Forthcoming in special issue of *Consciousness and Cognition* on "Social Perception"
Available online: http://dx.doi.org/10.1016/j.concog.2015.04.001

Direct Social Perception and Dual Process Theories of Mindreading

Mitchell Herschbach
Department of Philosophy
California State University, Northridge
18111 Nordhoff Street; ST-522
Northridge, CA 91330-8253
mitchell.herschbach@csun.edu

Abstract:
The direct social perception (DSP) thesis claims that we can directly perceive some mental states of other people. The direct perception of mental states has been formulated phenomenologically and psychologically, and typically restricted to the mental state types of intentions and emotions. I will compare DSP to another account of mindreading: dual process accounts that posit a fast, automatic "Type 1" form of mindreading and a slow, effortful "Type 2" form. I will here analyze whether dual process accounts' Type 1 mindreading serves as a rival to DSP or whether some Type 1 mindreading can be perceptual. I will focus on Apperly and Butterfill's dual process account of mindreading epistemic states such as perception, knowledge, and belief. This account posits a minimal form of Type 1 mindreading of belief-like states called registrations. I will argue that general dual process theories fit well with a modular view of perception that is considered a kind of Type 1 process. I will show that this modular view of perception challenges and has significant advantages over DSP's phenomenological and psychological theses. Finally, I will argue that if such a modular view of perception is accepted, there is significant reason for thinking Type 1 mindreading of belief-like states is perceptual in nature. This would mean extending the scope of DSP to at least one type of epistemic state.

Keywords:
direct social perception, mindreading, dual process theory, dual systems theory, perception, modularity, belief-like states

Direct Social Perception and Dual Process Theories of Mindreading

1. Introduction

Mindreading is the ability to understand and respond to the mental states of other agents. The nature of the cognitive processes enabling mindreading has been an important research question for decades, but recent research has focused on providing a more fine-grained analysis of the phenomena and mechanisms of mindreading. This includes characterizing the phenomenological experience of mindreading, delineating the various environmental contexts in which we mindread and the task demands of those different contexts, and determining the representational, architectural and processing characteristics of our mindreading mechanism(s).

The *direct social perception* (DSP) thesis is one recent attempt to recharacterize the phenomena of mindreading (e.g., Bohl & Gangopadhyay, 2014; Gallagher, 2008a,b; Gallagher & Varga, 2014; Gangopadhyay & Miyahara, in press; Krueger, 2012; Krueger & Overgaard, 2012; McNeill, 2012; Smith, 2010, in press; Zahavi, 2007, 2008, 2011). DSP rejects the idea that most mindreading involves an inferential process from perceptual information about a person's bodily and verbal behavior to a cognitive representation of the mental states underlying that behavior. While admitting that we do sometimes make these sorts of inferences from behavior to mental states, DSP claims that there is also a type of mindreading involving the *direct perception* of others' mental states. DSP advocates have generally focused on intentions and emotions as the mental state types capable of being directly perceived. The following passage by Max Scheler captures this position about emotion:

> For we certainly believe ourselves to be directly acquainted with another person's joy in his laughter, with his sorrow and pain in his tears, with his shame in his blushing, with his entreaty in his outstretched hands. . . If anyone tells me that this is not 'perception', for it cannot be so …I would beg him to … address himself to the phenomenological facts. (Scheler, 1954, p. 260)

Similar claims are made about directly perceiving intentions: For example, if I reach toward the cup to my left, others can directly perceive my intention to drink from the cup.

In order to help articulate the commitments of and evaluate DSP, I will compare it to another recent type of mindreading account: *dual process* accounts (e.g., Apperly, 2011; Apperly & Butterfill, 2009; Butterfill & Apperly, 2013a,b). Dual process accounts of mindreading claim we engage in two main types of mindreading, one relatively fast, automatic, unreflective, and cognitively efficient, the other relatively slow, controlled, reflective, and cognitively effortful. Further, they often posit that physically and/or functionally distinct psychological mechanisms enable these two mindreading types. This approach follows within the larger tradition in psychology of *dual process theories* that distinguish two general types of psychological processes, *Type 1 and 2 processes*, and often further claim these process types are enabled by distinct psychological mechanisms, often labeled *System 1 and 2* (e.g., Evans, 2010; Evans & Frankish, 2009; Evans & Stanovich, 2013; Frankish, 2010; Kahneman, 2003, 2011).

An initial motivation for comparing DSP and dual process mindreading accounts is that they both treat slow, deliberate, effortful mindreading and fast, automatic, effortless mindreading as distinct psychological kinds. Dual process accounts are explicit about this and often posit

separate psychological mechanisms responsible for these two types of mindreading phenomena. DSP advocates have been less focused on characterizing psychological mechanisms, but clearly treat perceptual and non-perceptual mindreading as distinct psychological types. This places DSP and dual process accounts in opposition to views that treat fast and slow mindreading as two modes of operation of the same basic type of psychological process or mechanism—for example, Carruthers (2013) argues that intuitive and reflective mindreading use the same basic representational resources.

But while both DSP and dual process accounts treat fast, automatic mindreading as psychologically distinctive, DSP is unique in characterizing it as *perceptual* in nature. The dual process literature certainly identifies perception as a paradigmatic example of a fast, automatic psychological process. But dual process accounts of mindreading tend not to even consider the possibility of perceptual mindreading as advocated by DSP. This invites the question of the relationship between DSP and dual process accounts of mindreading. Should they be seen as rival accounts of fast, automatic mindreading, with only one offering an accurate characterization of these mindreading phenomena? Or are they actually identifying distinct subtypes of fast, automatic mindreading, some perceptual and others non-perceptual?

One reason for thinking DSP and dual process accounts may be compatible accounts of distinct mindreading phenomena is that they often focus on different mental states types. As mentioned above, DSP advocates usually only claim that *intentions* and *emotions* can be directly perceived; they do not make this claim about *perceptions, beliefs*, or *thoughts* (e.g., Gallagher & Varga, 2014, p. 190). There are, however, dual process mindreading accounts of the fast, automatic understanding of epistemic states such as belief—one of the most influential and well-articulated being put forward by Ian Apperly and Stephen Butterfill (Apperly, 2011; Apperly & Butterfill, 2009; Butterfill & Apperly, 2013a,b). As we'll see below, Apperly and Butterfill argue that use of the full-blown concept of belief as a propositional attitude is likely too cognitively demanding to enable fast, automatic mindreading. They instead explain Type 1 belief understanding in terms of a "minimal" mindreading system operating with a non-propositional concept of "belief-like" states called "registrations," rather than the concept of belief proper. The point for now is that perhaps DSP and dual process accounts can be made compatible by treating DSP as a thesis about intentions and emotions, leaving dual process accounts to characterize fast, automatic attribution of epistemic states.

Why have DSP advocates in this way restricted the scope of their thesis to intentions and emotions? The main motivation for doing so seems to be the claim that mental state types that are closely connected (perhaps constitutively connected) with particular behaviors seem the best candidates for the DSP thesis. This is because everyone agrees we can perceive people's bodily behavior. So it is less controversial to treat mental states closely connected to behavior as being perceived. Intentions and emotions seem to be mental state types fitting this description. For example, as Spaulding (in press) puts it, "it is part of the concept of intention that an intention to Φ is correlated strongly with Φ-ing" (p. 3). Similarly, emotions are rather tightly associated with particular behavioral expressions—think of the prototypical facial expressions of the basic emotions. Epistemic mental states such as perception and belief, however, are much less tightly connected with any particular behaviors. As critics of behaviorism famously emphasized, how one acts will depend on not only one's beliefs, but also a host of other mental states. Accordingly, it is less plausible that a person's beliefs could be directly perceived.

It is worth considering, however, whether this restriction of the scope of the DSP thesis stands up to further scrutiny. Indeed, Butterfill and Apperly (2013b, p. 7) themselves have

recently mentioned the issue of the observability of mental states, remaining open minded about the DSP thesis and its application to their mindreading account. I will here explicitly take up the task of analyzing whether Apperly and Butterfill's account of fast, automatic attribution of belief-like epistemic states should be considered perceptual or non-perceptual mindreading. This task of determining what mental states types *should* fall within the scope of DSP thesis requires investigating what properties are essential to categorizing a mindreading process as perceptual or non-perceptual. The way I will address this issue is by investigating the dual process literature's own discussions of perception, and comparing it to other discussions of perception in the philosophy and psychological literature and the account of perception given by DSP advocates.

In sum, my primary motivation for comparing DSP with dual process accounts of mindreading is that both address fast, automatic forms of mindreading. Given that dual process accounts of Type 1 understanding of epistemic states are not formulated in terms of DSP, it is possible that fast, automatic mindreading will subdivide into perceptual and non-perceptual forms based on a distinction between mental state types that are more or less intimately connected to behavior. But it is a relatively unexplored question whether this possibility is empirically and theoretically motivated. This paper will make some initial inroads into this issue, focusing on whether Apperly and Butterfill's account of fast attribution of belief-like states should fall within the scope of DSP.

