

Crossmodal Basing

ZOE JENKIN 

*Department of Philosophy and Philosophy-Neuroscience-Psychology Program,
Washington University in St. Louis, St. Louis, MO, USA*
zoejenkin@wustl.edu

What kinds of mental states can be based on epistemic reasons? The standard answer is only beliefs. I argue that perceptual states can also be based on reasons, as the result of crossmodal interactions. A perceptual state from one modality can provide a reason on which an experience in another modality is based. My argument identifies key markers of the basing relation and locates them in the crossmodal Marimba Illusion (Schutz & Kubovy 2009). The subject's auditory experience of musical tone duration is based on the reason provided by her visual representation of the length of the musician's gesture and other stored perceptual principles.

1. Introduction

When you hear loud crunching sounds while eating potato chips, the chips feel especially crispy (Spence 2015). When you smell strawberries, sugar tastes sweeter (Frank & Byram 1988). These are crossmodal interactions. In crossmodal interactions, activity in one sensory modality influences activity in another modality (O'Callaghan 2019). As a result, our perceptual experiences are altered. In the potato chip example, auditory processing influences tactile processing, leading to an experience of increased crunchiness.

Crossmodal interactions have the potential to challenge traditional philosophical views, including the standard taxonomy of five senses (Macpherson 2011b), the idea that phenomenology is always unisensory (Macpherson 2011a, Smith 2015, O'Callaghan 2017a), and the view that perceptual evidence always originates for a single modality (O'Callaghan 2020). I argue for another significant implication of crossmodal interactions: they upend the traditional epistemic division between perception and belief. Standardly, only beliefs are epistemically evaluated as justified or unjustified, in virtue of whether and how they are based on reasons. Perceptions, in contrast, are not typically thought of as based on reasons or as epistemically evaluable. Misperceptions are considered mere unfortunate perceptual happenings, bad luck without a deeper normative impact.

Against this picture, I argue that a perceptual state in one modality can be based on a reason provided by a perceptual state in another modality. Crossmodal interactions instantiate the epistemic basing relation. Such crossmodally based perceptual experiences are epistemically evaluable.¹ My argument proceeds in three stages. In §2, I describe key markers of the epistemic basing relation. In §3, I argue that these key markers are present in the crossmodal Marimba Illusion (Schutz & Kubovy 2009). In §4, I discuss the nature of the reasons we have for this illusory experience. In §5, I argue that the Marimba Illusion is sufficiently person-level to involve epistemic basing.

2. Epistemic Basing

When the basing relation holds between a reason and a mental state, it links them both psychologically and epistemically.² The basing relation transmits epistemic support (or lack of support) from a reason to a mental state. These reasons are typically provided by prior mental states. When belief A is based on belief B, A is formed and/or maintained in response to the reason(s) B provides. Sherlock Holmesian deduction exemplifies epistemic basing. Consider a scene from Conan Doyle's *The Final Problem*, in which Watson receives a letter calling him away from a hike with Holmes to Reichenbach Falls (Conan Doyle 1893/1993). Holmes forms the belief that the letter is a fake (call this 'the letter belief'). On the basis of the reason the letter belief provides, Holmes forms the belief that Professor Moriarty is setting a trap at the falls (call this 'the trap belief'). This basing relation determines the trap belief's epistemic status. The trap belief is justified because the letter belief provides a good reason for the trap belief.

A reason is a consideration in favour of acting, believing, or mentally representing in a certain way (Scanlon 2014). Epistemic reasons are roughly truth-related considerations. They differ in kind from moral, prudential, and aesthetic reasons. There may be corresponding forms

¹ Other proponents of the view that perception is epistemically evaluable include McGrath (2013), Siegel (2017), Munton (2019), and Teng (2021).

² Other hybrid normative-psychological relations include inference (Audi 2008, Siegel 2017, Quilty-Dunn & Mandelbaum 2018, Neta 2019, Railton 2020), knowledge (Kim 1988, Sosa 1999, cf. Sylvan 2018), and moral responsibility (Fischer & Ravizza 1998, Scanlon 1998, Levy 2014).

of moral, prudential, and aesthetic basing relations, but I focus on epistemic basing here.³

The reasons involved in epistemic basing are an agent's motivating reasons. My arguments are neutral as to whether motivating reasons are mental states (e.g., Davidson 1963, Turri 2009), facts (e.g., Hyman 1999, Williamson 2000, Littlejohn 2012), or propositions (e.g., Fantl & McGrath 2009). Either way, the reasons that serve as epistemic bases must be *the agent's*. If reasons are mental states, the reasons that serve as epistemic bases must be the agent's mental states rather than those of other agents. If reasons are facts or propositions, the reasons that serve as epistemic bases must be *possessed* by the agent—that is, she must be epistemically related to them. It is commonly thought that our mental states epistemically relate us to facts or propositions and so enable us to possess them as reasons (e.g., Dancy 2000, Littlejohn 2012).⁴ I use the terminology of a mental state 'providing reasons' to remain neutral as to whether reasons are mental states themselves or something external such as facts or propositions. I return to this issue in §4.

Motivating reasons can be good or bad normative reasons. When reasons are both good and properly used, epistemic basing results in justified mental states. When reasons are bad and/or improperly used, epistemic basing results in unjustified mental states. If after learning the letter was a fake Watson had jumped to the conclusion that every letter ever written to him was a fake, this belief would be unjustified because his belief that the letter was a fake provides insufficient reason for it.

The basing relation determines not only the justificatory status of mental states, but also which mental states are ripe for epistemic evaluation in the first place. States that are based on reasons can be epistemically evaluated as justified or unjustified, depending on whether they result from good or bad basing.⁵ States that are instead formed by mere causal responses are not epistemically evaluable in the same way. It is inappropriate to epistemically evaluate the state of thermoreceptor activation in your palm because it is a purely physical state, independent of reasons. Similarly, if your friend shows you her new red slippers and you

³ I use the phrases 'epistemic basing' and 'basing on epistemic reasons' interchangeably. For my purposes, a reason is what plays the basis role in the basing relation, so all basing is basing on reasons.

⁴ On some views, possessing a reason also requires direct acquaintance with the relevant fact. For example, hallucinating that the cat is on the desk is not sufficient for possessing the reason that the cat is on the desk because hallucinations do not acquaint us with facts.

⁵ Some mental states may also be epistemically evaluable *because* they are not based on reasons, such as when emotions eclipse reasoning.

think ‘We’re not in Kansas anymore’, your thought is neither justified nor unjustified. It is associatively formed, orthogonal to reasons.

Because the basing relation is central to epistemic evaluation, it is important to determine its scope. The basing relation is not always introspectively transparent, so we cannot locate it through reflection alone. Sometimes our mental states are based on reasons unbeknownst to us. For example, in *Wuthering Heights* the teenage Heathcliff convinces himself he should no longer be with Catherine because he no longer loves her (Bronte 1847/1995). In fact, his affection remains. While he cannot admit it even to himself, his belief that he should no longer be with Catherine is unconsciously based on the reason that he is not good enough.

Given this lack of introspective transparency, we require indicators for the presence of the basing relation. While necessary and sufficient conditions on basing are notoriously elusive, diagnostic psychological markers of basing are within closer reach. Here, I outline some key markers of basing that are particularly relevant to crossmodal interactions. Whether reasons are mental states, facts, or propositions, these markers can help us identify when reasons are possessed by an agent and used in her reasoning.

The four key markers of basing are: 1) mental representations of causal relations, 2) mental representations that are available for use by other psychological systems, 3) epistemic support relations between the contents of mental states 4) transition rules that are sensitive to epistemic support. The first two are markers of mental states providing reasons, while the second two are markers that those reasons are used in basing. Having key markers of both types helps to capture the basing relation comprehensively. I do not claim these markers are necessary or sufficient for basing. Rather, they are diagnostic of basing in the same way that a red itchy rash and a recent walk in the woods are diagnostic of poison ivy.

The first key marker of basing is that the mental transition involves representations of causal relations.⁶ Causal representations are indicative of basing because causal relations are the backbone of reasoning. When we reason about which brand of ice cream to buy, we consider which one caused the most gustatory pleasure in the past. When we reason about who left muddy footprints in the hallway, we consider who

⁶ I use the terms ‘representations of causal relations,’ ‘causal representations,’ and ‘states with causal contents’ interchangeably. These terms pick out a kind of content, not any special causal efficacy.

has feet of the right size to cause the tracks. Causal representations tell us which features are indicative of which other features, allowing the mind to move between causes and effects. When we represent that A causes B, and we encounter A, we can infer B.⁷

While causal contents are the currency of much reasoning in cognition, they can also be visually represented (Michotte 1963, Leslie 1987, Siegel 2010). Evidence for this claim comes from studies demonstrating retinotopic adaptation for causal launching (Rolf, Dambacher, & Cavanagh 2013). Visual processing of causality is also relatively encapsulated from central cognition, which is characteristic of perceptual systems (Michotte 1963, Leslie 1988, Schlottmann & Shanks 1992, Scholl & Tremoulet 2000). Yet visual representations of causality are often considered conceptual or proto-conceptual due to their rich inferential role (Leslie 1988, Carey 2009). They are widely integrated with representations of spatial relations, quantity, agency, and morality. They also guide action. For example, infants use their visual representations of causal interactions between animate-looking blocks to infer which are helpers and which are hinderers, and to determine their preference for interacting with helpers (Hamlin, Wynn, & Bloom 2007). Visual representations of repeated contact causality generate conditional expectations about motion and state changes, such as an object's colour or the onset of music (Muentener & Carey 2010). Visual representations of causality depend on whether the entities involved are categorized as agents or inanimate objects (Kotovsky & Baillargeon 2000, Kosugi & Fujita 2002, Saxe & Carey 2006). When we are looking for perceptual states that provide reasons, visual representations of causality are particularly good contenders because they are well-integrated in the cognitive economy.

