

## VISUALLY PERCEIVING THE INTENTIONS OF OTHERS

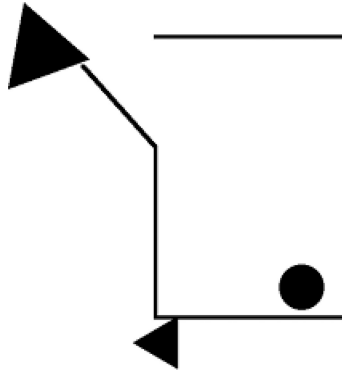
BY GRACE HELTON

*I argue that we sometimes visually perceive the intentions of others. Just as we can see something as blue or as moving to the left, so too can we see someone as intending to evade detection or as aiming to traverse a physical obstacle. I consider the typical subject presented with the Heider and Simmel movie, a widely studied ‘animacy’ stimulus, and I argue that this subject mentally attributes proximal intentions to some of the objects in the movie. I further argue that these attributions are unrevisable in a certain sense and that this result can be used to as part of an argument that these attributions are not post-perceptual thoughts. Finally, I suggest that if these attributions are visual experiences, and more particularly visual illusions, their unrevisability can be satisfyingly explained, by appealing to the mechanisms which underlie visual illusions more generally.*

**Keywords:** visual perception, high-level perception, social cognition, dual-systems theory, action, intention.

In a ground-breaking 1944 study, the psychologists Fritz Heider and Marianne Simmel presented subjects with a short film depicting a small triangle, a large triangle, and a circle moving about a stationary, partially filled-in square (see Figure 1). When asked to describe the events in the film, most subjects produced narratives which ascribed intentions and desires to the two-dimensional shapes. For instance, in one scene of the film, the large triangle moves to the right of the screen, a short distance behind the small triangle and circle, which also move to the right. Subjects described this scene as one in which the large triangle *chases* the other shapes and one in which the smaller shapes *flee* from the triangle (Heider and Simmel 1944). Though subjects might have described this and other scenes in purely geometric terms, their descriptions consistently treated the two-dimensional figures as agents performing actions.

It is difficult for subjects to interpret the movie in terms that do not involve actions and aims. Even when subjects are instructed to provide a narrative in terms of purely physical, geometric terms, they struggle to do so (Hashimoto 1966). Viewing the movie for oneself makes it clear why this is so. The figures just *seem* to be fleeing, hiding, chasing, and even fighting; this reading is effortless and immediate.



**Figure 1.** Screenshot from the Heider & Simmel movie.

These results suggest that the typical subject viewing the Heider and Simmel movie *experiences* the two-dimensional shapes as fleeing, hiding, and chasing, in some sense of ‘experiences’. The question remains what kind of experience this is. What kind of mental attitude or attitudes make up this experience? To keep things concrete, I will focus on the representation of the smaller shapes fleeing the large triangle, which I will refer to as *the ‘fleeing’ mental state*. On one view, the subject in this state *visually perceives* the smaller shapes as fleeing the triangle (Carruthers 2015; Gallagher 2008; Scholl and Gao 2013; Scholl and Tremoulet 2000). On another view, the subject in this state merely perceives shape, colour, and movement features, and the perception of these features quickly and automatically generates a post-perceptual *thought* that the smaller shapes are fleeing the triangle (Butterfill 2009: 408–9).<sup>1</sup> These views don’t exhaust possible explanations of the ‘fleeing’ state, but I take them to be more plausible than other explanations (I defend this presumption in Sections I and V.2).

I will argue that the typical subject viewing the Heider and Simmel movie visually perceives the smaller shapes as fleeing the triangle. Moreover, in visually perceiving the smaller shapes as fleeing the triangle, this subject perceives the smaller shapes as acting on a certain *proximal* intention, the intention of evading the large triangle. Thus, just as we can visually perceive something as blue or as moving to the left, so too can we visually perceive the intentions of others. In the case of the Heider and Simmel movie, the visual experience

<sup>1</sup> Prinz (2005, 2007) rejects the view that we perceive features beyond ‘intermediate-level’ features such as shape, colour, and motion. Spaulding (2015) permits that we might perceive motor intentions but denies that we perceive other kinds of intentions. Shea (2014: 85–6), following Carey (2009), places the representation of animacy in a borderline case, of a kind that is neither determinately cognitive nor determinately perceptual.

as of the intentions of others is a visual illusion, on a par with more familiar illusions such as the Müller-Lyer Illusion.

Though I will ultimately reject *the cognitive explanation* of the 'fleeing' mental state, on which this state is a thought, one of the aims of this paper is to show that the cognitive explanation is much more resilient than its opponents have supposed. In particular, a version of the cognitive explanation which appeals to automatically and quickly generated thoughts—what are sometimes called *System 1* thoughts—can accommodate the phenomenology of the 'fleeing' state.

The argument I offer in favour of the view that we can visually perceive others' intentions is rooted in the claim that the 'fleeing' mental state is, in a certain sense, *unrevisable*. This is the claim that the subject who tokens the 'fleeing' state is psychologically incapable of rationally revising that state in response to at least some kinds of conflicting evidence. (In Section IV, I clarify this claim and distinguish it from the stronger claim that the 'fleeing' state is *cognitively impenetrable*.) Drawing on empirical evidence, I mount an inductive argument that the view that the 'fleeing' state is a thought cannot accommodate the fact that the 'fleeing' state is unrevisable. This is because thoughts have a capacity to be rationally revised in response to conflicting evidence, though this capacity is often not realized. (Not all thoughts have this capacity; delusions lack it. In Section V.3, I offer different grounds for thinking that the 'fleeing' state is not a delusion.) In contrast, the view that the 'fleeing' state is a visual experience can satisfyingly explain this state's unrevisability.

The strategy I employ to argue that the 'fleeing' mental state is a visual experience reflects a deeper suggestion about what, in general, might distinguish perceptual states from the thoughts they elicit. The suggestion is that at least some perceptual states are not rationally revisable in response to counterevidence, whereas thoughts are rationally revisable in response to such evidence.

Here is the plan for the paper. In Section I, I clarify the thesis and suggest some reasons it matters. In Section II, I argue that the visual explanation of the 'fleeing' mental state is initially plausible. In Section III, I argue that a certain construal of the cognitive explanation is also initially plausible. In Section IV, I adduce empirical evidence in favour of the view that the 'fleeing' state is unrevisable. In Section V, I argue that the cognitive explanation of the 'fleeing' state cannot account for the fact that the 'fleeing' state is unrevisable. In Section VI, I show how the visual explanation can satisfyingly explain the unrevisability of the 'fleeing' state, and I conclude in favour of this explanation.

## I. THE THESIS AND WHY IT MATTERS

The thesis of this paper is that we sometimes visually perceive the intentions of others. In the relevant sense, visual perception is the kind of environmentally

triggered experience whose paradigm instances include seeing something as blue and seeing something as spherical. If the present thesis is true, then there is a kind of parity between the paradigm instances of visual perception and the visual perception of the intentions of others: if and insofar as a completed cognitive science must account for the visual experience of (say) colour, so too must a completed cognitive science account for the visual experience of the intentions of others.

Relatedly, the thesis that we visually perceive the intentions of others will not turn out to be true if we merely see others' intentions in an extended sense of 'see', e.g. in the sense of 'see' in which one might read pre-election polls and *see* who the election's victor will be.