I will begin in section 2 by summarizing the commitments of DSP. Then in section 3 I will describe dual process theory, focusing in particular on Apperly and Butterfill's account of epistemic-state attribution, particularly their account of the Type 1 attribution of belief-like states. In section 4, I will explicitly compare the two accounts. I will first examine what the dual process literature has to say about the nature of perception. I will show that the dual process perspective adopts a roughly modular account of perception that can be classified as a kind of Type 1 process and that is commonly adopted within the philosophical and psychological literature on the nature of perception. I will then compare this modular view of perception with DSP's account of perception. I will argue that the modular view of perception challenges DSP's standard phenomenological and psychological theses, but still allows for DSP's central claim that mental states can be perceived. Specifically, I will argue that Apperly and Butterfill's account of Type 1 mindreading of belief-like states largely fits this modular view of perception. Thus, I will contend that, if we assume the truth of the modular view of perception (which admittedly is not uncontroversial), there is significant reason for treating belief-like states as falling within the scope of the DSP thesis.

2. Direct Social Perception

As Michael & De Bruin (this issue) articulate, following Bohl and Gangophadyay (2014), DSP can be read as making metaphysical, epistemological, phenomenological, or psychological claims about the perceivability of (some) mental states. I will focus on the phenomenological and psychological theses, which make the most contact with dual process theories.

2.1. DSP's Phenomenological Thesis

DSP understood as a *phenomenological* thesis claims that we can become aware of others' mental states by "directly" experiencing them. This direct experience of others' mental states is

positively characterized as a form of perception similar to the perceptual experience of physical objects; it is negatively characterized as being unlike the experience of forming a belief about something via inference from perceptual information about something else. For example, Gallagher (2008a) compares DSP to visually perceiving his car:

> At the personal or conscious level, I do not have to perceptually piece together the shape and the color and the mass in order to get my car. Even if the sub-personal processes are complex (and I do not deny that they are), the perception that I have of my car is direct—I see it right there in front of me. I do not have to glue anything together, add an interpretation or add an inference. (p. 537)

Gallagher contrasts this kind of perceptual experience with experiential phenomena where "something [is] added to perception, e.g., an inference or interpretation that goes beyond what is perceived" (p. 537). For instance, "if my car was terribly totaled in an accident, I may not recognize it at first and I may have to use certain clues about its appearance to infer that it is my car" (p. 537). DSP theorists in this sense treat perception as involving "direct" phenomenological experience of something present in our environment. As Zahavi (2011) puts it, our experience of another's mental state "can be said to be direct in the sense that that state is my primary intentional object… the state is experienced as actually present to me" (p. 548). DSP advocates contrast this direct experience with *indirect* mental state understanding, where we first perceive something other than the target object/property, and afterward experience at least one additional psychological process in order to, in the end, become aware of the target (Zahavi, 2011, p. 548).

The core of this phenomenological account of direct perception thus seems to be the following: X is directly perceived if X is part of the content of our conscious experience, X is experienced as present, and this content is generated in a phenomenologically immediate way. It is key that what is directly perceived is itself "my primary intentional object," as Zahavi put it. If I look at a thermometer, the temperature itself is not my primary intentional object. The thermometer is *experienced* as a representation or sign of the temperature. So on this account the temperature would not itself by directly perceived. Thus for mental states to be directly perceived, another person's mental state would need to be consciously experienced as present to me, and this experience must not be generated by first experiencing something else that we use to become aware of the mental state.

DSP does not deny that we sometimes become aware of others' mental states indirectly via conscious reasoning, whether by the imaginative simulation posited by the simulation theory (ST) (e.g., Goldman, 2006) or the theoretical inference proposed by the theory theory (TT) (e.g., Gopnik & Wellman, 1992). For example, a friend might tell you that Rosario was not hired for her dream job, and from this infer that Rosario is unhappy. In such a case you don't directly experience Rosario's unhappiness, but instead infer it from the linguistically conveyed information about Rosario's not getting the job. Similarly, you might perceive David sitting quietly, and based on this behavioral evidence plus your conscious belief that David's girlfriend just broke up with him, consciously infer that David is sad. In this case David's sadness is not experientially present to you, but is something you consciously infer from perceiving his behavior plus conscious background knowledge. Without denying such experiences of indirect, non-perceptual mindreading do sometimes occur, DSP does reject the idea that *all* mindreading experiences are indirect in this way. DSP instead claims that we often can directly perceive people's intentions and emotions.

Note that DSP theorists do distinguish the direct experience we have of *other people's* mental states from the direct experience we have of *our own* mental states (see, e.g., Gangopadhyay & Miyahara, in press; Zahavi, 2007, 2008). Determining how exactly to characterize this phenomenological difference, as well as the difference between social perception and object perception, are important projects for DSP's phenomenological thesis (see Gangopadhyay & Miyahara, in press).

In sum, direct social perception is phenomenologically defined as mindreading where others' mental states are experientially present to us in an unmediated, immediate way. This is contrasted with experientially indirect mental state understanding, where conscious reasoning (e.g., mental simulation or theoretical inference) is used to generate a conscious mental state attribution.

2.2. DSP's Psychological Thesis

DSP is also sometimes framed non-experientially, in terms of the psychological processes enabling mental state understanding. The main focus here is the distinction between *perceptual* processes and post-perceptual *cognitive* processes.

According to DSP's account of the literature (which I do not endorse; see Herschbach, 2008), traditional mindreading accounts such as TT and ST treat perception as providing information about people's words and deeds, not their mental states. To understand others' mental states thus requires *cognitive* processes that use perceptual or verbal information about a person's behavior and their environment to infer their mental states. For example, TT claims that the mindreader represents a body of theoretical information about the causal/rational relations between mental states, behavior, and environmental conditions (e.g. "people who don't get what they want tend to be upset"). For TT, such theoretical information about minded beings must be used to *infer* the mental states possessed by a particular person at a particular time. While ST denies that mindreaders must possess theoretical knowledge about minds, it (according to DSP advocates) similarly claims that perception alone cannot allow us to understand people's minds. According to DSP's interpretation of ST, mental state understanding comes from mentally simulating the mental states likely exhibited by another person, and making an inference from the simulated mental state to an attribution of that mental state to the target person.

DSP's psychological thesis denies that post-perceptual cognition is always needed to access other minds; it instead claims that some mindreading occurs via perception alone. To distinguish DSP's psychological thesis from its phenomenological thesis, one cannot simply appeal to an *experiential* difference between perceptual and non-perceptual mindreading. Unfortunately DSP advocates have not been especially clear on this point. Typically they define perceptual processes negatively, by saying what they do not involve. One such negative thesis is that DSP does not involve theory- or simulation-based inference, since these are defined by DSP advocates as cognitive processes (e.g., Gallagher & Varga, 2014).

Sometimes DSP's psychological thesis is more baldly stated by associating cognition with inference, and thus negatively defining perceptual mindreading in terms of *non-inferential* psychological processes (Gallagher, 2008b). As Michael & De Bruin (this issue) mention, however, mainstream cognitive psychology and neuroscience characterize perception as an inferential process: sensory transducers generate low-level informational states that go through several stages of information-processing to generate higher-level perceptual representations of the environment (e.g., Pylyshyn, 1999). Even if these processes don't involve logical

relationships between propositionally structured representations, most psychologists characterize these information-processing operations on representational states as inferences in a looser sense. Gallagher & Varga (2014) have explicitly backed off the strong anti-inference view by allowing perception to involve such "Helmholtzian" inferences; but they still deny that DSP involves theory- or simulation-based inference, which are treated as non-perceptual, cognitive processes.

What grounds this distinction between perceptual, Helmholtzian inferences and non-perceptual, cognitive inferences? Gallagher & Varga (2014) offer a few arguments. They first argue that Helmholtzian inferences "are not rich enough to underpin mindreading" (p. 193): Helmholtzian inferences are characterized as simpler processes required for basic object recognition, and do not make use of information about mental states. Thus, Helmholtzian inferences could not support mindreading. But this seems to simply assume without argument that perceptual inference cannot operate with higher-level contents such as mental states. Many have advocated for the general view that perception is "theory laden" (e.g., Churchland, 1979; Kuhn, 1962) or "cognitively penetrable" (e.g., Siegel, 2010, 2011). Several recent authors have applied such an approach to argue that if we define DSP phenomenologically, TT and ST can be interpreted as subpersonal-level accounts of the psychological processes enabling such perceptual experiences (Bohl & Gangopadhyay, 2014; Carruthers, 2013, p. 144; Herschbach, 2008; Lavelle, 2012; Spaulding, 2010).

Why do DSP advocates believe it is wrong to treat simulation- or theory-based inferences as part of perceptual processes themselves? Gallagher & Varga's (2014) argument here is less clear. But they are definitely resistant to saying perceptual processes are made more complex by the addition of theory- or simulation-based inferences. They instead propose that social and cultural factors (e.g., implicit racial biases) can "shape" or "transform the perceptual process itself," appealing to the notion of neural plasticity as a mechanism for how this occurs without the addition of inferential processes (p. 196). Given that their "transformation" account is not especially well fleshed out, it is difficult to evaluate its merits relative to the argument that theoretical inferences can "cognitively penetrate" and thus be part of a perceptual process. But by offering this alternative account, Gallagher & Varga have not directly addressed why they think it is wrong to characterize simulation- and theory-based inferences as constitutive parts of perceptual processes.

In sum, DSP's psychological thesis seems to come down to the claim that mindreading need not involve cognition, and sometimes can involve just perception. DSP advocates do not, however, offer a detailed discussion of the differences between cognition and perception, mostly relying on the claim that mental simulation and folk psychological theorizing are inferential cognitive processes, and thus not involved in perceptual mindreading.