The second key marker of basing is that states that provide reasons are available for use across the mind. This is a key marker of basing because an agent's reasons typically serve a range of epistemic purposes. For example, Holmes uses the reason that the letter is a fake to conclude that Moriarty is setting a trap, but he could also use it to calculate the ratio of real-to-fake mail Watson received that week. Reasons are also typically available for use by multiple psychological subsystems, such as belief formation, action guidance, and perception. For example, Holmes might use the reason that the letter is fake to guide his visual examination of the handwriting, or to motivate calling Inspector Lestrade.

⁷ Causal contents are indicative of providing reasons but not necessary for it. States that provide reasons despite lacking causal contents include many visual experiences and many beliefs used as premises in categorical syllogisms.

In contrast, states that are embedded within a subsystem and only usable for a narrow set of purposes do not typically provide reasons. For example, early visual representations of texture gradients are primarily used to compute visual distance. They are not directly available for reasoning, action guidance, or use by other subsystems. While texture gradients are also used for a few select other tasks such as computing height and surface slant, these tasks are confined to vision.⁸ Unlike states that provide reasons, these representations are not available for wider use by other parts of the mind.⁹

Availability for use by multiple psychological subsystems is a key marker of basing because it indicates that a state serves the whole individual rather than merely her subsystems, and thus can provide *her* with a reason.¹⁰ This is important because when a state is based on a reason, the state's epistemic status redounds on the individual. An irrational belief makes the believer pro tanto irrational. States that serve the whole individual are relevant to her epistemic status.

The third key marker of basing is the presence of an epistemic support relation between the contents of mental states. When a mental state that P is properly based on a reason provided by a mental state that Q, Q epistemically supports P. Consider Holmes, whose belief that there is a trap at Reichenbach Falls is properly based on the reason provided by his belief that the letter is a fake. Even before considering Holmes' belief formation process, we can see that the content of his letter belief epistemically supports the content of his trap belief. While some poorly based states lack epistemic support, in the good case epistemic support is a fundamental element of the basing relation.

Basing is characterized not only by epistemic support relations but also by epistemic dependence relations. To illustrate, consider a situation in which Holmes' brother Mycroft hears of the fake letter but fails to connect it to Moriarty. Mycroft coincidentally daydreams that Moriarty is setting a trap at Reichenbach Falls and endorses this daydream. Here,

⁸ I do not deny that any visual representations of texture participate in crossmodal interactions (e.g., [Heller 1982](#)), although the evidence for such interactions is limited ([Whitaker 2008](#)). My point here is that the particular early monocular cue of texture gradient (i.e., the way texture density and detail gradually change from coarse to smooth) feeds only into proximal visual computations of depth, height, and slant and is not used beyond visual processing.

⁹ Thanks to an anonymous referee for help in clarifying this key marker of basing.

¹⁰ I leave open whether basing can ever occur within a subsystem. Because key markers are diagnostic rather than necessary, the possibility of basing within a subsystem is compatible with my arguments. For further discussion of epistemic basing within a subsystem, see [Lewis \(1982\)](#), [Egan \(2008\)](#), and [Jenkin \(2020\)](#).

Mycroft's trap belief is not based on the reason provided by the letter belief because there is no dependence relation between them.

Given the importance of epistemic dependence relations, the fourth key marker of basing is that based states are formed according to transition rules that are sensitive to reasons. Psychological evidence can reveal the nature of transition rules between mental states. For example, data on subjects' competency with arguments and proofs reveal the deductive principles used in reasoning (e.g., [Rips 1994](#)).¹¹ Similarly, data on a state's sensitivity to changes in the reasons provided can shed light on whether mental process are driven by reasons or arational factors.

Taken together, these key markers equip us with the tools to identify which mental states are based on reasons. These markers are particularly useful for inquiring whether states beyond belief, such as perceptions, are ever based on reasons. In the next section, I turn to this task.

3. Crossmodal Basing

In one particularly philosophically interesting form of crossmodal interaction, music perception is influenced by non-auditory senses ([Davidson 1993](#), [Vines et al. 2005a](#), [Vines et al. 2005b](#), [Huang et al. 2012](#)). Crossmodal music perception has been taken to show that music is not an exclusively auditory artform ([Bergeron & Lopes 2009](#), [O'Callaghan 2019](#)). Crossmodal music perception also provides a strong case for epistemic basing in perception.

I focus here on the crossmodal Marimba Illusion.¹² Examining illusions can help uncover the principles by which perception operates, including those governing crossmodal interactions ([O'Callaghan 2019](#)). Crossmodal interactions do not exclusively lead to illusions. They can also increase the overall accuracy of perception by amplifying environmental information ([Shams & Kim 2010](#), [Spence 2011](#), [O'Callaghan 2019](#)). While my focal example is illusory, my arguments can be extended to veridical perceptions. I return to this point in §4.

In the Marimba Illusion, the length of a visible gesture influences the perceived duration of an audible tone.¹³ Subjects are shown a

¹¹ For alternate views see [Johnson-Laird \(1997\)](#) and [Oaksford and Chater \(2007\)](#).

¹² For a video of this illusion, visit <https://maplelab.net/overview/audio-visual-integration/illusion/>.

¹³ The Marimba Illusion is an illusion according to most standard definitions because it results in a non-veridical perception (e.g., [Smith 2002](#), [Macpherson 2009](#), [Fish 2010](#)). It also satisfies [Prettyman's \(2019\)](#) definition because the subject is rendered unable to precisely discriminate between lengths of Marimba tones.

musician playing a marimba using striking gestures of varying lengths (Schutz & Lipscomb 2007). The tones produced have the same duration regardless of gesture length. When the musician uses long gestures, subjects hear the tone as lasting longer than when the musician uses short gestures. The subject's auditory experience is dependent on her visual experience.¹⁴

The states involved in this psychological process are as follows:

Auditory Input

Content: There is a tone starting at time T1 and location L1.

Visual Input

Content: There is a long gesture occurring from time T1 to T2 and from location L1 to L2.

Crossmodal Causal Representation

Content: The gesture causes the tone (due to spatial and temporal congruence of inputs).

Stored Crossmodal Information

Content: When a gesture causes a tone, the length of the gesture is typically proportional to the length of the tone.

Output Auditory Experience

Content: There is a long tone.

This process starts off with auditory and visual inputs. These states are already the result of significant processing and have been output from their respective modalities. The subject has conscious experiences with these contents. The experimental data do not determine whether the states that drive the crossmodal effect are these conscious experiences or unconscious perceptions just prior in the chain of processing, so I remain neutral on that question here.¹⁵ The auditory input represents the starting time and location of the marimba tone. The visual input represents the starting and ending times and locations of the gesture. In the long gesture condition, most of the gesture occurs *after* the key is hit and the tone begins, as if the musician were physically drawing out the tone. T1 is before the Marimba is struck, when the musician begins

¹⁴ The Marimba Illusion is a rare case of visual domination (i.e., stronger influence) in a temporal task. Typically, vision dominates in spatial tasks while audition dominates in temporal tasks (Schutz & Lipscomb 2007). This is a useful perceptual rule because vision has greater spatial precision while audition has greater temporal precision (Welch & Warren 1980).

¹⁵ See de Gelder et al. (2000) and de Gelder et al. (2002) for evidence from subjects with prosopagnosia and blindsight that some types of crossmodal interactions only occur in response to conscious experience. The Marimba Illusion has not been formally tested on subjects with deficits in auditory or visual consciousness.

the gesture, and T2 is after the Marimba is struck, when the musician finishes the gesture.

The two inputs generate a crossmodal causal representation. The perceptual system relies on principles that guide transitions from representations of spatiotemporal features to representations of causal relations. Because of the spatial and temporal coincidence of the gesture and tone, we represent the gesture as causing the tone.