Notice that the claim that we visually perceive others' intentions is neutral on the question of whether the relevant visual experiences are themselves partly constituted by cognitive elements.<sup>2</sup> The paper's thesis is that others' intentions are sometimes represented by visual experiences, not merely by *post*-perceptual thoughts.

The claim that we visually perceive the intentions of others is the claim that we sometimes enjoy visual experiences *as of* the intentions of others. These experiences are not necessarily accurate. Just as one might inaccurately perceive a red object *as* blue or a two-dimensional object *as* three-dimensional, so too might one inaccurately perceive some entity which lacks a certain intention *as* possessing that intention. Relatedly, the locution 'perception *as of*  $\Phi$ ' is conventionally non-committal as to whether the relevant perceptual state is accurate, and the locution 'perception *of*  $\Phi$ ' is conventionally reserved for accurate perception. I generally adhere to this convention but occasionally abandon it for the sake of naturalness. Throughout, all talk of perception should be understood to be talk of 'as of' perception.

The present thesis is the claim that our visual experiences sometimes ascribe to others the feature of *in fact* possessing a certain intention. This thesis will not turn out to be true if visual experiences only ever ascribe to others the perceiver-dependent feature of *seeming* to possess a certain intention. Relatedly, I am presuming a substantive, if widely held, view about perceptual content: that perception sometimes ascribes perceiver-independent features to objects, such as *being spherical* or *being grooved*. This view contrasts with the view that perception only ever ascribes perceiver-dependent features to objects, such as *seeming spherical* or *seeming grooved* (Glüer 2009; Cf. Brogaard 2014).

The thesis should further be clarified as the claim that *typical* humans sometimes perceive the intentions of others. If humans have this capacity, they have it contingently. Such humans might have been otherwise, as demonstrated by the fact that many humans *are* otherwise. For instance, subjects who are autistic

<sup>2</sup> For defences of this view, see: Byrne (2009), Carruthers (2015), Glüer (2009), and Quilty-Dunn (2015).

or who have legions of the amygdala do not visually perceive the intentions of others (Congiu, Schlottmann and Ray 2010; Heberlein and Adolphs 2004).

The argument I will make in favour of the thesis draws on cases of inaccurate representation and thus, can at most establish that there are visual illusions as of the intentions of others. One might reasonably wonder whether we ever accurately visually perceive others' intentions. I am presuming that if we ever inaccurately perceive others' intentions, then it is overwhelmingly likely that we also sometimes accurately perceive others' intentions. For it would be very surprising if the inaccurate representation as of the smaller shapes fleeing the triangle were a visual experience, but the accurate representation as of (say) a cat fleeing an overly friendly toddler were not a visual experience. Given the subjective and functional similarities of these representations, we can safely assume that they belong to the same mental kind.

Finally, the present thesis is restricted to the claim that we can visually perceive proximal intentions that are bound up with paradigm prey/predator activities, such as fleeing, chasing, fighting, seeking, and hiding. I take it to be an open question whether other intentions are perceived. My reason for focusing on these intentions is that their representations are cross-culturally invariant and diachronically stable, and my argument exploits these features.

Likewise, the present thesis is neutral on the question of whether mental states other than intentions—such as emotions and thoughts—are perceived. It is feasible that the present argument might be suitably modified to establish the perception of other mental states, but determining whether the argument might be so extended would require an examination of several issues which are beyond the scope of this paper.<sup>3</sup>

What is the import of the claim that we perceive intentions that are bound up with prey/predator activities? One reason this claim matters is that it might explain how typical subjects form at least some of their beliefs about others' mental states. The thesis would at least clear ground for a *direct perceptual view* of acquiring beliefs about others, on which subjects at least sometimes form beliefs about others' mental states via perception, without any mediating inference (see, e.g. Gallagher 2008).

Another reason it matters whether typical subjects perceive others' intentions is that this claim might illuminate certain *disorders of theory of mind*, such as autism spectrum disorders, since these involve deficits in attributing intentions (Congiu, Schlottmann and Ray 2010). Of course, such disorders do not merely involve deficits in the attributions of prey/predator intentions. Nevertheless, the result that non-autistic subjects perceive such intentions, whereas autistic

<sup>3</sup> One such issue is whether emotions or other mental states partly constitute some of the actions we perceive; for my argument relies on the claim that certain intentions partly constitute some of the actions we perceive.

subjects do not, raises the possibility that some of the other deficits associated with autism are rooted in a more basic perceptual deficit.

Another reason it matters whether we can visually perceive others' intentions is that if we can, this would tell in favour of the disputed claim that we can perceive at least some *high-level features*. High-level features are any features other than *low-level features*, whose paradigmatic instances in the visual domain include shape, colour, and motion. It has recently been proposed that we can visually perceive a range of high-level features, including artefactual features (e.g. *being a guitar*), absence features, event-causal features, natural kind features (e.g. *being a lemur*), social features (e.g. *being masculine*), and moral features (Audi 2013; Bayne 2008, 2009; Block 2014; Butterfill 2009; Chappell 2008; Cullison 2010; Farennikova 2013, Fish 2012; Nanay 2011a; Scholl and Tremoulet 2000; Siegel 2005, 2010; Sorensen 2012; cf. Prinz 2005, 2007; Rips 2011).<sup>4</sup> My argument that we visually perceive the intentions of others might in principle be used to show that we visually perceive other kinds of high-level features.

Before proceeding to the visual and cognitive explanations of the 'fleeing' state, it may be worth justifying why I focus on these views. Why, for instance, do I not take seriously the view that the 'fleeing' state is a merely entertained thought, of the sort the subject does not endorse?

Merely entertained thoughts typically cease to exist when certain of their subject's aims are fulfilled (Shah and Velleman 2005: 498). For instance, the geometry student tasked with proving that no triangle has four sides might, for the purpose of proving otherwise, entertain the thought that there exists some triangle that has four sides. Once this student has derived a contradiction, she will no longer retain this assumption, since it has served its purpose. In contrast, the 'fleeing' mental state persists so long as the relevant stimulus is present, whether or not this state has satisfied its subject's aims. So, I take it to be unlikely that the 'fleeing' state is a merely entertained thought.

## II. THE INITIAL PLAUSIBILITY OF THE VISUAL EXPLANATION

On the visual explanation of the 'fleeing' mental state, this state is a visual experience. But what is it to see someone *fleeing*? In this section, I argue that the visual experience as of someone fleeing is, at least in part, a visual experience as of someone acting on a proximal intention of evasion. There are three steps in this argument: first, the activity of fleeing is partly made up of a proximal intention of evasion. Second, the fact that fleeing constitutively involves such an intention is an easily conceptually accessible fact. Third,

<sup>4</sup> For a discussion of the relevance of high-level perception for epistemology and philosophy of action, see Helton (2016: 853).

a certain *perceptual closure principle* holds: if some feature  $\Phi$  is at least partly constituted by some feature  $\Psi$  and this relation is easily conceptually accessible, then to perceive  $\Phi$  is thereby to perceive  $\Psi$ .

### II.1. *The 'fleeing' mental state as attributing intention*

It is essential to my argument that subjects who view the Heider and Simmel movie represent the figures as performing actions, not merely as moving. Actions necessarily involve agents and, at least very often, aims. Recall that subjects describe the figures' movements in richly agential terms, such as *fleeing*, *hiding*, *fighting*, and *chasing*. Recall further that subjects struggle to describe these movements in purely geometric terms, such as *moving to the right* or *moving in front of* (Hashimoto 1966; Heider and Simmel 1944). The most straightforward explanation of these results is that subjects represent the figures as performing actions, such as fleeing and chasing, not merely as moving.