2.3. Phenomenological and Psychological Theses About the Nature of Perception

In sum, DSP's phenomenological and psychological theses offer two ways of characterizing a perceptual mindreading. The phenomenological version says a mental state is directly perceived if a mental state is experienced as present and this conscious content is generated in a phenomenologically immediate way. DSP's psychological thesis attempts to define the perceptual form of mindreading in non-phenomenological terms. But this is where DSP advocates have been less clear about what they take to be characteristic of perception versus cognition. Perception-based mindreading is largely defined negatively, in terms of what kinds of

cognitive processes it purportedly does *not* involve, namely, simulation-based or theory-based inference.

As we'll see in the next section, dual process accounts also make phenomenological and psychological claims about mindreading.

3. Dual Process Theories of Mindreading

In this section I will first describe general dual process theories in psychology, and then describe theories of mindreading that fit this general picture—in particular, Apperly and Butterfill's dual process account of the attribution of epistemic states.

3.1. Dual Process Theories in Psychology

Philosophers and psychologists have for centuries proposed that the mind is not unitary. But modern dual process theories have developed over the last 30-40 years across various subfields of psychology, including learning, reasoning, decision making, and social cognition (see Frankish & Evans, 2009). In the last 15-20 years researchers have tried to unite these approaches, developing domain-general dual process theories (e.g., Kahneman, 2003, 2011; Stanovich, 1999, 2005). To illustrate their two kinds of psychological processes, compare how you would solve the following two math problems (Kahneman, 2011, pp. 20-21):

2 + 2 = ?
17 x 24 = ?

As you read the first problem, the answer comes to your mind almost immediately, with little to no effort required. But to answer the second problem would require (for most people) a conscious, effortful process, probably using pen and paper. These two math problems typify the contrast between, respectively, *Type 1 and 2 processes*.

Researchers generally appeal to a cluster of features when defining Type 1 and 2 processes (Evans, 2008; Evans & Stanovich, 2013; Frankish, 2010; Frankish & Evans, 2009). Some of the more commonly emphasized features concern *speed*, *automaticity*, and *consciousness*. The title of Kahneman's (2011) recent book prioritizes the characteristic of speed, contrasting *fast* Type 1 processes and *slow* Type 2 processes. Sometimes the emphasis is placed on whether or not a psychological process is under the control of the agent, contrasting *automatic* processes that occur whenever they are triggered by the appropriate stimuli, with *controlled* processes that can be decoupled from the immediate stimulus conditions, initiated at will, and can override automatic responses. Speed and automaticity can be understood as purely subpersonal-level processing traits, but they are often given a phenomenological characterization as well. Accordingly, Type 1 processes are often characterized as *unconscious* or *preconscious*, while Type 2 processes are *conscious*. Similarly, the contrast between *low effort* and *high effort* processes seems to be characterized phenomenologically as much as in terms of subpersonal-level properties. Given the lack of agreement about how to define and study consciousness, however, dual process theories tend to emphasize the kinds of subpersonal-level processing traits studied by cognitive psychologists and neuroscientists: e.g., the capacity of information a process can handle (high vs. low), the types of operations performed on that information (associative vs.

rule-based), how many such operations can occur at any given time (parallel vs. sequential processing), and whether or not working memory is required.

Dual process theories have developed in the last 10-15 years into accounts of dual *systems*, usually called, following Stanovich (1999), *System 1* and *System 2*. Dual system theories contend that the mind is physically divided into two distinct systems/mechanisms so as to account for the clustering of processing features into two types. System 1, which enables fast, intuitive Type 1 processes, is generally considered an evolutionary older system we humans share with other animals. The slow, reflective Type 2 processing enabled by System 2 is thought to be a more recent evolutionary development distinctive to humans (Evans, 2008; Evans & Stanovich, 2013; Frankish, 2010; Frankish & Evans, 2009). Through this paper I will use "dual process theory" to refer to both process- and system-based theories, unless otherwise specified.

3.2. *Apperly & Butterfill's Dual Process Theory of Mindreading*

Most contemporary theories of social cognition posit multiple processes and systems, but only some both make explicit use of dual process theory and apply it to mindreading phenomena. One of the most influential and fully articulated is Ian Apperly and Stephen Butterfill's (Apperly, 2011; Apperly & Butterfill, 2009; Butterfill & Apperly, 2013a,b) dual system theory of how we understand others' epistemic states such as perception, knowledge, and belief—states not typically included within the scope of the DSP thesis.

Apperly and Butterfill characterize our mindreading abilities as exhibiting competing cognitive demands for efficiency and flexibility (e.g., Apperly, 2011, pp. 8-9). Fast social interactions (e.g., playing sports) seem to require acting quickly in light of what others want, can and cannot see, etc. But to achieve this speed, these mindreading processes likely need to be cognitively efficient. In comparison, consider the task of a jury determining the guilt or innocence of a defendant. This involves mindreading the mental states of the defendant and others, but through a slow, deliberate, careful examination of the evidence. This also requires greater flexibility about the informational resources that could be relevant to making such mindreading attributions (p. 8). Such flexibility surely involves greater demands on memory and attention, and which means less cognitive efficiency.

Apperly and Butterfill focus in particular on our ability to appreciate other agents' perceptions, knowledge states, and beliefs, and how these epistemic mental states affect their goal-oriented behavior. On their view, these two types of mindreading phenomena are explained in terms of two types of mindreading systems: an early developing system for Type 1 mindreading, and a later developing system for Type 2 mindreading. The early developing system's defining feature is its cognitive efficiency. This enables its speed and automaticity, but comes at the price of exhibiting signature limits in the "kinds of input" it can process and the "kinds of operations performed on that input" (Apperly, 2011, p. 144). According to Apperly and Butterfill (2009; Butterfill & Apperly, 2013a), this early-developing mindreading system does not operate with full-blown concepts of mental states as propositional attitudes. Instead, it is a "minimal theory of mind" operating with simpler concepts of goals, perceptions, and "belief-like" states. Minimal mindreading is argued to enable success on, for example, standard change-of-location false belief tasks, where agent's beliefs about the location of objects must be tracked. But such minimal mindreading would not enable appreciating that agents can represent the same object in different ways, via different visual appearances, descriptions or concepts (for recent supporting evidence, see Low & Watts, 2013; Low, Drummond, Walmsley, & Wang, 2014).

"Full-blown" mindreading, where beliefs, desires, and intentions are represented as such, is thought to involve a separate, Type 2 style mechanism that develops later in childhood. This system is more flexible in its inputs, operations, and outputs, which lets it capture the complexity of the abductive inferences involved in full-blown mindreading. In particular, this flexibility is required to represent beliefs and desires as propositional attitudes that "form complex causal structures, have arbitrarily nestable contents, interact with each other in uncodifiably complex ways and are individuated by their causal and normative roles in explaining thoughts and actions" (Butterfill & Apperly, 2013a, pp. 609-610). To accomplish this requires a cognitive system that is not "informationally encapsulated," but rather can access all information available elsewhere within the mind. Accordingly, this system must interface with a variety of other cognitive resources, particularly, language abilities, attention, memory, and executive control. It is these additional cognitive demands that would make System 2 mindreading more likely to be conscious, slow, controlled, and effortful (Apperly, 2011). Further, since these cognitive capacities develop across children's first several years of life, Type 2 mindreading will develop later in childhood than Type 1.

3.3. Apperly & Butterfill's Type 1, "Minimal" Mindreading

Given the goals of this paper, I will examine in greater detail Apperly and Butterfill's account of Type 1 "minimal" mindreading involving the attribution of belief-like states.

3.3.1. Concepts and Principles of Minimal Mindreading

I will begin by explaining the concepts and principles of Apperly and Butterfill's minimal mindreading account. Butterfill and Apperly (2013a) start with a teleological concept of *goal* as a simpler, non-representational version of the full-blown concepts of *goal-directed actions* and *intentions.* A minimal mindreading system represents agents as having goals in the sense of having outcomes toward which their body movements are directed. In other words, actions are understood as having the function of producing certain outcomes (e.g., grasping an object). To build an appreciation of why an agent might act toward a particular goal, Butterfill and Apperly introduce a minimal analog to the concept of *perception*, particularly, *seeing*. They define an agent's *field* at a particular time as a certain spatial area around the agent, demarcated in terms of physical proximity to the agent, lighting conditions, the agent's orientation, posture, and possibly eye direction. An agent is said to *encounter* an object when it falls within their field. Encountering is thus an agent-object relation that approximates the full-blown concept of perception, without capturing the representational, perspectival nature of genuine perception. Encountering is represented as a causal constraint on goal-directed action. Analogous to the principle that one cannot act on what one does cannot see, the minimal mindreading system uses the principle that "one cannot goal-directedly act on an object unless one has encountered it" (p. 615). With such an understanding of goals and encountering, a minimal mindreader could predict that the agent will seek out only the objects they have perceived/encountered and act in light of this behavioral prediction.

The concept of encountering is then used to define *registration*. Registration is a mental state analogous to belief in that it plays a certain functional role: it is causally related to the agent's encounters and goal-directed actions. Registration is an agential state capturing a relation between an agent, an object, and a location, analogous to a belief's being as an agent's attitude

toward a content (see Butterfill & Apperly, 2013b, pp. 7-9). One principle governing the notion of registration mimics the connection between perception and belief: usually an agent registers an object at a location "if and only if she most recently encountered it at that location" (2013a, p. 617). Correctly registering an object is considered a condition for successful action (p. 617). Registration is also understood as a cause of action, via the principle that "when an agent performs a goal-directed action with a goal that specifies a particular object, the agent will act as if the object were in the location she registers it in" (p. 619). Accordingly, the concept of registration can function as a simplified way of tracking an agent's true or false belief about an object's location.