The principle that guides this process belongs to the family of the Unity Assumption. The Unity Assumption says that when sensory events have high congruence they are perceived as having a common cause (Welch & Warren 1980). Congruence cues include spatial and temporal coincidence, shape, size, orientation, and texture. The Unity Assumption helps perceptual systems determine the cause of sensory signals by combining information across modalities. Other crossmodal effects that exploit the Unity Assumption include the McGurk Effect (McGurk & MacDonald 1976), the Rubber Hand Illusion (Botvinick & Cohen 1998), and the Tone-Flash Illusion (Shams, Kamitani, & Shimojo 2002). The Unity Assumption governs representations of common cause while the principle used in the Marimba Illusion governs representations of one event causing another, but they share the same basic structure.

Once the crossmodal causal representation is produced, it combines with stored information that says, when a gesture causes a tone, the length of the gesture is typically proportional to the length of the tone. While percussionists know that Marimba tones are insensitive to gesture-lengths, perceptual systems do not have access to this information. Instead, they exploit general principles connecting the proportions of causes and effects. The scope and contours of these principles depend on an immense range of empirical data. For simplicity I formulate the stored information as the specific proportional relation between gestures and tones that is entailed by more general principles of proportionality.

This stored information is restricted to direct impact gestures. The perceptual system would not represent the length of pressing a television remote's play button as proportional to the length of an episode. The stored information is also sensitive to whether the tone is percussive or sustained. Impact gestures such as striking a Marimba are taken to be proportional to percussive sounds but not sustained sounds such as cello tones (Vatakis & Spence 2008, Chuen & Schutz 2016).¹⁶ When there is appropriate coherence of gesture and tone, the application of

¹⁶ I thank an anonymous referee for bringing this point and the cited studies to my attention.

the stored information results in an auditory experience as of a falsely elongated tone.

Crossmodal interactions are sometimes described as crossmodal mental imagery (Spence & Deroy 2013, Nanay 2020).¹⁷ Crossmodal mental imagery occurs when sensory stimulation in one modality triggers mental imagery in another modality (Spence & Deroy 2013). A crossmodal mental imagery view of the Marimba Illusion would say that visual stimulation caused by the long gesture triggers auditory imagery of a long tone. Depending on how one defines ‘mental imagery’, this understanding of the process may be compatible with my description. Nanay defines mental imagery as ‘perceptual processing that is not triggered by corresponding sensory stimulation in the relevant sense modality’ (Nanay 2020, p. 83). On this permissive definition, the Marimba Illusion is crossmodal mental imagery. On stricter definitions that require e.g., voluntary control or a phenomenal difference between imagery and perception (Reid 1764/1997, Ichikawa 2009), the Marimba Illusion is not mental imagery. My arguments here proceed independently of classification as mental imagery, as they are grounded in psychological features of the process itself. If one classifies this process as crossmodal mental imagery, then my arguments show that not only perceptions, but also mental images can be based on reasons.¹⁸

At first glance, one might think that the Marimba is a merely causal process. Some crossmodal interactions do operate independent of reasons. When you press on your eyeball, you see spots called ‘phosphenes.’ This interaction between touch and vision is entirely explained by physical interactions without appeal to reasons or basing.¹⁹ Similarly, some forms of synesthesia are mediated by association rather than reason. Chromesthetes see certain colours when they hear certain sounds. While sounds causally trigger the visual experience of colours, the colours are not epistemically based on the sounds.²⁰

In the rest of this section, I argue that the Marimba Illusion differs from phosphenes and chromesthesia in that it bears all four key markers

¹⁷ I thank an anonymous referee for raising this issue.

¹⁸ For additional arguments for the view that mental images can be based on reasons, see Teng (2021). For further discussion of the epistemology of crossmodal mental imagery from a Reliabilist perspective, see Nanay (2020).

¹⁹ Phosphene generation is a form of crossmodal interaction, but it is not a constitutively crossmodal interaction in which one modality’s content depends on another’s. For discussion of constitutively crossmodal interactions, see O’Callaghan (2008), Bayne (2014), and Connolly (2014).

²⁰ I leave open whether other forms of synesthesia involve basing. For discussion of various forms of synesthesia, see Macpherson (2007), Deroy and Spence (2013), Auvray and Deroy (2015), and O’Callaghan (2017b).

of basing: mental states with causal contents (§3.1), widely available mental states (§3.2), epistemic support relations (§3.3), and epistemically sensitive transition rules (§3.4). These key markers indicate that the auditory experience of the long tone is based on the reason jointly provided by the auditory input, visual input, crossmodal causal representation, and stored information about proportional lengths. I treat these states as a providing a single reason rather than multiple reasons because absent any one state, the others would not provide a consideration in favour of the auditory experience of the long tone.

3.1 Causal Contents

In the Marimba Illusion, two of the states have causal contents: the crossmodal causal representation (the gesture causes the tone) and the stored crossmodal information (when a gesture causes a tone, the length of the gesture is typically proportional to the length of the tone). These causal representations enable the illusion to obtain. The auditory and visual inputs and the output all only represent spatiotemporal properties. Without appeal to any stored crossmodal information, the transition is left unexplained.

Additional evidence for causal contents in the Marimba Illusion comes from a version of the experiment in which the causal connection between the gesture and tone was made implausible (Schutz & Kubovy 2009). The marimba player's gestures were paired with sustained (e.g., clarinet) tones rather than percussive marimba tones. The tone elongation effect disappeared. Without the characteristic spatiotemporal properties of percussive tones, the mediating causal representation was not generated because the perceptual system encodes the information that impact events do not cause sustained sounds. These data indicate that the Marimba Illusion depends on causal representations.

The presence of causal representations fends off worries that the Marimba Illusion occurs too early in visual processing to involve epistemic basing. While we do not yet know the Marimba Illusion's neural location, some crossmodal interactions occur in the primary visual and auditory cortices (Schroeder & Foxe 2005, Watkins et al. 2006, Driver & Noesselt 2007), which would be a surprising locus of basing.²¹ While the primary visual and auditory cortices are traditionally associated with low-level perceptual processes such as orientation and frequency detection, there is increasing evidence that these cortices also house highly sophisticated computations. Primary visual cortex plays an important

²¹ I thank an anonymous referee for raising this issue.

role in computing size constancy (Sperandio, Chouinard, & Goodale 2012), and primary auditory cortex represents and classifies phonemes (Mesgarani et al. 2008). Given these results, we should remain open-minded about the neural localizations of epistemic basing.

The Marimba Illusion's dependence on causality indicates that wherever its neural location, its representations and transitions are highly sophisticated. The representation that the gesture caused the tone is sensitive to prior representations of temporal synchrony, spatial congruence, and the instrument-sound pairing (Kubovy & Schutz 2010). If any of these prior representations is absent (e.g., there is temporal asynchrony, or a clarinet sound is paired with a marimba), the causal representation is not generated. As in other causally mediated crossmodal interactions, these representations reflect a careful synthesis of prior perceptual states (Keetels, Stekelenburg, & Vroomen 2007, Stekelenburg & Vroomen 2007, Sekuler, Sekuler, & Lau 1997).²²

Not all crossmodal interactions involve rich causal structures. The production of phosphenes is a response to physical stimulation rather than to any mental states at all. Chromesthesia is a response to mental states, but not ones with causal contents. The inputs are auditory experiences of phonemes or musical tones, and the outputs are visual experiences of colours. The transition is well explained by a simple associative structure with no intervening states.²³ The Marimba Illusion contrasts with these more minimal forms of processing both in its rich causal representations and the complex inferential structures they mediate.

3.2 *Availability*

One might worry that states within perceptual systems are too task-specific to provide reasons. For example, luminance maps and edge representations are used to compute shape in early vision. These states lack the paradigmatic availability of states that provide reasons. Without states that provide reasons there is no basing on reasons. This worry is particularly pressing for modularists (e.g., Fodor 1983, Pylyshyn 1999), who hold that sensory input systems are domain specific, informationally encapsulated subsystems with proprietary information databases and processing rules. On modular architectures, the sensory modalities

²² See Green (2019) for further discussion of the unimodal processes that precede crossmodal interactions.

²³ Abstract forms of synesthesia such as magnitude-colour linkages (Ramachandran & Hubbard 2001) are more sophisticated yet still lack causal contents.

are typically each taken to break down into one or more functionally specified modules (e.g., vision might house separate modules for shape and colour processing). One might worry that states embedded within modules are insufficiently available to provide agents with reasons.

Crossmodal interactions avoid this worry. In the Marimba Illusion, the auditory and visual inputs interface across subsystems and are not confined to modules. The auditory representation of the tone's onset is not restricted to auditory tonal processing, and the visual representation of the long gesture is not restricted to visual motion processing. If these representations are initially produced by modules, they exit those modules when they enter the combinatory process of the Marimba Illusion.²⁴

The auditory representation of the tone's onset and the visual representation of the long gesture are available for a wide variety of uses. They are outputs of perception and so can be used for countless types of cognition and action. In response to hearing the tone's onset, someone working nearby might form the belief that the Marimbaist's performance had started or close the door to block outside noise. In response to seeing the long gesture, a music student might form the belief that playing the Marimba is tiring or mimic the long gesture herself. Just as these states provide reasons for belief and action, they also provide reasons when they participate in crossmodal interactions.