At least some actions, such as fleeing, are partly constituted by proximal intentions. Before arguing for this claim, it may be useful to illustrate the difference between proximal and *distal* intentions: my six-year-old nephew sometimes catches my eye, grins maniacally, and then runs away from me. He knows that when he does this, I'm likely to chase him, catch him, and tickle him. In running away from me, my nephew acts on the distal intention of getting caught, since his ultimate aim is to get me to catch him and tickle him. But in fleeing, he acts on the proximal intention of evading me. My nephew's distal and proximal intentions thus come apart.<sup>5</sup>

The following example suggests that a proximal intention of evasion is necessary for some act to count as fleeing: suppose you go on a jog in a wooded area. Unbeknownst to you, a curious squirrel scampers after you. Intuitively, you do not flee the squirrel, even though you do self-propel away from the squirrel. Plausibly, the reason your self-propulsion away from the squirrel does not count as an act of fleeing the squirrel is that you lack a proximal intention to *evade* the squirrel. This suggests that such an intention is a necessary condition on fleeing.

The fact that fleeing is partly made up of a proximal intention of evasion is a conceptually accessible fact; this fact can be accessed just by reflecting on cases such as the 'squirrel' case, without empirical investigation. In this way, this fact is unlike the fact that the sun is larger than the earth and unlike the fact that water is composed of H<sub>2</sub>O. Moreover, the fact that fleeing is made up of a proximal intention of evasion is *easily accessed* by reflecting on relevant cases. Discovering this fact does not require extensive or effortful deliberation, nor does it require theoretically sophisticated background assumptions beyond those demanded by ordinary reasoning.

<sup>5</sup> This distinction roughly follows Mele (1992).

Why does it matter that the relation between fleeing and an intention of evasion is easily conceptually accessible? The relevance is that perceptual content adheres, I submit, to the following *perceptual closure principle*:

If: S perceives  $\Phi$ ,  $\Phi$  is at least partly constituted by some feature  $\Psi$ , and this relation is easily conceptually accessible, then: in perceiving  $\Phi$ , S thereby perceives  $\Psi$ .

The reason to accept the perceptual closure principle is that it gives the intuitively correct result in a wide range of cases. For instance, it permits us to say that the subject who perceives some entity as square thereby perceives that entity as a closed geometric figure; that the subject who perceives some entity as azure thereby perceives that entity as blue; and that the subject who perceives some entity as moving thereby perceives that entity as changing in location.

At the same time, the perceptual closure principle does not force us to accept certain less-than-intuitive results. For instance, it does not force us to say that subjects perceive all features which are metaphysically entailed by those they perceive. Suppose that as it turns out, every feature metaphysically entails the feature *not being a square circle*. The closure principle does not predict that the subject who perceives (say) redness thereby also perceives the feature *not being a square circle*. This is because the feature *not being a square circle* does not constitute the feature *being red*, and the closure principle pertains only to features which at least partly constitute perceived features.

Likewise, the perceptual closure principle is neutral on the controversial question of whether we perceive natural kinds. For instance, it is neutral on the question of whether the subject who perceives a glass of water thereby perceives  $H_2O$  and on whether the subject who perceives a patch of light thereby perceives electromagnetic radiation. Since neither the relation between water and  $H_2O$  nor the relation between light and electromagnetic radiation is conceptually accessible, the principle is silent on such cases.

Finally, the perceptual closure principle does not force us to say that we can perceive arcane aspects of perceived features. For instance, there is a lively philosophical literature on the metaphysics of holes. One controversy is whether holes are immaterial (Casati and Varzi 2014). Suppose that as it turns out, holes are immaterial. Suppose further that this fact is in principle conceptually accessible. The perceptual closure principle would not force us to say that the subject who perceives some entity as a hole thereby perceives that entity as immaterial. The perceptual closure principle pertains only to easily accessible facts, and the fact that holes are immaterial particulars is presumably not easily accessible, as evidenced by the large and contentious philosophical literature devoted to settling this fact.<sup>6</sup>

<sup>6</sup> The question of whether we perceive holes' arcane traits is from Siegel (2009).



So far, I've suggested that fleeing is partly comprised of a proximal intention of evasion and that this fact is easily conceptually accessible. By combining these claims with the perceptual closure principle, we obtain the result that to visually perceive some entity as fleeing is thereby to perceive that entity as proximally intending to evade someone or something else. Going forward, I will assume that this claim is true and will focus on defending the claim that subjects viewing the Heider and Simmel figure *visually perceive* the figures as fleeing, and do not merely judge, think, or otherwise non-perceptually represent these figures as fleeing.

## *II.2. An initial defence of the visual explanation*

Initial evidence for the view that the 'fleeing' mental state is a perceptual experience comes from the fact that the 'fleeing' state is generated quickly, effortlessly, and apparently non-inferentially. The view that the 'fleeing' state is a *visual* experience—not an auditory or other sort of perceptual experience—is supported by the fact that the 'fleeing' state is dependent on certain features present in the subject's low-level visual experience, such as directionality and speed (Gao, Newman and Scholl 2009; Gao and Scholl 2011; Scholl and Gao 2013).

One worry about the visual explanation of the 'fleeing' mental state stems from the fact that the nature of intentions is a fraught theoretical issue. It is controversial, for instance, whether intentions are causally efficacious in action and whether unforeseen consequences can count as intentional (Setiya 2015).

There are two worries in this vicinity. The first is the methodological concern that we will not be able to determine whether we see intentions prior to settling on a theory of the nature of intentions. The second is the concern that if intentions turn out to have a conceptually sophisticated nature, then it is implausible that they are perceived, since our ordinary visual experiences do not seem to attribute conceptually sophisticated features to the world.

Both of these concerns can be addressed simultaneously: if we perceive intentions at all, we perceive them without thereby perceiving arcane aspects of their nature. In general, perception as of  $\Phi$  does not typically (if ever) involve the perception as of the arcane features of  $\Phi$ . This point is cogently argued by Susanna Siegel, who makes it by reflecting on the question, already discussed, of what it is to perceive a hole:

Suppose you see some cheese with a hole in it . . . According to one theory of the metaphysics of holes, they are immaterial particulars. According to an opposing theory, they are material but negative parts of material particulars. When the cheese appears to have a hole in it, does it appear to host an immaterial particular, or does it appear to have a material part – or neither, or both? . . . The best answer seems to be 'neither: '

visual experience just seems neutral on whether the hole in the cheese is a material but negative part or an immaterial part hosted by a material particular. (Siegel 2009: 539)<sup>7</sup>

Just as we can perceive the hole in a slice of Swiss cheese without perceiving it as (say) a negative part of a material particular, so too can we perceive someone as proximally intending to evade someone or something else without thereby perceiving that intention as (say) necessarily causally efficacious. Our experience of others' intentions is simply silent about some of the more obscure metaphysical questions about intentions. Moreover, this result is consistent with the previously defended claim that easily accessible aspects of intentions are perceived.

### III. THE INITIAL PLAUSIBILITY OF THE COGNITIVE EXPLANATION

On the cognitive explanation, the 'fleeing' mental state involves visual perception as of shape, colour, and movement features, combined with a post-perceptual thought, 'the smaller shapes are fleeing the triangle'. Visual experience itself does not represent the shapes as fleeing.