By design, the minimal concept of registration is not a perfect analog for belief as a propositional attitude. As an extensional relation between agents, objects, and locations, a registration cannot capture the intensionality or perspectival nature of belief, i.e., the fact that beliefs represent objects in a particular way, using particular concepts or modes of presentation. So a minimal mindreader would not be able to track agent's false beliefs involving mistakes of identity (e.g., failing to recognize that Superman and Clark Kent are distinct appearances of the same individual person). But ignoring the intensional nature of belief is exactly the kind of simplification that is supposed to make a minimal mindreading system cognitively efficient enough to enable fast, automatic mindreading of epistemic states.

3.3.2. Is Minimal Mindreading Really Mindreading?

One may ask whether the concepts of encountering and registration are similar enough to the concepts of perception and belief to be considered genuine mental state concepts. This is necessary if Apperly and Butterfill's account is to be considered a minimal form of actual mindreading, and thus worth comparing to DSP.

Butterfill and Apperly (2013a, p. 261, 2013b) explicitly argue that *registration* should be understood as a mental state concept because the minimal mindreader represents it as an "intervening variable" between environmental inputs and behavioral outputs. This is a common way philosophers and psychologists distinguish mindreading from non-mentalistic "behavior-reading" strategies that only represent environmental inputs and behavioral outputs. Full-blown mindreading represents beliefs as agential states that are caused by perceptual states, and that causally produce (in combination with other mental states, such as desires) goal-directed actions. Analogously, registrations are states that are (often but not necessarily) caused by encounters with objects, and that, in light of an agent's goals, causally lead to actions. The concept of registration thus differs from the full-blown concept of belief in terms of having a simpler functional role. Registrations also have simpler contents than beliefs: registrations do not have propositions as contents like beliefs do. But registration does share with belief the key features of being an intervening variable in a causal model and possessing contents. Registration is thus a simpler mental state concept than belief—but a mental state concept nonetheless.

Admittedly, defining the nature of mental states is notoriously controversial. But I believe this is a persuasive argument for considering registrations to be genuine mental states, and thus minimal mindreading to be genuine mindreading. While there is more to say about the nature of mental state concepts, the argument appeals to a well-motivated criterion for theoretically distinguishing mindreading from behavior-reading, which is commonly appealed to in the psychological and philosophical literature on mindreading. It should be noted, however, that Apperly and Butterfill make no attempt to apply this reasoning to *encounters.* Representing

an agent's *encounter* with an environmental object is just a way of describing which environmental objects the agent perceives, rather than a mental state representing those objects. This is a further way Apperly and Butterfill's characterize Type 1 mindreading as simpler than full-blown mindreading.

3.3.3. Applications of Type 1, Minimal Mindreading of Belief-Like States

Apperly and Butterfill developed their account of minimal mindreading in order to capture the fast, automatic nature of Type 1 mindreading phenomena. To analyze it with respect to the DSP thesis, it will be helpful to see in greater detail how Type 1 mindreading of belief-like registration states is supposed to work. To do this, we can examine the experimental studies that Apperly and Butterfill interpret in terms of their early-developing, minimal mindreading system.

Many of these experiments use change-of-location false belief tasks. In such tasks, participants are shown the following kind of scene, usually containing an agent, an object (e.g., a ball), and two containers located in front of the agent, where the object can be placed inside. The target agent first looks at the object being placed inside one of the containers—i.e., the agent's head and eyes are directed toward it and no other objects obstruct their line of sight. Then the agent leaves the scene or looks away (i.e., turns their head and/or body sufficiently so the containers are not in their line of sight). While the agent is not looking, the object moves to the other container (on its own, or through the action of another person). The agent returns to look at the containers, with the object out of sight inside one. Since the target agent didn't see the object move, upon their return they will continue to believe, falsely, that it is to be found in the original location. These false-belief scenes are typically interspersed with true-belief scenes: here the agent does observe the object's placement in a new location, so has a true belief about its location throughout the scene. It is assumed in such studies that the agent *wants* the target object, and upon their return to the scene will act on the goal of seeking it out. The scene usually ends then with the agent reaching into a container to grab the object. Whether this seeking behavior is successful depends on whether they had a true or false belief about the object's location. With studies involving children, the agent's goal is sometimes made explicit by initially presenting scenes of the agent repeatedly reaching for that object, before presenting a true- or false-belief scene.

One prototypical example of an "implicit" false belief task (see Low & Perner, 2012) given to children is Southgate et al. (2007). Their study presented 2-year-olds with videos of change-of-location scenarios involving a ball and two boxes. This was done after a familiarization phase to make explicit the agent's goal of grasping the ball. The experimenters did not give children any explicit instructions about engaging in mindreading, but instead measured children's looking behavior after the agent returned to looking at the boxes but prior to the agent's reaching into a box. They recorded which of the two boxes children first looked at and spent the most time looking at during that time period. This anticipatory looking behavior was consistent with children's representing the agent's true or false belief about the ball's location. Apperly and Butterfill's minimal mindreading account posits instead that watching the beginning of the video leads children to automatically represent that the observed agent *encountered* the ball in the first box, and thus that they *registered* the ball as being at that location. This registration state will persist while the agent looks away from and then looks back at the boxes. Why? Because during that time the ball is not within the agent's field and thus not encountered. Building in the assumption the agent has the goal of grasping the ball, this

information leads the child to automatically expect the agent to reach into the first box, rather than the second box where the ball was actually located. This automatically generated behavioral prediction is expressed in the children's anticipatory looking.

Studies with adults attempt to more directly address the automaticity and implicitness of the attribution of belief-like states. They often do so by presenting the kind of mindreading-inducing stimuli described above either (a) without any instructions to mindread, or (b) explicitly requiring participants to engage in a non-mindreading task. Examples of the latter are studies (e.g., Kovács et al., 2010; Schneider et al., 2012a, 2012b, 2014) that instruct participants to track the location of the object moved around during a change-of-location false belief scenario. If participants' nonverbal behavior (specifically, looking behavior) conveys that they've represented that the observed agent is mistaken about the object's location, despite this having nothing to do with their explicit task, this suggests participants are automatically, implicitly engaging in attribution of belief-like states. Schneider's work suggests this Type 1 mindreading of belief-like states occurs without any conscious awareness (Schneider et al., 2012a), and is unintentional and uncontrollable (Schneider et al., 2014).[1]

Other automaticity studies (e.g., Cohen & German, 2009, 2010) use a similar setup, having participants explicitly engage in a non-mindreading, object-tracking task. But they instead use a verbal task: they measure participants' reaction times in responding to questions about the object's location at the end of the video (e.g., "It is true that the object is in the location on the left") vs. questions about the agent's belief (e.g., "She thinks the object is in the location on the left"). If they respond just as quickly to both questions types, despite not being asked to track the observed agent's mental states, it suggests participants were automatically, implicitly mindreading while watching the video. Butterfill and Apperly (2013a) are not that specific about the uses to which minimal mindreading representations can be put, but do explicitly mention that representations of belief-like states could "be verbalized in terms of what an agent 'thinks'" (p. 627).[2] So this suggests automaticity studies such as these using verbal test stimuli could also be explained in terms of Type 1 minimal mindreading of belief-like states.

4. Type 1 Processes and Perception

For the purposes of this paper, I will be assuming Apperly and Butterfill's two-system account is well motivated theoretically and empirically (for recent critiques, see Carruthers, 2013; Thompson, 2014). My concern here is whether or not the sort of fast, automatic, Type 1 mindreading of belief-like states they describe should be considered perceptual in nature; in other words, I want to determine whether dual process theories of mindreading such as this one fit with DSP or provide an alternative way of describing the phenomena identified by DSP advocates.

To address this issue, I will first examine how dual process theorists themselves talk about perception. Then I will address how such an account of perception compares to DSP's

---

[1] One problematic piece of evidence is Schneider et al.'s (2012b) dual-task study, which found that minimal mindreading is disrupted by high cognitive load, an indicator that it depends on executive resources (e.g., working memory). This is contrary to Apperly and Butterfill's account of System 1 mindreading's cognitive efficiency.

[2] This appears to be a modification of their view over time, since Apperly (2011, p. 134) seemed to categorize verbal tasks as Type 2 phenomena.

phenomenological and psychological theses, and whether Apperly and Butterfill's Type 1 attribution of belief-like registration states should count as perceptual in nature.

4.1. Dual Process Theorists and Others on Perception vs. Cognition

Although dual process theorists often use perception to help define Type 1 processes, they tend to treat perception as distinct from and providing input to Type 1 and 2 psychological processes. For example, Kahneman (2003) makes a tri-part distinction between perception, System 1 "intuition," and System 2 "reasoning." System 1 is defined as being similar to perception with regard to several "operating characteristics": "fast, parallel, automatic, effortless, associative, slow learning, and emotional" (p. 698). But perception and System 1 are distinguished by their contents and functional inputs. Kahneman claims that perception operates with "percepts" while System 1 operates with both percepts and "conceptual representations." The only analysis given of the distinction between perceptual and conceptual representations is functional. Percepts are "stimulus-bound" in the sense of being generated by "current stimulation" (p. 698). In contrast, System 1 operations can be evoked by this type of perceptual information, as well as conceptual representations not tied to current stimulation, e.g., representations of past, present, or future. In addition, Kahneman makes clear that System 1 can take as input linguistic representations. So Kahneman's general dual process theory treats perception as functionally distinct from and providing one type of input to Type/System 1 and 2 processes.