The contrast cases of phosphenes and chromesthesia also involve multiple sensory modalities, so the underlying states are available to a degree. In phosphenes, the pressure on the eyeball initiates processing in both tactition and vision. However, these two uses are far fewer than the numerous ways perceptual outputs can be used for belief and action. Moreover, the physical pressure that produces phosphenes is not a mental state and so cannot provide a reason.

Chromesthesia starts from an auditory experience such as the simple sound 'ah,' which is available for a variety of forms of reasoning and action. This auditory experience might be used to form the belief that someone nearby is relaxing, or the belief that someone is about to sneeze (anticipating an impending '-choo'). Chromesthesia does seem to bear the availability marker of basing, but it lacks the other three. Availability alone is insufficient to demonstrate basing.

²⁴ Some modules span multiple sensory modalities (e.g., speech processing integrates visual and auditory inputs (Rosenblum 2005)), so the mere fact that these representations participate in crossmodal interactions does not always demonstrate availability outside a module. In the case of the Marimba Illusion, there is no independent evidence for a gesture-tone module so postulating one would be ad-hoc.

3.3 *Epistemic Support Relations*

When mental state A is properly based on a reason provided by mental state(s) B, A's content is epistemically supported by B's content(s). In the Marimba Illusion, the content of the output is epistemically supported by the contents of the inputs, crossmodal representation, and stored crossmodal information. To see this support relation, consider an analogous inference with beliefs:

- (P1) There is a long gesture.
- (P2) There is a tone.
- (P3) The gesture causes the tone.
- (P4) When a gesture causes a tone, the length of the gesture is typically proportional to the length of the tone.
- (C) The tone is long.

This is a valid inference. Although the contents of the states in the Marimba Illusion are more complex than these premises, they share a similar form. The Marimba Illusion and this inference also share similar epistemic support relations. Just as premises 1-4 in the above inference epistemically support its conclusion, the contents of the inputs, causal representation, and stored information in the Marimba illusion epistemically support the content of its output. In contrast, in phosphenes pressure on the eyeball does not epistemically support the presence of spots, and in chromesthesia sounds do not epistemically support colours. Unlike in the Marimba Illusion, these processes are not naturally reconstructed as rational inferences.

There is a salient difference between the Marimba Illusion and the analogous inference, though. In the Marimba Illusion the subject does not 'take' the premises to support the conclusion because some of the states are unconscious, while in the inference she does. On some views, there is a taking condition on inference, meaning that for a mental transition to be an inference the subject must take the premises to support the conclusion, and draw the conclusion because of this taking (Boghossian 2014, Chudnoff 2014, c.f. Wright 2014, McHugh & Way 2016, Siegel 2017). 'Taking' is often thought to involve an independent propositional attitude such as a belief or intuition that represents the epistemic support relation. While my arguments focus on the basing relation rather than inference, one might wonder whether there is also a taking condition on basing.²⁵ If so, it seems the Marimba Illusion cannot instantiate epistemic basing

²⁵ I thank the Editors of *Mind* for raising this issue.

because the subject is unaware of the underlying epistemic support relation.

A taking condition on basing runs into the problem of over-intellectualization. Restricting basing to relations in which the subject takes the reasons to support the conclusion excludes intuitive and important examples of basing. Consider a woman who has experienced consistent but subtle sexual harassment in her office for years. She firmly believes that she has been sexually harassed, but she cannot access the reasons for her belief because she has detected all the cues subconsciously, or because the memories are too traumatic.²⁶ Her belief that she has been sexually harassed seems based on reasons (and justified by those reasons), even though she does not take those reasons to support her belief. This over-intellectualization worry extends to young children, who intuitively seem able to base their beliefs on reasons yet lack the concept of epistemic support, and so cannot take their premises to support their conclusions.²⁷

These examples illustrate that basing does not require taking. Nonetheless, two core ideas behind the taking condition are central to basing: connection to the agent and epistemic dependence. The taking condition is one way of assuring a process is connected to the agent and that its conclusion is epistemically dependent on its premises, but it is not the only way. The Marimba Illusion does so in other ways. I take up the issue of connection to the agent in §5. I now turn to epistemic dependence.²⁸

3.4 Epistemically Sensitive Transition Rules

Basing on reasons requires not only epistemic support relations between contents but also epistemic dependence of the based state on its basis. The Marimba Illusion involves such epistemic dependence. We can see this in the epistemic sensitivity of the process's transition rules.

²⁶ For discussion of similar examples, see Fricker (2007) and Srinivasan (2020).

²⁷ One might wonder whether weaker versions of the taking condition are apt for basing, e.g., the subject must *be able* to take their reason(s) to support their conclusion. If this weaker requirement is understood as about the subject in the moment she instantiates basing, the above examples also show it to be inapt because the subjects are immediately unable to take their reasons to support their conclusions yet intuitively base their beliefs on reasons. If this weaker requirement is understood as about a counterfactual version of the subject in which she has different mental states, then it is met by the Marimba Illusion (e.g., if the subject became aware of her perceptual processing, she would take the presence of a long gesture and a tone apparently caused by that gesture to support the auditory experience of a long tone).

²⁸ One might also see the epistemic dependence argued for in §3.4 as meeting a weaker version of the taking condition on basing according to which taking does not require an independent propositional attitude but is instead embodied in the subject's response to her reasons.

The Marimba Illusion is not a stark on/off coupling between long gestures and long tones, and short gestures and short tones. Instead, the influence varies by degree, in proportion to the plausibility of the causal connection. When the gesture and tone are slightly out of sync, the illusion becomes weaker (Schutz & Kubovy 2009). The tone is still heard as elongated, but to a lesser extent than when the spatiotemporal cues are aligned.

Additional evidence for the epistemic sensitivity of the transition rule comes from the fact that the illusion does not obtain when marimba-striking gestures are paired with sustained clarinet tones (Schutz & Kubovy 2009). The perceptual system registers that impact gestures produce percussive tones and not sustained tones, making the causal connection implausible (Vatakis & Spence 2008, Chuen & Schutz 2016). When the reason the inputs provide does not support the presence of a causal connection, the crossmodal causal representation is not produced.²⁹ The process is governed by transition rules that operate over epistemic reasons.

In contrast, the transition rules governing phosphenes and chromesthesia are not epistemically sensitive. In phosphenes, an increase in optic pressure correlates with an increase in spots seen, but the pressure increase does not provide more (or any) reason in favour of seeing spots. In chromesthesia, as sounds become closer to the word 'green' or more statistically likely to indicate the presence of green, the visual experience of green does not become stronger. Unlike these cases, the Marimba Illusion displays basing's characteristic epistemic dependence.

4. Reasons and Illusion

The presence of all four key markers in the Marimba Illusion indicates that it is an instance of the epistemic basing relation. I have argued that the auditory input, visual input, crossmodal causal representation, and stored information jointly provide a reason for the output. What exactly is this reason?³⁰

The answer to this question depends on one's background ontology of reasons. If reasons are mental states, the reason is simply the states themselves. If reasons are propositions, the reason is the proposition that corresponds to the contents of the mental states: 1) the tone starts

²⁹ Another possibility is that the auditory system produces a representation of the tone as sustained, which functions as a psychological defeater for the crossmodal causal representation.

³⁰ I thank the Editors of *Mind* for posing this question.

at T1 and L1, 2) the gesture occurs from T1 to T2 and L1 to L2, 3) the gesture causes the tone, and 4) when a gesture causes a tone, the length of the gesture is typically proportional to the length of the tone. While 1, 2, and 4, are true, 3 is false. A gesture does cause the tone, but not the gesture that occurs from T1 to T2, which is after the key is struck. This renders the compound proposition false. While false propositions are implausible normative reasons, they are plausible motivating reasons, which are the kinds of reasons involved in basing.³¹

If reasons are facts, the natural account of the reason in the Marimba Illusion would be the fact corresponding to the above compound proposition.³² However, given that the gesture does not cause the tone, there is no such fact. What then is the reason on which the experience is based? This question is not unique to my example. All reasons-as-facts theorists must make sense of error cases in which beliefs are formed in response to false considerations.

Several strategies are available. A first option is to say that there is no reason and hence no basing relation, only 'putative basing'.³³ On this view, the Marimba Illusion does not involve the basing relation. However, this view is unsatisfying because many error cases, including the Marimba Illusion, have all the central features of the basing relation. Not only do error cases often bear the key markers of basing, but their false considerations seem to transmit epistemic support and determine (putatively) based states' epistemic statuses.

A second strategy says that error cases involve basing on reasons that are 'putative facts'.³⁴ Putative facts might be mental states or false propositions. On this view, the reason in the Marimba Illusion is the putative fact constituted by the inputs, crossmodal causal representation, and stored information, or the false proposition to which these states' contents correspond.

A third strategy maintains that reasons are always facts, so error cases involve basing without reasons.³⁵ Instead, the basing is on false mental states that are treated as reasons. This view allows that the relations of the basing relation need not always be a reason (or set of reasons)

³¹ For a defense of the claim that false propositions can be motivating reasons, see [Fantl and McGrath \(2009, ch. 4\)](#).