Because the 'fleeing' mental state is automatic, felt to be immediate, and unmediated by conscious inference, it might seem implausible that the 'fleeing' state is a thought. Such features are typically associated with perceptual states, not with thoughts. Indeed, some advocates of the view that we visually perceive the intentions of others have taken the presence of these features to tell against a cognitive explanation of experiences like the 'fleeing' state (Gallagher 2008: 536–8; Scholl and Gao 2013, Scholl and Tremoulet 2000: 299). In this section, I show that a version of the cognitive explanation that appeals to *System 1* thoughts can account for these aspects of the 'fleeing' state.

System 1 thoughts are formed automatically, quickly, and as the result of non-conscious, heuristic processing. Though the product of non-conscious processing, System 1 thoughts are often themselves conscious. System 1 thoughts contrast with *System 2* thoughts, which are formed deliberately, slowly, and as the result of explicit inference (Evans 2008). The following problem can be used to help illustrate System 1 thoughts. Try to answer it as quickly as you can:

If it takes 5 machines 5 minutes to make 5 widgets, how long would it take 100 machines to make 100 widgets?

If you are like most people, the answer *100 minutes* probably popped into your head (Thompson 2009). However, the correct answer to the question is

<sup>7</sup> Nanay (2011b: 305–6) makes a similar point.

5 minutes. The fact that most participants give an incorrect answer is somewhat surprising, given that the correct answer can be arrived at using a bit of simple algebra.<sup>8</sup>

By what process do subjects generate the incorrect answer? It appears that subjects rely on a *matching heuristic*. Subjects notice the pattern 5:5:5 and the partially completed pattern 100:100: --, and they generate 100 minutes as the correct answer (Thompson 2009).

Though subjects appear to be relying on a matching heuristic, it is very unlikely that subjects are aware that they are doing this. That is, this answer does not appear to be the result of conscious inference. Conscious inference is time consuming, and the incorrect answer is produced very quickly. Moreover, had subjects engaged in explicit inference, it is unlikely they would have opted for a matching heuristic, since a moment's consideration shows that the problem is not a matching task. Thus, on the best explanation, the attractive but incorrect response is the conscious output of a process which is itself non-conscious. Relatedly, the incorrect answer to the problem is not associated with a feeling of effort. One simply 'sees' the answer. In these respects, the incorrect response to the 'widgets' problem is a paradigm System 1 thought.

A version of the cognitive explanation which appeals to System 1 thoughts can accommodate the phenomenology of the 'fleeing' state. According to such an explanation, it is because the 'fleeing' thought is generated very quickly after the onset of the visual perception of shape and colour that there is no experienced temporal gap between the perception of these features and the thought, 'the smaller shapes are fleeing the triangle'. Further, the 'fleeing' thought is formed without conscious initiation or effort, which is why the subject does not experience initiation or effort. Finally, because the 'fleeing' thought is the conscious output of a non-conscious process, subjects are not consciously aware of forming an inference. In sum, the view that the 'fleeing' state is a System 1 thought can account for the fact that it just *seems* as though the smaller shapes are fleeing the triangle.

#### IV. THE UNREVISABILITY OF THE 'FLEEING' MENTAL STATE

In this section, I argue that the 'fleeing' mental state is, in a certain sense, *unrevisable*. I will ultimately argue that the cognitive explanation cannot accommodate this fact, whereas the visual explanation can. The 'fleeing' state is unrevisable in the following sense:

<sup>8</sup> Since 5 machines work at a rate of 1 widget per minute, 100 machines work at a rate of 20 widgets per minute. So, it will take 100 machines 5 minutes to complete 100 widgets. This task is originally from the Cognitive Reflection Task (Frederick 2005).

UNREVISABILITY (dfn.) Let  $m$  be a mental state which represents  $p$  and which is tokened in some subject  $S$ . Then,  $m$  is unrevisable just in case there exists at least some kind of evidence  $e$  such that  $m$  is not rationally revised in response to  $e$ , in all worlds in which:

- (i)  $S$ 's core psychological mechanisms are held fixed,
- (ii)  $e$  is among  $S$ 's evidence, and
- (iii)  $e$  is all-things-considered evidence that  $p$  is false.

Unrevisability is a kind of nomic incapacity and more particularly, a kind of psychological incapacity. Thus, for some mental state to be unrevisable, it is not necessary that it be unrevised in all metaphysically possible scenarios. It need only be unrevised in all psychologically possible scenarios, and more particularly, in all scenarios in which the relevant subject's core psychological mechanisms are held fixed.<sup>9</sup>

Some psychological mechanism is a core psychological mechanism for some subject just in case that mechanism is counterfactually robust and governs the revision of at least some of that subject's attitudes. Different kinds of attitudes are likely governed by different mechanisms of revision—e.g. emotions are likely governed by different mechanisms than those which govern thoughts or perceptual states. All of these are core mechanisms.

Even if we restrict our attention to typical adults, any attempted list of core psychological mechanisms will be vast and inevitably incomplete. Nevertheless, we can illustrate the notion of a core mechanism by considering the phenomenon of *cognitive dissonance*, which plays an important role in attitude regulation. Cognitive dissonance includes a mechanism which triggers a feeling of discomfort in response to inconsistent attitudes and further mechanisms which dispose subjects in a dissonant state to resolve their attitudes in the direction of greater mutual consistency (Harmon-Jones and Harmon-Jones 2007). All of these mechanisms are, for typical adults, core mechanisms, since they are counterfactually stable mechanisms which govern the revision of attitudes.

For a mental state to count as unrevisable, it must be incapable of *rational* revision. For a mental state to be rationally revised in response to counterevidence, that state must be reduced in strength or extinguished altogether; that state must not increase in strength or remain unchanged in strength (See

<sup>9</sup> One might wonder why I do not focus instead on the claim that the 'fleeing' state can exist at the same time as a thought that conflicts with it. If thoughts seldom or never conflict, this claim gives us some reason to think that the 'fleeing' state is not a thought. There is a reason to suppose that thoughts can and often do conflict (Byrne 2009; Evans 2008; Gertler 2011; Helton 2015; Huddleston 2012; Mandelbaum 2014; Mylopoulos 2015; Quilty-Dunn 2015). Superstitious views can conflict with one's better judgement (Huddleston 2012). Certain highly intuitive but incorrect responses to mathematical puzzles can persist even after one learns those responses to be incorrect (Helton 2015: 90–1; Mylopoulos 2015: 768). So, I take the claim that the 'fleeing' state can exist at the same time as a thought that conflicts with it to be insufficient to establish that the 'fleeing' state is not itself a thought.

footnote 9). Further, this state must be revised because of the evidence that conflicts with it, not because of non-evidential factors, such as prudential factors. Finally, the revision must occur via a non-deviant route. Suppose you should read in the paper that a generally favoured candidate is behind in exit polls and the shock causes you to fall and hit your head. Should the blow fortuitously cause you to lose your belief that the favoured candidate will win, your belief is revised by relevant evidence, but the route to revision is deviant, so the revision is not rational.

It is important to note that unrevisability is weaker than the familiar trait of *cognitive impenetrability*. Some mental state is cognitively impenetrable only if that mental state is rationally insusceptible to *all* sorts of counterevidence (Macpherson 2012). In contrast, some mental state is unrevisable just in case that mental state is rationally insusceptible to *at least some* sorts of counterevidence. Thus, every cognitively impenetrable mental state is unrevisable, but not every unrevisable mental state is cognitively impenetrable.

