belief or action cannot be distinguished, at least directly by the subject, from a correct guess. Subjects in this situation may still have a justified belief that can form their further thoughts or actions. This is because a standard epistemological position holds that thoughts or actions only need to be caused in the right way by states with the right content to be justified. However, the subject’s systems that

form reports or phenomenal judgments operate based on a state that is as good as a guess to those systems. So, from the point of view of these systems, we can distinguish between beliefs or actions that are justified simpliciter and those that are also justified from the subject's point of view.

Justification from a subject's point of view is not justification simpliciter, at least when justification simpliciter is treated in externalist ways (Bergmann 2006). Two mentally identical twins, for example, one in the actual world and one in a world that seems like the actual world but is not, will both have the same reportable reasons for their actions or beliefs, at least from their points of view when they are not in the know and are not guessing about their actual situations.

This does not entail that both are equally justified simpliciter in their beliefs and actions (Bergmann 2006, p. 60). The twins simply have the same justification from their points of view. The notion that distinguishes conscious and unconscious mental states is not justification simpliciter, since both twins, when not in the know, will have the same conscious and unconscious mental states. That is, they will have the same justification for thought or action only from their point of view. Justification from the subject's point of view does not entail that a subject is actually justified in their beliefs about the world.

The notion of justification from a subject's point of view is about the mental states that a subject can report as causes of their thoughts and actions, when the subject is not in the know and when they are not guessing. Subject's phenomenal judgments are justified from their point of view when the judgment is justified and when the subject does not have to guess and when they are not in the know.

So, phenomenal judgments are only justified simpliciter when subjects "directly" know, or

know from their point of view, about the states causing their phenomenal judgments. In such cases, subjects do not have to guess what state caused their phenomenal judgment. We can tell which circuits are active when subjects are lying or simply guessing independent of criteria for conscious states (Lee et. al. 2002). So, justification from a subject's point of view is not a circular way to demarcate a subject's conscious mental states.

Importantly, this criterion leaves an explanatory role for unconscious mental states. A subject's conscious mental states are not thought to be exhaustive of all of the subject's mental states. Subjects do not seem able to form justified phenomenal judgments from their point of view for all of their mental states at a time, nor for some at any time. This data suggests that a subject's unconscious mental states are not currently causing, or cannot cause, justified phenomenal judgments *from the subject's point of view*. That is, when the subject is not in the know and is not guessing about those states.

### 5.2.3 Criterion #3: Expressibility Criterion

Assessing which of a subject's mental states cause phenomenal judgments is often done by way of the content expressed in the subject's behavior. Often this is accomplished through a subject's verbal utterances. Two classes of a subject's behaviors express the content of the subject's mental states in different ways. Data suggest that the mental states that cause a subject's justified phenomenal judgments control only some of the subject's behaviors. So, these particular behaviors express mental states that can be considered conscious. The behaviors of a subject that express other mental states express the subject's unconscious mental states (Rey 1988).

Therefore, a third criterion that helps support an explanatory role for a subject's unconscious mental states is related to expression. Not all of the mental content attributed to subjects by researchers to explain the subject's behaviors can be expressed by the subject in the same way.

Researchers and subjects treat only some kinds of expressive behaviors as indicating the subject's conscious mental states.

Determining which of a subject's expressive behaviors indicate conscious mental states requires some independent theory regarding the subject's conscious mental states. So, expression is not a foundational criterion for demarcating a subject's conscious mental states. However, some behavior is necessary to determine that a subject understands the notion of phenomenal judgments (Cowey and Stoerig 1995). This means expression of some kind is required for researches to determine a subject's conscious mental states. The discussion in this section follows Finkelstein (1999), Moran (1997), and Rey (1988). Rey, for example, holds that subjects are able to express conscious mental states only in very particular ways. Other behaviors of a subject express the subject's unconscious mental states.

Think of the case where a human subject expresses the grammar of their native language simply by speaking the language. Human subjects generally have no articulate knowledge about the mental states involved in the processing of the grammar in their native language. Subjects also do not report phenomenal judgments for these mental states. However, subjects can obviously behave based on the mental states involved in the processing of their grammatical utterances. For example, they can utter some content in their language that is grammatical. This behavior expresses the mental states involved in the processing that causes grammatical utterances.

The states involved in this behavior are mental states because they function in the way required for mental status (see chapter 3). For example, these states can represent and misrepresent focused information regarding various parts of the linguistic structure. So, in one sense subjects express the mental states involved in the processing that causes their grammatical

utterances. This is true even when subjects do not know about, and do not issue phenomenal judgments for, these states.

The mental states expressed in the form of a subject's utterances are not like the mental states expressed in the content of their utterances. Chomsky, for example, may someday figure out English Grammar and go on to express this content in some utterance. In such a case, the content of the utterance likely expresses mental states that Chomsky would judge to be conscious. However, the mental states that are involved in causing the form of Chomsky's own grammatical utterances are not thereby expressed by the *content* of this utterance. This is true even if the content of the token mental states that caused the grammatical form of that utterance have the same content as the token mental states that caused the uttered content. Expressing mental states involved in the processing of grammar is not like expressing mental states in verbal utterances.

What this means is that mental states play different causal roles. Chomsky does not express the token contents of the mental states involved in his own linguistic processing in the content of his utterance even when expressing type-identical content in the utterance. The token mental states that cause the content of his utterance are not the same as the token mental states that cause the grammatical form of that utterance. The right causal relation is not present for certain mental states to be expressed in the content of a subject's utterances. For example, this right causal relation is not present for mental states involved in the processing of a subject's grammar.

Data suggest that the information-processing architecture in humans makes some mental states that are not current causes of, or makes states that cannot be causes of, expressions in certain ways. Interestingly, these special ways of expressing content seem to align with the states that subjects judge to be conscious. Expression in the content of verbal utterances is not

exhaustive of the ways subjects can report conscious mental states. However, these other ways of expressing conscious content are also not exhaustive of all of a subject's behaviors that express the subject's mental states. So, this criterion also leaves an explanatory role for unconscious mental states.

#### 5.2.4 Summary of the Role of Criteria for Conscious Mental States

Making sense of conscious mental states is important for establishing the contrasting notion of unconscious mental states (Manson 2000, p. 149). If something like the above criteria are not good, and if subjective qualities are not the way to go, then the explanatory value of conscious mental states is threatened. If the explanatory value of conscious mental states is threatened then so, too, is the explanatory value of unconscious mental states.

However, most information-processing models of the human mind demarcate only some of a subject's mental states as conscious (Hirst 1995; Lahav 1993). Certain criteria, like those above, seem the best way to demarcate a subject's conscious mental states in information-processing accounts. Although these criteria are very plausible, the important point is that the criteria for conscious mental states, in all or most naturalistic information-processing models, demarcate only a subclass of a subject's mental states as conscious.

For naturalistic theories, an empirical demarcation is necessary for causal explanatory value. Potential conceptual identifications between conscious mental states and mental states, or some kind of dependence of mental states on conscious mental states, will not have causal explanatory value. So, demarcating a subclass of a subject's mental states as their conscious mental states remains a plausible overall picture of human subject's mental states. Objections to unconscious mental states, such as dependence relations or identity relations between a subject's conscious and unconscious mental states, will be addressed in chapter six.

Importantly, even if the above criteria fail, most or all naturalistic criteria for a subject's

conscious mental states will still leave an explanatory role for unconscious mental states. In this way, unconscious mental states remain a part of most or all information-processing theories of the human mind.

### 5.3 Unconscious Mental States: Explanatory Success

The final consideration that supports the use of unconscious mental states in explanations for some behaviors of human subjects is the success of empirical psychological and neuroscience models that employ unconscious mental states. The scientific literature speaks for itself in this regard.

We have seen the explanatory value of mental states throughout the above discussion. Furthermore, the explanatory value of conscious mental states is illustrated in certain phenomena that seem best explained by conscious states, at least as we have demarcated them above. For further illustration, there seems to be a role for conscious states in certain adaptation effects, certain kinds of perceptual effects, such as illusions or “filling-in”, certain forms of binocular rivalry, certain kinds of action guidance, certain kinds of learning, and possibly extended serial processing such as certain kinds of logical deduction (Moradi, Koch, and Shimojo 2005; Mayr 2004; Shapiro, Lu, Knight, and Ennis 2009; Pessoa, Thompson, and Nöe 1998; Blake 2001; O’Shaughnessy 1992; Clark 2001; Milner and Goodale 1995; Lovibond and Shanks 2002; Newell 1982; O’Reilly 2006).

To spell out the various roles conscious states have in explanations of the behaviors involving the above phenomena will take us too far from our main task. Suffice it to say that the explanatory role of conscious mental states is held by most cognitive scientists to be valuable. This is true whether or not subjective qualities actually exist.

The crucial point is that the value of conscious mental-state explanation for a subject is not exhaustive of the value of mental-state explanation for that subject. Most researchers hold that

the explanatory role of conscious states does not coincide with the explanatory role for all of the mental states of a subject. Other behavior driven by perceptual and cognitive mental states in human subjects are valuably treated as behavior driven by unconscious mental states. On this model, a subject's unconscious mental states underlie, impact, or parallel the subject's conscious mental states.

For example, theorists posit unconscious mental states that help create a subject's conscious states. Many mental states underlie and lead to what subjects judge to be conscious perceptual or cognitive states. Perceptual and cognitive illusions are clear examples where unconscious mental states are employed in the explanation of the processing leading up to what subjects judge to be conscious states (Rock 1983; Piattelli-Palmarini 1994). The states that subjects judge to be conscious states produced when viewing the Ebbinghaus Illusion, for example, are created by states that represent visual shape and visual context. In this illusion, a circle is surrounded by smaller, or larger, circles and the central circle appears, respectively, larger or smaller than its actual size. Based on subject's phenomenal judgments, the representational state for visual context is often considered an unconscious mental state that makes for the illusory perception.

At other times, unconscious mental states can impact the states that subjects judge to be conscious. Subjects judge certain mental state impacts to be outside the stream of states they judge to be conscious. Memories, for example, often "pop" into the stream of states that subjects judge to be conscious (Mandler 1994). When mnemonic states occur in this way subjects often judge them to be unrelated to the conscious states occurring before this impact. Subjects also often judge the mnemonic states to not have been conscious before this time. Because the state is a memory, we can assume that the mnemonic state existed prior to its effect on what subjects judge to be their conscious mental states. So, these states fit the profile of a subject's currently

unconscious mental states. In this way, unconscious mental states can be seen as impacting upon the stream of states that subjects judge to be conscious.

Other examples of such impacts include subjects choosing words when speaking and planning speech articulation. Subjects often speak without being conscious of the particular words they choose. Likewise, speech articulation for the beginning of a word is influenced by later parts of the word. Subjects often do not judge the states in these processes to be conscious, even when the subject is cued. These examples also show how unconscious mental states impact what subjects judge to be their conscious mental states (Jackendoff 1994). There are many examples that fit this kind of influence of unconscious mental states on the states that subjects judge to be conscious.

Finally, some unconscious mental states seem to bypass, or run in parallel to, the states that subjects judge to be their conscious mental states. This happens, for example, in certain forms of motor processing, such as when mental states guide certain aspects of a subject's reaching and grasping behavior (Milner and Goodale 1995; Clark 2007; Castiello et. al. 1991). For example, subjects form an appropriate grip size when reaching for various objects and often do not judge the states that guide this behavior to be conscious, even when they are cued to attend to such states. The unconscious mental states involved in these processes are often considered strictly unconscious mental states. These states directly control some of a subject's behaviors without the possibility of making for, or interacting with, the subject's conscious mental states.

These kinds of psychological and neurological explanations show the broad range and success of models employing unconscious mental states when accounting for certain human behaviors. The broad range of successful applications of models employing unconscious mental states supports the claim that unconscious mental states are explanatorily valuable.

The explanatory success of employing both conscious and unconscious mental states, based on coding in an information-processing architecture, is not just a fluke. Given current neuropsychological theory and explanation, the most likely information-processing architecture for explaining some human behaviors makes some of a subject's mental states unconscious at a time and others unconscious all of the time. The important point is that a significant number of successful explanations use models with unconscious mental states and this testifies to the explanatory value of unconscious mental states (Kihlstorm 1990).

#### 5.4 Summary of Defense

The claim that some of a subject's behaviors, driven by perceptual and cognitive states, are best explained by unconscious mental states, of one kind or another, is well supported. There is a coherent information-processing, or coding, architecture, independent criteria, and a broad range of explanatory success that back up this claim. These considerations suggest that a subject's conscious and unconscious mental states do not overlap in the behaviors that they explain. Both strictly unconscious, and currently unconscious, mental states contribute to various explanations of human behavior. Together, these considerations argue in favor of an explanatory role for, and therefore an explanatory value of, various kinds of unconscious mental states.

A simple idea that supports the use of unconscious mental states is that the explanatory role of mere mental states, without further special features, such as being conscious or having features that could make them become conscious, is the norm in psychological explanation. Another simple idea that supports the use of unconscious mental states is that the explanatory role of mental states does not coincide with that of conscious mental states. If these ideas are correct, then in humans there are behaviors best explained by unconscious mental states.

Some mental states that make for, or run in parallel to, a subject's conscious states cannot become conscious mental states for that subject. Other mental states are simply not currently

coded in the way associated with conscious mental states. The information processing required for coding mental states as a conscious state is costly. This cost limits both the stage of processing at which a subject's conscious mental states occur and the number of conscious states that can occur in a subject at once. This architecture is supported by various criteria and the success of employing such a model.

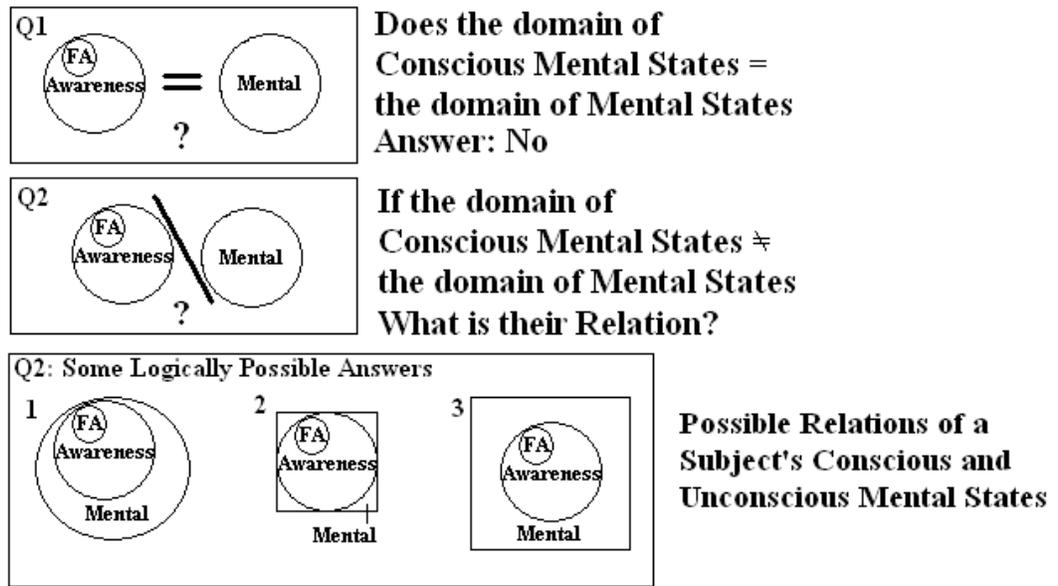
In the best current information-processing model, only a sub-set of a subject's mental states can be coded as conscious states at once and this sub-set cannot coincide with all of a subject's mental states over time. A coding architecture also shows how conscious states need not be processed within a special 'conscious-making' mental system. So, the currently successful information-processing model, and several independent criteria, give good reason, in terms of limited resources, to employ various kinds of unconscious mental states when accounting for certain human behaviors.

#### 5.5 Size, Kinds, and Unity: Varieties of Cognitive Unconscious

There are two main questions regarding unconscious mental states. The first question is whether the domain of a subject's mental states is exhausted by the domain of the subject's conscious mental states. Given the above analysis, a subject's conscious states seem to be made up of states the subject is focally attending and a set of states that the subject can justifiably judge to be, and to have been, conscious when given appropriate cues. This second set of mental states is often called a subject's awareness. The main issue has been whether this picture of a subject's conscious states coincides with the subject's mental states. There is clear reason to think that the explanatory roles of a subject's conscious and unconscious mental states do not overlap.

A second question arises if we hold that the explanatory roles of a subject's conscious and unconscious mental states do not overlap. The second question is about the relation between

these two domains of a subject's mental states. Most of the literature on unconscious mental states begins with this second question. Establishing the explanatory value of unconscious mental states leaves this second question open. These two main questions, and a variety of logically possible answers to the second, are illustrated in figure 4.



Summary of the two main questions and a few possible answers to question two.

Figure 4

Q1 in figure 4 illustrates the main question of whether the explanatory role of a subject's conscious mental states completely overlaps with the explanatory role of the subject's mental states. A subject's focally attended mental states are represented as a circle that can move anywhere within the set of states a subject can justifiably judge to be, and to have been, conscious, here labeled the subject's 'awareness'. The first parts of the present chapter address this issue and provide reason to think the answer to Q1 is 'no'.

If the answer to Q1 is no, then another question arises. Namely, what is the relation between a subject's conscious and unconscious mental states? Is the right view that a subject's conscious mental states can overlap with the domain of the subject's unconscious mental states at least at some time or other (model 1 of figure 4). This would deny the existence of strictly unconscious

mental states, or mental states that could not become conscious states of the subject. Strictly unconscious mental states are defended in section 5.4 and again in section 6.4.1.

For now, we can focus on models where unconscious mental states are of explanatory value for either only a few or for fairly many of the subject's behaviors (models 2 and 3 in figure 4). Overall, variations of model 2 suggest that only a small number of unconscious mental states are valuable when accounting for some of a subject's behaviors. Alternately, model 3 leaves a larger number of unconscious mental states explanatorily valuable when accounting for some of a subject's behaviors. Section 5.6 argues that something like model 3 is most likely correct.

Other unresolved issues include what kinds of unconscious mental states are of explanatory value. We have seen that unconscious mental states can potentially be perceptual or cognitive states which can be strictly or currently unconscious. Although generally supported, this issue has not been fully addressed in detail. A further unresolved issue is how unified unconscious mental states are within an individual subject. Generally, these last two issues address whether a subject's unconscious mental states constitute a dis-unified set of mental states of only one kind or something more unified and composed of various kinds of states.

With various numbers of, kinds of, and unifications of a subject's unconscious mental states there are quite a few pictures that are possible regarding human subjects' unconscious mental states. At one end of the spectrum are pictures that endorse human subjects' unconscious mental states as being only a small, highly dis-unified, set of states of only one kind. At the other end of the spectrum are pictures that suggest human subjects' unconscious mental states are fairly unified and numerous, as well as being composed of several kinds. With the large variety of pictures available the next few sections will cover just a few of these possibilities.

It is important not to confuse answers to the first main question (Q1) with answers to the

second main question (Q2). Endorsing the general explanatory value of unconscious mental states is not the same as endorsing any particular picture of unconscious mental states in human subjects. So, even if the view advocated for below turns out to be incorrect the explanatory value of unconscious mental states remains unthreatened.

#### 5.5.1 Merely Dis-Unified Unconscious Mental States

The first picture regarding the explanatory role of unconscious mental states in human subjects holds that these states are a small, merely dis-unified, set of mental states, likely made up of only one kind. This picture is the minimal endorsement of an explanatory role for unconscious mental states in human subjects. For example, only a few strictly unconscious perceptual states might be included in such a picture. Although interesting, this picture is merely a logical possibility. Very few, if any, actual researchers hold this view of the explanatory role of unconscious mental states in human subjects. Most researchers hold a subject's unconscious to be made up of fairly numerous, interrelated, and various kinds of mental states. So, the current, dis-unified, view is unlikely to be true.

However, establishing the explanatory value of unconscious mental states does not assume much detail regarding the kinds of states, their unity, or the number of states that make up a subject's unconscious mental states. So, on the present view, the unity, kind, and number of a subject's unconscious mental states is as minimal as possible. The simplest picture, therefore, regarding unconscious states supports that only a few such states could occur in widely separate and non-interacting places within the subject's mental system. Likewise, this view can hold that unconscious mental states are generally only low-level perceptual states, such as the states that occur in the first few stages of perceptual processing.

The only way in which this picture treats a subject's "unconscious" as a singular thing is that, explanatorily, we can say the subject's "unconscious" partly explains some form of behavior.

This kind of explanation indicates only that the states that were involved in certain information processes were mental, unconscious, and belonged to the subject. In such a picture, a disparate set of unconscious mental states is all that need constitute a subject's "unconscious" (Power and Brewin 1991; Hirst 1995). Obviously speaking about "an" unconscious in this way can be misleading. For example, this way of speaking can be taken as indicating a singular system accomplishing certain tasks. On this view, there is no such thing. Talking in this way is still explanatory, although nothing terribly unified, various, or numerous is meant in this kind of explanation.

One example of a minimal picture may be illustrated by Baars (1988). He holds that unconscious states are used only for "specialized tasks [that] have relatively limited domains, [and] are relatively isolated and autonomous" (Baars 1988, p. 117). Baars' picture can be illustrated by data showing that pupil size can influence the attractiveness of another person (Hess 1965; Zajonc 1980; Tombs and Silverman 2004). Such influence does not require various kinds of, a large number of, or any unity of, unconscious mental states. To be fair, Baars does suggest that *when taken together* a large amount of information processing in humans can be explained by unconscious mental states. However, this does not suggest a very unified, or numerous, set of states with much variety.