³² If these states have pictorial rather than propositional content, there is nonetheless a proposition to which their contents correspond.

³³ I thank the Editors of *Mind* for suggesting this possibility.

³⁴ For defenses of the view that reasons can be putative facts, see [Dancy \(2000\)](#), [Setiya \(2007\)](#), [Hornsby \(2008\)](#), [Schroeder \(2008\)](#), and [Comesaña and McGrath \(2014\)](#).

³⁵ For analogous views in the philosophy of action, see [Parfit \(1997\)](#), [Alvarez \(2010\)](#), and [Williamson \(2017\)](#).

and a mental state—they can instead be two mental states. On this view, there is no reason for the experience of the long tone, but it is nonetheless based on the inputs, crossmodal causal representation, and stored crossmodal information. The presence of the four key markers of basing evince that these mental states are treated as reasons by the perceptual systems.

The above discussion demonstrates that multiple views of the ontology of reasons are compatible with my arguments. While the auditory experience in the Marimba Illusion is not based on a fact, one can hold that good cases involve basing on facts while error cases involve another form of basing.

Even if one is convinced by the above arguments that illusory experiences can be based on reasons, one might question whether this conclusion generalizes to veridical perceptions. This worry is particularly pressing for disjunctivists who hold that illusions and veridical perception are of fundamentally different types (e.g., [Hinton 1973](#), [McDowell 2013](#)).³⁶ In response to this worry, my arguments can be extended to veridical forms of crossmodal interaction. The illusory nature of the Marimba Illusion highlights the crossmodal processing structure, so it is a useful focal example, but illusion is inessential to the presence of the four key markers of basing.

Consider a non-illusory marimba experience in which the tone is in fact long. However, the auditory input is slightly degraded so that the tone's ending time and location are unclear, perhaps due to background noise or a head cold. All other variables are held constant. In this scenario, the contents of all states are identical to those of the original Marimba Illusion:

Auditory Input

Content: There is a tone starting at time T1 and location L1.

Visual Input

Content: There is a long gesture occurring from time T1 to T2 and from location L1 to L2.

Crossmodal Causal Representation

Content: The gesture causes the tone (due to spatial and temporal congruence of inputs).

Stored Crossmodal Information

Content: When a gesture causes a tone, the length of the gesture is typically proportional to the length of the tone.

³⁶ Some disjunctivists hold instead that illusions are of the same type as veridical perception, while hallucinations alone are of a different type (e.g., [Snowdon 1990](#), [Brewer 2008](#)).

Output Auditory Experience

Content: There is a long tone.

The only differences between this scenario and the original Marimba Illusion are 1) the cause of the limited auditory input is degraded sensation rather than neglect of available cues, and 2) the output is veridical. The four key markers of basing are still present. The crossmodal causal representation and the stored crossmodal information still have causal contents. The auditory and visual inputs still interface across perceptual subsystems. The output is still epistemically supported by the inputs, crossmodal causal representation, and stored crossmodal information. The process is guided by the same epistemically sensitive transition rules. The disjunctivists' worry is assuaged by this veridical version of the argument for crossmodal basing.

While this is a hypothetical example, veridical crossmodal interactions have been well-studied (e.g., [Vroomen & de Gelder 2000](#), [Ernst & Banks 2002](#), [Alais & Burr 2004](#)). For example, subjects recognize low-contrast images of dogs more accurately when they are paired with barking sounds than when they are paired with meowing sounds ([Yuval-Greenberg & Deouell 2009](#)). Crossmodal interactions are often hypothesized to increase overall perceptual accuracy and reliability, while occasionally resulting in illusory errors ([Shams & Kim 2010](#), [O'Callaghan 2019](#)).

5. The Person-Level

One might worry that my arguments so far have neglected an important criterion on basing—that reasons used in basing must be person-level ([Brewer 1999](#), [Smithies 2019](#), [Teng 2021](#)).³⁷ This worry is well-placed, given that the basing relation underlies epistemic evaluation. Epistemic evaluation is applicable to mental states and processes that are in some sense attributable to the person, rather than to mere whirrings and grindings of mechanisms.

The force of this worry depends on exactly what one means by 'person-level.' The term has taken on a variety of meanings since its introduction ([Drayson 2014](#)). I will address three of these meanings in turn.

The term 'person-level' was originally used to describe a type of psychological explanation. According to [Dennett \(1969\)](#), person-level explanations ascribe mental states and processes to whole persons. For

³⁷ Thanks to an anonymous referee for raising this issue.

example, ‘Irene Adler fled London because she believed Holmes was on her trail’, is a person-level explanation. In contrast, subpersonal explanations ascribe mental states and processes to subsystems, such as the language faculty or visual edge detection. For example, ‘The language faculty parsed the sentence “Colourless green ideas sleep furiously” as syntactically correct because of the phrase structure rules’ is a subpersonal explanation (Chomsky 1975). While Dennett’s distinction is primarily about types of explanation, for simplicity I will refer to states and processes that function in person-level explanations as person-level, and states and processes that only function in subpersonal explanations as subpersonal.³⁸

According to the Dennettian distinction, the states in the Marimba Illusion are person-level because they are ascribed to the individual rather than to her subsystems. The relevant states are 1) the auditory representation of the tone’s onset, 2) the visual representation of the long gesture, 3) the representation that the gesture causes the tone, and 4) the stored information that when a gesture causes a tone, their lengths are proportional. These states all interface across subsystems, rather than functioning solely within one. The auditory and visual inputs combine outside of vision and audition to generate the representation that the gesture causes the tone. The representations that the gesture causes the tone and the stored information about proportional lengths draw on information from both vision and audition. These representations coordinate across subsystems and transcend the divisions of the mind, so they are best ascribed to the individual. On the classic Dennettian meaning of ‘person-level’, the Marimba Illusion meets the person-level criterion on basing.

A related meaning of ‘person-level’ is serving individual functioning, rather than functioning of a subsystem (Burge 2010). Such states and processes may guide action and belief, or coordinate processing across subsystems. Crossmodal interactions serve individual functioning by resolving conflicts across senses, weighing information, and improving the accuracy and reliability of the individual’s perception (O’Callaghan 2019). So, on this definition the Marimba Illusion also meets the person-level criterion on basing.

A more restricted meaning of ‘person-level’ is the availability of a state to guide belief-formation and action.³⁹ The auditory perception of

³⁸ See Wu (2014), Drayson (2017), Teng (2021) for similar terminology.

³⁹ For variations of this view, see e.g., Evans (1982) and Phillips (2020).

the tone's onset and the visual perception of the long gesture are available to guide beliefs such as that a Marimba is being played and actions such as approaching to hear better. Similarly, the representation that the gesture caused the tone might be used to form beliefs about how to play a Marimba, or to guide one's own Marimba-playing. Given the significant psychological evidence that causation is visually perceived (Michotte 1963, Leslie 1987, Rolfs, Dambacher, & Cavanagh 2013), this causal representation should be available for the same kinds of usages as a typical perception.

It is less clear whether the stored information about the proportional lengths of gestures and tones can directly guide belief-formation and action. This state interfaces with perceptual subsystems rather than with belief and action. Yet this point does not threaten the view that the Marimba Illusion involves basing. First, most of its states are person-level, which may be sufficient for basing. Once person-level states and epistemic dependence are involved, the core of the basing relation is present. Second, the availability of a state to guide belief and action does not map onto the sense of 'person-level' that is relevant to the basing relation. While this sense may be useful in other contexts, the central concept of the basing relation as an epistemic dependence relation between a mental state and a reason does not come with a restriction to belief-formation and action. The notion of the person-level that is relevant to basing is one that picks out states and processes that are attributable to the individual because when states are based on reasons, they normatively redound on the individual. In addition to beliefs and actions, perceptions and states that coordinate the individual's subsystems are also attributable to the individual.

In summary, the criterion that epistemic bases must be person-level is either met by the Marimba Illusion or inapt, depending on what one means by 'person-level.' In either case this criterion does not threaten the idea that the Marimba Illusion is an instance of the basing relation. A final point is that the person-level criterion most naturally applies to the based state rather than to the underlying bases (Sylvan 2016). The based state is the only state I claim is epistemically evaluable. In the Marimba Illusion, the based state (the auditory experience of the long tone) is uncontroversially person-level, in virtue of being the individual's conscious experience (Dennett 1969, McDowell 1994, Burge 2010, Phillips 2020), so this version of the person-level criterion is easily met.

6. Conclusion

I have argued that in the Marimba Illusion, a perceptual experience is based on a reason. The next natural question is, is this good epistemic basing, leading to a justified perceptual experience, or poor epistemic basing, leading to an unjustified perceptual experience?