The difference between cognitive impenetrability and unrevisability can be illustrated as follows: suppose that some subject visually perceives a pineapple's shape. Suppose further that this subject's experience of the pineapple's shape is not rationally responsive to any kind of counterevidence—whether testimony, previously observed environmental patterns, conflicting information from other sensory modalities, or anything else. In this case, the visual perception of the pineapple's shape is cognitively impenetrable. It is therefore also unrevisable.

But now, suppose that the subject's visual experience of the pineapple's shape is rationally responsive to some kinds of counterevidence, but not other kinds. For instance, suppose the experience is rationally responsive to previously observed patterns in the environment but not to reliable testimony. In this case, the visual experience of the pineapple's shape is unrevisable, because there are some kinds of counterevidence which do not rationally influence it. At the same time, this experience is not cognitively impenetrable, since there are other kinds of counterevidence which do rationally influence it.

To show, then, that the 'fleeing' mental state is unrevisable, we need only establish that there are at least some kinds of counterevidence which have no rational impact on that state. We needn't additionally establish that the 'fleeing' state is rationally insusceptible to all forms of counterevidence, i.e. that it is cognitively impenetrable.

#### *IV.1. An argument for the 'fleeing' mental state's unrevisability*

In this section, I argue that the 'fleeing' mental state is unrevisable. I will ultimately draw on this result to argue that the 'fleeing' state is perceptual. I do not claim that all perceptual experiences are unrevisable. But for those that are, their unrevisability can be used as part of an argument to establish that they are perceptual.

Here is an inductive argument that the ‘fleeing’ mental state is unrevisable. Let *S* be a typical subject viewing the Heider and Simmel movie under normal viewing conditions:

- (1) *S* represents the smaller shapes as fleeing the triangle.
- (2) *S* enjoys all-things-considered evidence that the smaller shapes are not fleeing the triangle.
- (3) *S*’s representation as of the smaller shapes fleeing the triangle does not diminish in strength or disappear in response to the evidence that that representation is inaccurate.
- (4) *S*’s representation as of the smaller shapes fleeing the triangle is such that it cannot be rationally revised in response to at least some kinds of counter-evidence, i.e. the ‘fleeing’ mental state is unrevisable.

First to consider (1), which states that the typical subject viewing the Heider and Simmel movie represents the smaller shapes as fleeing the triangle. The reasons for accepting this claim have been previously enumerated: subjects describe the movie in terms that ascribe intentions to the animated figures, and this interpretation is irresistible. Introspection further supports the view that in viewing the Heider and Simmel movie, it just *seems* as though the shapes are fleeing, chasing, hiding, and so on.

Why accept (2), that the typical subject who tokens the ‘fleeing’ mental state possesses all-things-considered evidence that this mental state is inaccurate? Consider that the typical subject viewing the Heider and Simmel movie is likely very familiar with the medium of two-dimensional animation. So, we can safely suppose that this subject enjoys all-things-considered evidence that the figures in the animation are not *really* acting on intentions. This subject understands that these shapes are inanimate geometric shapes, incapable of acting or intending.

One concern about (2) is that it seems possible that the typical subject might *selectively* access her evidence that the figures are inanimate, flipping between this evidence and the fantasy that the figures are alive. On this view, this subject never accesses the relevant evidence *at the same time* that she tokens the ‘fleeing’ state. If this is right, then there is never a simultaneous conflict between the subject’s accessed evidence and her representation.

Here is a reason to think that subjects at least sometimes access evidence that the figures are not animate at the very same time that subjects represent the figures as fleeing and engaging in other actions. Most subjects viewing the movie for the first time find it surprising and, relatedly, amusing. This suggests that subjects are aware that what they are experiencing is inaccurate even while they are in the grip of the illusion. For if this were not the case, it would be difficult to explain why the ‘fleeing’ representation should be surprising, since surprise necessarily involves a violation of expectations. Subjects should

find the movie no more surprising than seeing a blade of grass as green or than feeling an unsanded bit of wood as rough.

A different objection to (2) stems from the fact that the creatures in the Heider and Simmel movie inhabit a fictional world. Evaluated relative to the context of the relevant fictional world, it is *true* that the smaller shapes flee the triangle. Similarly, evaluated relative to the fictional world of *Pride and Prejudice*, it is *true* that Miss Bennet wishes to marry Mr Darcy. If the representation as of the smaller shapes fleeing the triangle is evaluated relative to the fictional world which these shapes inhabit, then the typical subject should lack evidence that this representation is inaccurate. For evaluated relative to the fictional context, this representation is accurate.

As it turns out, the fact that the Heider and Simmel movie describes a fictional world does not explain why it elicits an unrevisable representation of others' intentions. For there are contexts which don't involve fiction in which subjects represent others as acting and in which these representations appear insusceptible to counterevidence.

For instance, in one study, subjects reported that certain robots seemed to them to have a capacity to plan and a capacity to exercise self-control, even though the robots in question did not appear human-like or biological in nature (Gray and Wegner 2012: 126–7). Since these robots were of a crude mechanical design, subjects were almost certainly aware that these robots did not in fact possess agency. Moreover, since these robots were three-dimensional entities, it seems unlikely that the subjects regarded them as inhabiting a fictional scenario.

Going forward, I will continue to talk in terms of the Heider and Simmel movie, since this stimulus has been empirically investigated in respects that are important to my argument. But I will help myself to the assumption that the fact that the Heider and Simmel movie depicts a fictional world is inessential to an explanation of why the movie gives rise to an unrevisable representation as of others' intentions.

Next, to defend (3), which states that the 'fleeing' mental state is neither reduced in strength nor extinguished in response to all-things-considered evidence that it is inaccurate. The support for (3) comes from a number of studies of the Heider and Simmel movie and of relevantly similar 'animacy' stimuli, which show that virtually all subjects represent the figures in these stimuli as animate (Scholl and Gao 2013). Indeed, the only subjects who do not experience the 'animacy' interpretation are those who have an autism spectrum disorder or lesions of the amygdala (Abell, Happé and Frith 2000; Heberlein and Adolphs 2004; Klin 2000; Rutherford, Pennington and Rogers 2006). If the typical subject's 'fleeing' state were in some cases revised in response to conflicting evidence, we would expect some subjects to report that the 'fleeing' state is diminished in strength or disappears after onset. But we do not find this.

The claim that the ‘fleeing’ mental state is unrevisable is further supported by the fact that the ‘fleeing’ state remains robust for subjects across multiple viewings. For instance, many researchers working on the psychology of animacy have seen the Heider and Simmel movie countless times. These researchers continue to experience the smaller shapes as fleeing the triangle, even after multiple viewings (Scholl and Gao 2013). So, permitting subjects additional time to assimilate the counterevidence does not seem to impact the ‘fleeing’ state.

I have argued that the typical subject viewing the Heider and Simmel movie has a counterfactually robust tendency to represent the smaller shapes as fleeing the triangle and that this tendency is not responsive to evidence that the shapes aren’t really fleeing the triangle. These claims together give us some reason to think that the typical subject’s ‘fleeing’ mental state is not capable of being rationally revised in response to at least some sorts of counterevidence, i.e. that the ‘fleeing’ state is unrevisable.

#### *IV.2. Objections to the unrevisability claim*

Since the argument presented is inductive, it purports to show that (1)–(3) together probabilify (4). The argument does not purport to show that (1)–(3) together entail (4). Thus, one might grant that (1)–(3) are true but maintain that there are other grounds for rejecting the conclusion that the ‘fleeing’ mental state is unrevisable. I will consider and reject two such grounds.