In his dual process theory of mindreading, Apperly (2011, pp. 119-125) similarly analogizes fast, Type 1 mindreading to perceptual processes (e.g., vision). He says both seem to be enabled by psychological "modules": "informationally encapsulated" mechanisms that are cognitively efficient in that they "perform specific operations using their own small set of knowledge and representational sources, and are receptive to only a small set of external inputs" (p. 120). Later he claims that slow, Type 2 mindreading "seems less like perception [compared to Type 1 mindreading] and more like reasoning" (p. 125). But Apperly does not go so far as to actually describe Type 1 mindreading as perceptual in nature; he, like Kahneman, only notes similarities between Type 1 mindreading and perception. Generally Apperly and Butterfill refer to mindreading as a type of "reasoning," where mental state representations are produced by processes of "inference" from information about behavior.

In sum, dual process theorists often treat perception as separate from the two processing types/systems they identify. It would seem on this view perception is characterized as generating fairly low-level representations of environmental objects and their properties, which can be further processed by Type 1 and 2 processes. Given the numerous connections dual process theorists draw between perception and Type 1 processes, one could ask whether it is more theoretically appropriate to categorize perception as a kind of Type 1 process, rather than a separate psychological kind distinct from Type 1 and 2 processes.

In light of various criticisms waged against dual process theories over the years, Evans and Stanovich (2013) offer a refined dual process theory that does re-categorize perception in this way. They define Type 1 processes in terms of a single essential feature: being *autonomous* processes. Processes are said to be autonomous when their execution "is mandatory when their triggering stimuli are encountered and … are not dependent on input from high-level control system," so make minimal demands on working memory (p. 236). This definition of autonomous processes thus combines the two notions of *automaticity* and *cognitive impenetrability* or *informational-encapsulation.* Departing from the idea that Type 1 processes are enabled by a

single physical system, Evans and Stanovich (2013, p. 236) identify a host of different psychological mechanisms that meet this definition, from traditional Fodorian modules (Fodor, 1983) to general processes of implicit learning and conditioning. On Evans and Stanovich's dual process account, while not all Type 1 processes would be perceptual, autonomous perceptual processes would count as Type 1 processes. This view has the advantage of recognizing the many similarities between perception and other Type 1 processes by treating them as being of the same psychological kind.[3]

Another virtue of Evans and Stanovich's account is that it is consistent with an account of perception that has been highly influential view across psychology and philosophy: the view that perceptual processes are *modular* in nature (e.g., Fodor, 1983; Pylyshyn, 1999; Scholl & Gao, 2013; Scholl & Tremoulet, 2000). Though many have departed from Fodor's full definition of modularity, this is usually taken to mean that perception meets the following four conditions: (a) it is *mandatory* or *automatic*, i.e., operates whenever presented with a particular, limited range of inputs, without the need for conscious control or effort; (b) it operates *unconsciously*, with only the final outputs being potentially accessible to consciousness; (c) it is *cognitively impenetrable* or *informationally encapsulated*, i.e., insensitive to information contained in other cognitive mechanisms; and (d) it is *fast* in speed. Given this definition, modular perception would be classified as autonomous and thus a Type 1 process.

These features of modular mechanisms, however, would not alone be sufficient to distinguish perception from other non-perceptual, modular Type 1 mechanisms. For example, a Type 1 process automatically producing the solution to "2+2" clearly does not lead us to *perceive* the answer. Beyond the above four features of modularity, we need to capture the fact that perception is based on our causal contact with the world. Kahneman (2003) captured this idea when claiming that perceptual processes use as their inputs the "current stimulation" of our sense receptors. Let's call this fifth condition *stimulus-sensitivity*.

Scholl and Gao (2013) offer a detailed articulation of such a modular view of perception, including the requirement of stimulus-sensitivity. They characterize this idea of stimulus-sensitivity by writing that "a hallmark feature of perception (vs. cognition) is its strict dependence on subtle visual display [or other sensory input] details; percepts seem to be irresistibly controlled by the nuances of the visual [or other sensory] input regardless of our knowledge, intentions, or decisions" (p. 209). This passage captures Kahneman's idea that perception must originate from sensory input. But it goes further by claiming that variations in perceptual states are "irresistibly controlled by" nuanced sensory cues, and thus largely unaffected by other psychological factors such as our beliefs and intentions. This stronger notion of stimulus-sensitivity thus incorporates traditional features of modularity such as automaticity and informational-encapsulation. Whether we bundle automaticity and informational-encapsulation into our definition of "stimulus-sensitivity," or use that phrase to refer just to the fact that perception's inputs are impingements of our sensory organs, seems just a terminological difference. The modular view of perception is well captured by this package of functional traits. In sum, the modular view characterizes perception as involving the fast, automatic, informationally-encapsulated processing of sensory inputs.

Note that the above description does not specify the nature of the *outputs* of perceptual processes, other than that they may be conscious or unconscious. Some in the modularity camp

---

[3] An earlier co-authored work of Kahneman's strongly emphasizes the "perception-like" nature of Type 1 "intuitive thinking," even saying "The boundary between perception and judgment is fuzzy and permeable" (Kahneman & Frederick, 2002, p. 50).

(e.g., Fodor, 1983) have defended the idea that perceptual outputs are rather "shallow," only representing simple "low-level" features. In the case of vision, this would include "shapes and other spatial properties or relations, textures, colors, lightness, and motion" (Burge, 2014, p. 575). This seems to be Kahneman's (2003) view. But there is increasing evidence that perception can include "higher-level" contents as well—even if one accepts the modular view that perception is not cognitively penetrable. For example, Burge (2010) surveys empirical research that the human visual system can perceptually represent *bodies* (i.e., three-dimensional shapes with connected boundaries). Block (2014) and Burge (2014) discuss evidence that humans also visually represent high-level properties like *faces*. Furthermore, Scholl and Gao (2013) explicitly adopt a modular view of perception when defending the idea that humans can visually perceive *animacy* and *goal-directed action* (see also Scholl & Tremoulet, 2000). They mainly focus on perceiving "chasing," i.e., one agent acting with the goal of catching another agent. That the modular view of perception could permit the perception of *goals*, even in a minimal sense, shows the potential compatibility between the modular view and the DSP thesis.

To recap: starting from the dual process literature's discussions of perception, I have motivated treating perception as a kind of Type 1 process, rather than as a kind of input to Type 1 and 2 processes. This is possible if Type 1 processes are characterized as autonomous, broadly modular processes. From this perspective, perception, if modular in nature, can be classified as a kind of Type 1 process. To be distinguished from other modular Type 1 processes, perception must exhibit the additional condition of stimulus-sensitivity. This modular account thus defines perception in terms of the functional processing characteristics of being fast, automatic/mandatory, informationally encapsulated, and stimulus-sensitive, with these processes operating outside conscious awareness, and only its products or outputs as potentially accessible to consciousness. This modular view of perception is not only compatible with dual process theory, but is a very common account of perception throughout philosophy and psychology. These are two reasons for this paper to take seriously the modular view of perception. Another is that the modular view of perception allows for not just low-level perceptual content, but higher-level features. This includes bodies, faces, and action goals—which means the modular view is compatible with the DSP thesis that mental states are perceivable.

4.2. Phenomenological vs. Psychological Definitions of Perception

How does this modular view of perception relate to the phenomenological and psychological theses of DSP? First, note that the modular view defines perception in terms of functional, subpersonal-level processing characteristics (speed, automaticity, information-encapsulation, and stimulus-sensitivity). The only mention of phenomenology is the claim that the processes generating percepts are introspectively opaque. But this view also allows for fully unconscious perception, where both the process and the end product of perception remains outside of conscious awareness (see Burge, 2014, p. 583). Further, this view's advocates explicitly reject the idea that phenomenology alone is sufficient to distinguish perception from cognition. As Pylyshyn (1999) puts it: "…phenomenology turns out to be an egregiously unreliable witness" about the nature of perception and cognition because "Our subjective experience of the world fails to distinguish among the various sources of this experience, whether they arise from the visual system or from our beliefs" (p. 362). That is, phenomenological methods may help identify the conscious content of an experience. But introspection does not identify whether this experience is generated from perception or, as Block (2014) puts it, is "primarily the 'cognitive

phenomenology' of a conceptual over-lay on perception" (p. 566). The empirical techniques of perceptual psychology and neuroscience are required to get at the sources of our experiential states.

For example, Block (2014) argues that adaptation studies can help to determine whether an experience is perceptual or cognitive. These studies appeal to the phenomenon of perceptual adaptation: when a neural system receives a certain type of sensory stimulation for a period of time, it adapts to this stimulation, making it easier to respond to other stimuli of that type. For example, Block (2014, pp. 563-564) describes an adaptation study (Butler et al., 2008) suggesting we possess higher-level visual representations of facial expressions like anger and fear. In the study, participants stare at, say, a fearful face. Because of perceptual adaptation, when they next look at a face ambiguous between fear and anger, they are biased to see it as a fearful face. When the same lower-level features (e.g., orientation, curvature, shape) are presented but not in the coherent structure constitutive of a recognizable facial expression, the adaptation effect is extinguished. This suggests the perceptual representation displaying adaptation is genuinely a higher-level representation of a facial expression. According to Block (see also Scholl & Gao, 2013, p. 206), there is no evidence that conceptual representations display such adaptation effects. So adaptation studies serve as a useful type of evidence to identify the contents of perception as distinct from cognition—notably, one that gives no direct role to phenomenological evidence.