Upon first inspection, it looks like good basing. The inputs, causal representation, and stored information together epistemically support the auditory experience of the long tone. It seems reasonable for the mind to move from representations of the gesture as long, of the gesture as causing the tone, and of such gestures and tones being proportional, to the conclusion that the tone is long. If this analysis is correct, the crossmodal illusion is a justified misfire, producing an experience that is on firm epistemic ground despite its inaccuracy.⁴⁰

However, the process also ignores a crucial piece of information: the auditory input that says the tone is short. The auditory system has perfectly good access to this information. The subject's ears are not blocked, and the sound quality is not degraded. Auditory information that could be used to directly calculate tone duration is registered but is overlooked or overridden due to visual dominance. This resembles evidential neglect, which typically renders beliefs unjustified.

Whether this form of neglect impugns the epistemic status of the experience depends on whether the neglected auditory input can provide a reason. This input is likely a state of sensory registration rather than a full-fledged mental representation. It has the potential to become a representation of tone duration if it is further processed. The system ignores something that could provide a reason if properly pursued, akin to a detective failing to follow up on a promising lead. The questions of whether and when we are rationally required to pursue potential reasons cuts across the domains of perception and belief.⁴¹ While answers to these questions are beyond the scope of this paper, the fact that they surface at this juncture demonstrates that crossmodal interactions fit congruously into the broader epistemic landscape.⁴²

⁴⁰ Other crossmodal illusions that may admit of this justified misfire analysis include the McGurk Effect (McGurk & MacDonald 1976), the Rubber Hand Illusion (Botvinick & Cohen 1998), the Tone-Flash Illusion (Shams, Kamitani, & Shimojo 2002), ambiguous motion (Maeda, Kanai, & Shimojo 2015), and bouncing/streaming objects (Sekuler, Sekuler, & Lau 1997).

⁴¹ For discussion of the availability of reasons, see Parfit (1997), Williamson (2000), Kelly (2008), and Schroeder (2008). For discussion of the norms of attention and inquiry, see Foley and Fumerton (1982), Siegel (2017), and Friedman (2017).

⁴² Other ways crossmodal illusions might be unjustified include misweighting of reasons and modulation by emotion rather than reasons (e.g., Wang, Wang, & Spence 2016).

The view that crossmodal experiences can be based on reasons extends beyond the Marimba Illusion. Recently, several philosophers and psychologists have argued that crossmodal perceptual processing is the norm rather than the exception (Macpherson 2011b, Spence 2011, O'Callaghan 2019, Nanay 2020). If so, much of perception may involve epistemic basing. Other crossmodal interactions that exploit the Unity Assumption, such as the Tone-Flash Illusion (Shams, Kamitani, & Shimojo 2002), the Ventriloquist Effect (Alais & Burr 2004), and the Rubber Hand Illusion (Botvinick & Cohen 1998), are good candidates. Crossmodal interactions with simpler associative or brute physical mechanisms like phosphenes and chromesthesia are not.

My arguments here destabilize the traditional epistemic divisions between perception and belief. The epistemic role of the senses is not only to gather information from the external world through sensory exploration (Matthen 2014), but also to base experiences on reasons provided across modalities. Basing can occur in unconscious, automatic, and perceptual parts of the mind. While crossmodal illusions give us distorted reflections of reality, they are often reasoned distortions rather than arational ones.⁴³

References

- Alais, David, and David Burr 2004, 'The Ventriloquist Effect Results from Near-Optimal Bimodal Integration', in *Current Biology* 14, pp. 257–262
- Alvarez, Maria 2010, *Kinds of Reasons: An Essay on the Philosophy of Action* (Oxford: Oxford University Press)
- Audi, Robert 2008, 'Intuition, Inference, and Rational Disagreement in Ethics', in *Ethical Theory and Moral Practice* 11(5), pp. 475–492
- Auvray, Malika, and Ophelia Deroy 2015, 'How Do Synesthetes Experience the World?', in M. Matthen (ed.), *Oxford Handbook of the Philosophy of Perception* (New York: Oxford University Press)
- Bayne, Timothy 2014, 'The Multisensory Nature of Perceptual Consciousness', in D.J. Bennett and C.S. Hill (eds.), *Sensory Integration*

⁴³ I thank Selim Berker, Ned Block, Frédérique de Vignemont, Todd Ganson, E.J. Green, Fiona Macpherson, Casey O'Callaghan, Jake Quilty-Dunn, Susanna Rinard, Susanna Siegel, Elizabeth Spelke, and Timothy Williamson for their feedback on the ideas in this article. I am also grateful to audiences at the Harvard Metaphysics and Epistemology Workshop, the Southern Society for Philosophy and Psychology, the Washington University Mind Group, and the MIT Works-in-Progress Group for their questions and discussion. I also thank the Editors of *Mind* and two anonymous referees for their helpful comments.

- and the Unity of Consciousness*, pp. 15–36 (Cambridge, MA: MIT Press)
- Bergeron, Vincent, and Dominic McIver Lopes 2009, ‘Hearing and Seeing Musical Expressions’, in *Philosophy and Phenomenological Research* 78(1), pp. 1–16
- Boghossian, Paul 2014, ‘What is Inference?’ in *Philosophical Studies* 169, pp. 1–18
- Botvinick, Matthew, and Jonathan Cohen 1998, ‘Rubber Hands ‘Feel’ Touch that Eyes See’, in *Nature* 391, p. 756
- Brewer, Bill 1999, *Perception and Reason* (Oxford: Oxford University Press)
- 2008, ‘How to Account for Illusion’, in A. Haddock and F. Macpherson (eds.), *Disjunctivism: Perception, Action, and Knowledge*, pp. 168–180 (Oxford: Oxford University Press)
- Bronte, Emily 1847/1995, *Wuthering Heights* (London: Penguin Books)
- Burge, Tyler 2010, *Origins of Objectivity* (Oxford: Oxford University Press)
- Carey, Susan 2009, *The Origin of Concepts* (New York: Oxford University Press)
- Chomsky, Noam 1975, *The Logical Structure of Linguistic Theory* (New York: Springer)
- Chudnoff, Elijah 2014, ‘The Rational Roles of Intuition’, in A. Booth and D. Rowbottom, (eds.), *Intuitions*, pp. 9–35 (Oxford: Oxford University Press)
- Chuen, Lorraine, and Michael Schutz 2016, ‘The Unity Assumption Facilitates Cross-Modal Binding of Musical, Non-Speech Stimuli: The Role of Spectral and Amplitude Envelope Cues’ in *Attention, Perception, & Psychophysics* 78, pp. 1512–1528
- Comesaña, Juan and Matthew McGrath 2014, ‘Having False Reasons’, in C. Littlejohn and J. Turri (eds.), *Epistemic Norms*, pp. 59–80 (Oxford: Oxford University Press)
- Conan Doyle, Arthur 1893/1993, *The Memoirs of Sherlock Holmes* (New York: Oxford University Press)
- Connolly, Kevin 2014, ‘Making Sense of Multiple Senses’, in R. Brown (ed.), *Consciousness Inside and Out: Phenomenology, Neuroscience, and the Nature of Experience* (Dordrecht: Springer)
- Dancy, Jonathan 2000, *Practical Reality* (Oxford: Oxford University Press)
- Davidson, Donald 1963, ‘Actions, Reasons, and Causes’, in *The Journal of Philosophy* 60(23), pp. 685–700

- Davidson, Jane W. 1993, 'Visual Perception of Performance Manner in the Movements of Solo Musicians', in *Psychology of Music* 21(2), pp. 103–113
- de Gelder, Beatrice, Gilles Pourtois, Jean Vroomen, and Anne-Catherine Bachoud-Lévi 2000, 'Covert Processing of Faces in Prosopagnosia is Restricted to Facial Expressions: Evidence from Cross-Modal Bias', in *Brain and Cognition* 44(3), pp. 425–44
- de Gelder, Beatrice, Gilles Pourtois, and Lawrence Weiskrantz 2002, 'Fear Recognition in the Voice is Modulated by Unconsciously Recognized Facial Expressions but Not by Unconsciously Recognized Affective Pictures', in *Proceedings of the National Academy of Sciences of the United States of America* 99(6), pp. 4121–4126
- Deroy, Ophelia, and Charles Spence 2013, 'Why We Are Not All Synesthetes (Not Even Weakly So)', in *Psychonomic Bulletin & Review* 20, pp. 1–22
- Dennett, Daniel 1969, *Content and Consciousness* (London: Routledge)
- Drayson, Zoe 2014, 'The Personal/Subpersonal Distinction', in *Philosophy Compass* 9(5), pp. 338–346
- 2017, 'Psychology, Personal and Subpersonal', in *The Routledge Encyclopedia of Philosophy* (Taylor and Francis)
- Driver, John and Toemme Noesselt 2007, 'Multisensory Interplay Reveals Crossmodal Influences on 'Sensory-Specific' Brain Regions, Neural Response, and Judgments', in *Neuron* 57(1), pp. 11–23
- Egan, Andy 2008, 'Seeing and Believing: Perception, Belief Formation, and the Divided Mind', in *Philosophical Studies* 140, pp. 47–63
- Ernst, Marc O. and Martin S. Banks 2002, 'Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion', in *Nature* 415(6870), pp. 429–433
- Evans, Gareth 1982, *The Varieties of Reference* (Oxford: Oxford University Press)
- Fantl, Jeremy and Matthew McGrath 2009, *Knowledge in an Uncertain World* (Oxford: Oxford University Press)
- Fischer, John Martin and Mark Ravizza 1998, *Responsibility and Control: A Theory of Moral Responsibility* (Cambridge, UK: Cambridge University Press)
- Fish, William 2010, *Philosophy of Perception: A Contemporary Introduction* (New York: Routledge)
- Fodor, Jerry 1983, *Modularity of Mind* (Cambridge, MA: MIT Press)
- Foley, Richard and Richard Fumerton 1982, 'Epistemic Indolence', in *Mind* 91(361), pp. 38–56