#### 5.5.2 The Limited Cognitive Unconscious

The above picture is hardly what most theorists mean when they say "the" cognitive unconscious has explanatory value (Kihlstrom 1987; Hassin et. al. 2005; Wilson 2002).

Alternate pictures that are generally seen in psychology and cognitive science characterize "the" cognitive unconscious as more than a dis-unified, limited, set of states made up of only one kind of mental state.

A second picture regarding unconscious mental states in human subjects can be seen in certain

current lines of research. This picture suggests more unification, various kinds, and a larger number of unconscious mental states in individual subjects. Although this view is partly correct, it is still too limited. The right picture is discussed in section 5.6.

The present view endorses some interrelation between a subject's unconscious mental states but endorses such interaction only within fully encapsulated processing. If such fully encapsulated mental processing exists, for example, in the operations within sensory modalities, the states involved in this kind of processing are often thought to be isolated from the rest of a subject's mental states. Such states are thought to be isolated both in respect to other sensory modality processing as well as higher-level cognitive processing. Such interactions need not even include higher-level states within encapsulated processing. So, on this view, interactions of unconscious mental states would be very limited.

The kinds of unconscious mental states that occur in encapsulated processing are often thought to be limited as well (Fodor 1983; Pylyshyn 2003). The amount of information processing required to code the states that occur in encapsulated processing is somewhat limited. For example, in certain priming studies, subjects were presented with stimuli that they did not judge to be conscious, but which influenced their subsequent behaviors. In these studies, only phonological parts of terms, rather than the meanings of the terms, seemed to have effect on the subject's behavior (Abrams and Greenwald 2000).

Phonological features of terms such as 'smut' and 'bile' primed a subject's behavior in regard to the term 'smile'. Clearly, in these priming conditions subject's unconscious mental states were not processed to the level of semantic meaning. This is because terms such as 'smut' and 'bile', with negative semantic meaning, primed behaviors related to terms such as 'smile', with positive semantic meaning (Abrams and Greenwald 2000). This indicates a low level of

information processing required for these unconscious mental states and implies the kinds of unconscious mental states are limited.

Some experimental paradigms have been thought to reveal that subjects acquire complex unconscious mental representations. This has been thought to occur, for example, when subjects are able to perform well regarding complex predictions for stimuli such as Reber Grammars. Reber Grammars are ways of stringing letters together, for example, which have various possibilities for what follows a particular string of letters. The rule for which letter can come next in a string is based on past letters that have been strung together. With training, subjects' behavior seems influenced by states representing the underlying rules forming the Reber Grammar (Reber 1967).

However, explanations for such behavior have been shown to potentially overestimate the level of sophistication of a subject's unconscious mental representations. For example, the mental states involved in this processing, which subjects can judge to be conscious, likely represent short strings of legal characters. This alternate kind of explanation suggests that any unconscious mental states which do seem involved in this processing have much less information than the full Reber Grammar, and so require much less processing than once thought (Cleermans, Destrebecqz, and Boyer 1998).

Support for a limited picture where low-level perceptual states of limited number and unity compose a subject's unconscious mental states can be found in sensory, as well as other potentially encapsulated, information processing. In this kind of processing there is a certain degree of causal interrelation between various low-level mental states. Although only composed of low-level perceptual states, such processing is often thought to be based on unconscious mental states (Fodor 1983; Pylyshyn 2003). This view suggests that a subject's unconscious

mental states, though simple, have a degree of unification that is beyond what is endorsed by the dis-unified view.

On the current picture, a subject's unconscious mental states are limited in number, have only a limited unification, and are made up of only limited kinds of mental states. Although still too limited, this picture does go beyond the dis-unified view (but see Carruthers 2006; Samuels 1998).

### 5.5.3 The Medium-Sized Cognitive Unconscious

Clearly, the limited picture of a subject's unconscious mental states is distinct from the dis-unified picture that holds only a small set of dis-unified unconscious mental states have an explanatory role for certain human behaviors. However, the limited picture is also not all that is generally meant by theorists who endorse the explanatory value of "the" cognitive unconscious. Limited and dis-unified pictures capture only a part of what many theorists endorse when they hold that the cognitive unconscious is of explanatory value.

The limited view discussed above is based largely around the notion of unconscious mental states occurring only in low-levels of encapsulated processing. The notion of encapsulated processing of mental states, however, has had problems from its inception (Marshall 1984; Coltheart 1999). The original formulation of modular processing, for example, required such processing to have nine different features (Fodor 1983). One of these features was full isolation of the mental states from both higher-level cognitive processing and other low-level perceptual processing. So, on the limited view, for example, interactions of unconscious mental states would not include higher-level cognitive states or interrelations with mental states in other low-level information processing.

The theory and data that supported fully encapsulated processing came under immediate attack from neuroscience and psychology. Although it seemed that many mental systems were

isolatable, for example, in lesion studies, such systems were shown to interact in normal processing (Milner and Goodale 1995; Churchland, Ramachandran, Sejnowski 1994). In this way, a less encapsulated view of many mental systems became popular (Shallice 1984).

The current, medium-sized, picture of unconscious mental states holds that a fair number, variety, and level of unified unconscious mental states beyond somewhat encapsulated processing have explanatory value. The medium-sized view of a subject's unconscious endorses interactions of unconscious mental states between various mental systems. This view, therefore, characterizes a subject's unconscious mental states as more interrelated with a larger number of states than previous pictures. Although in part correct, this current picture is also too limited as is discussed in section 5.6.

Certain kinds of visual illusions illustrate the creation of fairly complex states in somewhat encapsulated processing. In the Ebbinghaus Illusion, for example, a circle is surrounded by smaller, or larger, circles. In each case, the central circle appears, respectively, larger or smaller than its actual size. The processing that gives rise to such illusions must relate mental states for visual shape and visual context. Visual context in this illusion requires higher-level visual processing than is required for creating states about visual shape. Visual context in the Ebbinghaus Illusion is composed of, at least, several visual shapes. So, interactions between higher-, and lower-, level visual states in this illusion suggest more interrelation between unconscious mental states than is endorsed in the limited view.

Although involving fairly complex states, perceptual illusions are isolated from cognitive processing. For example, certain illusory percepts are impervious to a subject's cognitive states, such as the subject's knowledge that the percepts are illusions, or the subject's desire to see, or hear, the stimuli in any other way (Rock 1983). Somewhat encapsulated processing seems to

involve interaction between higher-, and lower-, level perceptual states within various perceptual domains (Pylyshyn 2003).

As an illustration, Pylyshyn makes a clear case for interactions of high-, and low-, level mental states within visual processing. These kinds of interactions form important constraints on how visual states are processed. For example, higher-, and lower-, level visual mental states that represent various kinds of visual features interact to help form stable visual interpretations of the incoming signal. Visual processing is not the only example of where such interactions occur. The states in such interactions often do not cause subjects to create phenomenal judgments. So, interaction between higher-, and lower-, level unconscious mental states, that occur in somewhat encapsulated processing, goes beyond the limited view.

Likewise, color processing is an example of interaction between mental states within various parts of visual processing. Color processing is a fairly independent mental system since color computation can be damaged in isolation from, say, shape processing. However, the computation of color is based on the computation of several other non-color properties as well (Hoffman 1998, p. 113). For example, the computation of the color of surfaces and the color of light sources, as well as depth and shading of shapes, all mutually constrain color processing (Hoffman 1998, p. 114). In this way, color processing shows a certain amount of interaction among various kinds of mental states within the visual domain that also goes beyond what the limited picture endorses.

Further examples of mental state interactions also clearly go beyond the limited view. For example, mental-state interaction can be found between larger mental systems. This kind of interaction is well beyond what the limited picture endorses. In processing occurring in what are termed the “what” and “where” (or “how”) pathways, there is a certain amount of separation of

the visual information (Milner and Goodale 1995). However, in certain cases mental states in these systems influence one another.

In a striking illustration, pointing is faster for certain visual illusions than for others. In the Ebbinghaus Illusion, when a circle is surrounded by smaller circles the central circle appears larger than its actual size. Pointing to this circle is faster than when the same sized circle is surrounded by larger circles, making the central circle appear smaller, at least, to certain information-processing systems within humans (Lee and van Donkelaar 2002). Pointing is generally attributed to the “where” pathway, whereas perception of size is attributed to the “what” pathway. An influence on pointing behavior due to the apparent size of the central circle suggests an interaction between the mental states within the “what” and “where” systems (Milner and Goodale 1995, p. 202; Churchland, Ramachandran, Sejnowski 1994). Such data provide evidence for more interaction between a subject’s unconscious mental states than is endorsed by the limited view.

Interactions between states in two different sensory systems, such as vision and audition, have also been shown. One example is the McGurk Effect. In this effect, auditory and visual percepts seem to interact so that the mental states that subjects judge to be conscious are combinations of both the auditory and visual information. In certain cases, subjects are presented with an auditory signal, such as a voice saying “*ba*”. Simultaneously, the subjects are also presented with a visual signal, such as lips moving to the word-form “*ga*”. What subjects report hearing is a sound that is influenced by both the auditory and visual signals. In these scenarios subjects often report hearing the sound “*da*” (McGurk & McDonald 1976). This result indicates interaction between mental states from different modalities. Such interaction is beyond what is endorsed by the limited picture.

Although not involving interactions with higher-level cognitive mental states, the kinds of interactions just described argue for a degree of unification between mental states beyond that endorsed in the limited picture. The kinds of mental states involved in such interactions are likely unconscious based on subject's reports and the level of processing at which the states occur. Subjects seem unable to directly justify judgments that the states in these kinds of interrelations are conscious.

Interactions across mental systems do not necessarily suggest that the states involved constitute a further mental system (Lyons 2001). What the above kinds of interrelations do suggest is a degree of unification beyond that which occurs in somewhat encapsulated mental systems. This kind of unification is therefore beyond the dis-unified or limited views of a subject's unconscious mental states. Furthermore, these interactions are one of the most common kinds of unification stressed by theorists who endorse the explanatory value of a subject's cognitive unconscious (Kihlstrom 1987). Likewise, the interactions involve a variety of kinds of unconscious mental states beyond that endorsed by the dis-unified, or the limited, pictures.

The medium-sized picture of a subject's cognitive unconscious endorses unconscious mental states created in later stages of perceptual information-processing. For example, percepts guiding pointing, or full percept interaction between sense modalities, as in the McGurk effect, suggest a fair degree of processing before the states interact. Given this fact the current picture not only endorses more unification, it also endorses a greater number of, and variety of, unconscious mental states than the dis-unified, or the limited, pictures addressed above.

#### 5.5.4 The Vast Cognitive Unconscious

All of the previous pictures of a subject's unconscious mental states are careful to separate off the subject's cognitive states from being included as part of the subject's unconscious mental

states. However, quite a few theorists endorse the idea that both cognitive as well as perceptual mental states are part of a subject's unconscious mental states. So, a vast picture of a subject's unconscious mental states endorses a greater number and variety of unconscious mental states than the pictures considered above.

In the vast picture, the unification of unconscious states in a subject does not likely go much beyond that of intersystem interaction seen in the medium-sized view. So, although the vast picture of a subject's unconscious mental states includes a greater number and kind of unconscious mental states, it endorses about the same level of unification as the medium-sized picture.

Theorists such as Piattelli-Palmarini and Damasio, among others, suggest that a subject's decision making and problem solving are influenced by unconscious mental states that are cognitive (Piattelli-Palmarini 1994; Damasio 1996; Payne et. al. 2008; Dijksterhuis and Norgren 2006). Likewise, Timothy Wilson, in his book *Strangers to Ourselves*, as well as many other social psychologists, defends the idea that a subject's social cognition, as well as other kinds of higher-level decision making processes, is influenced by unconscious mental states that are cognitive. Wilson suggests that subject's cognitive systems are "regulating a good deal of high-level, sophisticated thinking to the unconscious" (Wilson 2002, p. 7; Bargh 2008, p. 73). Fodor and Pylyshyn both think that the mental states involved when subjects solve the Frame Problem, or those involved in abductive reasoning, are unconscious mental states that are cognitive (Fodor 1983; 2000; Pylyshyn 1987). Poincare's discovery of Fuchsian functions, and other sudden cognitive level insights, are classically given as examples of unconscious mental states that are cognitive (Smith 1999). Finally, Hofstadter and Boden think that even a large part of a subject's general creativity, as well as their analogical thinking, is mediated by unconscious mental states

that are, at least sometimes, cognitive (Hofstadter 1995; Boden 1994; Chalmers, French, and Hofstadter 1995). A subject's unconscious mental states, on these views, include both low-level perceptual states and higher-level cognitive states.

The current picture of a subject's unconscious mental states as vast in number, kind, and with a fair amount of unification is becoming more popular (Lewicki, Hill, and Czyzewsky 1992). For example, race cognition is being studied in which the explanatory value of unconscious mental states, ranging from perceptual to cognitive mental states, is being revealed. Subjects who report no racial bias still behave in ways that are sometimes strongly influenced by race. For example, subjects will take longer to press a button that associates a picture of a person of color with a positive word (Ito and Cacioppo 2000). The mental states involved here would be, in part, high-level interpretations of stimuli, such as those required for cognitive states, and so would involve more kinds of, and more interrelation between, mental states than other views endorse.

The current view of a subject's unconscious mental states does not endorse that such states form a single mental system (Lyons 2001). However, on the current view, interactions of unconscious mental states include higher-level cognitive states as well as various perceptual states. This is not a part of any of the previous views of a human subject's unconscious mental states. Although cognition and perception may be quite dis-unified mental processes, some mental state interaction between higher-, and lower-, level mental systems will buy a further degree of unity, number of, and kinds of unconscious mental states than any of the previous pictures endorse. This view encompasses much that is correct, as is shown in the next section.

#### 5.6 What is the Right Picture?

Roughly speaking, the vast unconscious seems to be the right picture. The right picture appears to be that unconscious mental states, of both perceptual and cognitive kinds, are many in

number and are fairly unified. Although not yet fully addressed, the right picture also seems to be that both currently and strictly unconscious mental states have explanatory roles. Various kinds, large numbers, and a fair amount of unity between unconscious mental states make up the correct picture regarding a subject's cognitive unconscious.

As noted, the right picture likely includes various *kinds* of unconscious mental states. Unconscious states in both perceptual and cognitive domains of processing have an explanatory role. Furthermore, explanatory value is gained by having some of a subject's perceptual and cognitive states occur as either currently unconscious states or states that cannot become conscious for that subject. Each of these kinds of unconscious states plays a different role in various explanations of human behavior and each kind of unconscious state helps researchers account for certain behaviors. Various kinds of unconscious mental states increase the explanatory value of the cognitive unconscious overall. Given the coding, or information-processing, architecture that is most likely correct, each of these kinds of unconscious mental states will retain some valuable explanatory role.

We have seen reason for treating certain of a subject's mental states as currently unconscious. For example, currently unconscious states are employed in explanations involving behavior driven by mnemonic mental states. Cognitive illusions are another illustration of a subject's currently unconscious mental states. Subjects can often become aware of the states that drive certain cognitive biases, for example (Piattelli-Palmarini 1994).

Likewise, there is reason to believe certain of a subject's mental states are sometimes strictly unconscious mental states of the subject as well. There is an architectural divide between the mental states *that make for* a subject's conscious states and the subject's conscious states. The states that make for a subject's conscious states are not in a position to ever occur as conscious

states, at least given the picture of the information-processing architecture advocated here (see also Jopling 1996, p. 158; Gallagher 2005). Some of a subject's unconscious mental states also seem to run parallel to, and do not just underlie, conscious mental states (Milner and Goodale 1995; Place 2000). These states are often considered strictly unconscious as well.

However, flexible pictures of the relation between a subject's conscious and unconscious states may argue against the explanatory value of a large number of strictly unconscious mental states in individual human subjects. Marcel's data, as we have seen, show that within single subjects and on single trials various ways of probing subjects about their conscious states can disagree (Marcel 1993). When instructed to press a button and to utter a verbal response when a stimulus is represented consciously, subjects will press the button and then say they were not conscious of the stimuli.

On flexible pictures, then, it seems no particular set of a subject's mental states are consistently counted as conscious or unconscious mental states independent of probes that determine if the states are conscious. Various experimental probes, for example, can allow very low-level mental content to gain the features we attribute to conscious states. Likewise, without the right probe, very high-level mental content can, at times, be considered unconscious content (Dennett and Akins 2008, p. 4325; Dennett 1991, p. 308; Tononi and Edelman 1998; Kinsbourne 1988; O'Brien and Opie 1998). Strikingly, flexible pictures of a subject's unconscious mental states suggest, for example, that high-level cognitive mental states can be considered unconscious. Perhaps more strikingly, on flexible pictures very low-level perceptual mental states might be considered conscious.

Flexible pictures do not argue against a vast number of unconscious mental states. Nor do they suggest that there is any kind of dependence between conscious states and unconscious

mental states. However, they do seem to argue against a large number of strictly unconscious mental states.

Whatever the likelihood of flexible pictures is, they are still compatible with the idea that some strictly unconscious mental states are explanatorily valuable. For example, many theorists who endorse flexible pictures hold that strictly unconscious mental states have an explanatory role in accounts of the lowest levels of encapsulated perceptual processing, such as sensory modular processing. So, a class of strictly unconscious mental states is still possible, and unlikely to vary within a single subject, even on very flexible pictures of a subject's cognitive unconscious.

Overall, then, various kinds of unconscious mental states retain their explanatory roles and so have explanatory value. Both perceptual and cognitive mental states, either as currently or strictly unconscious states, help researchers account for various behaviors in human subjects.

The right picture of the human cognitive unconscious also seems to include a large *number* of unconscious mental states. Certainly most pictures do not suggest a subject's unconscious mental states amount to just a handful. The same is true for a subject's conscious mental states (Lamme 2006; Sperling 1960; see also section 4.2.1.2 above). Further, the right picture does not suggest that a subject's conscious and unconscious states are evenly divided. The ratio of unconscious to conscious mental states in a subject is likely close to 70/30 or even 80/20 in favor of unconscious mental states, given the number of behaviors explained by either conscious or unconscious mental states. Pictures that endorse an exactly equal division between conscious and unconscious mental states would be hard to defend. At the very least, a partially skewed ratio of states is more likely than pictures that endorse only a few conscious or unconscious states or pictures that endorse an equal division of these states.

Many stages of information processing in human subjects seem to involve unconscious mental states. For example, explanations ranging from low-level perceptual processing to those involved in high-level processes, such as abduction, seem to involve unconscious mental states. This suggests that some states from low-level perceptual processing to high-level cognitive processing are unconscious mental states of one kind or another. If these considerations are correct they suggest that a large number of a subject's mental states are unconscious.

Why divide a subject's current and strictly unconscious mental states in the way depicted in figure 3? In figure 3 a subject's perceptual and cognitive unconscious mental states are unevenly divided. Figure 3 suggests that a subject's strictly unconscious perceptual states likely outnumber the subject's strictly unconscious cognitive states. Figure 3 also suggests that these numbers are reversed for a subject's currently unconscious cognitive states compared to the subject's currently unconscious perceptual states.

It will take more work than there is room for to show that such a picture is actually correct. However, briefly addressing these numbers is important. It has been argued that the coding processes required for making conscious states are composed of a variety of late-stage information-processing operations. The idea is that cognitive states also occur fairly late in human information processing.

If this is true, then it seems likely that late stage mental states, like cognitive states, will have more opportunity to be coded as conscious states. Early information-processing states, such as certain perceptual states, may then have less opportunity to be coded as conscious states. Encapsulated stages of information-processing in human subjects would make mental states in early perceptual processing more likely to be strictly unconscious states of the subject. So, a subject's strictly unconscious mental states are plausibly composed of more perceptual states

than cognitive states.

Alternately, a subject's currently unconscious mental states are likely composed of more cognitive states than perceptual states. Mnemonic states make up a large portion of the cognitive states of a particular subject. All or most of a subject's autobiographical mnemonic states are best treated as mental states that can become conscious if the subject is given the right cue. The same is true of mental states involved in creating unconscious cognitive biases. From these examples, it seems likely that a subject's currently unconscious mental states are composed of more cognitive states than perceptual states.

Whatever the actual numbers of a subject's currently and strictly unconscious mental states turns out to be there is good reason to think there a large number of unconscious mental states that have an explanatory role to play in accounts of human subjects' behavior. Importantly, a subject's unconscious mental states likely outnumber their conscious states by a fair margin, given the number of behavioral accounts that unconscious mental states currently figure in.

Finally, the right picture suggests there is a fair amount of *unity*, at least at mid-levels of processing, among unconscious mental states. The kinds of mental states that seem to be unconscious have a fair amount of interrelation both within single mental systems as well as between mental systems. Inter- and intra encapsulated processing interrelations suggest that unconscious mental states do not just occur in low-level encapsulated processes.

The amount of interrelation likely does not constitute a further mental system. So, using the term "the cognitive unconscious" must be done carefully. Although not a single system, certain characterizations of a subject's unconscious mental states do seem to underestimate the unification of these states. For example, Wilson describes human subjects' unconscious mental states as composing a mere "collection of modules" (Wilson 2002, p. 7). This characterization

does not capture the amount of interaction that occurs between a subject's unconscious mental states.

The above considerations clearly show there is an explanatory role for a subject's unconscious mental states to play beyond the dis-unified, limited, or medium-sized views of a subject's unconscious mental states. The above considerations suggest a picture with a large number of various kinds of unconscious mental states with a fair amount of unity as the best explanatory model for many human subjects' behaviors.

### 5.7 Summary and Conclusions

The particular picture of unconscious mental-state explanation just advocated may or may not be correct. Certainly lots of considerations point in the direction of a vast picture of a subject's unconscious mental states. Other characterizations of human subjects' cognitive unconscious, currently popular in the literature, also endorse the vast picture (Kihlstrom 1987; Hassin et. al. 2005).