I believe this offers a substantial challenge to DSP's phenomenological thesis, which distinguishes experientially direct, perceptual mindreading from non-perceptual mindreading that is experientially indirect because it involves conscious reasoning. If the modular view of perception is correct, a mental state attribution could, in principle, be experientially direct because (a) it is generated by a perceptual process, or (b) a top-down cognitive process operates outside of awareness and outputs a conscious mental state attribution. Accordingly, phenomenological contents would be inadequate to define such experiences as perceptual or non-perceptual in nature.

While the modular view of perception challenges DSP's phenomenological thesis, it does not mean a deathblow for DSP. The modular account offers a psychological definition of perception that is more robust and empirically motivated than DSP's negative psychological thesis (which simply define DSP as not involving theory- or simulation-based inference). In addition, the modular account admits the existence of higher-level perceptual contents—indeed, some of its adherents have already marshaled evidence suggesting some mental states should be included amongst the list of perceptual contents. As already mentioned, Scholl and Gao (2013) explicitly argue that humans can visually perceive actions as animate and goal-directed. They remain open what concept of "goal" is supported by this research; but Apperly and Butterfill's minimal teleological concept seems a plausible option. This means there are at least some major adherents of the modular view that appear to endorse the DSP thesis with regard to perceiving goals. This is not quite the direct perception of *intentions* as such, but it is nonetheless an endorsement of DSP. The modular view of perception may also be compatible with the perception of other mental states that already fall within the scope of the DSP thesis: emotions. As just mentioned, Block (2014) discusses evidence that humans visually perceive emotion-related facial expressions (e.g., angry vs. fearful faces). Depending how one fills out DSP, this is a short step away from saying we directly perceive emotions themselves. Some DSP advocates argue that bodily movements are literal constituents of emotions. Accordingly, they argue that by seeing a proper part of an emotion (a facial expression) we are directly seeing the emotion itself,

rather than simply an external sign of an emotion (Gallagher & Varga, 2014; Krueger & Overgaard, 2012). This argument is still open to them with the modular view of perception.

From the perspective of DSP, another benefit of the modular view of perception is that it at least partially supports DSP's treatment of TT and ST as accounts of cognition and not perception. Recall that a core feature of the modular view is that perceptual systems are informationally encapsulated from an agent's beliefs, desires, and intentions. If TT treats a folk psychological theory as just a subset of the beliefs we have about the world, then such beliefs would be part of "central cognition" (Fodor, 1983) and not be accessible by modular perceptual systems. Similarly, if ST says we use our own beliefs, desires, and other mental states to simulate the mental states of another person, these representations would similarly be cut off from our encapsulated perceptual modules. Thus, the modular view of perception supports treating certain versions of TT and ST as cognitive and not perceptual. The modular view does, however, appear compatible with the idea that our perceptual modules themselves contain theoretical information about the mind so as to automatically generate perceptual representations of other agents' mental states. So the modular view of perception does not completely endorse DSP negative psychological thesis, but is still consistent with a good portion of it.

In sum, I believe DSP ultimately has nothing too significant to fear and much to gain from the modular view of perception. While the modular view challenges the use of phenomenology to define perception, its psychological definition of perception has significant advantages over DSP's negative psychological thesis. For one, it more positively characterizes the processing and representational properties characteristic of perception and cognition. But it does so while allowing for mental states to be high-level perceptual contents. The modular view even partially supports DSP's resistance to treating TT and ST as parts of perception. DSP should welcome this sort of attempt to use perceptual psychology to define the psychological characteristics of perception.

The final step in my argument will be to consider what the modular view should say about Apperly and Butterfill's account of Type 1 minimal mindreading of belief-like states.

4.3. Is Type 1 Mindreading of Belief-Like States a Form of DSP?

Assuming the modular view of perception is endorsed, what should it say about Apperly and Butterfill's account of Type 1 minimal mindreading of belief-like states?

Recall the modular account's definition of perception: perception involves fast, automatic, informationally-encapsulated, stimulus-sensitive processes, whose products may or may not be consciously accessible. According to Apperly and Butterfill's dual process account, does Type 1 belief-like state attribution fit this definition? On their account, the attribution of belief-like states is fast, automatically generated by a relatively narrow-range of stimuli, and informationally encapsulated. The experiments using change-of-location false-belief tasks described in section 3.3.3 provide empirical support for this interpretation of minimal mindreading as fast, automatic, and informationally encapsulated. Further, if Schneider's findings prove correct, Type 1 belief-like state representations are not consciously accessible. This is not a problem for the modular account, which permits unconscious perceptual states. But as an empirical claim, it should be noted that it is inconsistent with the interpretation of Cohen and German's (2009, 2010) verbal tasks, which seem to require conscious awareness of one's belief-like state representations.

That covers four of the five requirements of the modular theory of perception. What about the requirement of stimulus-sensitivity? Consider the interpretation of the stimulus-sensitivity condition which requires that percepts are generated in a bottom-up fashion from physical impingements upon our sensory transducers, encapsulated from top-down influence by one's beliefs, desires, and intentions. Apperly and Butterfill's account of Type 1 belief-like state attribution fits this definition of the stimulus-sensitivity condition. They claim minimal mindreading is informationally encapsulated, with belief-like state representations automatically generated in a bottom-up fashion from sensory stimuli. The automaticity studies described in section 3.3.3 (e.g., Kovács et al., 2010; Schneider et al., 2012a, 2012b, 2014) provide empirical support for this interpretation: regardless of a person's explicit goals (e.g., those defined by the instructions given to participants), these studies found evidence that visually observing the relevant stimuli automatically produced attributions of belief-like states. This would mean Type 1 mindreading of belief-like states meets all the requirements of the modular view of perception, and is thus an extension of the scope of the DSP thesis to a kind of epistemic state.

Scholl and Gao's (2013) interpretation of the stimulus-sensitivity condition is, however, more robust. In addition to being insensitive to our beliefs and intentions, they propose that perceptual processes display a "dramatic dependence on subtle visual display details" (p. 209). Block's (2014) discussion of adaptation studies similarly emphasizes the importance of a careful examining of how observers respond to different stimuli. Scholl and Gao (2013), however, indicate that adaptations studies are not appropriate for sensory cues that are "so necessarily dynamic in both space and time," such as the animated displays of moving geometric figures used in their studies to cue the perception of chasing (p. 206). This suggests adaptation studies would also be unhelpful for studying whether Type 1 mindreading of belief-like states is robustly stimulus-sensitive.

One argument Scholl and Gao (2013) make with respect to visually perceiving goals is that performance on *visuomotor tasks*, rather than verbal tasks, is a good indicator of a genuine perceptual process and not a cognitive judgment. For example, their studies of chasing perception (Gao, McCarthy, & Scholl, 2010; Gao, Newman, & Scholl, 2009; Gao & Scholl, 2011) require participants to watch a computer display filled with moving geometric shapes identical in color and shape, one of which ("the wolf") is chasing one other ("the sheep"). One task given to participants was to control the sheep's movements to avoid getting caught by the wolf. This task thus requires perceiving the wolf as having the goal of chasing the sheep. These studies manipulated in various ways the visual cues underlying chasing. For example, Gao et al. (2009) varied the wolf's "Chasing Subtlety," defined as "the maximal angular deviation of the wolf's heading compared to perfect heat seeking" (Scholl & Gao, 2013, p. 211). "Perfect heat seeking" means the wolf is always heading directly toward the sheep; this is described as a Chasing Subtlety value of 0°. A Chasing Subtlety value of 30° means the wolf is always heading within 30° to the left or to the right of the direction of the sheep. They found that participants were better at perceiving the wolf when Chasing Subtlety was 0° as opposed to 30°. In addition, participants had a very hard time detecting the wolf when Chasing Subtlety was as large as 90°. Although the wolf was *actually* chasing the sheep in all cases, varying this visual (motion) cue dramatically affected an observer's ability to *perceive* the chasing. As they put it, these studies "reveal stark limits on the ability of observers to simply 'decide' what counts as chasing. Rather, the implicit performance measures seem to be tapping into an underlying ability whose limits cannot be influenced merely by decisions about what features should matter for detecting animacy; rather, only those factors that actually *do* matter will facilitate detection and avoidance"

(p. 214). For this reason, Scholl and Gao argue that such visuomotor measures are good ways of determining whether a response is robustly stimulus-sensitive and thus truly a perceptual phenomenon.

Existing studies of belief understanding have failed to examine the precise sensory cues that trigger Type 1 attribution of belief-like states (for initial discussions of such a cue-based approach to the study of mindreading, see German & Cohen, 2012; Wertz & German, 2013). So there is not the kind of research needed to say whether Type 1 attribution of belief-like states is robustly stimulus-sensitive. But there are a few things we can say to motivate thinking it might well be.