A first reason to doubt that the ‘fleeing’ mental state is unrevisable comes from the fact that tested subjects tend to be highly educated members of industrially developed countries. These subjects are likely to have been extensively exposed to cartoons and graphically rich video games. It is not unreasonable to suppose that this exposure has trained such subjects to suspend reality when presented with animations. Thus, it seems at least plausible that such subjects have a masked capacity to revise the ‘fleeing’ state in response to counterevidence, a capacity which is obscured by their culturally entrained preference to indulge fantasy over reality.

Cross-cultural research on animacy stimuli refutes the hypothesis that subjects’ familiarity with animated media explains their tendency to experience the smaller shapes as fleeing the triangle. Both Germans and hunter-horticulturist Shuar subjects token ‘animacy’ interpretations in response to ‘animacy’ stimuli (Barrett *et al.* 2005). The fact that the ‘animacy’ interpretation arises for subjects very seldom exposed to animated media suggests that the tendency of subjects to represent the Heider and Simmel figures as animate is not the product of culturally specific training. So, there is no reason to think that subjects could suspend their representations of the figures as animate, were they to choose to.



A second reason to doubt that the 'fleeing' mental state is unrevisable comes from the suspicion that the 'fleeing' state might be destroyed by means of a broadly Pavlovian process of conditioning. For instance, perhaps electrically shocking subjects at appropriate times would destroy their tendency to represent the Heider and Simmel figures as animate.

If it should turn out that a process of conditioning could extinguish the 'fleeing' mental state, this result would not show that the 'fleeing' state is revisable in the relevant sense. Recall that unrevisability is a kind of incapacity to be *rationally* revised. Pavlovian interventions do not involve evidence; rather, they forge positive or negative associations with relevant stimuli. Hence, such interventions cannot bring about a rational revision.

I conclude that the 'fleeing' mental state is unrevisable. In the next section, I argue that the fact that the 'fleeing' state is unrevisable can be used to show that the 'fleeing' state is not a thought or a delusion.

## V. IS THE 'FLEEING' MENTAL STATE A THOUGHT?

In Section III, I argued that the claim that the 'fleeing' mental state is a System 1 thought is more attractive than other cognitive explanations. In this section, I argue that the System 1 explanation ultimately fails because it cannot accommodate the fact that the 'fleeing' state is unrevisable. I also argue that the view that the 'fleeing' state is a delusion cannot explain the fact that the 'fleeing' state is unrevisable.

To be clear, I will argue that System 1 thoughts have a certain capacity for rationality, not that System 1 thoughts are in fact rational. I will also not claim that System 1 thoughts are typically or often rational. Thus, the observation that System 1 thoughts sometimes conflict with other thoughts is not a counterexample to my claim.

Furthermore, my claim is that *already-tokened* System 1 thoughts can be revised. This claim is neutral on whether System 1 thoughts can be *formed* on rational grounds. Since unrevisability concerns whether already-tokened attitudes can be revised, it is irrelevant whether System 1 thoughts are capable of being rationally formed.<sup>10</sup>

### V.1. *System 1 thoughts as revisable*

The question of how and whether System 1 thoughts can be revised is a matter of ongoing empirical inquiry. Nevertheless, a recent study provides us some reason to think that System 1 thoughts are revisable. Subjects were presented with a series of cognitive problems of the kind that tend to elicit System 1

<sup>10</sup> Mandelbaum (2014) argues that thoughts are never rationally formed.

responses. In some of these problems, the intuitive, System 1 response was correct. In others, the intuitive, System 1 response was incorrect. Subjects who provided an incorrect System 1 response reported a relatively weak feeling of confidence in their response. Further, when given the opportunity to re-evaluate their responses, subjects who reported this relatively weak feeling of confidence spent a longer time producing a second response than other subjects. These subjects were also more likely to change their answer than other subjects. This longer response time suggests that in forming their second response, these subjects switched to an analytic, System 2 mode of thinking (Thompson and Johnson 2014).

This study suggests that the following mechanisms govern the revision of System 1 thoughts: a mechanism which detects conflict between System 1 thoughts and other available evidence, a mechanism which triggers a weak feeling of confidence in response to the detection of such a conflict, and a mechanism which facilitates increased analytic engagement in response to a weak feeling of confidence.

This cluster of mechanisms plausibly renders System 1 thoughts capable of being rationally revised. For if System 1 thoughts lacked such a capacity, it would be difficult to explain why this cluster of mechanisms should exist. If not to respond to conflicts between System 1 thoughts and evidence derived from other sources, it is hard to see why a mechanism for detecting such conflicts should exist. If not to trigger re-evaluation of the evidence, it is hard to see why the detection of such conflicts should trigger in subjects a weak feeling of confidence. And if not to minimize such conflicts, it is hard to see why this feeling should in turn prompt subjects to re-evaluate the evidence in a more analytical mode of thought.<sup>11</sup>

We have some reason, then, to think that System 1 thoughts are capable of being rationally revised. Moreover, there is no reason to think that this capacity is selectively responsive to some sorts of counterevidence but not others. Thus, it would appear that System 1 thoughts are revisable. Since the ‘fleeing’ state is unrevisable, it is unlikely that the ‘fleeing’ state is a System 1 thought.<sup>12</sup>

## V.2. *Is the ‘fleeing’ state a delusion?*

So far, I’ve argued that it is unlikely that the ‘fleeing’ mental state is a System 1 thought. This result leaves open the possibility that the ‘fleeing’ state is an aberrant thought, such as a delusion. Since delusions are characteristically

<sup>11</sup> Subjects’ changed responses were not always in the direction of increased accuracy, which might be explained by the difficulty of the problems used (Thompson and Johnson 2014: 217).

<sup>12</sup> There is some evidence that System 1 thoughts are not extinguished in response to counterevidence but are rather moved to the level of the non-conscious in response to such evidence (Svedholm-Häkkinen 2015). If this view is correct, then the ‘fleeing’ state is unlikely to be a System 1 thought, since the ‘fleeing’ state remains conscious for most subjects.

unrevisable, the delusion-based view of the 'fleeing' state is at least consistent with the result that the 'fleeing' state is unrevisable (Byrne 2011: 120, 2012: 106–7). Nevertheless, I will argue that the delusion-based explanation cannot *explain* why the 'fleeing' state is unrevisable. This is because the mechanisms which underpin delusional persistence do not figure in the 'fleeing' state.

Recent results suggest that delusions are unresponsive to counterevidence because their subjects have atypical evidence-gathering tendencies and—in the case of persecutory delusions—atypically high levels of worry (Favrod *et al.* 2014; Freeman *et al.* 2015; Freeman and Garety 2014; Garety *et al.* 2014; Moritz *et al.* 2013). However, the 'fleeing' state is reliably elicitable in most subjects, including those who have typical evidence-gathering tendencies and typical levels of worry. Thus, the mechanisms which explain delusional persistence are simply absent in many subjects who token the 'fleeing' state. So, the view that the 'fleeing' state is a delusion cannot explain why the 'fleeing' state is unrevisable in these typical subjects.

In light of the failure of extant cognitive attitudes to account for the 'fleeing' mental state, one might propose that the 'fleeing' state is a novel cognitive attitude. This state might be a *sui generis* attitude, such as a 'seeming,' or it might be a novel variant of one of the familiar attitudes, such as a novel sort of thought or delusion.