However, the most important result from the preceding discussion is that many psychological and neurological explanations incorporate unconscious mental states. This is true regardless of whether such states play a very unified, or numerous, role in explanations of human subjects' behaviors or whether such states are made up of a variety of kinds.

Successful models that employ unconscious mental states are standard explanations of many human behaviors in psychology and neuroscience. Such success gives a burden of proof to any theorists who try to deny the explanatory value of unconscious mental states. Simply put this just means that mental-state explanations are valuable even when the mental states in question do not have special features that make them conscious mental states, or even make them states that can become conscious mental states of the subject who has them. For this reason, explanations positing unconscious mental states should be retained in psychological theorizing.

Outside of psychological theory there is a potentially important moral lesson regarding unconscious mental states as well. Knowing about, and therefore being able to partly compensate for, behavior driven by unconscious mental states may be a way to have a more healthy “mental hygiene” (Piattelli-Palmarini 1994, p. 4). This is true, for example, in individual or national decision making. In this way, unconscious mental states may be important explanatory constructs for non-psychologists as well.

## Chapter 6: Objections and Replies

There are several kinds of objections to the explanatory value of unconscious mental states and even to the possibility that unconscious mental states have explanatory value. This chapter addresses the arguments for each of these objections and gives replies to each objection. The replies show that some human behaviors are still best explained by employing unconscious mental states. The explanatory value of various kinds of unconscious mental states is shown to withstand these objections.

Four main objections will be considered in the following sections. The first kind of objection denies that certain explanatory states are still *mental* states while the state is not conscious. A second kind of objection denies that unconscious mental states have mental status *independent* of links to conscious mental states. A third kind of objection denies that mental states are ever *unconscious* mental states. A final kind of objection grants the independent explanatory value and mental status of unconscious mental states but denies *strictly* unconscious mental states. This last objection holds that all of a subject's mental states have a potential to become conscious mental states, at least at some time or other, for that subject.

The sections below show that current arguments for each of these objections fail. The replies to these objections will reveal the fairly standard picture where the explanatory value of unconscious mental states is maintained. This will be true both for mental states that *can* become conscious mental states as well as for strictly unconscious mental states that cannot become conscious.

### 6.1 Objection #1: Deflation

The first objection to unconscious mental states holds that the explanatory states in empirical psychological accounts, when not conscious states, are no longer mental states at all (James 1890; Descartes 1641; see also Searle 1992; Bennett and Hacker 2003, p. 269). This objection

holds that the conscious mental states of a subject's cognitive system are the only mental states of that system. Other explanatory states are merely physical states, or information-bearing states. This is sometimes called "The Brain-Track Argument" because it attempts to reduce unconscious mental states to the function of the brain, or mere information-bearing states and processes in the brain (James 1890, p. 166). The argument for this position is just that non-mental states explain all that is needed in psychology or neuroscience, at least when conscious mental states are not involved.

As an illustration, take the folk example of a forgotten desire. Desires are often explained as conscious mental states that sometimes help motivate a subject's behavior. When desires seem to return, for example, when a subject knows about, or thinks about, a desire for a second time, the explanation is often done in terms of a second conscious mental state with no intervening mental state. This is a deflationist explanation in that one conscious episode of desire is posited to explain behavior and then another. Deflationists deny that the states connecting two conscious episodes are ever mental states (Searle 1994, p. 854).

This view explains forgotten and then remembered desires in the following way. At the time of the first conscious occurrence of desire in a subject there is a mental state. When this first conscious episode of desiring ends so does the mental status of any states related to explaining this aspect of the subject's behavior. Other conscious states may explain other behaviors of the subject, but there is no longer a mental state related to this desire. In the interim between a subject's conscious events, where two conscious states of desire have the same content, there are physical and mere information-bearing states in the subject. On the second occurrence of a conscious state of desire with the same content the second conscious desire in the subject is not a token-identical desire to the first.

So, on this account, the desire does not return in any sense other than *another* conscious state directed toward the same thing occurs as a conscious state in the subject. That is, a different *type-identical* content occurs in the subject as a *second* conscious experience. On this account, continuity of any unconscious mental states in the subject is not required to mediate between these two occurrences of conscious experiences of desire in the subject.

Brain states, or mere information-bearing states, alone are not thought to be either conscious or unconscious mental states. They are simply thought to be non-conscious. This is the way that the states internal to trees, for example, are considered non-conscious but not *unconscious* (Bennett and Hacker 2003, p. 270). If deflationary accounts such as the above are correct, then the states connecting certain events described as a subject's two conscious desires need not be mental. Likewise, explanations for the processing leading up to, or surrounding, single conscious events in a subject would not require mental states, on this view, as well. We can see from this how the current objection would deny the need for any unconscious mental states.

#### 6.1.1 First Reply to Deflation

There are two replies to the above deflationist objection. The first reply argues that unconscious mental states sometimes link two conscious events in a subject. The second reply argues that unconscious mental states are required in the processing leading up to, and surrounding, a subject's single conscious events.

The first reply to deflationists suggests that unconscious mental states are often valuable explanations for states linking a subject's conscious mental episodes (Wakefield 2001, p. 346). If the same token mental state is explanatory during, and then between, a subject's conscious episodes then, based on subjects' phenomenal judgments during the interim between conscious events, these states would be unconscious mental states. Subjects do not justifiably judge states intervening between conscious episodes to be conscious or to have been conscious even when

they are given various kinds of cues.

During the interim between conscious events the states mediating the subject's behavior could be recorded. The linking states are likely to remain coded as mental states during this time but not as conscious mental states. If this is true, the states linking certain conscious episodes in a subject cannot be deflated to *mere* physical, or information-bearing, states since they are still coded as mental states. Likewise, they cannot be considered conscious states based on subjects' justified phenomenal judgments. This means the states in question will be unconscious mental states of the subject.

It is generally agreed in the debate over unconscious states that *something* must link certain kinds of behavior explained by two conscious states. For example, when a subject has a memory of an earlier perception, where both the memory currently occurs as a conscious state and the perception had occurred as a conscious state, a state that links the conscious memory with the past conscious perception is required (Searle 1994; Alston 1971; Smith 1999). Without such linking states the memory could not carry information about the earlier perception.

To begin the illustration of this idea, think of an example of the same type of content occurring as conscious content in a single subject at different times. In one kind of case, subjects have a particular desire for ice cream, and one year later they have "the same desire" return as a conscious state. During the second conscious occurrence the explanatory state most likely only has the same content as the first. The linking states involved need not have retained the same content, or been the same token mental state, for this explanation to work.

Alternately, imagine a person walking along a line of shops who repeatedly averts their attention from a conscious desire for ice cream. In this case, the same conscious desire can be said to return to focal attention, all the while remaining a conscious mental state in some non-

focal way. No unconscious mental state is needed to explain this case. In this second case the causal process that underlies the different focal episodes of attention can be the same conscious mental state. When this is the case we can say that the focal conscious episodes are of “the same mental state”, or that they are the same desire, in a stronger sense than merely that they have the same content (Wakefield 2001, p. 346). The causal state linking focal attention episodes is likely the same token mental state as well. This state is just not focally attended during certain times.

Finally, cases of complete forgetting, for short periods, are not like forgetting for longer periods or refocusing attention. When subjects completely forget for short periods, for example during a conversation while walking along a line of shops that reminds them of their desire for ice cream, a single token mental state can explain both the conscious episodes of desire and the states that link such episodes. The argument for this claim is provided just below. In such cases the mediating state will be mental but not a conscious mental state.

To see this better, think of various desire-generating processes that can continue to generate type-identical mental states in a single person’s mind. Sometimes a desire-generating causal process is external. For example, a person walks down the line of shops where various shapes that look like ice cream cones cause various type-identical motivational states related to ice cream. Other times the desire-generating causal process can be internal but not based on a continuous token-identical mental content. For example, intermittent epileptic seizures could cause someone to have recurrent conscious desires for ice cream. In neither of these cases is the explanation helped by positing a continuous token-identical unconscious mental state that links the subject’s conscious episodes.

However, sometimes the desire-generating process continuously causes an internal state that

is helpfully described as the same token-identical mental state. In these cases, explanation is helped by positing a single internal state about ice cream that makes a subject have several conscious episodes featuring that desire for ice cream. In some cases, the desire-generating process for various conscious episodes is an internal mental state that is continuously generated and causes several conscious events.

When this state has features best described as mental content the same desire can be said to return as a conscious states in a stronger sense than merely that two independent conscious episodes of the same content occurred. In such cases we can explain the subject having the same token desire return as a conscious event. So, it seems some conscious episodes can be explained as being linked by unconscious mental states (Smith 1999; Fodor 1994, p. 9; see also Mandler 1994; Kvavilashvili and Mandler 2004; Brown 1991; however, see Wheeldon 1954 on the return of headaches).

If a state's continuity as a mental content is sometimes explanatorily valuable then it shows that the states in certain psychological or neurological explanations cannot be reduced to *merely* physical, or mere information-bearing, states with *no* mental status. If the semantic continuity of token-identical mental contents is helpful, then, at least for these explanations, unconscious mental states will have an explanatory role (Smith 1999, p. 66).

#### 6.1.2 Second Reply to Deflation

The second reply to deflationists argues that unconscious mental states have an explanatory role in the processing leading up to, and surrounding, single conscious events. If the processing that leads to, or surrounds, single conscious states are valuably described as employing mental states then it seems that these, *prima facie pre-conscious* or *post-conscious*, mental states will be valuably described as *unconscious* mental states.

To see this reply better, think of a commonly accepted explanation for conscious vision. The

internal processing that allows a system to become visually aware is often thought to occur in several stages (see chapter two). Later stages build on processing that occurs in earlier stages. Some of the states involved in these stages of processing are thought to be valuably described as mental states.

For example, some theorists posit stages of processing where two possible visual interpretations interact. These interactions are thought to be explained by unconscious mental states. As an illustration, bi-stable conscious visual perception, as in the case of the Necker Cube, seems best explained by two competing percepts about the orientation of the cube (Palmer 1999; O'Reilly and Munakata 2000). Necker Cubes are often simple line drawings of all the outside edges of a cube. Such drawings can appear to be oriented in two directions. The percepts for either interpretation of orientation seem to compete for conscious status because subjects report the image switching between orientations and judge each switch in orientation to be a conscious state. However, subjects rarely report both interpretations occurring as conscious states at once.

This suggests that at various stages of visual processing, operations sensitive to the two different interpretations seem to employ mental states. For example, whole interpretations of the cube's orientation seem to compete without subjects judging both of these interpretations to be conscious states at once. In this sense, the best explanation for these stages of visual processing seems to involve mental states leading up to the single conscious visual perceptions that switch back and forth. This explanation would involve unconscious mental states.

Unconscious mental states are valuable parts of explanations for cognitive phenomena outside of vision as well. Mental states are explanatorily valuable in stages of processing illustrated, for example, in Craik and Lockhart's model of memory (1972). This model suggests that the

strength of a memory trace is the result of the type of processing done at encoding. So, for example, if a subject is presented with a word and asked merely to count the number of “e’s” in it, the subject does not need to sound out the word or even discern its meaning. Orthographic processing such as counting “e’s” is not a very “deep level” of processing, according to this model.

Stronger mnemonic traces can be established if subjects are asked to determine whether the word rhymes with another. Still stronger memory traces occur if subjects judge the word meaning. In this last case, subjects have to process the word’s orthography, the word’s sound, as well as its meaning. These extra stages of processing allow the items to be remembered more readily because they have passed through more processing. States representing the number of “e’s”, for example, would be unconscious when subjects process meaning or the sound of the word.

The states in each stage of the above mnemonic processing are about features of each word. So, representational states about the various features of the stimuli are thought to help explain processing leading up to certain conscious mnemonic experiences. For example, when subjects determine meaning they must process the number of “e’s” but often do not judge the states involved in such processing to be conscious.

Memory and vision are not the only cases where explanation of human behavior seems helped by employing mental-state explanation in processing prior to a subject’s conscious experience (Hirst 1995; Gallagher 2005). For example, pre-attentive processing, where certain sensory items stand out among others, are also thought to illustrate mental state processing leading up to conscious experiences (Pashler 1998; Treisman 1985).

Furthermore, processing *surrounding* certain conscious experiences of a subject is valuably

explained by unconscious mental states. This is illustrated in the case of certain linguistic judgments. As Jackendoff notes, “language continually involves unconscious [states], both prior to and during conscious phases” (1992, p. 90). Subjects form judgments about legal grammatical constructs while hearing complex sentences. Many of the states leading up to and surrounding such judgments are valuably described as unconscious mental states.

For example, when subjects attempt to make sense of, and parse, sentences such as “John appealed to each of the boys to like the other” there seem to be states involved that point to various parts of the sentence (Jackendoff 1992, p. 84). These states are mental states because they have focused information about and can misrepresent various parts of the linguistic structure. The states that surround and influence conscious judgments regarding such sentences are not judged by the subject parsing the sentence to be conscious. So, these states seem to be explanatorily valuable uses of unconscious mental states.

Other linguistic explanations illustrate the value of unconscious mental states leading up to and surrounding a subject’s conscious experience as well. For example, when subjects parse garden path sentences, quick re-interpretations are required. These re-interpretations seem to involve unconscious mental states. The states that influence new conscious judgments of what the sentence means are not reported by subjects to be, or to have been, conscious. Likewise, these states are likely to remain coded as mental states but not as conscious mental states. This seems to indicate that the states in question are unconscious mental states.

Sentences such as “The horse raced past the barn fell” are illustrative examples. Linguistic processing here seems governed largely by states both leading up to and surrounding subjects’ single conscious experiences. The states leading to a subject’s conscious judgment in these cases seem best described as unconscious mental states. When subjects process garden-path sentences

subjects' linguistic processing is described as forming an interpretation of one word as the main verb. After this interpretation, subjects are pressed to give conscious reinterpretations of this unconscious linguistic processing.

These kinds of processes seem to involve unconscious mental states because subjects rarely judge the states involved, which make it difficult for them to reinterpret the sentence, to be conscious. When subjects do report such states they do not report that the mental states involved were conscious. Likewise, such states likely retain their coding as mental states. In this way, unconscious mental states seem to valuably explain the processing surrounding certain single conscious experiences.

From the above we can see that unconscious mental states are explanatorily valuable for several kinds of empirical psychological processes leading up to, and surrounding, single conscious experiences of human subjects. The failure of *merely* neurophysiological, or non-mental information processing, explanations for these phenomena suggests that mental-state explanations are correct. The above considerations mean that deflationary accounts are wrong in regard to both explanations employing mental states that link certain conscious episodes of a subject as well as explanations employing mental states that lead up to, and surround, a subject's single conscious episodes.

#### 6.2 Objection #2: Dependence

The second objection to unconscious states grants the mental status of such states. However, this objection holds that the mental status of such states is dependent on the possibility of a subject's unconscious mental states to become the subject's conscious states (Searle 1992, p. 155). For theorists who hold this view, the mental status of a subject's psychological states, while unconscious, is not original to such states but derived from the potential of such states to become the subject's conscious states with real mental status (Searle 1992; Ludwig 1996;

Strawson 1994; but see Flanagan 1995).

We saw the distinction between real and derived mental content in chapter three when discussing the content of states in computer programs (see section 3.1.1). There we saw that without a natural, and non-mental, process to focus the information in a state there is no way to tell what a computer state is about without interpretation from an actual mental system, such as the computer programmer. Computer states could mean quite different things without an interpreter.

The current objection holds that unconscious mental states are like the syntactic states of a computer. A subject's unconscious mental states, on this view, do not have real content without the possibility of becoming conscious states. Conscious mental states are held to be interpreted states that have mental content.

Dependence theorists link mental properties to conscious states. Theorists often call these positions "Connection Principles" (CPs). In general, theorists who hold CPs posit that a state must either *be* a conscious state or possess the *possibility of becoming* a conscious state to possess mental properties. Dependence theses rely on the commonly recognized distinction between original and derived content (but see Dennett 1987). This distinction is used to reject a realist metaphysical reading of the mental status of states that are not currently conscious states (Searle 1992; Ludwig 1996). According to this objection, the non-conscious states that explain psychological behaviors may act "as if" they had mental content but do not possess real content while not conscious states (Searle 1992, p. 155; Ludwig 1996; Archard 1984, p. 125; Horgan and Tienson 2002; Strawson 1994). This means the explanatory value of unconscious mental states is dependent on links to conscious states and so objects to the independence of content in unconscious mental states.

Searle lays out the general position regarding dependence like this, “to the extent that [unconscious states] are genuinely mental, they must in some sense preserve their [mental status] even when unconscious, but the only sense that we can give to the notion that they preserve their [mental status] when unconscious is that they are possible contents of consciousness” (1992, pg 159-160).

CPs formulated in this general or abstract way are ambiguous. Fodor and Lepore (1994), for example, wonder what kind of “possibility” is at stake in Searle’s position. The claim that a state must have the possibility of becoming conscious can mean several things (Wakefield 2001, p. 848). For example, it is unclear whether conceptual possibility or what Fodor and Lepore call ‘metaphysical possibility’ is at stake (Fodor and Lepore 1994, p. 838). One question is whether the possibility at stake concerns holding the actual laws of nature and the conditions of the human organism fixed.

Searle responds that his position is based on causal possibility (1994, p. 850). However, a nontrivial *positive* formulation of a CP seems difficult. Searle’s position appears to be that he just cannot think of anything *other than* conscious status that would make the information in psychologically explanatory states determinate enough to be considered mental (Searle 1992, p. 155). However, Searle says little about just how conscious status makes the information in a state determinate enough for mental status (Fodor and Lepore 1994). Furthermore, Searle’s lack of positive formulation is common in the literature (Horgan and Tienson 2002; Strawson 1994; Ludwig 1996). Therefore, Fodor and Lepore say that, at least Searle’s, “CP itself simply isn’t clear enough to be worth the bother of arguing about” (1994, p. 845).

Fodor and Lepore may be right. As Wakefield (2001, p. 848) notes, Searle’s notion of becoming “in principle” conscious may be so weak as to allow for a total philosophical zombie,

a creature functionally like us but somehow without conscious states, to have mental states. For example, as long as the philosophical zombie's mental states could become conscious states of *some other creature* a general CP could be fulfilled! This effectively grants the value of unconscious mental states for a substantial number of human behaviors, maybe even for all of a subject's cognitively and perceptually driven behaviors, and renders the explanatory value of CPs dubious at best.

We can best think of CPs as requiring concrete links between a single subject's mental states and that subject's conscious mental states. This avoids the ambiguity of CPs formulated abstractly. Since there are several attempts to hold CPs in the literature it seems worth examining a few representatives of this type of objection, and the replies to them, before moving on (Strawson 1994; Malpas 1992; Horgan and Tienson 2002; Searle 1992; Ludwig 1996).

#### 6.2.1 Searle's Version of the Dependence Objection

Searle gives one kind of argument for a CP (1992; see also Strawson 1996; Malpas 1992).

His position is that the mental states explaining human behavior must be conscious states or be states with the potential to become conscious mental states of the subject who has them. This is because Searle believes real mental status is dependent on the fine-grained contents he thinks can only exist as conscious states. For Searle, this establishes dependence between mental states and conscious mental states. His argument goes something like this:

- 1) We need determinate content in mental-state explanations.
- 2) Some conscious content has the requisite determinate content.
- 3) Conscious content is the only way to have the needed determinate content.
- 4) So, mental states with determinate content are either conscious states or need a link to potentially becoming conscious states to have this determinate level of content.
- 5) Therefore, no strictly unconscious state, that is, a state without potential to become a conscious state, has determinate content. Any unconscious mental states that are explanatorily valuable, then, are so because of a link to becoming conscious states of the subject who has them.

To see Searle's Connection Principle better we can take his example of a subject looking at a

glass of water. Subjects can direct their attention to water or they can direct their attention to H<sub>2</sub>O while the causal link from the world remains the same. On Searle's view, for a person to do this the causal relations between the subject's conscious states and the world do not have to change. Searle holds that such a change in mental directedness is entirely dependent on a change in conscious experience.

To see why this might be the case, think of Searle's position on real mental content. For Searle, there is nothing in the head of a dreamless sleeping person, say, that individuates the particular way a state is directed to the world. Neurophysiology, for Searle, does not differentiate between fine-grained ways of pointing, such as pointing to H<sub>2</sub>O rather than water. So, for Searle, neurophysiological states alone do not have original determinate content. Within virtually all naturalistic theories, when subjects point to, or attend to, water and then to H<sub>2</sub>O one state and then another becomes active. These states have different content and their activation determines whether the subject is pointing to, or attending, water or H<sub>2</sub>O. However, for Searle, if these states are not potential conscious states they will not be fine-grained enough to pick out the two different determinate contents in question (Searle 1992).

Searle thinks that, in the case of water and H<sub>2</sub>O, an informational state in the brain may be equally causally correlated with both water and H<sub>2</sub>O at once. This is at odds with most naturalistic theories of content. When a person entertains the two different contents, there are two different internal states with different content. Searle thinks, however, that from the informational state, or the brain state, alone we will not be able to distinguish between the content 'water' and the content 'H<sub>2</sub>O'. Searle believes the two states will be equally causally correlated with the liquid in the glass.

Searle believes something more is needed for the causal correlation of an informational state,

or a brain state, to become as finely individuated as the two possible contents. For a subject's merely neurophysical, or merely informational, states to have a *derived* difference between the content 'water' and the content 'H<sub>2</sub>O', on Searle's view, it must be dispositionally linked to becoming one of the subject's conscious states that has one or the other of these contents (Searle 1992, p. 152).