First, note that many of the existing studies of Type 1 belief attribution follow Scholl and Gao's recommendation of using visuomotor tasks. Specifically, as described in section 3.3.3, many studies use nonverbal measures of looking behavior. Admittedly, looking behavior can be a conscious, intentional action. But Scholl and Gao argued that visuomotor tasks are good test of what participants automatically perceive, rather than what they choose to believe about the world. The nonverbal mindreading tasks similarly test participants' implicit looking behavior. They do not give participants any explicit instructions to engage in looking behavior; some even explicitly engage participants in a task unrelated to mindreading. Nonetheless, participants display anticipatory looking behavior indicative of attributing belief-like states based on the visual stimuli presented to them. This suggests that participants can't help but *seeing* the observed agents as possessing belief-like states about object locations, just like the participants in Scholl and Gao's studies couldn't help seeing chasing. This interpretation is consistent with Apperly and Butterfill's view that Type 1 mindreading is inflexible in the *uses* to which its mental state representations can be put, with "implicit" behavioral tasks such as predictive eye movements included and slow, conscious reasoning excluded. Such inflexibility suggests the outputs of Type 1 mindreading process are not fully part of "central" cognition in Fodor's (1983) sense. This limited inferential role for the mental state representations generated by Type 1 mindreading processes makes it more plausible that they are modality-specific perceptual processes fairly directly influencing behavioral output.

In addition, note that most of the existing studies of Type 1 belief-like state attribution use remarkably similar visual stimuli. This is consistent with Apperly and Butterfill's view that Type 1 mindreading is relatively inflexible and efficient in terms of being generated by a limited range of input conditions. As described in section 3.3.3, most of these studies use nearly identical videos of change-of-location false-belief scenarios. Further, the visual stimuli used to induce representations of belief-like registration states seem even less dynamic than the studies of chasing perception. All that's presented to cue a representation of a belief-like registration state is a depiction of an agent encountering an object within their field—i.e., an agent directs their head and eyes toward an object with nothing obstructing their line of sight. This stimulus is temporally rather short, and is relatively simple in the number of objects and relations to be tracked. The false-belief scenario continues after this, but then the main requirement of the mindreader is *not* to update their initial representation of where the agent registers the object as being located, because the agent does not again encounter it. Perhaps subtle changes in the visual cues would lead to disruptions of the minimal mindreading system's ability to generate representations of encountering and, subsequently, belief-like registration states. For instance, perhaps small modifications in the orientation of the eyes and head of the observed agent, or the introduction of distractor objects, could cause such disruptions in mindreading of belief-like states. Without further studies, we simply do not know whether Type 1 belief-like state

attribution is robustly stimulus-sensitive in this way. Recall that Apperly and Butterfill's point in developing their account of minimal mindreading was to characterize a way to track others' beliefs in a cognitively efficient manner. It seems possible that this efficiency involves even further "signature limits" in its range of inputs than they initially thought.

Another way to go is to argue that Scholl and Gao's notion of robust stimulus-sensitivity is simply too strong a requirement on perception. Perhaps some but not all perceptual phenomena are robustly stimulus-sensitive. If that's the case, all we need for genuine perception of belief-like states is that belief-like state representations are automatically generated from sensory input, while being informationally encapsulated from our beliefs, intentions, and other higher-order psychological states. I already argued above that Apperly and Butterfill's characterization of their theory fulfills this weaker interpretation of stimulus-sensitivity, as well as the other functional requirements of the modular view of perception (speed and automaticity), and that the existing experimental research provides decently strong empirical support for this position.

Before I end, it is worth considering a type of objection commonly raised against DSP: *are belief-like states the kind of state that is even capable of being perceived?* This objection is grounded in questions about the metaphysical relation between mental states and physical behavior. If mental states are literally "inner causes" of physical behavior, it seems that only behavior is literally perceivable. DSP advocates have given different answers to this worry; one is that behavior is a constitutive part of some mental states, so that the mental state is perceived by perceiving the behavior (see Krueger & Overgaard, 2012). As discussed earlier, this is a main reason that DSP has been defended mainly for emotions and intentions, which have more tight connections to behavior than epistemic mental states.

I believe the modular account offers a helpful response to this worry, which does not depend upon addressing the metaphysics of the mind-body relation. For the modular account, perception occurs when the appropriate stimulus cues automatically trigger a perceptual representation of that feature, in a bottom-up, informationally encapsulated fashion. Technically, the represented feature need not actually be present in the environment at all. For example, in Scholl and Gao's studies of chasing perception, there are no real animate agents present—only animated videos of geometric figures displaying the appropriate cues to trigger percepts of goals (see Scholl & Gao, 2013, p. 206). Similarly, consider "amodal completion," which is the "capacity to perceptually represent an entity as whole or completed, even though less than the whole entity causally affects the sensory apparatus" (Burge, 2010, p. 417). An example often used by DSP advocates is that even if only the front side of an object causally impinges upon our retinas, the object's backside is experienced as present. These are cases where our perceptual systems generate perceptual representations that go beyond the information present in the sensory stimuli (e.g., Burge, 2010, pp. 417, 448). Thus, whether belief-like states are perceivable is, for this modular account, can't simply be an issue of whether mental states physically impinge upon our sense organs, or are in some other way "directly present" to our senses. Rather, what matters is whether our perceptual systems are set up to be automatically triggered by specific cues to form percepts representing other people's belief-like states. That is, it is a matter of the content of our perceptual states that are produced by modular perceptual systems. As I argued above, Apperly and Butterfill's theoretical account of an early-developing, minimal mindreading system seems to fit this requirement, and has a growing amount of empirical support behind it. The kind of empirical research described by Block (2014) and Burge (2014) seems most relevant to more adequately addressing the empirical merits of this claim.

In sum, the account of Type 1 belief-like state attribution offered by Apperly and Butterfill largely fits the modular definition of perception, with the stimulus-sensitivity condition being the main theoretical and empirical sticking point. If the modular view of perception is adopted by DSP, then there is good reason for extending the scope of the DSP thesis beyond goals and emotions to belief-like epistemic states.

5. Conclusion

My examination of the relationship between DSP and dual process theories of mindreading has reached several, admittedly tentative, conclusions. One is that dual process theory has good reason to treat perceptual processes as a subset of Type 1, fast, automatic processes, rather than a separate psychological kind that provides some of the inputs for Type 1 and 2 processes. This entails adopting a modular view of perception that has widespread support across philosophy and psychology. Further, I have argued that this view of perception provides substantial advantages over existing phenomenological and psychological formulations of DSP. The modular view recommends completely abandoning a phenomenological definition of perception, and instead focusing on functional and other processing characteristics of psychological processes to define perception.

If such a modular account of perception is adopted, I have argued that DSP remains a viable account of at least some modular Type 1 mindreading phenomena. Specifically, the existing scope of the DSP thesis to the mindreading of goals and emotions can arguably be retained, and perhaps given even greater empirical support by studies from vision science. Further, I have argued that the modular view of perception can seems able to widen the scope of the DSP thesis to include some epistemic states—specifically, Apperly and Butterfill's account of Type 1 mindreading of belief-like states.

This conclusion that the scope of the DSP thesis may be extended to belief-like states is, however, entirely conditional upon the acceptance of the modular view of perception. But the modular view of perception, while influential, is certainly controversial. Many critics reject the modular view's depiction of perception as a wholly bottom-up, informationally-encapsulated process. They instead endorse the *cognitive penetrability* of perception, contending that perception can be influenced by our beliefs and intentions in a top-down fashion (e.g., Siegel, 2010, 2011). What would the rejection of the modular view of perception and adoption of the cognitive penetrability thesis about perception mean for DSP? I only have room here to offer a few speculative conclusions about this important possibility.

First, I think that admitting the cognitive penetrability of perception would not save DSP's original psychological thesis. Recall that DSP's psychological thesis negatively defines direct perception of mental states as not involving theory- or simulation-based inferences. But if perception can be penetrated in a "top-down" way by cognitive processes, and folk psychological theorizing and simulation are cognitive processes, what prevents these cognitive processes from being some of the ones that cognitively penetrate perception?

More significantly, rejecting the modular account's picture of our cognitive architecture may generate a more fundamental threat to DSP. As Shea (2014) argues, accepting a non-modular view of perception may have the consequence of blurring the divide between perception and cognition. How so? Because the modular view seems to define perception in terms of "bottom-up" processes that build up from sensory input, and cognition in terms of "top-down" processes that depend on representational resources that are "higher" in the sense of being further

in the processing hierarchy away from sensory input. And if we reject the modular view's definition of perception and cognition in terms of bottom-up vs. top-down processes, we'd find ourselves with a continuum of psychological processes that vary in how much they involve bottom-up vs. top-down influences. But we may lack a principled reason for calling some perceptual and others cognitive. Accordingly, rejecting the modular, bottom-up view of perception risks losing traction on how to demarcate perception from cognition. And without a strict distinction between perception and cognition, it could be difficult to even formulate DSP as a psychological thesis.

If these speculations hold true, adopting the cognitive penetrability of perception may be problematic for DSP, while its rival, the modular view, may be more accommodating. But adjudicating the debate about the cognitive penetrability or impenetrability of perception is itself a huge project for philosophers and psychologists. One point I hope we can all agree about, however, is that we need to devote more attention to accurately identifying the phenomenological and subpersonal-level processing- and representational-properties of mindreading phenomena—in line with German and Cohen's (2012) cue-based approach to mindreading research. Only with such data will we be able to make empirically informed claims about the architecture of the psychological processes underlying mindreading.