- Frank, Robert A. and Jennifer Byram 1988, 'Taste-Smell Interactions are Tastant and Odor Dependent', in *Chemical Senses* 12(3), pp. 445–455
- Fricker, Miranda 2007, *Epistemic Injustice: Power and the Ethics of Knowing* (Oxford: Oxford University Press)
- Friedman, Jane 2017, 'Inquiry and Belief', in *Noûs* 53(2), pp. 296–315
- Green, E. J. 2019, 'Binding and Differentiation in Multisensory Object Perception', in *Synthese* 198, pp. 4457–4491
- Hamlin, J. Kiley, Karen Wynn, and Paul Bloom 2007, 'Social Evaluation by Preverbal Infants', in *Nature* 450, pp. 557–560
- Heller, Morton A. 1982, 'Visual and Tactual Texture Perception: Intersensory Cooperation', in *Perception and Psychophysics* 31, pp. 339–344
- Hinton, J. M. 1973, *Experiences: An Inquiry into Some Ambiguities* (Oxford: Clarendon Press)
- Hornsby, Jennifer 2008, 'A Disjunctive Conception of Acting for Reasons', in A. Haddock and F. Macpherson (eds.), *Disjunctivism: Perception, Action, Knowledge*, pp. 244–261 (Oxford University Press)
- Huang, Juan, Darik Gamble, Kristine Sarnlertsophon, Xiaoqin Wang, and Steven Hsiao 2012, 'Feeling Music: Integration of Auditory and Tactile Inputs in Musical Meter', in *PLoS ONE* 7(10), p. e48496
- Hyman, John 1999, 'How Knowledge Works', in *The Philosophical Quarterly* 49(197), pp. 433–451
- Ichikawa, Jonathan 2009, 'Dreaming and Imagination', in *Mind and Language* 24(1), pp. 103–121
- Jenkin, Zoe 2020, 'The Epistemic Role of Core Cognition', *Philosophical Review* 129(2), pp. 251–298.
- Johnson-Laird, Philip 1997, 'Rules and Illusions: A Critical Study of Rips' The Psychology of Proof', in *Minds and Machines* 7(3), pp. 387–407
- Keetels, Mirjam, Jeroen Stekelenburg, and Jean Vroomen 2007, 'Auditory Grouping Occurs Prior to Intersensory Pairing: Evidence from Temporal Ventriloquism', in *Experimental Brain Research* 180(3), pp. 449–456
- Kelly, Thomas 2008, 'Evidence: Fundamental Concepts and the Phenomenal Conception', in *Philosophy Compass* 3(5), pp. 933–955
- Kim, Jaegwon 1988, 'What is "Naturalized Epistemology?"' in *Philosophical Perspectives* 2, pp. 381–405
- Kosugi, Daisuke and Kazuo Fujita 2002, 'How Do 8-Month-Old Infants Recognize Causality in Object Motion and That in Human action?' in *Japanese Psychological Research* 44(2), pp. 66–78

- Kotovskiy, Laura and Renée Baillargeon 2000, 'Reasoning About Collisions Involving Inert Objects in 7.5-Month-Old Infants', in *Science* 3(3), pp. 344–359
- Kubovy, Michael and Michael Schutz 2010, 'Audiovisual Objects', in *The Review of Philosophy and Psychology* 1, pp. 41–61
- Leslie, Alan M. 1987, 'Do Six-Month-Old Infants Perceive Causality?' in *Cognition* 25, pp. 265–288
- 1988, 'The Necessity of Illusion: Perception and Thought in Infancy', in L. Weiskrantz (ed.), *Thought Without Language*, pp. 185–210 (Oxford, UK: Oxford Science Publications)
- Levy, Neil 2014, 'Consciousness, Implicit Attitudes, and Moral Responsibility', in *Noûs* 48(1), pp. 21–40
- Lewis, David 1982, 'Logic for Equivocators', in *Noûs* 16, pp. 431–441
- Littlejohn, Clayton 2012, *Justification and the Truth Connection* (Oxford: Oxford University Press)
- Macpherson, Fiona 2007, 'Synesthesia, Functionalism, and Phenomenology', in M. Marraffa, M. de Caro, and F. Ferretti (eds.), *Cartographies of the Mind: Philosophy and Psychology in Intersection, volume 4 of Studies in Brain and Mind*, pp. 65–80 (Dordrecht: Springer)
- 2009, 'Perception, Philosophical Perspectives', in T. Bayne, A. Cleeremans, P. Wilken (eds.), *The Oxford Companion to Consciousness*, pp. 502–508 (Oxford: Oxford University Press)
- 2011a, 'Cross-Modal Experiences', in *Proceedings of the Aristotelian Society* 111, pp. 429–468
- 2011b, 'Taxonomising the Senses', in *Philosophical Studies* 153(1), pp. 123–142
- Maeda, Fumiko, Ryota Kanai, and Shinsuke Shimojo 2015, 'Changing Pitch Induced Visual Motion Illusion', in *Current Biology* 14(23), pp. R990–R991
- Matthen, Mohan 2014, 'How to Be Sure: Sensory Exploration and Empirical Certainty', in *Philosophy and Phenomenological Research* 88(1), pp. 38–69
- McDowell, John 1994, 'The Content of Perceptual Experience', in *The Philosophical Quarterly* 44(175), pp. 190–205
- 2013, 'Acting in the Light of a Fact', in D. Bakhurst, B. Hooker, and M.O. Little (eds.), *Thinking About Reasons: Themes from the Philosophy of Jonathan Dancy*, pp. 13–28 (Oxford: Oxford University Press)
- McGrath, Matthew 2013, 'Phenomenal Conservatism and Cognitive Penetration: The Bad Basis Counterexamples', in C. Tucker (ed.), *Seemings and Justification: New Essays on Dogmatism and Phenomenal Conservatism*, pp. 225–247 (Oxford: Oxford University Press)

- McGurk, Harry, and John MacDonald 1976, 'Hearing Lips and Seeing Voices', in *Nature* 264264(5588), pp. 746–748
- McHugh, Conor and Jonathan Way 2016, 'Against the Taking Condition', in *Philosophical Issues* 26(1), pp. 314–331
- Mesgarani, Nima, Stephen V. David, Jonathan B. Fritz, and Shihab A. Shamma 2008, 'Phoneme Representation and Classification in Primary Auditory Cortex', in *Journal of the Acoustic Society of America* 123(2), pp. 899–909
- Michotte, Albert 1963, *The Perception of Causality* (Oxford: Basic Books)
- Muentener, Paul and Susan Carey 2010, 'Infants' Causal Representations of State Change Events', in *Cognitive Psychology* 61(2), pp. 63–86
- Munton, Jessie 2019, 'Perceptual Skill and Social Structure', in *Philosophy and Phenomenological Research* 99(1), pp. 131–161
- Nanay, Bence 2020, 'Multimodal Mental Imagery and Perceptual Justification', in B. Brogaard and D. E. Gatzia (eds.), *The Epistemology of Non-Visual Perception* (New York: Oxford University Press)
- Neta, Ram 2019, 'The Basing Relation', in *The Philosophical Review* 128(2), pp. 179–217
- O'Callaghan, Casey 2008, 'Seeing What You Hear: Cross-Modal Illusions and Perception', in *Philosophical Issues* 18(1), pp. 316–338
- 2017a, 'Grades of Multisensory Awareness', in *Mind & Language* 32(2), pp. 155–181
- 2017b, 'Synesthesia vs. Crossmodal Illusions', in O. Deroy (ed.), *Sensory Blendings: New Essays on Synesthesia*, pp. 45–58 (Oxford: Oxford University Press)
- 2019, *A Multisensory Philosophy of Perception* (Oxford: Oxford University Press)
- 2020, 'Multisensory Evidence', in *Philosophical Perspectives* 30(1), pp. 238–256
- Oaksford, Mike and Nick Chater 2007, *Bayesian Rationality: The Probabilistic Approach to Human Reasoning* (Oxford: Oxford University Press)
- Parfit, Derek 1997, 'Reasons and Motivation', in *Proceedings of the Aristotelian Society Supplementary Volume* 71, pp. 99–130
- Phillips, Ian 2020, 'Object Files and Unconscious Perception: A Reply to Quilty-Dunn', in *Analysis* 8(2), pp. 293–301
- Prettyman, Adrienne 2019, 'Perceptual Precision', in *Philosophical Psychology* 32(6), pp. 923–944
- Pylyshyn, Zenon 1999, 'Is Vision Continuous with Cognition? The Case for Cognitive Impenetrability of Visual Perception', in *Behavioral and Brain Sciences* 22, pp. 341–423