Horgan and Tienson hold a CP much like Searle's (Horgan and Tienson 2002). Horgan and Tienson believe that the indeterminacy found, for example, when trying to translate a foreign speaker's use of the term "gavagai" in the presence of a rabbit is solved by conscious experience. In such cases, the foreign speaker could mean rabbit, or undetached-rabbit-parts, or several other things that coincide with the presence of rabbits (Quine 1960). Horgan and Tienson say that indeterminacy of content

"seems plainly false – *and false for phenomenological reasons*...because there is something that it is like to have the occurrent thought that rabbits have tails, and what it is like is different from what it would be like to have the occurrent thought that collections of undetached rabbit parts have tail-subsets" (Horgan and Tienson 2002, p. 522, emphasis added).

So, Horgan and Tienson also believe the fine individuation of mental content is grounded by conscious states. They are clear about this when they say, "the overall [conscious] character of such a state *comprises*... its specific mental content" (Horgan and Tienson 2002, p. 524, emphasis added). This makes mental content, on their view, dependent on conscious content and so objects to standard treatments of unconscious mental states.

#### 6.2.1.1 Reply to Searlean Versions of Dependence

There are several problems with Searle's position. These problems entail that the states explaining certain human behaviors can still be accomplished with unconscious mental content that is independent of conscious mental states. We will take each problem in turn.

The first thing to ask about Searle's position is why he thinks nothing in the third-person point of view can differentiate fine-grained mental content. This is at odds with a causal perspective

on mental states outlined in chapter three (see also Fodor 1987; 1990a,b; 2009). Many theorists hold that differences in fine-grained mental content would be individuated by causal relations in the primary case and a network of other behaviors and interactions with other internal states in the case of higher-level content (Pylyshyn 2009).

Why does Searle reject a causal approach to mental content? He believes that he, for example, has mental states intrinsically. So, from the inside, as it were, for Searle, there is a *sense* of directedness that is not accounted for from the third-person perspective (Searle 1992; see also Horgan and Tienson 2002). For example, some theorists hold that conscious states can be directed in a fine-grained way without relation to a real world (Meixner 2006). For such theorists subjects do not have to have causal relations with a world for their states to be directed in this sense. This means that Searle is endorsing a classical view and denying the causal approach to mental content. However, having determinate *seeming* content, from the first-person point of view, may not be the same as having actual content. In this way, Searle may be onto a different topic than representational content. This shows that Searle is rejecting unconscious mental states for the wrong reasons.

A further problem for Searle's argument for his CP arises when examining the physical features of the states that make for, or cause, a subject's conscious states. If the complexity, say, of psychologically explanatory states that cause conscious states is comparable to the complexity of the conscious states themselves, these states will be real mental states. A state coded so as to cause a conscious state in a subject requires a sufficient degree of complexity to be able to cause the specific content of the conscious states. The complexity of the state coded so as to cause a conscious state will mean the unconscious state has mental content of its own. Since this is true, unconscious mental states are required in Searle's position as sufficient causes of determinate

conscious content.

To see this more clearly we want to look at how two conscious states with different content are *physically* individuated. There must be physical differences that individuate one conscious state in a subject being directed to the world in one way from another being directed in a different way (Kriegel 2003). Differences in physical complexity, for example, must underlie the physical individuation of the two different contents of the two conscious states.

Other physical states must then have the power to cause the conscious states that are finely individuated in this physical way. These other states must have this power so they can preserve the differences in fine-grained content of the conscious states. This requires that the *causes* of conscious contents be at least as finely individuated as the conscious contents themselves. How else could the causes be disposed to bring about just those contents if this were not the case (Smith 1999)? Each of the finely individuated conscious contents requires its own finely individuated internal cause having comparable physical individuating features.

This leads to a critique of Searle's argument for a dependence thesis. This critique is given by several philosophers (Armstrong 1991; Smith 1999; Kriegel 2003; Van Gulik 1995). These theorists each suggest that if two different unconscious states cause fine grained differences in a subject's conscious contents then "there must be corresponding detail in the...unconscious [states as well]" (Armstrong 1991, p. 70). So, corresponding differences in the features of a subject's unconscious states is *required* by positions such as Searle's.

As one author notes, because of the role these states play in explaining certain human behaviors, "there is no reason why this subtle relation should not also differentiate *unconscious* beliefs" (Kriegel 2003, p. 280). Armstrong also concludes that, "once this point has been seen [we] should... attribute... intrinsic mentality to the unconscious belief" (1991, p. 70; see also

Smith 1999, p. 153).

Armstrong goes on to press this complexity objection further (1991). He suggests that the causal structure of finely individuated unconscious states could be divorced from any potential to become conscious states. Such states, with much the same physical complexity as conscious states, could still function in the subject's mental system as mental states. This means that these states can be treated as mental states independent of *any* tie to conscious states.

Revisiting the objection, then, Horgan and Tienson's initial thought does seem partly correct (Horgan and Tienson 2002). Semantic indeterminacy seems false *from* a subject's conscious perspective. However, this alone does not establish dependence between mental content and a subject's conscious states. It is not enough, for example, to show that indeterminacy is false *only based on* a subject's conscious mental states (Horgan and Tienson 2002).

For example, it seems there are both differences in directedness of certain of a subject's mental states and a difference in the subject's conscious content. Perhaps neither is more fundamental than the other. In fact, they may be conceptually distinct and so unrelated in any way (see Kriegel 2003, p. 286; Crane 1998; but see McGinn 1988). Worse, for dependence theorists such as Searle, or Horgan and Tienson, it may be the case that the subject's unconscious mental states are explaining the differences in the subject's conscious content. For example, thoughts about rabbits, rather than undetached-rabbit-parts, could be delivered *to* conscious status individuated in this way rather than delivered *from* conscious status in this way.

For several reasons, then, we can be skeptical of Searlean style arguments for a dependence thesis. This leaves unconscious mental states as independently explanatorily valuable for explanations of many human behaviors.

### 6.2.2 Ludwig's Version of Dependence

Searlean arguments are just one way to attempt to defend a dependence thesis, or CP. Ludwig

argues for dependence of mental states on conscious content in another way (Ludwig 1996; see also Strawson 1994; Horgan and Tienson 2002; Pitt 2004). Ludwig holds that unless a psychological explanatory state can become a conscious state then it does not belong to the subject *qua* mental (1996; Jacob 2003). For Ludwig, conscious states belong to their subjects in special ways that make them mental. This is because of the unique knowledge a subject has of the content of their conscious states. Without this special knowledge, or the potential to gain this knowledge about a state, then the explanatory states of a system do not belong to the system as *mental* states (Ludwig 1996). Ludwig's notion of a mental state is essentially tied to a subject's knowledge of the content of that state. He holds this view because he thinks there is a "centrality of the first-person point of view in our conception of mental phenomena" (1996, p. 31).

Ludwig is dissatisfied, as he says, "both with Searle's formulations of the connection principle and his argument for it" (1996, p. 31). Ludwig thinks that Searle does not focus on the correct aspects of representation in formulating his arguments for dependence, namely the centrality of the first-person perspective in mental phenomena (1996).

Ludwig holds that only states "presented-*to*" a conscious subject, in the unique way that conscious states are, can be mental states. This requires a view of a subject's self such that content can be presented to that "thing". In terms of systems internal to a person, however, Ludwig does not give much detail regarding what this means. He just holds that only first-person states belong to the subject in a way that has mental content because this is a central part of our *conception* of mental phenomena (Ludwig 1996).

Metzinger clarifies something like Ludwig's position (Metzinger 2000). Metzinger notes that conscious mental states have, "two highly interesting characteristics: centeredness and perspectivalness...[that give] a global, structural, property [to conscious] space as a whole... This

simply amounts to the fact that under standard conditions, the dynamics of conscious experience unfolds in a space that is centered” (2000, p. 288-9; see also Kriegel 2007; Levine 2001; Clark and Kiverstein 2007). The idea is that only those states unfolding in a centered space of informational states are mental states *of a subject*. Auditory echolocation and visual foveal perception during movement are two illustrative examples.

Ludwig’s position is that only states that have features such as centeredness and perspectivalness are first-person conscious states and only these states are mental in an original way. He grants that other states may act as-if they were mental, but only when they have a *potential* to become states that are part of a centered space with features such as centeredness and perspective. With this potential such states only gain a *derived* mental status (Ludwig 1996, p. 20; Jacob 2003). Features like centeredness and perspectivalness, or the potential to gain such features, gives a subject special knowledge of the state such that it can have real mental content.

This argument for a Connection Principle is like the position McGinn offers on conscious states (1988). McGinn offers the objection against unconscious mental states that they do not present their content to an experiencing subject (1988). In McGinn’s terminology they have no “inward-pointing face”, or subject-*presenting* aspect, of content and so cannot be mental. Theorists such as Ludwig and McGinn believe that to have original mental properties a state must represent *to* some part of the cognitive system, often called the subject, in a first-person way (see Bennett and Hacker 2003, p. 68).

#### 6.2.2.1 Reply to Ludwig’s Version of Dependence

The reply to Ludwig’s formulation of a dependence thesis objection to unconscious mental states begins by noting that not all representations attributed to human subjects, either by researchers or by the subjects themselves, are attributed in the way Ludwig requires. That is, the explanatory value of some mental states does not require that they are coded into a centered

space or “presented-*to*” a subject in first-person ways. Furthermore, as we will see, states *with* the qualities of being centered and having perspective are not necessarily conscious states.

To begin the reply to Ludwig, we can see that when explaining behavior based, for example, on echolocation “[psychologists] ascribe to [the system] representations of the speed of sound and of the distance between the ears” (Evans 1982, p.104 his note 22, also see p. 154). When doing this, psychologists attribute a split between content that is attributed to sub-systems and the content that is attributed to the cognitive system as a whole. Many explanatory mental states of particular representational systems are not “presented-*to*” the whole cognitive system.

Although not all representational states are “presented-*to*” the cognitive system as a whole, what is interesting to ask, in respect to Ludwig’s type of dependence, is whether all content which *is* presented in a centered and perspective-like way is dependent on being, or even having the potential to become, a conscious state. If Ludwig happened to be right then a subject’s unconscious mental states would be limited to states that are not centered and do not have perspective. This might not be a bad result, but it is not a necessary requirement that centered and perspectival states are conscious mental states. Likewise, holding that unconscious mental states are never centered or perspective-like rejects some views of the cognitive unconscious that include these kinds of higher-level states (Shevrin 1992; Armstrong 1991; Kriegel 2003).

All centered and perspectival content is not dependent on conscious states or the potential to become a conscious state. To see this we must separate the properties that make a state a conscious state from the properties that make content centered and perspectival. If we can do this we can break Ludwig’s argument for dependence between certain kinds of mental states and conscious states. If centered and perspectival content can occur in representations that are not conscious states, then even centered and perspectival representations can be unconscious mental

states.

This response rests on the idea that we can distinguish between two dimensions of certain mental states. There is, for example, the supposed dimension of ‘what it is like’ for the subject to have certain mental states. This dimension is supposed to be how the mental state feels to that subject. There is also the dimension of the centeredness and perspectivalness of certain representations, the dimension of ‘who it is for’ (Kriegel 2007; Metzinger 2000; Ludwig 1996). This second dimension is the global structure, or perspective, of certain mental states. For example, these features of mental states represent how the various parts of a visual or auditory field will flow for the whole organism when it moves.

To see why Ludwig’s arguments for a dependence thesis fail we must ask which dimension of mental states grounds mental content with features such as centeredness and perspective. The content of these kinds of states depends, in part, on the dimension of ‘who it is for’. This does not, however, entail that these states are dependent on the supposed dimension of ‘what it is like’ for a system to have such states. This is because the special content of these states comes from their information flow, specifically information flow that can be described as centered and perspectival.

So, it seems we can distinguish the dimensions of ‘who’ a mental state is for from the supposed dimension of ‘what it is like’ for that system to have that mental state. Two recent research programs, that of Gibsonian psychology and of particular connectionist networks, show how certain states can fulfill the ‘who it is for’ dimension by having perspective and centeredness without being conscious mental states. To see this possibility better we must first look at Bermudez’s notion of Gibsonian self-specifying representations (1998, p. 106 and 108).

There is, for any animal in motion, a dynamic flow in, say, their optical or auditory field,

which specifies a particular location that gives the states representing the animal's location and motion special content. That is, the information in certain representations defines a center point where the creature in the scene is and how they are moving through that scene. The representation will show that *from this perspective* one particular location is special for the subject. Such representational flow, for Bermudez, is then 'self-specifying' in this respect.

Is this self-specifying content the kind of first-person "presentation-to" required by Ludwig? One reason for thinking that these representations are the kind of special content Ludwig specifies is the spatial processing that a system must reserve just "for" the creature. For example, there is a location, or perspective, that is created in linear-perspective drawings. Just as there is a vanishing-point in linear-perspective drawings, so, too, optical, or auditory, flow patterns define a point "for" the subject viewing the scene (Mandik 2001; Mollon 1997). Such a "viewing-point" is reserved as a special location in space, and so special processing must be reserved for that point. This kind of representation seems like a basis for the special knowledge Ludwig's position relies on.

Notably, this kind of first-person "presentation-to" does not require such states to be conscious states. We can strengthen this intuition if we combine Bermudez's basic notion with the work of Arbib and Liaw (1995). The work these researchers do shows the processing of 'optical' flow patterns in neural networks. As an illustration we can think of the visual display in a flight simulator. Flight simulation, as well as the patterns that Arbib and Liaw simulate, seem to employ representations that have centeredness and perspective. There is a flow of information that makes it appear that the viewer is passing through a scene.

For example, the set of vectors formed by the movement of the simulated plane in a flight simulator can be represented graphically by arrows for the speed and movement of the simulated

plane. When this is done a kind of center and perspectivalness is created within this set of vectors. The information about the centered perspective would be represented by the overall set and orientation of the arrows as the ‘viewer’ passes through the scene. In this way, a system’s spatial behavior could be represented as if passing through a visual or auditory field. This set of vectors, or their graphical representation in terms of arrows, can have the information flow similar to what Bermudez identifies as self-specifying (Mandik 2001; Mollon 1997; Shani 2006). Here, too, a viewing-point is well defined in the overall representational flow. This will give the representation a centeredness and perspective without these states needing to be described as conscious states.

These networks show a physical realization of the centeredness and perspective of representations. From the self-specifying aspects of these neural networks alone there is no reason why this content need *also* be treated as conscious content or content that can become conscious (but see Bermudez 1998). Properties of conscious states, or of potentially becoming conscious states, seem to do no work. This breaks Ludwig’s argument for dependence of these kinds of mental states being, or having the potential to become, conscious states. Such a break means unconscious mental states can still have an explanatory role, even in the case centered and perspectival representations.

### 6.3 Objection #3: Identity of Mental States and Conscious Mental States

The third objection to unconscious mental states holds that mental states explaining cognitive phenomena are always conscious states, at least of some form or other (Hanna 2008; Lind 1986; Carel 2006; Janet 1886; Brentano 1874; Descartes 1641; James 1890). This objection is like the first objection in its conclusion that all of a subject’s mental states are conscious. Here, instead of denying that unconscious states are mental the current objection denies they are unconscious.

This objection is best illustrated by the distinction between focal and peripheral, or even

diplopic, visual content. Diplopic visual content is demonstrated, for example, in the double vision humans have when looking at an object that is far off but attend to an object that is close (Tye 2008). Some diplopic content, in particular, is minimal conscious content in that it is almost always in the background of a subject's visual attention. Likewise, at least one of the images in diplopia is often faint.

This objection holds that all mental content is either fully conscious content, such as in focal vision, or background, or minimal, conscious content, such as in diplopia. More content than certain kinds of visual content is considered minimal conscious content on this view. For example, unattended representations of tactile, or other sensation, are often considered background or minimal conscious content. Most mental states explaining human behaviors are simply treated as background conscious content on this objection (Hanna 2008; James 1890; Janet 1886; Brentano 1874).

Several examples support the idea that there are fully and minimally conscious states in a single person's overall cognitive system at a time. The first example, as we have seen, is in vision. Further, though, focally attended states in general suggest many of a subject's conscious states are in the background, at least in relation to focal attention. Tactile and auditory states, when they are considered conscious states, are not always the focus of a subject's attention. However, just because some of subject's conscious content is not focally attended does *not* entail that any of the subject's other mental states are unconscious. There is the possibility that non-attended states are merely minimally conscious states.

As an illustration, when a person is fully concentrating on a difficult math problem they do not become blind to their peripheral visual states, numb for many of their tactile states, or deaf to auditory stimulation. The case of attention seems to give clear examples of a division between

focal and background conscious states in a single person's mental system at one time.

Generally, examples such as these fit a 'two-structure' model of conscious experience (Dilworth 2005). As Mangan states, "contemporary cognitive research supports the hypothesis that consciousness has a two-part structure: a focused region of articulated experience surrounded by a field of relatively unarticulated, vague experience" (1993, p. 89). This less articulated region of conscious experience is sometimes known as "fringe" mental content (Galin 1994; James 1890).

Feelings of familiarity, correctness, or the feel of the relative ease of retrieving information, when these feelings are not directly attended, form part of the background of conscious mental states in a subject's mental system. The above feelings are often 'fringe content' of conscious awareness, but are still considered conscious mental states nonetheless. This is generally because subject's either judge them to be conscious states or, more commonly, to have been conscious states when they are cued in a way that they attend to them (Mangan 1993; Galin 1994; Schwartz et. al. 1991). One take on this objection treats many of a subject's mental states as 'fringe' conscious content.

### 6.3.1 Hanna's Version of the Identity Thesis

Hanna, for example, presses the distinction between focal and minimal conscious states to the logical extreme (2008). He suggests that *all* mental states are conscious states of one form or another. From this he concludes there is no such thing as the cognitive unconscious (2008, p. 59). For example, he states that, "consciousness goes all the way... to the ground floor of cognition" (2008, p. 59). He uses the example of divided attention as a model for how to treat all mental content (2008, p. 60; see also Kohut 1971; Ricoeur 1974; Rauhala 1969). All mental states, on Hanna's view, are thought to be conscious states, or at least non-focal or minimal, conscious states.

Hanna believes that conscious status, of one form or another, is essential to mentality. His reasons for this are quite complex, linking bodily states, the metaphysics of telling right from left, and a notion of mental content that is not tied to conceptualization (2008). As an illustration, he believes that information regarding which hand is a right hand in a universe with only a single right hand, is essentially embodied. Hanna believes the knowledge of such information requires having such a hand. Hanna further believes that this content is non-conceptual. He also holds that such content is essentially tied to being a conscious state, at least of one form or another. Suffice it to say here, he believes many of the states that help explain psychological behavior to have this kind of non-conceptual content. Furthermore, for Hanna, these kinds of content are essentially conscious, often of a non-focal, or minimal, form of consciousness.

How can Hanna make a position such as this plausible? Think of cases of blindsight. Blindsight patients report an inability to see certain visual stimuli, but can accurately respond to the stimuli based on visual processing. These results are often thought to demonstrate the presence of unconscious mental content guiding blindsight patient's visual behavior.

However, sometimes blindsight patients report that they have a sense, or a feeling, that a stimulus is present or that it is moving (Weiskrantz 1997). Hanna uses reports such as these to argue that typical conscious visual experience has, at least, two aspects. One aspect is the visual image. Another aspect might be called the "visual feel". Hanna, for example, suggests that cases of blindsight are simply the loss of the conscious visual *image* and not a loss of the conscious visual *feel*. In this way, the visual but "blind" states guiding action, in cases of blindsight, would actually be minimally conscious states (Hanna unpublished). Hanna bases this idea on data regarding the reports blindsight subjects sometimes give regarding "visual feel". For example,

blindsight patients sometimes report having a vague feeling as of motion when they are able to detect motion in their blind-field.

With this distinction in hand Hanna treats results from any patients, relevantly like blindsight patients, in the same way. That is, he treats any patient data that seems to illustrate unconscious mental states as illustrating minimally conscious mental states, instead (Hanna 2008; Galin 1994). With this distinction, he tries to hold that *all* mental states are some kind of conscious mental state (Hanna 2008; Horgan and Tienson 2002). For Hanna, if a subject does not judge a state to be conscious it is simply because it is very minimally conscious.

A potential justification for Hanna's generalization of attributing some for of consciousness to all mental states is his belief that mental states must be conscious states to be mental. Why would a theorist hold such a position? Sometimes, this belief is due to the idea that mental states alone just *cannot* "build up" to fully-integrated conscious experience (Jopling 1996, p. 155; Bennett and Hacker 2003; McDowell 1994a; Gallagher 2005).

### 6.3.2 Lind's Version of the Identity Thesis

There is another way to formulate this kind of objection. This second formulation does not invoke minimal phenomenal qualities. This formulation suggests that mental states explaining certain human behaviors are actually *fully* conscious mental states but are simply quickly forgotten. This objection is given, in part, by William James and Rene Descartes among many others (James 1890, p. 165; see Descartes 1641 on distance judgments; Smith 1999, p. 60; Simmons 2001, p. 35; Lind 1986).

Lind, for example, says "[unconscious] acts and activities are actually conscious processes so *short-lived* that they are not reflectively retrievable" (1986, p. 334). He goes on to state that the "continuously shifting patterns of variegated phenomenal prominence [during a slam dunk in basketball, say]... are too swift and subtle for anything but hypothetical reconstruction" (1986, p.

331). Although Lind should be worried over the details of how these states actually function in a physical system, he is not worried here about his position being either ad hoc or unverifiable. He is merely supporting the logical possibility that all mental states are fully conscious mental states. He says, “the [unconscious states] in question are actually *directly* conscious [states] to which we do not have *reflective* access” (Lind 1986, p. 334).