References

Apperly, I. A. (2011). *Mindreaders: The cognitive basis of "theory of mind."* Hove, England: Psychology Press.

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*(4), 953–970.

Block, N. (2014). Seeing-as in the light of vision science. *Philosophy and Phenomenological Research*, *89*(3), 560–572.

Bohl, V., & Gangopadhyay, N. (2014). Theory of mind and the unobservability of other minds. *Philosophical Explorations*, *17*(2), 203–222.

Burge, T. (2010). *The origins of objectivity*. Oxford: Oxford University Press.

Burge, T. (2014). Reply to Block: Adaptation and the upper border of perception. *Philosophy and Phenomenological Research*, *89*(3), 573–583.

Butler, A., Oruc, I., Fox, C. J., & Barton, J. J. (2008). Factors contributing to the adaptation aftereffects of facial expression. *Brain Research*, *1191*, 116–126.

Butterfill, S. A., & Apperly, I. A. (2013a). How to construct a minimal theory of mind. *Mind & Language*, *28*(5), 606–637.

Butterfill, S. A., & Apperly, I. A. (2013b, November 8). Replies to three commentaries on minimal theory of mind. *The Brains Blog.* Retrieved February 1, 2015, from http://philosophyofbrains.com/wp-content/uploads/2013/11/replies-to-commentaries-on-minimal-theory-of-mind.pdf

Carruthers, P. (2013). Mindreading in infancy. *Mind & Language*, *28*(2), 141–172.

Churchland, P. M. (1979). *Scientific realism and the plasticity of mind*. Cambridge: Cambridge University Press.

Cohen, A. S., & German, T. C. (2009). Encoding of others' beliefs without overt instruction. *Cognition*, *111*(3), 356–363.

Cohen, A. S., & German, T. C. (2010). A reaction time advantage for calculating beliefs over

public representations signals domain specificity for "theory of mind." *Cognition*, *115*(3), 417–425.
Evans, J. S. B. T. (2008). Dual-processing accounts of reasoning, judgment, and social cognition. *Annual Review of Psychology*, *59*(1), 255–278.
Evans, J. S. B. T. (2010). *Thinking twice: Two minds in one brain*. Oxford: Oxford University Press.
Evans, J. S. B. T., & Frankish, K. (2009). *In two minds: Dual processes and beyond*. Oxford: Oxford University Press.
Evans, J. S. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science*, *8*(3), 223–241.
Fodor, J. A. (1983). *The modularity of mind*. Cambridge, MA: MIT Press.
Frankish, K. (2010). Dual-process and dual-system theories of reasoning. *Philosophy Compass*, *5*(10), 914–926.
Frankish, K., & Evans, J. S. B. T. (2009). The duality of mind: An historical perspective. In J. S. B. T. Evans & K. Frankish (Eds.), *In two minds: Dual processes and beyond* (pp. 1-29). Oxford: Oxford University Press.
Gallagher, S. (2008a). Direct perception in the intersubjective context. *Consciousness and Cognition*, *17*(2), 535–543.
Gallagher, S. (2008b). Inference or interaction: Social cognition without precursors. *Philosophical Explorations*, *11*(3), 163–174.
Gallagher, S., & Varga, S. (2014). Social constraints on the direct perception of emotions and intentions. *Topoi*, *33*(1), 185–199.
Gao, T., McCarthy, G., & Scholl, B. J. (2010). The wolfpack effect: Perception of animacy irresistibly influences interactive behavior. *Psychological Science*, *21*(12), 1845–1853.
Gao, T., Newman, G. E., & Scholl, B. J. (2009). The psychophysics of chasing: A case study in the perception of animacy. *Cognitive Psychology*, *59*(2), 154–179.
Gao, T., & Scholl, B. J. (2011). Chasing vs. stalking: Interrupting the perception of animacy. *Journal of Experimental Psychology: Human Perception and Performance*, *37*(3), 669–684.
Gangopadhyay, N., & Miyahara, K. (in press). Perception and the problem of access to other minds. *Philosophical Psychology*.
German, T. C., & Cohen, A. S. (2012). A cue-based approach to "theory of mind": re-examining the notion of automaticity. *British Journal of Developmental Psychology*, *30*(1), 45–58.
Goldman, A. (2006). *Simulating minds: The Philosophy, psychology, and neuroscience of mindreading*. Oxford: Oxford University Press.
Gopnik, A., & Wellman, H. M. (1992). Why the child's theory of mind really is a theory. *Mind & Language*, *7*(1-2), 145–171.
Herschbach, M. (2008). Folk psychological and phenomenological accounts of social perception. *Philosophical Explorations*, *11*(3), 223–235.
Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist*, *58*(9), 697–720.
Kahneman, D. (2011). *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.
Kahneman, D., & Frederick, S. (2002). Representativeness revisited. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases* (pp. 49–81). New York.
Kovács, A. M., Teglas, E., & Endress, A. D. (2010). The social sense: Susceptibility to others' beliefs in human infants and adults. *Science*, *330*(6012), 1830–1834.
Krueger, J. (2012). Seeing mind in action. *Phenomenology and the Cognitive Sciences*, *11*(2),

149–173.
Krueger, J., & Overgaard, S. (2012). Seeing subjectivity: Defending a perceptual account of other minds. *ProtoSociology: Consciousness and Subjectivity*, *47*, 239–262.
Kuhn, T. S. (1962). *The structure of scientific revolutions*. Chicago: University of Chicago Press.
Lavelle, J. S. (2012). Theory-theory and the direct perception of mental states. *Review of Philosophy and Psychology*, *3*(2), 213–230.
Low, J., & Perner, J. (2012). Implicit and explicit theory of mind: State of the art. *British Journal of Developmental Psychology*, *30*(1), 1–13.
Low, J., & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mind-reading system. *Psychological Science*, *24*(3), 305–311.
Low, J., Drummond, W., Walmsley, A., & Wang, B. (2014). Representing how rabbits quack and competitors act: limits on preschoolers' efficient ability to track perspective. *Child Development, 85*(4), 1519–1534.
McNeill, W. E. S. (2012). On seeing that someone is angry. *European Journal of Philosophy*, *20*(4)*,* 575–597.
Michael, J., & De Bruin, L. (this issue). How direct is social perception? *Consciousness and Cognition.*
Pylyshyn, Z. (1999). Is vision continuous with cognition? The case for cognitive impenetrability of visual perception. *Behavioral and Brain Sciences*, *22*(3), 341–365.
Scheler, M. (1954). *The nature of sympathy* (P. Heath, Trans.). London: Routledge & Kegan Paul.
Schneider, D., Bayliss, A. P., Becker, S. I., & Dux, P. E. (2012a). Eye movements reveal sustained implicit processing of others' mental states. *Journal of Experimental Psychology: General*, *141*(3), 433–438.
Schneider, D., Lam, R., Bayliss, A. P., & Dux, P. E. (2012b). Cognitive load disrupts implicit theory-of-mind processing. *Psychological Science*, *23*(8), 842–847.
Schneider, D., Nott, Z. E., & Dux, P. E. (2014). Task instructions and implicit theory of mind. *Cognition*, *133*(1), 43–47.
Scholl, B. J., & Gao, T. (2013). Perceiving animacy and intentionality: Visual processing or higher-level judgment? In M. D. Rutherford & V. A. Kuhlmeier (Eds.), *Social perception: Detection and interpretation of animacy, agency, and intention* (pp. 197–229). Cambridge, MA: MIT Press.
Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, *4*(8), 299–309.
Shea, N. (2014). Distinguishing top-down from bottom-up effects. In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and its modalities* (pp. 73–91). New York: Oxford University Press.
Siegel, S. (2010). *The contents of visual experience*. New York: Oxford University Press.
Siegel, S. (2011). Cognitive penetrability and perceptual justification. *Nous*, *46*(2), 201–222.
Smith, J. (2010). Seeing other people. *Philosophy and Phenomenological Research*, *81*(3), 731–748.
Smith, J. (in press). The phenomenology of face-to-face mindreading. *Philosophy and Phenomenological Research*.
Southgate, V., Senju, A., & Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, *18*(7), 587–592.
Spaulding, S. (2010). Embodied cognition and mindreading. *Mind & Language*, *25*(1), 119–140.

Spaulding, S. (in press). On whether we can see intentions. *Pacific Philosophical Quarterly.*
Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: L. Erlbaum Associates.
Stanovich, K. E. (2005). *The robot's rebellion: Finding meaning in the age of Darwin*. Chicago: University of Chicago Press.
Thompson, J. R. (2014). Signature limits in mindreading systems. *Cognitive Science, 38*(7), 1432–1455.
Wertz, A. E., & German, T. C. (2013). Theory of mind in the wild: Toward tackling the challenges of everyday mental state reasoning. *PloS One*, *8*(9), e72835.
Zahavi, D. (2007). Expression and empathy. In D. D. Hutto & M. Ratcliffe (Eds.), *Folk Psychology Re-Assessed* (pp. 25–40). Doredrecht: Springer.
Zahavi, D. (2008). Simulation, projection and empathy. *Consciousness and Cognition*, *17*(2), 514–522.
Zahavi, D. (2011). Empathy and direct social perception: A phenomenological proposal. *Review of Philosophy and Psychology*, *2*(3), 541–558.