- Quilty-Dunn, Jake and Eric Mandelbaum 2018, 'Inferential Transitions', in *Australasian Journal of Philosophy* 96(3), pp. 532–547
- Railton, Peter 2020, 'Comment on Susanna Siegel, "The Rationality of Perception"', in *Philosophical and Phenomenological Research* 101(3), pp. 735–754
- Ramachandran, Vilayanur S. and Edward M. Hubbard 2001, 'Synesthesia—A Window into Perception, Thought and Language', in *Journal of Consciousness Studies* 8(12), pp. 3–34
- Reid, Thomas 1764/1997, *An Inquiry into the Human Mind on the Principles of Common Sense*, ed. D. Brookes (University Park, PA: Pennsylvania State University Press)
- Rips, Lance 1994, *The Psychology of Proof* (Cambridge, MA: MIT Press)
- Rolfs, Martin, Michael Dambacher, and Patrick Cavanagh 2013, 'Visual Adaptation of the Perception of Causality', in *Current Biology* 23, pp. 250–254
- Rosenblum, Lawrence D. 2005, 'Primacy of Multimodal Speech Perception', in D. B. Pisoni and R. E. Remez (eds.) *The Handbook of Speech Perception*, pp. 51–78 (Malden, MA: Blackwell)
- Saxe, Rebecca and Susan Carey 2006, 'The Perception of Causality in Infancy', in *Acta Psychologica* 123, pp. 144–165
- Scanlon, T.M. 1998, *What We Owe to Each Other* (Cambridge, MA: Harvard University Press)
- 2014, *Being Realistic About Reasons* (Oxford: Oxford University Press)
- Schlottmann, Anne and David R. Shanks 1992, 'Evidence for a Distance Between Judged and Perceived Causality' in *Quarterly Journal of Experimental Psychology* 44A, pp. 321–342
- Scholl, Brian J. and Patrice D. Tremoulet 2000, 'Perceptual Causality and Animacy', in *Trends in Cognitive Science* 4, pp. 299–309
- Schroeder, Charles E. and John Foxe 2005, 'Multisensory Contributions to Low-Level, 'Unisensory' Processing', in *Current Opinion in Neurobiology* 15(4), pp. 454–458
- Schroeder, Mark 2008, 'Having Reasons', in *Philosophical Studies* 139(1), pp. 57–71
- Schutz, Michael and Michael Kubovy 2009, 'Causality and Crossmodal Integration', in *Journal of Experimental Psychology: Human Perception and Performance* 35(6), pp. 1791–1810
- Schutz, Michael and Scott Lipscomb 2007, 'Hearing Gestures, Seeing Music: Vision Influences Perceived Tone Duration', in *Perception* 36, pp. 888–897

- Sekuler, Robert, Allison Sekuler, Renee Lau 1997, 'Sound Alters Visual Motion Perception', in *Nature* 385, pp. 308
- Setiya, Kieran 2007, *Reasons Without Rationalism* (Princeton, NJ: Princeton University Press)
- Shams, Ladan, Yukiyasu Kamitani, and Shinsuke Shimojo 2002, 'Visual Illusion Induced by Sound', in *Cognitive Brain Research* 14, pp. 147–152
- Shams, Ladan and Robyn Kim 2010, 'Crossmodal Influences on Visual Perception', in *Physics of Life Reviews* 7, pp. 269–84
- Siegel, Susanna 2010, *The Contents of Visual Experience* (Oxford: Oxford University Press)
- 2017, *The Rationality of Perception* (Oxford: Oxford University Press)
- Smith, A.D. 2002, *The Problem of Perception* (Cambridge, MA: Harvard University Press)
- Smith, Barry 2015, 'The Chemical Senses', in M. Matthen (ed.) *The Oxford Handbook of Philosophy of Perception*, pp. 314–353 (New York: Oxford University Press)
- Smithies, Declan 2019, *The Epistemic Role of Consciousness* (New York: Oxford University Press)
- Snowdon, Paul F. 1990, 'The Objects of Perceptual Experience', in *Proceedings of the Aristotelian Society Supplementary Volume* 64, pp. 121–150
- Sosa, Ernest 1999, 'How to Defeat Opposition to Moore', in *Philosophical Perspectives* 13, pp. 141–153
- Spence, Charles 2011, 'Crossmodal Correspondences: A Tutorial Review', in *Attention, Perception, & Psychophysics* 73(4), pp. 971–995
- 2015, 'Eating with Our Ears: Assessing the Importance of the Sounds of Consumption on our Perception and Enjoyment of Multisensory Flavour Experiences', in *Flavour* 4(3)
- Spence, Charles and Ophelia Deroy 2013, 'Crossmodal Mental Imagery', in S. Lacey and R. Lawson (eds.), *Multisensory Imagery*, pp. 157–183 (New York: Springer)
- Sperandio, Irene, Phillippe A. Chouinard, and Melvin A. Goodale 2012, 'Retinotopic Activity in V1 Reflects the Perceived and Not the Retinal Size of an Afterimage', in *Nature Neuroscience* 15, pp. 540–542
- Srinivasan, Amia 2020, 'Radical Externalism', in *The Philosophical Review* 129(3), pp. 395–431

- Stekelenburg, Jeroen J. and Jean Vroomen 2007, 'Neural Correlates of Multisensory Integration of Ecologically Valid Events', in *Journal of Cognitive Neuroscience* 19(12), pp. 1964–1973
- Sylvan, Kurt 2016, 'Epistemic Reasons II: Basing', in *Philosophy Compass* 11(7), pp. 377–389
- 2018, 'Knowledge as a Non-Normative Property', in *Philosophy and Phenomenological Research* 97(1), pp. 190–222
- Teng, Lu 2021, 'Cognitive Penetration: Inference or Fabrication?' in *Australasian Journal of Philosophy* 99(3), pp. 547–563
- Turri, John 2009, 'The Ontology of Epistemic Reasons' in *Noûs* 43, pp. 490–512
- Vatakis, Argiro and Charles Spence 2008, 'Evaluating the Influence of the 'Unity Assumption' on the Temporal Perception of Realistic Audiovisual Stimuli', in *Acta Psychologica* 127(1), pp. 12–23
- Vines, Bradley W., Carol L. Krumhansl, Marcello M. Wanderly, and Daniel J. Levitin 2005a, 'Cross-Modal Interactions in the Perception of Musical Performance', in *Cognition* 101(1), pp. 80–113
- Vines, Bradley W., Carol L. Krumhansl, Marcello M. Wanderly, Ioana M. Dalca, and Daniel Levitin 2005b, 'Dimensions of Emotion in Expressive Musical Performance', in *Annals of the New York Academy of Sciences* 1060, pp. 462–466
- Vroomen, Jean and Beatrice de Gelder 2000, 'Sound Enhances Visual Perception: Cross-Modal Effects of Auditory Organization on Vision' in *Journal of Experimental Psychology: Human Perception and Performance* 26, pp. 1583–90
- Wang, Qian J., Sheila Wang and Charles Spence 2016, "'Turn up the Taste": Assessing the Role of Intensity and Emotion in Mediating Crossmodal Correspondences Between Basic Taste and Pitch', in *Chemical Senses* 41(4), pp. 345–356
- Watkins, Susanne, Ladan Shams, Shigeki Tanaka, John-Dylan Haynes, and Geraint Rees 2006, 'Sound Alters Activity in Human V1 in Association with Illusory Visual Perception', in *NeuroImage* 31(3), pp. 1247–1256
- Welch, Robert B. and David H. Warren 1980, 'Immediate Perceptual Response to Intersensory Discrepancy' in *Psychological Bulletin* 88(3), pp. 638–667
- Whitaker, T. Aisling, Christina Simões-Franklin, and Fiona N. Newell 2008, 'Vision and Touch: Independent or Integrated Systems for the Perception of Texture?', in *Brain Research* 1242m pp. 59–72

- Williamson, Timothy 2000, *Knowledge and Its Limits* (Oxford: Oxford University Press)
- 2017, ‘Acting on Knowledge’, in J. Adam Carter, Emma C. Gordon, and Benjamin W. Jarvis (eds.), *Knowledge-First*, pp. 163–181(Oxford: Oxford University Press)
- Wright, Crispin 2014, ‘Comment on Paul Boghossian, “What is Inference”’, in *Philosophical Studies* 169(1), pp. 1–11
- Wu, Wayne 2014, *Attention* (London: Routledge)
- Yuval-Greenberg, Shlomit, and Leon Y. Deouell 2009, ‘The Dog’s Meow: Asymmetrical Interaction in Cross-Modal Object Recognition’, in *Experimental Brain Research* 193, pp. 603–614