For some of a subject’s conscious mental states Lind’s conclusion does seem to be true, at least in practical situations. For example, when subjects are not artificially cued many of their conscious states are not remembered or reflected upon. Lind’s illustration involving the slam dunk seems common enough. Variegated patterns of conscious states do seem to occur for certain subjects during almost any movement. Likewise, the explanatory apparatus he uses to explain these patterns are general enough to be used for other cases.

For example, we can use his methodology to explain some human behaviors often thought to involve unconscious states. Change *blindness* results, when subjects do not notice salient changes in a visual scene, have sometimes been interpreted to be mediated by fully conscious states that are simply unnoticed or rapidly forgotten by subjects (Dulany 2000; Wolfe 1999; Block 2007). Likewise, pre-attentive processing could also be explained by fully conscious but quickly forgotten states (Lamme 2003, p. 13, fig. 2).

Given possibilities like the above, Lind suggests that it is more parsimonious to account for all of a subject’s mental states in terms of a single form of conscious state, rather than posit minimally conscious states or unconscious states (Lind 1986, p. 342). He generalizes his conclusion to all of a subject’s mental states because he thinks consciousness is essential to mental status.

This position does endorse two kinds of conscious state, one kind that is not reflectively

accessed and another that is reflectively accessed. However, this position does not entail any states are minimally conscious states. In this way, the states in question could simply be rapidly forgotten and so un-reportable and unreflectively detected (Lind 1986, p. 335; Simmons 2001, p. 35; Descartes 1641; James 1890). Again, Lind is not worried about the mechanisms that carry-out the control of such behavior. If he were, his appeal to parsimony would be less appealing.

Lind's position may be less plausible than Hanna's. Both Lind's and Hanna's positions hold that there are no unconscious mental states. Both theorists may be right that some human behaviors thought to be explained by unconscious states may be better explained by minimally conscious states, or even rapidly forgotten, but still fully conscious states.

However, positions that posit minimal, or quickly forgotten, conscious states for *all* of a subject's mental states may be empirically unverifiable. This partly depends on what counts as an empirical measure of conscious states (see Block 2007; Dennett 1991). These positions need not be just unverifiable dodges, even if unverifiable dodges are popular with some theorists who attempt to save such proposals. What drives such theorists is the idea that conscious states are what controls and guides our actions. This idea is intuitive and philosophically popular, whether or not the states guiding some human behaviors are fully, or minimally, conscious states (O'Shaughnessy 1992; Wallhagen 2007; Kelly 2007; Ricouer 1974; Kriegel 2007).

### 6.3.3 Reply to Identity of Mental States and Conscious Mental States

The reply to the objection that all mental states of a system are actually conscious mental states, at least in some form or other, needs to break the identity that some theorists think holds between mental states and conscious mental states. Any supposed identity relation between conscious states and mental states is a stronger relation than that proposed by dependence theorists.

To show how this supposed identity is broken we must note that there is no conceptual

confusion, such as an obvious category mistake, or a contradiction, that directly follows from accounting for some kinds of human behaviors with mental states that are not conscious states of any form. This means there is no essential conceptual identity between human subjects' mental states and conscious mental states. Worse, positions which identify conscious mental states and mental states seem to entail their own conceptual difficulties. These are good reasons to abandon such an identity thesis.

Take the case of dreamless sleep. We attribute autobiographical mnemonic content, beliefs, and even limited representation of sensory information about the local environment to human subjects even when they are in dreamless sleep (Goldman 1999). Positions that identify all mental content with some conscious state, in whatever form, entail that a particular human subject would have conscious experience of all of its mental states, even during dreamless sleep. If this entailment is not incoherent it seems to unnecessarily stretch the meaning of the term "conscious experience". We do attribute mental content to humans in dreamless sleep and these states are not generally considered to be conscious states. This makes the current objection sound stipulative and without independent motivation. Worse, identification between a subject's mental states and their conscious states seems explanatorily unnecessary for most accounts that employ mental states.

Other cases are almost as bad. For example, on these views, all of a subject's mnemonic mental states would be conscious at all times. Even if all of a subject's mnemonic mental states can become conscious at some time or other, identity theorists are not positing a dependence of mental states on the potential to become conscious mental states. Identity theorists hold that a subject's mnemonic mental states are actually conscious before they are remembered. Not only does this seem unnecessary, and unfeasible, but it seems to confuse the notion of memory with

some notion of perception as well. A mental system, such as a human subject, escaping from predation, requires a highly concentrated focus of attention. On identity views, such a system would also have conscious states regarding all of its other mental states at this time as well. Again, we can attribute beliefs, memories, and sensory representations to subjects even when they are highly concentrated on other tasks (Goldman 1999). There is no explanatory reason for all, or even most, of such a subject's mental states to be conscious at this time.

Many of the above entailments seem unnecessary or worse. Although it is important to grant minimal, as well as quickly forgotten, conscious states for some psychological explanations this alone does not show that all of a subject's mental states *must* be conscious or even that all of the subject's mental states *can be* considered conscious states at a time. Furthermore, no necessary explanatory work seems to be done by various kinds of conscious states in many accounts of human behavior. Even granting minimal, or rapidly forgotten, conscious states, for some cases, a substantial number of human behaviors can still be valuably explained using unconscious mental states.

#### 6.4 Objection #4: All Mental States Can Become Conscious States

A final type of objection to unconscious mental states addresses only strictly unconscious mental states. Strictly unconscious mental states are the mental states of a subject that cannot become conscious to that subject, at least if the laws of nature and the information-processing architecture of the subject are held fixed.

The current objection grants the explanatory value of many unconscious mental states. This objection also grants that unconscious mental states have mental status independently of any tie to conscious states and so rejects dependence theses and identity theses relating conscious, and unconscious mental, states. However, this objection does hold that all of a subject's mental states can become the subject's conscious states of some form or other *at some time or other*. In

this way, the present objection rejects strict unconscious mental states. For example, theorists who hold this position suggest that with the right sorts of cues and in the right kinds of circumstances any of a subject's unconscious mental states can become conscious states for that subject.

The current objection denies a popular aspect of the standard picture of a subject's unconscious mental states. The standard view of the cognitive unconscious holds that certain mental states of a subject cannot become conscious mental states to the subject they are in, in any way at any time. For example, Jackendoff states that a language user's "ability to understand and utter an indefinitely large number of sentences that [they have] never heard before – is governed by complex principles of which one is not *and cannot* be consciously aware" (1982, p. 83-4, emphasis added; Kihlstrom 1987; 1990; 2004). The idea that certain mental states cannot become conscious mental states for the subject who has them is endorsed by many current supporters of the explanatory value of unconscious mental states (Hassin et. al. 2005; Wilson 2002; Kihlstrom 1987).

The objection to strict unconscious mental states requires making sense of the idea that *all* of a subject's mental states can become conscious states of one kind or another at some time or other. There are at least two general accounts for how unconscious mental states can become conscious states. These two accounts are related to first-order and second-order accounts of a subject's mental states becoming conscious states of that subject.

Why would a theorist believe that *all* of a subject's mental states could become conscious states for that subject, at least in some way or other at some time or other? This is important for phenomenologists who hold that conscious states are still a *central* aspect of mind, even while granting the existence of independent unconscious mental states. For example, as Carel notes a

subject's mental states may all be accessible as conscious states with the right cues (2006, p. 180). He believes this because he thinks a subject's mental states are essentially tied to the subject's ability to access these states consciously. This notion of accessed mental states is tied, for Carel, to making a mental state some form of conscious state. This means that the mental states of a subject are not totally outside the reach of their conscious access.

One common, second-order, account of mental states becoming conscious states holds that for a subject's mental state to become a conscious state the subject's attention is needed. This is a second-order theory in that an attentional state must be directed to a mental state for the mental state to become a conscious state (Rosenthal 2005). Several theorists suggest that all of a subject's unconscious mental states could become conscious states via some form of directed attention by the subject. Of course, this may require a certain amount of practice but the possibility is thought to be, in principle, attainable by certain subjects (Kupperman 1984; Van Dusen 1960, p. 39; see also Carel 2006, p. 180; Ricouer 1974). This possibility is not entailed by the second-order account of conscious states. Most second-order theories of conscious states hold that many of a subject's states will not become conscious no matter how the subject's attention is directed.

This is just one way of accounting for how mental states could become conscious mental states. It presumes a higher-order state, such as the subject's attention, needs to be directed at a lower-order mental state for the lower-order state to become a conscious mental state of the subject (Rosenthal 2002; Byrne 2007; Pitt 2004). Alternately, there are first-order theories of a subject's mental states becoming conscious mental states of that subject. On first-order views, it is not necessary to posit the involvement of other mental states of any order, such as higher-order attentional states, for a subject's mental states to become conscious mental states of that subject.

For example, a theorist could hold that only when a mental state interacts with certain operations, or controls certain behaviors, does that state become a conscious mental state in a first-order way for the subject. Only the mental states of a mental system that are *actively* interacting with certain operations, or controlling certain behaviors, would be considered conscious states on these views (see Dretske 2006 for a similar theory). In this form, the current objection would suggest that any of a subject's mental states can enter the relations that make them conscious mental states (Thomasson 2000, p. 203; Gennaro 2006; see also Brentano 1874). The radically flexible pictures of a subject's conscious mental states that we discussed in chapter three may also allow for this possibility (Dennett 1991; Dennett and Akins 2008).

#### 6.4.1 Reply to All Mental States Can Become Conscious States

The reply to the objection that all of a subject's mental states can become some kind of conscious state for that subject, at least at some time or other, reemphasizes the explanatory value of strict unconscious states. Much of the defense for strictly unconscious mental states was presented in section 5.6.

The standard view of a subject's unconscious mental states holds that some of a subject's mental states are unable to ever become conscious states of any kind for that subject at any time. Strictly unconscious mental states are, generally, presupposed in many explanations of various human behaviors (Fodor 1983; Lahav 1993; Dennett 1978; Shallice 1988; Bisiach 1988; Leman 1981; Hirst 1995; Place 2000; Jackendoff 1982; Chomsky 1972; Dehaene et. al. 2006). A proponent of the cognitive unconscious can grant that *some* of a subject's mental states can become conscious states of that subject in some way or other at some time or other. Granting this does not impact the thesis that a substantial number of human behaviors can still be explained by strictly unconscious mental states.

What the reply to this objection needs to make sense of is an architectural divide which makes

it the case that certain mental states of a subject cannot become conscious mental states of that subject. As long as the laws of nature and the information-processing architecture are held fixed, then the architectural divide would prevent certain of a subject's states from becoming conscious states of that subject. An architectural divide of this kind would separate certain mental states within a subject's mental system. One way this is commonly done is to suggest that certain of a subject's mental systems lack access to particular mental states within the subject. This lack of access would be what makes the mental states strictly unconscious mental states of that subject.

However, the suggestion that one mental system is where all of a subject's conscious states exist risks endorsing Cartesian Materialism. No such 'consciousness system' seems to exist. For example, no lesion studies have been able to selectively abolish all and only conscious states. Endorsing that a single system in a subject is where all and only conscious mental states exist in a subject's mental system is not necessary. As we have noted above (chapter five), conscious states need only be available *as* certain kinds of mental states (Jackendoff 1987, p. 18; Jopling 1996, p. 161).

A subject's strictly unconscious mental states would be states that cannot be coded as conscious mental states for that subject. Jopling holds, for example, that if mental states are involved in making conscious states possible in the first place, then those mental states that underlie a system's conscious experience are states that cannot become conscious states themselves, at least for that subject. The architectural divide between mental states that *make for* conscious states shows that some of a subject's unconscious mental states will never be able to occur as conscious states (Jopling 1996, p. 158; Gallagher 2005). This means that some mental states of systems capable of having conscious states are strictly unconscious mental states.

Are the unconscious states that make for a subject's conscious states the only strictly

unconscious mental states of a subject's mental system? This may be the case on very flexible pictures (Dennett 1991; Tononi and Edelman 1998). However, the standard picture of a subject's cognitive unconscious holds that there are various kinds of strictly unconscious mental states. For example, some unconscious mental states run parallel to, and do not just underlie, a subject's conscious mental states. Some of the states in these parallel processes are thought to be valuably described as strictly unconscious mental states of the subject in question.

Data regarding subjects reaching for and grasping stimuli that look to be different sizes show that their reaching hand has the *correct* grip size, sometimes even to a subject's own surprise. Many theorists have taken these results to indicate strictly unconscious mental states guiding subjects' motor behavior (Place 2000; Milner and Goodale 1995; Clark 2001; Castelli et. al. 2005; however see Wallhagen 2007). One reason to believe that these motor states are a subject's strictly unconscious mental states is that they run fully parallel to processes that make states available as conscious states. Likewise, subjects are hard pressed to notice such states. Furthermore, even when subjects are informed about the existence of these states they do not judge these states to be, or to have been, conscious.

From these brief examples of mental states that underlie and parallel the processes that make for a subject's conscious states we can see that even strictly unconscious mental states will have a valuable explanatory role for a substantial number of human behaviors. Including states that are strictly unconscious is a standard view of a subject's cognitive unconscious.

#### 6.5 Summary of Objections and Replies

Mental states valuably explain a substantial number of human behaviors. Unconscious mental states are generally involved in the explanation of processing leading up to, surrounding, and impacting what subjects judge to be their conscious experience (Jackendoff 1992). Many of these explanations do not require that the explanatory mental states employed have any other

special properties which might make them conscious mental states, or have the potential to become conscious mental states.

Furthermore, views that suggest a dependence or identification of a subject's mental states and their conscious mental states do no required explanatory work. Current explanations in cognitive science show the independence of mental states in the explanatory framework needed when accounting for a substantial number of human behaviors (Armstrong 1991). Likewise, deflation, dependence, and identity objections, as well as the objection to strictly unconscious mental states, can lead to unnecessary consequences (Wakefield 2001). In the above ways, the explanatory value of various kinds of unconscious mental states remains well defended.

#### 6.6 Final Summary

The defense of the explanatory value of unconscious mental states and the responses to objections against such states leads to a standard view regarding a subject's cognitive unconscious. Various kinds of unconscious mental states are explanatorily valuable for a substantial number of human behaviors. These kinds of explanatory states are of two kinds: 1) perceptual and cognitive unconscious mental states able to become a subject's conscious states, in some form or other at some time or other and 2) strictly unconscious perceptual and cognitive mental states that cannot become conscious states at any time for the subject who has them.

### Bibliography

- Abrams, R. and Greenwald, A. (2000). "Parts Outweigh the Whole (Word) in Unconscious Analysis of Meaning". *Psychological Science*. vol. 11: 118-124.
- Adams, F. (2003). "The Informational Turn in Philosophy". *Minds and Machines*. vol. 13: 471-501.
- Aizawa, K. (2003). "Cognitive Architecture: The Structure of Cognitive Representations". In: *Blackwell Guide to Philosophy of Mind*. p. 172-188. Stick, S. and Warfield, T. (eds.). Oxford: Blackwell Publishing.
- Akins, K. (1996). "Lost the Plot? Reconstructing Dennett's Multiple Drafts Theory of Consciousness". *Mind and Language*. vol. 11(1): 1-43.
- Akins, K. (2002). "A Question of Content". In: *Daniel Dennett*. p. 206-246. Brook, A. and Ross, D. (eds.). Cambridge: Cambridge University Press.
- Alston, W. (1971). "Dispositions and Occurrences". *Canadian Journal of Philosophy*. vol. 1: 125-154.
- Anderson, R. and Lebiere, C. (1998). *Atomic Components of Thought*. NJ: Lawrence Erlbaum Associates, Inc.
- Antony, M. (2008). "Are Our Concepts of Conscious States and Conscious Creature Vague?". *Erkenntnis*. vol. 68(2): 239-263.
- Antony, L. and Levine, J. (1991). "The Nomic and the Robust". In: *Meaning in Mind: Fodor and His Critics*. p. 1-16. Lower, B. and Rey, G. (eds.). Oxford: Blackwell.
- Aquila, R. (1989). "Intentionality, Content, and Primitive Mental Directedness". *Philosophy and Phenomenological Research*. vol. 49: 583-604.
- Arbib, M. and Liaw, J. (1995). "Sensorimotor Transformation in the Worlds of Frogs and Robots". *Artificial Intelligence*. vol. 72: 53-79.
- Archard, D. (1984). *Consciousness and the Unconscious*. Illinois: Open Court.
- Armstrong, D. (1991). "Searle's Neo-Cartesian Theory of Consciousness". *Philosophical Issues*. vol. 1: 67-71.
- Armstrong, D. (1999) *The Mind-Body Problem: An Opinionated Introduction*. Westview Press: Boulder, Colorado.
- Atkins, K. (1996). "Of Sensory Systems and the 'Aboutness' of Mental States". *The Journal of Philosophy*. vol. 93: 337-372.

- Audi, R. (1994). "Dispositional Beliefs and Dispositions to Believe". *Nous*. vol. 28(4): 419-434.
- Aydede, M. (2000). "On the Type/Token Relation of Mental Representations". *Facta Philosophica*. vol. 2(1): 23-49.
- Baars, B. (1988). *Cognitive Theory of Consciousness*. Cambridge: Cambridge University Press.
- Baars, B., Fehling, M., LaPolla, M., McGovern, K. (1997a). "Consciousness Creates Access: Conscious Goal Images Recruit Unconscious Action Routines, But Goal Competition Serves To 'Liberate' Such Routines, Causing Predictable Slips" In: *Scientific Approaches to Consciousness*. Cohen and Schooler (eds.). New Jersey: L. Erlbaum Associates.
- Baars, B. (1997b). "Some Essential Difference between Consciousness and Attention, Perception, and Working Memory". *Consciousness and Cognition*. vol. 6: 363-371.
- Bach, K. (1997). "Do Belief Reports Report Beliefs?". *Pacific Philosophical Quarterly*. vol. 78: 215-241.
- Bargh, J. and Morsella, E. (2008). "The Unconscious Mind". *Perspectives on Psychological Science*. vol. 3(1): 73-79.
- Barlow, H. (1982). "General Principles: The Senses Considered As Physical Instruments". In: *The Senses*. p. 1-32. Barlow, H. and Mollon, J. (eds.). Cambridge: Cambridge University Press.
- Bayne, T. (2007). "Conscious States and Conscious Creatures: Explanation in the Scientific Study of Consciousness". *Philosophical Perspectives*. vol. 21: 1-22.
- Beck, L. (1966). "Conscious and Unconscious Motives". *Mind*. vol. 75(298): 155-179.
- Bennett, M. and Hacker, P. (2003). *Philosophical Foundations of Neuroscience*. Oxford: Blackwell.
- Benoist, J. (2007). "Two (or Three) Conceptions of Intentionality". *Tijdschrift voor Filosofie*. vol. 69: 79-103.
- Bergmann, M. (2006). *Justification without Awareness*. New York: Oxford University Press.
- Bermudez, J. (1994). "Peacocke's Argument against the Autonomy of Nonconceptual Representational Content". *Mind and Language*. vol. 9(4): 402-418.
- Bermudez, J. (1995a). "Nonconceptual Content: From Perceptual Experience to Subpersonal Computational States". *Mind and Language*. vol. 10(4): 333-369.
- Bermudez, J. (1995b). "Syntax, Semantics, and Levels of Explanation". *Philosophical Quarterly*. vol. 45(180): 361-367.

- Bermudez, J. (1998). *The Paradox of Self Consciousness*. Cambridge: MIT Press.
- Bermudez, J. (2000). "Personal and Sub-Personal: A Difference without a Distinction". *Philosophical Explorations*. vol. 1: 63-82.
- Bermudez, J. (2003). "Nonconceptual Mental Content". *The Stanford Encyclopedia of Philosophy (Spring 2003 Edition)*. Edward N. Zalta (ed.). URL = <<http://plato.stanford.edu/archives/spr2003/entries/content-nonconceptual/>>.
- Betzler, M. (2007). "Making Sense of Actions Expressing Emotions". *Dialectica*. vol. 61(3): 447-466.
- Biederman, I. (1987). "Recognition-by-Components: A Theory of Human Image Understanding". *Psychological Review*. vol. 94(2): 115-47.
- Bilgrami, A. (1994). "On McDowell on the Content of Perceptual Experience". *Philosophical Quarterly*. vol. 44(175): 206-213.
- Bisiach, E. (1988). "The Haunted Brain and Consciousness". In: *Consciousness in Contemporary Science*. p. 101-120. Marcel, A. and Bisiach, E. (eds.). Cambridge: Cambridge University Press.
- Blake, R. (2001). "A Primer on Binocular Rivalry, Including Current Controversies". *Brain and Mind*. vol. 2: 5-38.
- Blakemore, S. and Frith, C. (2003). "Self-Awareness and Action". *Current Opinion in Neurobiology*. vol. 13: 219-224.
- Block, N. (1981). "Psychologism and Behaviorism". *Philosophical Review*. vol. 90(1): 5-43.
- Block, N. (1990). "The Computer Model of the Mind". In: *An Invitation to Cognitive Science (vol. 3)*. p. 247-289. Smith, E. and Osherson, D. (eds.). Cambridge: MIT Press.
- Block, N. (1995). "On a Confusion about a Function of Consciousness". *Behavioral and Brain Sciences*. vol. 18(2): 227-287.
- Block, N. (2005). "Two Neural Correlates of Consciousness". *Trends in Cognitive Science*. vol. 9(6): 46-52.
- Block, N. (2007). "Consciousness, Accessibility, and the Mesh between Psychology and Neuroscience". *Behavioral and Brain Science*. vol. 30: 481-548.
- Block, N. (2008). "Consciousness and Cognitive Access". *Proceedings of the Aristotelian Society*. vol. 108: 289-317.
- Boden, M. (1994). "Precis of The Creative Mind: Myths and Mechanisms". *Behavioral and*

*Brain Sciences*. vol. 17: 519-570.