The suggestion that the 'fleeing' mental state is a *sui generis* attitude cannot be altogether ruled out. Indeed, if there were no better explanation available, we should endorse this hypothesis. However, as I will now show, there is a better explanation, one which does not force us to enlarge our ontology of mental attitudes. On this view, the 'fleeing' state is a visual illusion.<sup>13</sup>

## VI. VISUALLY PERCEIVING THE INTENTIONS OF OTHERS

On the visual explanation, the 'fleeing' mental state is a visual experience. More precisely, the 'fleeing' state is a visual illusion, since it misrepresents inanimate entities as acting on intentions. Since the view that the 'fleeing' state is a visual illusion posits that the state is an established mental kind, the view is less optically extravagant than the view that the 'fleeing' state is a novel kind of attitude. At the same time, the view that the 'fleeing' state is a visual illusion can explain the unrevisability of the 'fleeing' state, succeeding where cognitive explanations failed.

Visual illusions are characteristically unresponsive to at least some kinds of counterevidence. There are two major explanations of why this is so. On the first account, at least some sorts of extravisual information never enter

<sup>13</sup> The view that the 'fleeing' state is an associative, arational *alief* warrants more consideration than I can give it here (Gendler 2008a,b). For concerns about positing aliefs, see Currie and Ichino (2012), Doggett (2012), Nagel (2012), and Mandelbaum (2013).

into the visual system as inputs, with the result that this information does not impact visual experience. On this view, the visual system lacks access to such information. (Fodor 1983, 1984, 2001: 63). On the second account, the visual system cannot process at least some sorts of extra-visual information, with the result that this information does not impact visual experience. On this view, even if the visual system were to enjoy access such information, it would not be able to interpret that information (Barrett 2006; Carruthers 2004, 2006, 2011).

For our purposes, it does not matter which of these explanations of visual illusions is correct. Either explanation can account for the unrevisability of the ‘fleeing’ mental state. On the first explanation, the reason the ‘fleeing’ state is unresponsive to conflicting evidence is that the ‘fleeing’ state is a visual experience, and the visual system lacks access to at least some kinds of extravisual evidence. On the second explanation, the reason the ‘fleeing’ state is unresponsive to conflicting evidence is that the ‘fleeing’ state is a visual experience, and the visual system lacks the capacity to process at least some kinds of extravisual evidence. Either way, by positing that the ‘fleeing’ state is a visual experience, and more particularly a visual illusion, we can satisfyingly explain its unrevisability by appealing to the mechanisms which underlie visual illusions more generally.<sup>14</sup>

## REFERENCES

- Abell, F., Happé, F. and Frith, U. (2000) ‘Do Triangles Play Tricks? Attribution of Mental States to Animated Shapes in Normal and Abnormal Development’, *Journal of Cognition and Development*, 15: 1–16.
- Audi, R. (2013) *Moral Perception*. Princeton: Princeton University Press.
- Barrett, H. (2006) ‘Modularity and Design Reincarnation’, in P. Carruthers, S. Stich and S. Laurence (eds) *The Innate Mind*, Vol. 2, 199–217. Oxford: OUP.
- Barrett, H. C. et al. (2005), ‘Accurate Judgments of Intention from Motion Cues Alone: A Cross-Cultural Study’, *Evolution and Human Behavior*, 26: 313–31.
- Bayne, T. (2008) ‘The Phenomenology of Agency’, *Philosophy Compass*, 3: 182–202.
- (2009) ‘Perception and the Reach of Phenomenal Content’, *The Philosophical Quarterly*, 59: 385–404.
- Block, N. (2014) ‘Seeing-As in the Light of Vision Science’, *Philosophy and Phenomenological Research*, 89: 560–72.
- Brogaard, B. (2014) ‘The Phenomenal Use of “Look” and Perceptual Representation’, *Philosophy Compass*, 9: 455–68.
- Butterfill, S. (2009) ‘Seeing Causing and Hearing Gestures’, *The Philosophical Quarterly*, 59: 405–28.
- Byrne, A. (2008) ‘Experience and Content’, *The Philosophical Quarterly*, 59: 429–51.
- (2011) ‘Knowing That I Am Thinking’, in A. Hatzimoysis (ed.) *Self-Knowledge*, 105–24. Oxford: OUP.

<sup>14</sup> For invaluable feedback on this paper, I would like to thank: Tom Avery, Ned Block, Chiara Brozzo, David Chalmers, Jeremy Dolan, Gabriele Ferretti, Don Garrett, Kris Goffin, Laura Gow, Zoe Jenkin, Angelica Kaufmann, Alex Kiefer, Tim Maudlin, Bence Nanay, Jake Quilty-Dunn, John Richardson, Adam Schneit, Maarten Steenhagen, Michael Strevens, Margot Strohminger, Neil Van Leeuwen, Henry Shevlin, and Nick Young, as well as audiences at: Colgate University, CUNY Graduate Center, and Hobart & William Smith Colleges. Special thanks to two anonymous referees.

- (2012) 'Knowing What I See', in D. Smithies and D. Stoljar (eds) *Introspection and Consciousness*, 183–210. Oxford: OUP.
- Carey, S. (2009) *The Origin of Concepts*. New York: Oxford University Press.
- Carruthers, P. (2004) 'The Mind is a System of Modules Shaped by Natural Selection', in C. Hitchcock (ed.) *Contemporary Debates in Philosophy of Science*, 293–311. Oxford: Blackwell.
- (2006) *The Architecture of the Mind*. Oxford: OUP.
- (2011) *The Opacity of Mind: An Integrative Theory of Self-Knowledge*. Oxford: OUP.
- (2015) 'Perceiving Mental States', *Consciousness and Cognition*, 36: 498–507.
- Casati, R. and Varzi, A. (2014) 'Holes', in: E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*. <<https://plato.stanford.edu/archives/spr2014/entries/holes/>> accessed 29 July 2017.
- Chappell, T. (2008) 'Moral Perception', *Philosophy*, 83: 421–37.
- Congiu, S., Schlotmann, A. and Ray, E. (2010), 'Unimpaired Perception of Social and Physical Causality, but Impaired Perception of Animacy in High Functioning Children with Autism', *Journal of Autism and Developmental Disorders*, 40: 39–53.
- Cullison, A. (2010) 'Moral Perception', *European Journal of Philosophy*, 18: 159–75.
- Currie, G. and Ichino, A. (2012), 'Aliens Don't Exist, Though Some of Their Relatives Do', *Analysis* 72: 788–98.
- Doggett, T. (2012) 'Some Questions for Tamar Szabo ó Gendler', *Analysis* 72: 764–74.
- Evans, J. S. B. (2008) 'Dual-Processing Accounts of Reasoning, Judgment, and Social Cognition', *Annual Review of Psychology*, 59: 255–78.
- Farennikova, A. (2013) 'Seeing Absence', *Philosophical Studies*, 166: 429–54.
- Favrod, J. et al. (2014) 'Sustained Antipsychotic Effect of Metacognitive Training in Psychosis: A Randomized-Controlled Study', *European Psychiatry*, 29: 275–81.
- Fish, W. (2012) 'High-Level Properties and Visual Experience', *Philosophical Studies*, 162: 43–55.
- Fodor, J. A. (1983). *The Modularity of Mind*. Bradford: Books/MIT Press.
- (1984) 'Observation reconsidered', *Philosophy of Science*, 51: 23–43.
- (2001) *The Mind Doesn't Work That Way: The Scope and Limits of Computational Psychology*. Cambridge, MA: MIT Press.
- Frederick, S. (2005) 'Cognitive Reflection and Decision Making', *The Journal of Economic Perspectives*, 19: 25–42.
- Freeman, D. et al. (2015) 'Effects of Cognitive Behaviour Therapy for Worry on Persecutory Delusions in Patients with Psychosis (WIT): A Parallel, Single-Blind, Randomised Controlled Trial with a Mediation Analysis', *The Lancet Psychiatry*, 2: 305–13.
- and Garety, P. (2014) 'Advances in Understanding and Treating Persecutory Delusions: A Review', *Social Psychiatry and Psychiatric Epidemiology*, 49: 1179–89.
- Gallagher, S. (2008) 'Direct Perception in the Intersubjective Context', *Consciousness and Cognition*, 17: 535–43.
- Gao, T. and Scholl, B. J. (2011) 'Chasing vs. Stalking: Interrupting the Perception of Animacy', *Journal of Experimental Psychology: Human Perception and Performance*, 37: 669–84.
- Newman, G. E. and Scholl, B. J. (2009) 'The Psychophysics of Chasing: A Case Study in the Perception of Animacy', *Cognitive Psychology*, 59: 154–79.
- Garety, P. et al. (2014) 'Cognitive Mechanisms of Change in Delusions: An Experimental Investigation Targeting Reasoning to Effect Change in Paranoia', *Schizophrenia Bulletin*, 41: 400–10.
- Gendler, T. S. (2008a) 'Alien and Belief', *The Journal of Philosophy*, 105: 634–63.
- (2008b) 'Alien in Action (and Reaction)', *Mind & Language*, 23: 552–85.
- Gertler, B. (2011) 'Self-Knowledge and The Transparency of Belief', in A. Hatzimosis (ed.) *Self-Knowledge*, 125–45. Oxford: OUP.
- Glür, K. (2009) 'In Defence of a Doxastic Account of Experience', *Mind & Language*, 24: 297–327.
- Gray, K. and Wegner, D. M. (2012) 'Feeling Robots and Human Zombies: Mind perception and the Uncanny Valley', *Cognition*, 125: 125–30.
- Harmon-Jones, E. and Harmon-Jones, C. (2007) 'Cognitive Dissonance Theory after 50 Years of Development', *Zeitschrift für Sozialpsychologie*, 38: 7–16.
- Hashimoto, H. (1966) 'A Phenomenal Analysis of Social Perception', *Journal of Child Development*, 2: 1–26.
- Heberlein, A. S. and Adolphs, R. (2004) 'Impaired Spontaneous Anthropomorphizing Despite Intact Perception and Social Knowledge', *Proceedings of the National Academy of Sciences of the United States of America*, 101: 7487–91.