Bouveresse, J. (1995). *Wittgenstein Reads Freud: The Myth of the Unconscious*. Cosman (trans.) Princeton: Princeton University Press.

Bowers, K. (1987). "Revisioning the Unconscious". *Canadian Psychology*. vol. 28: 93-104.

Braitenberg, V. (1986). *Vehicles: Experiments in Synthetic Psychology*. Cambridge: MIT Press.

Brentano, F. (1874). *Psychology from an Empirical Point of View*. London: Routledge.

Brown, A. (1991). "Review of the Tip of the Tongue Experience". *Psychological Bulletin*. vol. 109(2): 204-223.

Brown, D. (1996). "A Furry Tile about Mental Representation". *Philosophical Quarterly*. vol. 46(185): 448-466.

Burston, D. (1986). "The Cognitive and Dynamic Unconscious: A Critical and Historical Perspective". *Contemporary Psychoanalysis*. vol. 22: 133-157.

Byrne, A. (2007). Course materials for 24.500 Topics in the Philosophy of Mind: Perceptual Experience, Spring 2007. MIT OpenCourseWare (<http://ocw.mit.edu>). Massachusetts Institute of Technology.

Carel, H. (2006). *Life and Death in Freud and Heidegger*. Amsterdam: Rodopi.

Carruthers, P. (2006). "The Case for Massively Modular Models of Mind". In: *Contemporary Debates in Cognitive Science*. p. 3-21. Stainton, R. (ed.). Oxford: Blackwell Publishing.

Carver, C. (2006). "When Mechanistic Models Explain". *Synthese*. vol. 153: 355-376.

Carver, C. and Bechtel, W. (2006). "Mechanism". In: *The Philosophy of Science: An Encyclopedia*. p. 469-478. Sarkar, S. and Pfeifer, J. (eds.). New York: Routledge.

Castelli, F., Happe, F., Frith, U., and Frith, C. (2005). "Movement and Mind: A Functional Imaging Study of Perception and Interpretation of Complex Intentional Movement Patterns". In: *Social Neuroscience: Key Readings*. p. 155-169. Cacioppo, J. and Berntson, G. (eds.). New York: Psychology Press.

Casti, R. (2008). "The Cognitive Science of Holes and Cast Shadows". *Trends in Cognitive Sciences*. vol. 10(2): 54-55.

Castiello, U., Pauligana, Y., and Jeannerod, M. (1991). "Temporal Dissociation of Motor Responses and Subjective Awareness: A Study in Normal Subjects". *Brain*. vol. 114: 2639-2655.

Chalmers, D. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. New York:

Oxford University Press.

Chalmers, D., French, R., and Hofstadter, D. (1995). "High-level Perception, Representation, and Analogy: A Critique of Artificial-Intelligence Methodology". In: *Fluid Concepts and Creative Analogies*. p.169-194. Hofstadter, D. (ed.). New York: Basic Books.

Cherniak, C. (1981). "Minimal Rationality". *Mind*. vol. 90: 161-183.

Chisholm, R. (1967). "Intentionality". In: *Encyclopedia of Philosophy*. p. 201-204. Edwards, P. (ed.) New York: Macmillian and Free Press.

Chomsky, N. (1972). *Language and Mind*. New York: Harcourt Brace Jovanovich.

Chomsky, N. (1976). "On the Nature of Language". In: *Origins and Evolution of Language and Speech*. p. 46-57. Steklis, H. and Lancaster, J. (eds.). New York: New York Academy of Sciences.

Chomsky, N. (1990). "On the Nature, Use, and Acquisition of Language". In: *Mind and Cognition*. p. 627-646. Lycan, W. (ed.). Oxford: Blackwell.

Chomsky, N. (2000). *New Horizons in the Study of Language and Mind*. Cambridge: Cambridge University Press.

Churchland, P. (1981). "Eliminative Materialism and the Propositional Attitudes". *The Journal of Philosophy*. vol. 78(2): 67-90.

Churchland, P., Ramachandran, V., Sejnowski, T. (1994). "A Critique of Pure Vision". In: *Large-Scale Neuronal Theories of the Brain*. p. 23-60. Koch, C. and Davies, J. (eds.). Cambridge: MIT Press.

Clapin, H. (2006). "Mediation, Information, Error: Comment on Markman-Dietrich on Representation-Mediation". *Psychology*. vol. 9(74).  
<http://www.cogsci.ecs.soton.ac.uk/cgi/psyc/psummary?9.74>

Clark, A. (2001). "Visual Experience and Motor Action: Are the Bonds Too Tight?". *Philosophical Review*. vol. 110(4): 495-519.

Clark, A. (2007). "What Reaching Teaches: Consciousness, Control, and the Inner Zombie". *British Journal for the Philosophy of Science*. vol. 58(3): 563-594.

Clark, A. and Chalmers, D. (1998). "The Extended Mind". *Analysis*. vol. 58(1): 7-19.

Clark and Kiverstein (2007). "Experience and Agency: Slipping the Mesh". *Behavioral and Brain Sciences*. vol. 30: 502-503.

Cleeremans, A., Destrebecqz, A., and Boyer, M. (1998). "Implicit Learning: News from the

- Front". *Trends in Cognitive Sciences*. vol. 2(8): 275-281.
- Collins, A. (1998). "Beastly Experience". *Philosophy and Phenomenological Research*. vol. 58(2): 375-380.
- Collins, A. (1969). "Unconscious Belief". *The Journal of Philosophy*. vol. 66(20): 667-680.
- Coltheart, M. (1999). "Modularity and Cognition". *Trends in Cognitive Sciences*. vol. 3(3): 115-120.
- Coulter, J. and Sharrock, W. (2007). *Brian, Mind, and Human Behavior in Contemporary Cognitive Science*. Edwin Mellen Press.
- Cowey, A. and Stoerig, P. (1995). "Blindsight in Monkeys". *Nature*. vol. 373: 247-249.
- Craik, F. and Lockhart, R. (1972). "Levels of Processing: A Model for Memory Research". *Journal of Verbal Learning and Verbal Behavior*. vol. 11: 671-684.
- Crane, T. (1998). "Intentionality as the Mark of the Mental". In: *Current Issues in Philosophy of Mind*. p. 229-251. O'Hear (ed.). New York: Cambridge University Press.
- Craver, C. (2002). "Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory," *Philosophy of Science*. Supplemental 69: S83-97.
- Cunningham, S. (1997). "Two Faces on Intentionality". *Philosophy of Science*. vol. 64(3): 445-460.
- Cussins, A. (1990). The Connectionist Construction of Concepts. In: *The Philosophy of Artificial Intelligence*. p. 369-440. Boden, M. (ed.). New York: Oxford University Press.
- Cussins, A. (2002). "Postscript: Experience, Thought and Activity". In: *Essays on Non-Conceptual Content*. p. 147-159. Gunther (ed.). Cambridge: MIT Press.
- Damasio, A. (1996). "The Somatic Marker Hypothesis and the Possible Functions of the Pre-Frontal Cortex". *Philosophical Transactions of the Royal Society of London, B*. vol. 351: 1413-1420.
- Davidson, D. (1982). "Paradoxes of Irrationality". In: *Philosophical Essays on Freud*. p. 289-305. Wollheim, R. and Hopkins, J. (eds.). Cambridge: Cambridge University Press.
- Davidson, D. (2001/1970). "Mental Events". In: *Essays on Actions and Events*. p. 207-224. Davidson, D. (ed.). Oxford: Clarendon Press.
- Davidson, D. (1987). "Knowing One's Own Mind". *Proceedings and Addresses of the American Philosophical Association*. vol. 60: 441-458.

- Davies, M. (2000a). "Interaction Without Reduction". *Mind and Society*. vol. 1: 87-105.
- Davies, M. (2000b). "Persons and Their Underpinnings". *Philosophical Explorations*. vol. 1: 42-62.
- Davies, M. (1995). "Consciousness and Varieties of Aboutness". In: *Philosophy of Psychology*. p. 356-392. MacDonald, C. and MacDonald, G. (eds.). Cambridge: Blackwell.
- d'Avila Garcez, A., Gabbay, D., Ray, O., and Woods, J. (2007). "Abductive Reasoning in Neural-Symbolic Systems". *Topoi*. vol. 26: 37-49.
- Dawson, M. (1991). "The How and the Why of What Went Where In Apparent Motion: Modeling Solutions to the Motion Correspondence Problem". *Psychological Review*. vol. 98(4): 569-603.
- Debanne, D. (2004). "Information Processing in the Axon". *Nature Reviews Neuroscience*. vol. 5: 304-316.
- Dehaene, S., Nacchache, L., Sackur, J., Sergent C. (2006). "Conscious, Preconscious, and Subliminal Processing: A Testable Taxonomy". *Trends in Cognitive Sciences*. vol. 10: 204-211.
- Demir, H. (2008). "Counter Factuals and Conditional Probabilities: A Critical Analysis of the Counter Factual Theory of Information". *Australasian Journal of Philosophy*. vol. 86(1): 45-60.
- Dennett, D. (1978). "Towards a Cognitive Theory of Consciousness". In: *Brainstorms*. p. 149-178. Dennett, D. (ed.). Cambridge: MIT Press.
- Dennett, D. (1987). *Intentional Stance*. Cambridge: MIT Press.
- Dennett, D. (1991). *Consciousness Explained*. Boston: Little, Brown, and Co.
- Dennett, D. (1995). *Darwin's Dangerous Idea*. New York: Simon and Schuster.
- Dennett, D. (1998). "Myth of Double Transduction". In: *Toward a Science of Consciousness II: The Second Tucson Discussions and Debates*. p. 97-107. Hameroff, S., Kaszniak, A., and Scott, A. (eds.). Cambridge: MIT Press.
- Dennett, D. (2000). "With a Little Help from My Friends". In: *Dennett's Philosophy*. p. 327-388. Ross, D., Brook, A., and Thompson, D. (eds.). Cambridge: MIT Press.
- Dennett, D. (2001). "Are We Explaining Consciousness Yet?". *Cognition*. vol. 79: 221-237.
- Dennett, D. (2005). *Sweet Dreams: Philosophical Obstacles to a Science of Consciousness*. Cambridge: MIT Press.
- Dennett, D. (2007). "What RoboMary Knows". In: *Phenomenal Concepts and Phenomenal*

- Knowledge: New Essays on Consciousness and Physicalism*. p. 15-31. Alter, T. and Walter, S. (eds.). New York: Oxford University Press.
- Dennett, D. and Akins, K. (2008). "Multiple Drafts Model". *Scholarpedia*. 3(4):4321-4326.
- Dennett, D. and Kinsbourne, M. (1992). "Time and the Observer". *Behavioral and Brain Sciences*. vol. 15(2):183-247.
- Descartes, R. (1641). *Meditations on First Philosophy*. Cress, D. (trans.) Indianapolis, Indiana: Hackett Publishing Company.
- De Sosa, R. (2002). "Fringe Consciousness and the Multifariousness of Emotions". *Psyche*. vol. 8(14): <http://psyche.cs.monash.edu.au/v8/psyche-8-14-desousa.html>
- Devitt, M. (1994). "The Methodology of Naturalized Semantics". *The Journal of Philosophy*. vol. 91: 545-72.
- Dietrich, E. and Markman, A. (2003). "Discrete Thoughts: Why Cognition Must Use Discrete Representations". *Mind and Language*. vol. 18(1): 95-119.
- Dilworth, J. (2005). "The Double Content of Perception". *Synthese*. vol. 146: 225-243.
- Dijksterhuis, A. and Norgren, L. (2006). "A Theory of Unconscious Thought". *Perspectives on Psychological Science*. vol. 1: 95-109.
- Dixon, N. (1971). *Subliminal Perception: The Nature of a Controversy*. New York: McGraw-Hill Inc.
- Doris, J. (2009). "Skepticism about Persons". *Philosophical Issues*. vol. 19(1): 57-91.
- Dretske, F. (1978). "The Role of the Percept in Visual Cognition". In: *Minnesota Studies in the Philosophy of Science: Perception and Cognition*. p. 107-127. Savage, W. (ed.). Minneapolis: University of Minnesota Press.
- Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge: MIT Press.
- Dretske, F. (1986). "Misrepresentation". In: *Belief: Form, Content and Function*. p. 17-36. Bogdan, R. (ed.). New York: Oxford University Press.
- Dretske, F. (1988). *Explaining Behavior: Reasons in a World of Causes*. Cambridge: MIT Press.
- Dretske, F. (1994). "Explanatory Role of Information". *Philosophical Transactions of the Royal Society of London, A*. vol. 349: 59-70.
- Dretske, F. (1995). *Naturalizing the Mind*. Cambridge: MIT Press.

- Dretske, F. (2006). "Perception without Awareness". In: *Perceptual Experience*. p. 147-180. Gendler, T. and Hawthorne, J. (eds.). New York: Oxford University Press.
- Dulany, D. (2000). "Inattentional Awareness". *Psyche*. vol. 7(5): <http://psyche.cs.monash.edu.au/v7/psyche-7-05-dulany.html>
- Ebdon, M. (1993). "Is the Cerebral Neocortex a Uniform Cognitive Architecture?". *Mind and Language*. vol. 8(3): 368-403.
- Egan, A. (2008). "Seeing and Believing: Perception, Belief Formation, and the Divided Mind". *Philosophical Studies*. vol.140(1): 47-63.
- Egan, F. (1995). "Computation and Content". *Philosophical Review*. vol. 104: 181-203.
- Eliasmith, C. (2003). "Moving Beyond Metaphors: Understanding the Mind for What It Is". *The Journal of Philosophy*. vol. 100(10): 493-520.
- Ellis and Young (1990). "Accounting for Delusional Misidentification" *British Journal of Psychiatry*. vol. 157: 239-248.
- Erdelyi, M. (1986). "Experimental Indeterminacies in the Dissociation Paradigm". *Behavioral and Brain Sciences*. vol. 9: 30-31.
- Ericsson, K. and Simon, H. (1980). "Verbal Reports as Data". *Psychological Review*. vol. 87(3): 215-251.
- Evans, G.. (1982). *The Varieties of Reference*. New York: Oxford University Press.
- Farah, M., and Feinberg, T. (2000). *Patient Based Approaches to Cognitive Neuroscience*. Cambridge: MIT Press.
- Feldman, R. (1988). "Having Evidence". In: *Philosophical Analysis: A Defense by Example*. p. 83-104. Austin, D. (ed.). Netherlands: Kluwer Academic Publishers.
- Finkelstein, D. (1999). "On The Distinction between Conscious and Unconscious States of Mind". *American Philosophical Quarterly*. vol. 36(2): 79-100.
- Flanagan, O. (1995). "Consciousness". *Oxford Companion to Philosophy*. p. 152-3. Honderich, T. (ed.). New York: Oxford University Press.
- Flew, A. (1956). "Motives and the Unconscious". *Minnesota Studies in the Philosophy of Science*. p. 155-173. Feigl, H. and Scriven, M. (eds.). Minneapolis: University of Minnesota Press.
- Fodor, J. (1975). *The Language of Thought*. New York: Thomas Y. Crowell.

- Fodor, J. (1983). *The Modularity of Mind*. Cambridge: MIT Press.
- Fodor, J. (1987). *Psychosemantics*. Cambridge: MIT Press.
- Fodor, J. (1990a). *A Theory of Content, and Other Essays*. Cambridge: MIT Press.
- Fodor, J. (1990b). "Information and Representation". In: *Information, Language, and Cognition*. p. 175-190. Hanson, P. (ed.). Vancouver: University of British Columbia Press.
- Fodor, J. (1994). *The Elm and the Expert*. Cambridge: MIT Press.
- Fodor, J. (2000). *The Mind Doesn't Work That Way*. Cambridge: MIT Press.
- Fodor, J. (2007). "Headaches Have Themselves". *London Review of Books*. May 24.
- Fodor, J. (2008). *LOT 2: The Language of Thought Revisited*. New York: Oxford University Press.
- Fodor, J. (2009). "Introduction: So What's So Good About Pylyshyn?". In: *Computation, Cognition, and Pylyshyn*. p. ix-xvii. Dedrick, D. and Trick, L. (eds.). Cambridge: MIT Press.
- Fodor, J. and Lepore, E. (1992). *Holism: A Shoppers Guide*. Oxford; Blackwell.
- Fodor, J. and Lepore, E. (1994). "What Is the Connection Principle?" *Philosophy and Phenomenological Research*. vol. 54(4): 837-845.
- Fodor, J. and Lepore, E. (2001). "Why Compositionality Won't Go Away: Reflections on Horwich's 'Deflationary' Theory". *Ratio*. vol. 14(4): 350-368.
- Fodor, J. and Pylyshyn, Z. (1981). "How Direct Is Visual Perception? Some Reflections on Gibson's 'Ecological Approach'". *Cognition*. vol. 9: 139-196.
- Fodor, J. and Pylyshyn, Z. (1988). "Connectionism and Cognitive Architecture: A Critical Analysis". *Cognition*. 28: 3-71.
- Fox, M. (1973). "On Unconscious Emotions". *Philosophy and Phenomenological Research*. vol. 34(2): 151-170.
- Freedman, D., Riesenhuber, M., Poggio, T., and Miller, E. (2001). "Categorical Representation of Visual Stimuli in the Primate Prefrontal Cortex". *Science*. vol. 291: 312-316.
- Freud, S. (1915). "The Unconscious", *S.E.*, 14. London: Imago Publishing Co.
- Field, H. (1978). "Mental Representation". *Erkenntnis*. vol. 13: 9-61.
- Frith, C. (1992). *The Cognitive Neuropsychology of Schizophrenia*. Hillsdale, NJ: Lawrence

Erlbaum Associates.

Frith, C., and Frith, U. (2005). "Theory of Mind". *Current Biology*. vol. 15(17): R644-645.

Frith, U. and Happe, F. (1999). "Theory of Mind and Self-Consciousness: What is it Like to be Autistic?". *Mind and Language*. vol. 14: 1-22.

Galin, D. (1994). "The Structure of Awareness: Contemporary Applications of William James' Forgotten Concept of the 'Fringe'". *Journal of Mind and Behavior*. vol. 15(4): 375-401.

Gallagher, S. (2005). *How the Body Shapes the Mind*. Oxford: Clarendon Press.

Gallistel, C. (1990). *The Organization of Learning*. Cambridge: MIT Press.

Gazzaniga, M., Ivry, R., and Mangun, G. (1998). *Cognitive Neuroscience: The Biology of the Mind*. New York: Norton and Company, Inc.

Gennaro, R. (1994). "Does Mentality Entail Consciousness?". *Philosophia*. vol. 24(3-4): 331-358.

Gennaro, R. (2006). "Between Pure Self-Referentialism and the Extrinsic HOT Theory of Consciousness". In: *Self-Representational Approaches to Consciousness*. p. 221-248. Kriegel, U. and Williford, K. (eds.). Cambridge: MIT Press.

Godfrey-Smith, P. (1992). "Induction and Adaptation". *Synthese*. vol. 92: 283-312.

Gold, I. and Stoljar, D. (1999). "A Neuron Doctrine in the Philosophy of Neuroscience". *Behavioral and Brain Sciences*. vol. 22: 809-869.

Goldberg, S. (2002). "Belief and its Linguistic Expression: Towards a Belief Box Account of First-Person Authority". *Philosophical Psychology*. vol. 15(1): 65-76.

Goldman, A. (1999). "Internalism Exposed". *The Journal of Philosophy*. vol. 6: 271-293.

Goodale, M. and Arbib, M. (1998). "The Cognitive Architecture of Vision and Action". In: *Constraining Cognitive Theories: Issues and Opinions*. p. 205-234. Pylyshyn, Z. (ed.). London: Ablex Publishing Corporation.

Gopnik, A. (1993). "Psychopsychology". *Consciousness and Cognition*. vol. 2: 264-280.

Grau, J. (2002). "Learning and Memory without a Brain". In: *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*. p. 77-88. Bekoff, M., Allen, C., Burghardt, G. (eds.). Cambridge: MIT Press.

Grice, P. (1957). "Meaning". *Philosophical Review*. vol. 66: 377-388.

- Hadley, R. (1995). "The Explicit-Implicit Distinction". *Minds and Machines*. vol. 5(2): 219-242.
- Hales, S. (2000). "The Problem of Intuition". *American Philosophical Quarterly*. vol. 37(2): 135-147.
- Halligan, P. and Oakley, D. (2000). "The Greatest Myth of All". *New Scientist*. Nov. 18: 34.
- Hamilton, W. (1865). *Lectures on Metaphysics*. vol. 1. Boston: Gould and Lincoln.
- Harnad, S. (1990). "The Symbol Grounding Problem". *Physica D*. vol. 42: 335-346.
- Harnad, S. (2005). "Cohabitation: Computation at Seventy, Cognition at Twenty" In: *Computation, Cognition, and Pylyshyn*. p. 245-258. Dedrick, D. and Trick, L. (eds.). Cambridge: MIT Press.
- Hanna, R. (2008). "Kantian Non-Conceptualism". *Philosophical Studies*. vol. 137: 41-64.
- Hanna, R. (unpublished). "Essentially Embodied Cognition and the Deep Consciousness Thesis". <http://philosophyk.files.wordpress.com/2008/11/hannaripoct08.pdf>
- Hassin, R., Uleman, J., Bargh, J. (2005). *The New Unconscious*. New York: Oxford University Press.
- Hatfield, G. (1988). "Representation and Content in Some (Actual) Theories of Perception". *Studies in the History and Philosophy of Science*. vol. 19(2): 175-214.
- Hatfield, G. (1992). "Color Perception and Neural Encoding: Does Metameric Matching Entail Loss of Information?". *Philosophy of Science Association*. vol. 1: 492-504.
- Hatfield, G. (2002). "Perception as Unconscious Inference". In: *Perception and the Physical World: Psychological and Philosophical Issue in Perception*. p. 115-143. Heyer, D. and Mausfeld, R. (eds.). New York: John Wiley & Sons, Ltd.
- Haugeland, J. (1991). "Representational Genera". In: *Philosophy and Connectionist Theory*. p. 61-89. Ramsey, W., Stich, S., and Rumelhart, D. (eds.). Lawrence Erlbaum Associates.
- Haugeland, J. (1998). "Mind Embodied and Embedded". In: *Having Thought: Essays in the Metaphysics of Mind*. p.207-237. Haugeland, J. (ed.). Cambridge: Harvard University Press.
- Heath, R., Carpenter, M., Mettler, F., and Kline, N. (1949). "Visual Apparatus: Visual Fields and Acuity, Color Vision, Autokinesis". In: *Selective Partial Ablation of the Frontal Cortex*. p. 489-491. Mettler, F. (ed.). New York: P. B. Hoeber.
- Heck, R. (2000). "Nonconceptual Content and the Space of Reasons". *Philosophical Review*. vol. 109: 483-523.

- Heil, J. (1989). "Minds Divided". *Mind*. vol. 98(392): 571-583.
- Hess, E. (1965). "Attitude and Pupil Size". *Scientific American*. vol. 212: 46-54.
- Hirst, W. (1995). "Cognitive Aspects of Consciousness" In: *The Cognitive Neurosciences*. p. 1307-1319. Gazzaniga, M. (ed.). Cambridge: MIT Press.
- Hobson, J. (2001). *The Dream Drugstore*. Cambridge: MIT Press.
- Hoffman, D. (1998). *Visual Intelligence*. Norton: London.
- Hofstadter, D. (1979). *Gödel, Escher, Bach: An Eternal Golden Braid*. New York: Basic Books.
- Hofstadter, D. (1985). "Waking up from the Boolean Dream, or, Subcognition as Computation" In: *Metamagical Themas*. p. 631-670. Hofstadter, D. (ed.). New York: Basic Books.
- Hofstadter, D. (1995). *Fluid Concepts and Creative Analogies*. New York: Basic Books.
- Horgan, T. and Tienson, J. (2002). "The Intentionality of Phenomenology and the Phenomenology of Intentionality". In: *Philosophy of Mind: Classics and Contemporary Readings*. p. 520-533. Chalmers, D. (ed.). London: Oxford University Press.
- Hornsby, J. (2000). "Personal and Sub-Personal". *Philosophical Explorations*. vol. 1: 6-24.
- Hull, C. (1934). "The Concept of the Habit-Family Hierarchy and Maze Learning: Part I". *Psychological Review*. vol. 41: 33-54.
- Hulse, D., Reed, C., and Schroeder, T. (2004). "The Impossibility of Conscious Desire". *American Philosophical Quarterly*. vol. 41(1): 73-80.
- Hunziker, H. (2006). *In the Eye of the Reader: Foveal and Peripheral Perception - From Letter Recognition to the Joy of Reading*. [In German: see [http://www.learning-systems.ch/multimedia/vis\\_e02.htm](http://www.learning-systems.ch/multimedia/vis_e02.htm) for English translation of relevant data]. Transmedia Zurich.
- Hunziker, M. (2006). "Intentionality Is Not Only a Mark of the Cognitive". In: *John Heil: Symposium on His Ontological Point of View*. p. 235-250. Esfeld, M. and Heil, J. (eds.). New Brunswick: Ontos.
- Hurley, S. (2003). "Animal Action in the Space of Reasons". *Mind and Language*. vol. 18(3): 231-256.
- Hurley, S. (1998). "Vehicles, Contents, Conceptual Structure, and Externalism". *Analysis*. vol. 58(1): 1-6.
- Hyslop, A. (2005). "Other Minds". *The Stanford Encyclopedia of Philosophy (Winter 2005 Edition)*. Edward N. Zalta (ed.). URL =

<<http://plato.stanford.edu/archives/win2005/entries/other-minds/>>.

Israel, D. and Perry, J. (1990). "What Is Information?" In: *Information, Language, and Cognition*. Hanson, P. (ed.). Vancouver: University of British Columbia Press.

Ito, T. and Cacioppo, J. (2000). "Electrophysiological Evidence of Implicit and Explicit Categorization Processes". *Journal of Experimental and Social Psychology*. vol. 36: 660-676.

Jackson, F. (1982). "Epiphenomenal Qualia". *Philosophical Quarterly*. vol. 32: 127-36.

Jackendoff, R. (1987). *Consciousness and the Computational Mind*. Cambridge: MIT.

Jackendoff, R. (1992). "Unconscious Information in Language and Psychodynamics". In: *Languages of Mind*. p. 83-98. Jackendoff (ed.). Cambridge: MIT Press.

Jackendoff, R. (1994). *Patterns in the Mind: Language and Human Nature*. New York: Basic Books.

Jacob, P. (1995). "Consciousness, Intentionality, and Function. What is the Right Order of Explanation?" *Philosophy and Phenomenological Research*. vol. 55(1): 195-200.

Jacob, P. (2003). "Consciousness and Intentionality". *The Stanford Encyclopedia of Philosophy (2003 Edition)*. Edward N. Zalta (ed.). URL = <<http://plato.stanford.edu/entries/consciousness-intentionality/>>.

Janet, P. (1886). "Les actes inconscients et le dedoublement de la personnalite pendant le somnambulisme provoque". *Revue Philosophique*. vol. 22(II): 577-792.

James, W. (1890). *Principles of Psychology*. vol. 1. New York: Dover Publications.

Jeannerod, M. (1997). *The Cognitive Neuroscience of Action*. Oxford: Blackwell.

Johnson-Laird, P. (1980). "Mental Models in Cognitive Science". *Cognitive Science*. vol. 4: 71-115.

Jopling, D. (1996). "Sub-Phenomenology". *Human Studies*. vol. 19: 153-173.

Jung, C. (1959). *The Archetypes and the Collective Unconscious*. New York: Princeton University Press.

Kanizsa, G. (1985). "Seeing and Thinking". *Acta Psychologica*. vol. 59(1): 23-33.

Kenny, A. (1984) "The Homoculus Fallacy". In: *Legacy of Wittgenstein*. p.125-136. Kenny (ed.). New York: Blackwell.

Kelly, S (2007). "Seeing Things in Merleau-Ponty". In: *The Cambridge Companion to Merleau-*

- Ponty*. p. 74-110. Carman, T. and Hansen, M. (eds.). Cambridge: Cambridge University Press.
- Kihlstrom, J. (1984). "Conscious, Subconscious, Unconscious: A Cognitive View". In: Bowers, K. and Meichenbaum, D. (eds.). *The Unconscious Reconsidered*. New York: Wiley-Interscience.
- Kihlstrom, J. (1987). "The Cognitive Unconscious". *Science*. vol. 237:1445-1452.
- Kihlstrom, J. (1990). "The Psychological Unconscious". In: *Handbook of Personality Theory and Research*. p. 445-464. Perrin, L. (ed). New York: Guilford Press.
- Kihlstrom, J. (2004). "Availability, Accessibility, and Subliminal Perception". *Consciousness and Cognition*. vol. 13: 92-100.
- Kihlstrom, J., Barnhardt, T., and Tataryn, D. (1992). "The Psychological Unconscious: Lost, Found, and Regained". *American Psychologist*. vol. 47(6): 778-791.
- Kihlstrom, J., Mulvaney, S., Tobias, B., Tobis, T. (2000). "The Emotional Unconscious". In: *Cognition and Emotion*. p. 30-86. Kihlstrom, J., Bower, G., Forgas, J., Niedenthal, P. (eds.). New York: Oxford University Press.
- Kinsbourne, M. (1988). "Integrated Cortical Field Model of Consciousness". In: *Consciousness in Contemporary Science*. p. 239-256. Marcel, A. and Bisiach, E. (eds.). New York: Oxford University Press.
- Kirsh, D. (1991). "When is Information Explicitly Represented". In: *Information, Language, and Cognition*. p. 340-365. Philippi, H. (ed.). New York: Oxford University Press.
- Kohut, H. (1971). *The Analysis of the Self*. New York: International Universities Press.
- Kriegel, U. (2003). "Is Intentionality Dependent upon Consciousness?". *Philosophical Studies*. vol. 116: 271-307.
- Kriegel, U. (2007). "The Phenomenologically Manifest". *Phenomenology and Cognitive Science*. vol. 6: 115-136.
- Kupperman, J. (1984). "Investigations of the Self". *Philosophy East and West*. vol. 34(1): 37-51.
- Kvavilashvili, L. and Mandler, G. (2004). "Out of One's Mind: A Study of Involuntary Semantic Memories". *Cognitive Psychology*. vol. 48: 47-94.
- Lacewing, M. (2007). "Do Unconscious Emotions Involve Unconscious Feelings?". *Philosophical Psychology*. vol. 20(1): 81-104.
- Lahav, R. (1993). "What Neuropsychology Tells Us About Consciousness". *Philosophy of Science*. vol. 60: 67-85.

- Lamme, V. (2003). "Why Visual Attention and Awareness are Different". *Trends in Cognitive Science*. vol. 7: 12-18.
- Lamme, V. (2006). "Towards a True Neural Stance on Consciousness". *Trends in Cognitive Sciences*. vol. 10(11): 494-501.
- Laycock, T. (1860). *Mind and Brain: or The Correlations of Consciousness and Organization*. London: Simpkin, Marschall.
- Le Doux, J. (1994). "Emotion, Memory and the Brain". *Scientific American*. vol. 270: 50-57.
- Le Doux, J. (1996). *The Emotional Brain*. New York: Touchstone.
- Lee, T., Liu, H., Tan, L., Chan, C., Mahankali, S., Feng, C., Hou, J., Fox, P., and Gao, J. (2002). "Lie Detection by Functional Magnetic Resonance Imaging". *Human Brain Mapping*. vol. 15: 157-164.
- Lee, J. and van Donkelaar, P. (2002). "Dorsal and Ventral Visual Stream Contributions to Perception-Action Interactions During Pointing". *Experimental Brain Research*. vol. 143: 440-446.
- Leibniz, G. (1704). *New Essays on Human Understanding*. Cambridge: Cambridge University Press.
- Leman, C. (1981). "Conscious and Unconscious Mental States". *Philosophy Research Archives*. vol. 13: 1-23.
- Lettvin, J., Maturana, H., McCulloch, W., and Pitts, W. (1959). "What the Frog's Eye Tells the Frog's Brain". *Proceedings of the Institute of Radio Engineers*. vol. 47: 1940-1951.
- Levine, J. (1983). "Materialism and Qualia: The Explanatory Gap". *Pacific Philosophical Quarterly*. vol. 64: 354-361.
- Levine, J. (2001). *Purple Haze: The Puzzle of Consciousness*. New York: Oxford University Press.
- Lewicki, P., Hill, T., and Czyzewska, M. (1992). "Nonconscious Acquisition of Information". *American Psychologist*. vol. 47: 796-801.
- Lewis, D. (1986). *In the Plurality of Worlds*. Oxford: Blackwell.
- Lind, R. (1986). "Does the Unconscious Undermine Phenomenology?". *Inquiry*. vol. 29: 325-344.
- Locke, J. (1689). *An Essay Concerning Human Understanding*. New York: Oxford University Press.

- Loftus, E. and Klinger, M. (1992). "Is the Unconscious Smart or Dumb?" *American Psychologist* vol. 47(6): 761-765.
- London, H. (2005). "Dendritic Computation". *Annual Review of Neuroscience*. vol. 28: 503-532.
- Lovibond, P. and Shanks, D. (2002). "The Role of Awareness in Pavlovian Conditioning: Empirical Evidence and Theoretical Implications". *Journal of Experimental Psychology*. vol. 28(1): 3-26.
- Ludwig, K. (1996). "Explaining Why Things Look the Way They Do". In: *Perception*. p. 18-60. Akins, K. (ed.). New York: Oxford University Press.
- Lycan, W. (1988). *Judgment and Justification*. Cambridge: Cambridge University Press.
- Lycan, W. (1998). "In Defense of the Representational Theory of Qualia" In: *Language, Mind, and Ontology*. p. 479-488. Tomberlin, J. (ed.). Atascadero, CA: Ridgeview Publishing.
- Lycan, W. (2002). "Explanation and Epistemology". In: *Oxford Handbook of Epistemology*. p. 408-434. Moser, P. (ed.). New York: Oxford University Press.
- Lyons, W. (1980). *Emotion*. Cambridge: Cambridge University Press.
- Lyons, J. (2001). "Carving the Mind at its (Not Necessarily Modular) Joints". *British Journal for the Philosophy of Science*. vol. 52: 277-392.
- Macphail, E. (1998). *The Evolution of Consciousness*. New York: Oxford University Press.
- Malach, R. (2007). "The Measurement Problem in Consciousness Research". *Behavioral and Brain Sciences*. vol. 30: 516-517.
- Malpas, J. (1992). "The Intentional-Horizontal Structure of the Psychological". In: *Donald Davidson and the Mirror of Meaning*. p. 114-135. Malpas, J. New York: Cambridge University Press.
- Mandik, P. (2001). "Mental Representation and the Subjectivity of Consciousness". *Philosophical Psychology*. vol. 14: 179-202.
- Mandler, G. (1994). "Hypermnesia, Incubation, and Mind Popping: On Remembering without Really Trying". In: *Attention and Performance XV: Conscious and Unconscious Information Processing*. p. 1-33. Umiltà, C. and Moscovitch, M. (eds.). Cambridge: MIT Press.
- Manson, N. (2000a). "'A Tumbling Ground of Whimsy?' The History and Contemporary Role of the Conscious/Unconscious Contrast". In: *The History of the Mind-Body Problem*. p. 148-168. Crane, T. and Patterson, S. (eds.). New York: Routledge.

- Manson, N. (2000b). "State Consciousness and Creature Consciousness: A Real Distinction". *Philosophical Psychology*. vol. 13(3): 405-410.
- Mangan, B. (1993). "Taking Phenomenology Seriously: The 'Fringe' and Its Implications for Cognitive Research". *Consciousness and Cognition*. vol. 2: 89-108.
- Massaro, D. and Cowan, N. (1993). "Information Processing Models: Microscopes of the Mind". *Annual Review of Psychology*. vol. 44: 383-425.
- Mathews, R. (1984). "Troubles with Representationalism". *Social Research*. vol. 51: 1065-1097.
- Marcel, A. (1993). "Slippage in the Unity of Consciousness". *Experimental and Theoretical Studies of Consciousness*. p. 168-186. Ciba Foundation Symposium.
- Marcus, G. (2001): *The Algebraic Mind: Integrating Connectionism and Cognitive Science*. Cambridge: MIT Press.
- Margetts, E. (1953). "Concept of the Unconscious in the History of Medical Psychology". *Psychiatric Quarterly*. vol. 27: 115-138.
- Markman, A. and Dietrich, E. (1998). "In Defense of Representation as Mediation". *Psychology*. vol. 9(48). <http://www.cogsci.ecs.soton.ac.uk/cgi/psyc/newpsy?9.48>
- Marr, D. (1982). *Vision*. San Francisco: W.H. Freeman.
- Marras, A. (1973). "Sellars on Thought and Language". *Nous*. vol. 7(2): 152-163.
- Marshall, J. (1984). "Multiple Perspectives on Modularity". *Cognition*. vol. 17(3): 209-242.
- Martin, M. (1992). "Perception, Concepts, and Memory". *Philosophical Review*. vol. 101(4): 745-63.
- Mayr, U. (2004). "Conflict, Consciousness, and Control". *Trends in Cognitive Sciences*. vol. 8(4): 145-148.
- McCarthy, J. (1979). "Ascribing Mental Qualities to Machines". In: *Philosophical Perspectives in Artificial Intelligence*. p. 93-118. Ringle, M. (ed). New Jersey: Humanities Press.
- McBride, R. (1999). "Consciousness and the State/Transitive/Creature Distinction". *Philosophical Psychology*. vol. 12(2): 181-196.
- McClosky, M. (1954). "Report On Analysis Problem No. 4: 'If a Distraction Makes Me Forget My Headache, Does It Make My Head Stop Aching, or Does It Only Stop Me Feeling It Aching?'". *Analysis*. vol. 14(2): 53-54.
- McConkie, G. and Rayner, K. (1973). "The Span of the Effective Stimulus during Fixations in

Reading”. Paper Presented at the Annual Meeting of the American Educational Research Association. New Orleans: Feb. 26-Mar. 1.

McDowell, J. (1994a). *Mind and World*. Cambridge: Harvard University Press.

McDowell, J. (1994b). “The Content of Perceptual Experience”. *Philosophical Quarterly*. vol. 44(175): 190-205.

McGurk, H. and McDonald, J. (1976). “Hearing Lips and Seeing Voices”. *Nature*. vol. 264: 746-748.

McGinn, C. (1988). “Consciousness and Content”. *Proceedings of the British Academy*. vol. 74: 225-245.

McKee, P. (1972). “Non-Conscious Seeing”. *American Philosophical Quarterly*. vol. 9(4): 319-326.

McKinsey, M. (1979). “Expressing Mental States”. *Philosophia*. vol. 8(4): 657-671.

McLaughlin, B. and Tye, M. (1998). “Is Content Externalism Compatible With Privileged Access?”. *The Philosophical Review*. vol. 107(3): 349-380.

McLoughlin, J. (1999). “Unwittingly Recapitulating Freud: Searle’s Concept of a Vocabulary of the Unconscious”. *Ratio*. vol. 12(1): 34-53.

Meixner, U. (2006). “Classical Intentionality”. *Erkenntnis*. vol. 65: 25-45.

Mel, B. (1994). “Information Processing in Dendritic Trees”. *Neural Computation*. vol. 6: 1031-1085.

Merikle, P. (2000). “Subliminal Perception”. In: *Encyclopedia of Psychology*. vol.7: 47-49.

Merikle, P. and Reingold, E. (1998). “On Demonstrating Unconscious Perception: Comment on Draine and Greenwald”. *Journal of Experimental Psychology: General*. vol. 127: 304-310.

Metzinger, T. (2000). “The *Subjectivity* of Subjective Experience: A Representationalist Analysis of the First-Person Perspective”. In: *Neural correlates of Consciousness: Empirical and Conceptual Issues*. Cambridge: MIT Press.

Miller, G. (1962). *Psychology: The Science of Mental Life*. New York: Harper and Row.

Millikan, R. (1989). “Biosemantics”. *The Journal of Philosophy*. vol. 86: 281-297.

Milner, D. and Goodale, M. (1995). *Visual Brain in Action*. New York: Oxford University Press.

Minsky, M (1981). “Jokes and the Logic of the Cognitive Unconscious”. In *Cognitive*

*Constraints on Communication*. p. 175-200. Vaina, L. and Hintikka, J. (eds.). Boston: Kluwer Academic Press.

Mitroff, S., Simons, D., and Franconeri, S. (2002). "The Siren Song of Implicit Change Detection". *Journal of Experimental Psychology: Human Perception and Performance*. vol. 28(4): 798-815.

Mollon, J. (1997). "... On the Basis of Velocity Clues Alone': Some Perceptual Themes 1946-1996". *Quarterly Journal of Experimental Psychology*. vol. 50A(4): 859-878.

Moradi F., Koch C., Shimojo S. (2005). "Face Adaptation Depends on Seeing the Face". *Neuron* vol. 45(1): 169-75.

Moran, R. (1997). "Self-Knowledge: Discovery, Resolution, and Undoing". *European Journal of Philosophy*. vol. 5(2): 141-161.

Moran, R. (2001). *Authority and Estrangement*. Princeton: Princeton University Press.

Mullane, H. (1965). "Unconscious Emotion". *Theoria*. vol. 31: 181-190.

Mumford, S. (1999). "Intentionality and the Physical: A New Theory of Disposition Ascription". *Philosophical Quarterly*. vol. 49(195): 215-224.

Naccache, L. (2006). "Is She Conscious?". *Science*. vol. 313: 1395-1396.

Nagel, T. (1974). "What is it Like to be a Bat?". *The Philosophical Review*. vol. 83(4): 435-50.

Nakayama K., He Z., and Shimojo S. (1995). "Visual Surface Representation: A Critical Link Between Lower-Level and Higher-Level Vision". In: *Visual Cognition*. p 1-20. Kosslyn, S. and Osherson, D. (eds.). Cambridge: MIT Press.

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

Newell, A. (1982). "The Knowledge Level". *Artificial Intelligence*. vol.18(1): 87-127.

Nisbett, R. and Wilson, T. (1977). "Telling More Than We Can Know". *Psychological Review*. vol. 84(3): 231-259.

Norman, K. and O'Reilly, R. (2003). "Modeling Hippocampal and Neocortical Contributions to Recognition Memory: A Complementary Learning Systems Approach". *Psychological Review*. vol. 110: 611-646.

Nørretranders, T. (1998). *The User Illusion: Cutting Consciousness Down to Size*. New York: Viking.