- Heider, F. and Simmel, M. (1944) 'An Experimental Study of Apparent Behavior', *The American Journal of Psychology*, 57: 243–59.
- Helton, G. (2015) *Two Worlds, One Mind: The Divide Between Perception and Belief*, unpublished doctoral dissertation. New York: New York University.
- (2016) 'Recent Issues in High-Level Perception', *Philosophy Compass*, 11: 851–62.
- Huddleston, A. (2012) 'Naughty Beliefs', *Philosophical Studies*, 160: 209–22.
- Klin, A. (2000) 'Attributing Social Meaning to Ambiguous Visual Stimuli in Higher Functioning Autism and Asperger Syndrome: The Social Attribution Task', *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, 41: 831–46.
- Mele, A. R. (1992) 'Recent Work on Intentional Action', *American Philosophical Quarterly*, 29: 199–217.
- Macpherson, F. (2012) 'Cognitive Penetration of Colour Experience: Rethinking the Issue in Light of an Indirect Mechanism', *Philosophy and Phenomenological Research*, 84: 24–62.
- Mandelbaum, E. (2013) 'Against Alief', *Philosophical Studies*, 165: 197–211.
- (2014) 'Thinking is Believing', *Inquiry*, 57: 55–96.
- Moritz, S. et al. (2013) 'Complementary Group Metacognitive Training (MCT) Reduces Delusional Ideation in Schizophrenia', *Schizophrenia Research*, 151: 61–9.
- Mylopoulos, M. I. (2015) 'Agentive Awareness Is Not Sensory Awareness', *Philosophical Studies*, 172: 761–80.
- Nagel, J. (2012) 'Gendler on Alief', *Analysis*, 72: 774–88.
- Nanay, B. (2011a) 'Do We See Apples As Edible?', *Pacific Philosophical Quarterly*, 92: 305–22.
- (2011b) 'Do We Sense Modalities with Our Sense Modalities?', *Ratio*, 24: 299–310.
- Prinz, J. J. (2005) 'A Neurofunctional Theory of Consciousness', in A. Brook and K. Akins (eds) *Cognition and the Brain: Philosophy and Neuroscience Movement*, 381–96. Cambridge: CUP.
- (2007) 'The Intermediate Level Theory of Consciousness', in M. Velms and S. Schneider (eds.) *The Blackwell Companion to Consciousness*, 247–60. Malden, MA: Blackwell.
- Quilty-Dunn, J. (2015) 'Believing in Perceiving: Known Illusions and the Classical Dual-Component Theory', *Pacific Philosophical Quarterly*, 96: 550–75.
- Rips, L. (2011) 'Causation from Perception', *Perspectives on Psychological Science*, 6: 77–97.
- Rutherford, M. D., Pennington, B. F. and Rogers, S. J. (2006) 'The Perception of Animacy in Young Children with Autism', *Journal of Autism and Developmental Disorders*, 36: 983–92.
- Scholl, B. J. and Gao, T. (2013) 'Perceiving Animacy and Intentionality: Visual Processing or Higher-Level Judgment?', in M. D. Rutherford and V. A. Kuhlmeier (eds.) *Social Perception: Detection and Interpretation of Animacy, Agency, and Intention*, 197–230. Cambridge, MA: MIT Press.
- and Tremoulet, P. D. (2000) 'Perceptual Causality and Animacy', *Trends in Cognitive Sciences*, 4: 299–309.
- Setiya, K. (2015), 'Intention', in E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy*, <http://plato.stanford.edu/archives/sum2015/entries/intention/>, accessed 1 May 2015.
- Shah, N. and Velleman, J. D. (2005) 'Doxastic Deliberation', *The Philosophical Review*, 114: 497–534.
- Shea, N. (2014) 'Distinguishing Top-Down from Bottom-Up Effects', in S. Biggs, M. Matthen and D. Stokes (eds.), *Perception and Its Modalities*, 73–91. Oxford: OUP.
- Siegel, S. (2005) 'The Phenomenology of Efficacy', *Philosophical Topics*, 33: 265–84.
- (2009) 'The Visual Experience of Causation', *The Philosophical Quarterly*, 59: 519–40.
- (2010) *The Contents of Visual Experience*. Oxford: OUP.
- Sorensen, R. (2012) 'The Twin Towers Riddle', *Philosophical Studies*, 162: 109–17.
- Spaulding, S. (2015) 'On Whether We Can See Intentions', *Pacific Philosophical Quarterly*, 98: 150–70.
- Svedholm-Häkkinen, A. M. (2015) 'Highly Reflective Reasoners Show No Signs of Belief Inhibition', *Acta Psychologica*, 154: 69–76.
- Thompson, V. A. (2009) 'Dual Process Theories: A Metacognitive Perspective', in J. S. B. T. Evans and K. Frankish (eds) *In Two minds: Dual Processes and Beyond*, 171–95. Oxford: OUP.
- and Johnson, S. C. (2014) 'Conflict, Metacognition, and Analytic Thinking', *Thinking & Reasoning*, 20: 215–44.

Princeton University, USA