- O'Brien, G and Opie, J. (1998). "The Disunity of Consciousness". *Australasian Journal of Philosophy*. vol. 76(3): 378-395.
- O'Regan, J. and Nöe, A. (2001). "A Sensorimotor Account of Vision and Visual Consciousness". *Behavioral and Brain Sciences*. vol. 24(5): 939-1031.
- O'Reilly, R. (2006). "Biologically Based Computational Models of High-Level Cognition". *Science*. vol. 314: 91-94.
- O'Reilly, R. and Munakata, Y. (2000). *Computational Explorations in Cognitive Neuroscience*. Cambridge: MIT Press.
- O'Shaughnessy, B. (1992). "The Diversity and Unity of Action and Perception". In *The Contents of Experience*. p. 216-66. Crane, T. (ed.). Cambridge: Cambridge University Press.
- Owings, D. (2002). "The Cognitive Defender: How Ground Squirrels Assess Their Predators". In: *The Cognitive Animal: Empirical and Theoretical Perspectives on Animal Cognition*. p. 19-26. Bekoff, M., Allen, C., and Burghardt, G. (eds.). Cambridge: MIT Press.
- Pacherie, E. (2000). "Levels of Perceptual Content". *Philosophical Studies*. vol. 100: 237-254.
- Palmer, D. (1975). "Unfelt Pains". *American Philosophical Quarterly*. vol. 12 (4): 289-298.
- Palmer, S. (1999). *Vision Science: Photons to Phenomenology*. Cambridge: MIT Press.
- Pantev, C., Okamoto, H., Ross, B., Stoll, W., Ciurlia-Guy, E., Kakigi, R., Kubo, T. (2004). "Lateral Inhibition and Habituation of the Human Auditory Cortex". *European Journal of Neuroscience*. vol. 19(8): 2337-2344.
- Papineau, D. (1996). "Doubtful Intuitions". *Mind and Language*. vol. 11: 130-132.
- Pascal, B. (1670) *Pensees* (iv). Nabu Press.
- Pashler, H. (1998). *Psychology of Attention*. Cambridge: MIT Press.
- Payne, J., Samper, A., Bettman, J., Luce, M. (2008). "Boundary Conditions on Unconscious Thought in Complex Decision Making". *Psychological Science*. vol. 19: 1118-1123.
- Peacocke, C. (1992). *A Study of Concepts*. MIT, press.
- Peacocke, C. (1994a). "Content, Computation, and Externalism". *Mind and Language*. vol. 9(3): 303-335.
- Peacocke, C. (1994b). "Nonconceptual Content: Kinds, Content, and Relations". *Mind and Language*. vol. 9(4): 419-430.

- Peacocke, C. (1998). "Nonconceptual Content Defended". *Philosophy and Phenomenological Research*. vol. 58(2): 381-388.
- Peacocke, C. (1999). "Computation as Involving Content: A Response to Egan". *Mind and Language*. vol. 14: 195-202.
- Pessoa, L., Thompson, E., and Nöe, A. (1998). "Finding out about Filling-in: A Guide to Perceptual Completion for Visual Science and the Philosophy of Perception". *Behavioral and Brain Sciences*. vol. 21(6): 723-748.
- Piaget, J. (1973). "The Affective Unconscious and the Cognitive Unconscious". *Journal of the American Psychoanalytic Association*. vol. 21: 249-261.
- Piattelli-Palmarini, M. (1994). *Inevitable Illusions: How Mistakes of Reason Rule Our Minds*. New York: John Wiley and Sons.
- Pinker, S. (1999). *How the Mind Works*. New York: Norton.
- Pitt, D. (2004). "What is it Like to Think that P?". *Philosophy and Phenomenological Research*. vol. 69(1): 1-36.
- Place, U. T. (1996). "Intentionality as the Mark of the Dispositional". *Dialectica*. vol. 50: 91-120.
- Place, U.T. (2000). "Consciousness and the Zombie Within" In: *Beyond Dissociation: Interactions between Dissociated Implicit and Explicit Processing*. Rossetti, Y. and Revonsuo, A. (eds.) p. 295-329. Philadelphia: J. Benjamins.
- Pollen, D. (2008). "Fundamental Requirements for Primary Visual Perception". *Cerebral Cortex*. vol. 18: 1991-1998.
- Posner, M. (1994). "Attention: The Mechanisms of Consciousness". *Proceedings of the National Academy of Sciences*. vol. 91(16): 7398-7403.
- Power, M. and Brewin, C. (1991). "From Freud to Cognitive Science: A Contemporary Account of the Unconscious". *British Journal of Clinical Psychology*. vol. 30(4): 289-310.
- Prinz, J. (2000). "A Neurofunctional Theory of Visual Consciousness". *Consciousness and Cognition*. vol. 9: 243-259.
- Pylyshyn, Z. (1980). "Computation and Cognition: Issues in the Foundations of Cognitive Science". *Behavioral and Brain Sciences*. vol. 3: 111-169.
- Pylyshyn, Z. (1984). *Computation and Cognition*. Cambridge: MIT Press.
- Pylyshyn, Z. (1987). *Robot's Dilemma: The Frame Problem in Artificial Intelligence*. New

Jersey: Ablex Publishing.

Pylyshyn, Z. (2001). "Visual Indexes, Preconceptual Objects, and Situated Vision". *Cognition*. vol. 80: 127-158.

Pylyshyn, Z. (2003). *Seeing and Visualizing: It's Not What You Think*. Cambridge: MIT Press.

Pylyshyn, Z. (2009). "Perception, Representation, and the World: The FINST That Binds". In: *Computation, Cognition, and Pylyshyn*. p. 3-48. Dedrick, D. and Trick, L. (eds.). Cambridge: MIT Press.

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

Ratliff, F. (1965). *Mach Bands: Quantitative Studies on Neural Networks in the Retina*. San Francisco: Holden Day.

Rauhala, L. (1969). *Intentionality and the Problem of the Unconscious*. Turku: Turun Yliopisto.

Rayner, K. and Bertera, J. (1979). "Reading Without a Fovea". *Science*. vol. 206:468-469.

Reber, A. (1992). "The Cognitive Unconscious: An Evolutionary Perspective". *Consciousness and Cognition*. vol. 1: 93-133.

Reber, A. (1967). "Implicit Learning of Artificial Grammars". *Journal of Verbal Learning and Verbal Behavior*. vol. 6: 855-863.

Reber, R., Wurtz, P., Zimmermann, T. (2004). "Exploring 'Fringe' Consciousness: The Subjective Experience of Perceptual Fluency and its Objective Bases". *Consciousness and Cognition*. vol. 13: 47-60.

Rensink, R. (2000). "When Good Observers Go Bad: Change Blindness, Inattentional Blindness, and Visual Experience". *Psyche*. vol. 6(09): <http://psyche.cs.monash.edu.au/v6/psyche-6-09-rensink.html>

Rescorla, R. (1988). "Pavlovian Conditioning: It's Not What You Think It Is". *American Psychologist*. vol. 43: 151-160.

Revonsuo, A. (1999). "Binding and the Phenomenal Unity of Consciousness". *Consciousness and Cognition*. vol. 8:173-185.

Rey, G. (1988). "Toward a Computational Account of *AKRASIA* and Self-Deception". In: *Perspectives on Self-Deception*. p. 264-296. McLaughlin, B., and Rorty, A. (eds.). Berkeley: University of California Press.

Rey, G. (1997). *Contemporary Philosophy of Mind: A Contentiously Classical Approach*. Cambridge: Blackwell.

- Ricoeur, P. (1974). *Conflict of Interpretation*. Evanston: Northwestern University Press.
- Rock, I. (1983). *Logic of Perception*. Cambridge: MIT Press.
- Rorty, R. (1979). *Philosophy and the Mirror of Nature*. Princeton: Princeton University Press.
- Rorty, R. (1991). "Freud and Moral Reflection". In: *Essays on Heidegger and Others*. p. 143-163. Rorty, R. (ed.). New York: Cambridge University Press.
- Rosenblueth, A., Wiener, N., and Bigelow, J. (1943). "Behavior, Purpose, and Teleology". *Philosophy of Science*. vol. 10: 18-24.
- Rosenthal, D. (1993). "Higher-Order Thoughts and the Appendage Theory of Consciousness". *Philosophical Psychology*. vol. 6: 155-166.
- Rosenthal, D. (2002). "How Many Kinds of Consciousness?". *Consciousness and Cognition*. 11(4): 653-665.
- Rosenthal, D. (2005). "Why Are Verbally Expressed Thoughts Conscious?". In: *Consciousness and Mind*. p. 282-305. Rosenthal, D. (ed.). Oxford: Clarendon Press.
- Rovane, C. (2005). "Not Mind-Body but Mind-Mind?". *Journal of Consciousness Studies*. vol. 7: 82-92.
- Rozin, P. (1976). "The Evolution of Intelligence and Access to the Cognitive Unconscious". *Progress in Psychobiology and Physiological Psychology*. vol 6: 245-280.
- Rumelhart, D., McClelland, J., and The PDP Research Group. (1987). *Parallel Distributed Processing*. Cambridge: MIT Press.
- Rupert, R. (1999). "The Best Test Theory of Extension: First Principle(s)". *Mind and Language*. vol. 14: 321-355.
- Rupert, R. (2008). "Causal Theories of Mental Content". *Philosophy Compass*. vol. 3: 353-380.
- Rupert, R. (2009). *Cognitive Systems and the Extended Mind*. New York: Oxford University Press.
- Ryle, G. (1954). "Report On Analysis Problem No. 4: 'If a Distraction Makes Me Forget My Headache, Does It Make My Head Stop Aching, or Does It Only Stop Me Feeling It Aching?'". *Analysis*. vol. 14(3): 51-52.
- Samuels, R. (1998). "Evolutionary Psychology and the Massive Modularity Hypothesis". *British Journal for the Philosophy of Science*. vol. 49: 575-602.

Schwarz, N., Bless, H., Strack, F., Klump, G., Rittenauer-Schatka, H., Simons, A. (1991). "Ease of Retrieval as Information: Another Look at the Availability Heuristic". *Journal of Personality and Social Psychology*. vol. 61: 195-202.

Schwitzgebel, E. (2004). "Belief". In: *The Stanford Encyclopedia of Philosophy*, Zalta, E. (ed.). URL = <<http://plato.stanford.edu/archives/fall2006/entries/belief/>>

Searle, J. (1990). "Consciousness, Explanatory Inversion, and Cognitive Science". *Behavioral and Brain Sciences*. vol. 14: 585-642.

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

Searle, J. (1994). "The Connection Principle and the Ontology of the Unconscious". *Philosophy and Phenomenological Research*. vol. 54(4): 847-855.

Searle, J. (1997). "The Explanation of Cognition". In: *Thought and Language*. p. 103-126. Preston, J. (ed.). Cambridge: Cambridge University Press.

Sears, C. and Pylyshyn, Z. (2000). "Multiple Object Tracking and Attentional Processing". *Canadian Journal of Experimental Psychology*. vol. 54: 1-14.

Segal, G. (1989). "Seeing What Is Not There". *The Philosophical Review*. vol. 98: 189-214.

Sellars, W. (1964). "Notes on Intentionality". *The Journal of Philosophy*. vol. 61: 655-665.

Sellars, W. (1969). "Language as Thought and Communication". *Philosophy and Phenomenological Research*. vol. 29: 506-527.

Sergent, C. and Dehaene, S. (2004). "Is Consciousness a Gradual Phenomenon? Evidence for an All-or-None Bifurcation during the Attentional Blink". *Psychological Science*. vol. 15: 720-728.

Shallice, T. (1984). "More Functionally Isolable Systems but Fewer 'Modules'". *Cognition*. vol. 17: 243-252.

Shallice, T. (1988). *From Neuropsychology to Mental Structure*. Cambridge: Cambridge University Press.

Shani, I. (2006) "Narcissistic Sensations and Intentional Directedness: How Second-Order Cybernetics Helps Dissolve the Tension between the Egocentric Character of Sensory Information and the (seemingly) World-Centered Character of Cognitive Representations". *Cybernetics and Human Knowing*. vol. 13: 87-110.

Shapiro, L. (1993). "Content, Kinds, and Individualism in Marr's Theory of Vision". *The Philosophical Review*. vol. 102(4): 489-513.

Shapiro, L. (1994). "What is Psychophysics?" *Philosophy of Science Association* vol. 2: 47-57.

Shapiro, A., Lu, Z., Knight, E., and Ennis, R. (2009). "The Breaking Curveball". *American*. vol. 60(2): 11.

Shevrin, H. (1992). "The Freudian Unconscious and the Cognitive Unconscious: Identical or Fraternal Twins?". In: *Interface of Psychoanalysis and Psychology*. p. 313-326. Barron, J., Eagle, M., and Wolitzky, D. (eds.). Washington, DC: American Psychological Association.

Shevrin, H. and Dickman, S. (1980). "The Psychological Unconscious: A Necessary Assumption for All Psychological Theory?" *American Psychologist*. vol. 35(5): 421-434.

Sievert, D. (1980). "Sellars and Descartes on the Fundamental Form of the Mental" *Philosophical Studies*. vol. 37: 251-257.

Simmons, A. (2001) "Changing the Cartesian Mind: Leibniz on Sensation, Representation, and Consciousness". *Philosophical Review*. vol. 110(1): 31-73.

Slater, C. (1994). "Discrimination Without Indication: Why Dretske Can't Lean on Learning". *Mind and Language*. vol. 9: 163-180.

Sligte, I., Scholte, H., Lamme V. (2008). "Are There Multiple Visual Short-Term Memory Stores?". *PLoS ONE*. vol. 3(2): 1-9 [e1699].

Smith, A. (1978). "Unconsciousness and Quasiconsciousness in Plotinus". *Phronesis*. vol. 23: 292-301.

Smith, D. (1999). *Freud's Philosophy of the Unconscious*. London: Kluwer Academic Publishers.

Snodgrass, M. (2004). "Dissociation Paradigm and its Discontents: How Can Unconscious Perception or Memory be Inferred? ". *Consciousness and Cognition*. vol. 13(1): 107-116.

Solomon, R. (1984). "I Can't Get It Out of My Mind': (Augustine's Problem)". *Philosophy and Phenomenological Research*. vol. 44(3): 405-412.

Speaks, J. (2006). "Is Mental Content Prior to Linguistic Meaning". *Nous*. vol. 40: 428-467.

Sperling, G. (1960). "The Information Available in Brief Visual Presentations". *Psychological Monographs: General and Applied*. vol. 74(11): 1-29.

Spivey, M. and Dale, R. (2006). "Continuous Temporal Dynamics in Real-Time Cognition". *Current Directions in Psychological Science*. vol. 15: 207-211.

Spohn, W. (2002). "The Many Facets of the Theory of Rationality". *Croatian Journal of Philosophy*. vol. 2(6): 249-264.

- Stabler, E. (1983). "How Grammars are Represented". *Behavioral and Brain Science*. vol. 6: 391-264.
- Sterelny, K. (1990). *The Representational Theory of Mind*. Cambridge: Basil Blackwell.
- Sterelny, K. (1995). "Basic Minds". *Philosophical Perspectives*. vol. 9: 251-270.
- Sterelny, K. (2003). "Charting Control-Space: Comments on Susan Hurley's 'Animal Action in the Space of Reasons'". *Mind and Language*. vol. 18(3): 257-265.
- Sterelny, K. (2006). "Folk Logic and Animal Rationality". In: *Rational Animals?* p. 293-312. Hurley, S. and Nudds, M. (eds.). New York: Oxford University Press.
- Stich, S. (1978). "Beliefs and Subdoxastic States". *Philosophy of Science* vol. 45: 499-518.
- Stich, S. (1983). *From Folk Psychology to Cognitive Science*. Cambridge: MIT Press.
- Stich, S. (1990). *The Fragmentation of Reason*. Cambridge: MIT Press.
- Strawson, G. (1994). *Mental Reality*. Cambridge: MIT Press.
- Strawson, G. (2003). "Mental Ballistics or the Involuntariness of Spontaneity". *Proceedings of the Aristotelian Society*. vol. 103(3): 227-256.
- Strawson, G. (2006). *Consciousness and its Place in Nature: Does Physicalism Entail Panpsychism*. Book News, Inc; Michigan.
- Sutherland, R. (2005). "Cognitive Processes". In: *The Behavior of the Lab Rat: A Handbook with Tests*. p. 422-435. Whishaw, I. and Kolb, B. (eds.). New York: Oxford University Press.
- Schwartz, N., Bless, H., Strack, F., Klumpp, G., Rittenauer-Schatka, H., and Simmons, A. (1991). "Ease of Retrieval as Information: Another Look at the Availability Heuristic". *Journal of Personality and Social Psychology*. vol. 61: 195-202.
- Swinburne, R. (1985). "Thought". *Synthese*. vol. 48: 153-171.
- Thomasson, A. (2000). "After Brentano: A One-Level Theory of Consciousness". *European Journal of Philosophy*. vol. 8(2): 190-209.
- Tolman, E. (1948). "Cognitive Maps in Rats and Men". *Psychological Review*. vol. 55: 189-208.
- Tombs, S. and Silverman, I. (2009). "Pupillometry: A Sexual Selection Approach". *Evolution and Human Behavior*. vol. 25: 221-228.
- Tononi, G. and Edelman, G. (1998). "Consciousness and Complexity". *Science*. vol. 282(5395): 1846 – 1851.

- Treisman, A. (1985). "Pre-Attentive Processing in Vision". *Graphics and Image Processing*. vol. 31(2): 156-177.
- Tulving, E. and Pearlstone, Z. (1966). "Availability versus Accessibility of Information of Memory for Words". *Journal of Verbal Learning and Verbal Behavior*. vol. 5: 381-391.
- Tye, M. (2008). "Blurry Images, Double Vision, and Other Oddities: New Problems for Representationalism?". In: *Mind and Cognition: An Anthology, 3<sup>rd</sup> Edition*. p. 707-724. Lycan, W. and Prinz, J. (eds.). Oxford: Wiley-Blackwell.
- Uttal, W. (1967). "Evoked Brain Potentials: Signs or Codes?" *Perspectives in Biology and Medicine*. vol. 10: 627-639.
- Vaina, L. (1990). "'What' and 'Where' in the Human Visual System: Two Hierarchies of Visual Modules". *Synthese*. vol. 83: 49-91.
- Van Gulick, R. (1995). "Why the Connection Argument Doesn't Work". *Philosophy and Phenomenological Research*. vol. 55(1): 201-207.
- Van Dusen, W. (1960). "Existential Analytic Psychotherapy". *American Journal of Psychoanalysis*. vol. 20(1): 35-40.
- Wakefield, J. (1991). "Why Emotions Can't Be Unconscious: An Exploration of Freud's Essentialism". *Psychoanalysis & Contemporary Thought*. Vol. 14(1): 29-67.
- Wakefield, J. (2001). "Do Unconscious Mental States Exist? Freud, Searle, and the Conceptual Foundations of Cognitive Science". Ph.D. Dissertation. University of California, Berkley.
- Wallhagen, M. (2007). "Consciousness and Action: Does Cognitive Science Support (Mild) Epiphenomenalism?". *British Journal of Philosophical Studies*. vol. 58: 539-561.
- Waltz, D. (1975). "Understanding Line Drawings of Scenes with Shadows". In: *The Psychology of Computer Vision*. p. 19-91. Winston, P. (ed.). New York: McGraw-Hill.
- Warmbrod, K. (1992). "Primitive Representation and Misrepresentation". *Topoi*. vol. 11: 89-101.
- Wegner, D. (1989). *White Bears and Other Unwanted Thoughts: Suppression, Obsession, and the Psychology of Mental Control*. New York: Guilford Press.
- Weiskrantz, L. (1997). *Consciousness Lost and Found: A Neuropsychological Exploration*. New York: Oxford University Press.
- Wheeldon, J. (1954). "Report On Analysis Problem No. 4: 'If a Distraction Makes Me Forget My Headache, Does It Make My Head Stop Aching, or Does It Only Stop Me Feeling It

Aching?”. *Analysis*. vol. 14(2): 55-56.

Wehr, M. and Zador, A. (2003). “Balanced Inhibition Underlies Tuning and Sharpens Spike Timing in Auditory Cortex”. *Nature*. vol. 426: 442-446.

Whyte, L. (1960). *The Unconscious Before Freud*. New York: Basic Books.

Williamson, J. (1966). “Realization and the Unconscious”. *Philosophy and Phenomenological Research*. vol. 27(1): 11-26.

Wilson, T. (2002). *Strangers to Ourselves*. Cambridge: Belknap Press of Harvard University Press.

Wilson, T. and Stone, J. (1985). “Limitations of Self-Knowledge. More on Telling More than we can Know”. In: *Self Situations and Social Behavior. Review of Personality and Social Psychology*. Shaver (ed.). vol. 6: 167-183.

Wolfe, J. (1999). “Inattentive Amnesia”. In: *Fleeting Memories*. p. 71-94. Coltheart, V. (ed.). Cambridge: MIT Press.

Wright, W. (2005). “Distracted Drivers and Unattended Experience”. *Synthese*. vol. 114: 41-68.

Yablo, S. (1992). “Mental Causation”. *Philosophical Review*. vol. 101: 245-280.

Zajonc, R. (1980). “Feeling and Thinking: Preferences Need No Inferences”. *American Psychologist*. vol. 35: 151-175.

Zajonc, R. (2001). “Mere Exposure: A Gateway to the Subliminal”. *Current Directions in Psychological Science*. vol. 10: 224-228.

Zeki, S. (2003). “The Disunity of Consciousness”. *Trends in Cognitive Science*. vol. 7: 214-218.

Zeki, S. (2007). “A Theory of Micro-Consciousness”. In: *The Blackwell Companion to Consciousness*. p. 580-588. Velmans, M. and Schnieder, S. (eds.). Oxford: Blackwell.

Zemach, E. (1987). “Unconscious Mind or Conscious Minds?”. *Midwest Studies in Philosophy*. vol. 10: 121-149.

Zimmerman, M. (1986). “Neurophysiology of Sensory Systems”. In: *Fundamentals of Sensory Physiology*. p. 31-80. Schmidt, R. (ed.). Berlin: Springer-Verlag.

Zimmerman, M. (1989). “The Nervous System in the Context of Information Theory”. In: *Human Physiology*. p. 166-173. Schmidt, R. and Thews, G. (eds.). Berlin: Springer-Verlag.