Journal of the American Philosophical Association (2021) 329–347 © The Author(s), 2021. Published by Cambridge University Press on behalf of the American Philosophical Association. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.
DOI:10.1017/apa.2020.29

Implicit Bias as Mental Imagery

ABSTRACT: *What is the mental representation that is responsible for implicit bias? What is this representation that mediates between the trigger and the biased behavior? My claim is that this representation is neither a propositional attitude nor a mere association (as the two major accounts of implicit bias claim). Rather, it is mental imagery: perceptual processing that is not directly triggered by sensory input. I argue that this view captures the advantages of the two standard accounts without inheriting their disadvantages. Further, this view also explains why manipulating mental imagery is among the most efficient ways of counteracting implicit bias.*

KEYWORDS: implicit bias, mental imagery, propositions, associations

I. Bias versus Attitude

Some of our behavior is biased. By this I mean that the behavior goes against our reported beliefs. And often we are not fully aware of these biases. I take bias to be an observable phenomenon: a feature of our observable behavior (see, for example, De Houwer 2019). The question I raise in this article is about the mental representation that is responsible for this biased behavior (often referred to as *implicit attitude*). This representation mediates between the trigger and the biased behavior (Holroyd 2016; Brownstein 2016; Del Pinal and Spaulding 2018; cf. Johnson 2019 for an alternative framing of implicit bias). And my claim is that this representation is neither a propositional attitude nor a mere association (as the two major accounts of implicit bias would claim). It is mental imagery.

I am interested in the representation that mediates between the trigger and the biased behavior, which I call *biasing representation*. But this picks out a lot of things. When people are reluctant to throw darts at the picture of the face of a loved one or to drink lemonade from a sterilized bedpan (Rozin et al. 1990; see also Gendler 2008), we also get behavior that goes against our beliefs (which represent these actions as harmless). The term *implicit bias* is more often used to describe a more specific phenomenon: representation about certain racial and gender groups that influences our behavior in such a way that it goes against our

This work was supported by a European Research Council Consolidator grant (no. 726251), Fonds voor Wetenschappelijk Onderzoek Odysseus grant (G.0020.12N), and Fonds voor Wetenschappelijk Onderzoek research grant (G0C7416N) and Fonds voor Wetenschappelijk Onderzoek research grant (G0E0218N). Special thanks for comments by Eric Mandelbaum, Ema Sullivan-Bissett, Beate Krickel, Dan Williams, Gerardo Viera, Kevin Lande, Nick Wiltsher, Kris Goffin, Chris McCarroll, Brad Saad, Laura Silva, and two anonymous referees. I presented an earlier version of this article at the Department of Psychology of the University of Ghent in September 2018, and I am grateful for feedback from Jan De Houwer, Agnes Moors, Frederick Verbruggen, Pieter Van Dessel, Marcel Brass, and Hugo Mercier.

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

beliefs (Greenwald, McGhee, and Schwartz 1998; Greenwald and Nosek 2009; Dunham, Baron, and Banaji 2008).

Here is an example. In an elevator, you might stand a little bit farther away from people whose skin color is different from yours. You may or may not be aware of this. But you think of yourself as someone who does not make distinctions between people because of their skin color. So your behavior is biased in the sense that it goes against your belief. And the mental representation (presumably about people with various skin colors) that is responsible for this behavior is what this article is about. Implicit bias is a genuinely heterogenic phenomenon (see Holroyd and Sweetman 2016 for a detailed taxonomy; see Johnson 2019 for even more heterogeneity). Because of this heterogeneity, it may be helpful to consider examples from the literature about cognitive biases in general alongside various (diverse) examples of implicit bias.

The two main candidates in the literature are the following. This biasing representation might be an association—between a specific skin color and a specific trait, say, being dangerous. Or this biasing representation might be a propositional attitude – an attitude toward the proposition that people with this specific skin color tend to be dangerous. My aim is to carve out and defend a third option, according to which this representation is mental imagery: perceptual processing that is not directly triggered by sensory input. I argue that this view captures the advantages of the two standard accounts without inheriting their disadvantages.

In short, empirical findings show that the biasing representation would need to be both sensitive to semantic content and insensitive to logical form. But associations are not sensitive to semantic content. And propositional attitudes are not insensitive to logical form. I argue that mental imagery is much better suited to fulfill the theoretical role of the biasing representation: it is sensitive to semantic content (unlike associations) and insensitive to logical form (unlike propositional attitudes). My aim is not to dismiss completely the associationist or the propositionalist account of implicit bias (or both), but rather to reframe the debate by adding an extra important ingredient to any explanation of implicit bias (be they associationist, propositionalist): mental imagery.

Some clarifications: there is emerging evidence that implicit bias may not be unconscious: we may be more aware of our implicit attitudes than the original bias findings might suggest (Nier 2001; Hahn et al. 2014; Ranganath, Smith, and Nosek 2008; Machery 2016; Toribio 2018; Berger 2020). Nothing I say in this article takes sides on this issue. Further, there have been some debates about just what the most highly publicized indicator of implicit bias, the online freely available Implicit Association Test shows or does not show (Forscher et al. 2019). As I take the Implicit Association Test to be only one of many experimental procedures that aims to demonstrate implicit bias, I will leave this debate to the side. If the reader is skeptical of the Implicit Association Test, this is not a reason to be skeptical of the broader phenomenon of biased behavior in general.

II. Association versus Proposition

The question about what these biasing representations are is crucial not just out of theoretical interest. If we want to try to eliminate implicit bias, very

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

different procedures would be needed depending on what these biasing representations are.

Suppose that you have an implicit bias that makes it more likely that when you think of a caregiver, you think of a woman. This is in fact a very widespread example of implicit bias. The question is: What is the underlying biasing representation? The classic candidate is association. You have probably seen more female caregivers than male caregivers. And, following the mechanism of classic conditioning, you formed an association between being a caregiver and being a woman. One way to think about associations is as some kind of connection strength in your mind between the concept of being a caregiver and the concept of being a woman. When one concept is activated, the other one is very likely to be also activated. So when you hear someone talk about a caregiver, this gives rise to your thinking of a woman. Association is supposed to be quick, not under our voluntary control and, according to many (see Mandelbaum 2016 for a summary), symmetrical (it goes both from caregiver to woman and vice versa).

The alternative view is that the underlying biasing representation is a propositional attitude (typically a belief). So you have a propositional attitude that caregivers are (likely to be) women. And it is this propositional attitude that explains your biased behavior. In the case of propositional attitudes, the relation between being a caregiver and being a woman is not symmetrical. The propositional attitude that caregivers are (or tend to be) women is different from the propositional attitude that women are (or tend to be) caregivers. (See especially Mandelbaum 2016; Levy 2015. Some psychologists also often talk about propositions in this context, although what they mean by this tends to be something very different, see De Houwer, Thomas, and Baeyens 2001; De Houwer 2009, 2011.).[1]

I argue that the biasing representation is neither an association nor a propositional attitude: it is mental imagery. The plan is the following: In section III, I outline what I take to be the most important empirical findings for and against both the associationist and the propositionalist approaches. In section IV, I introduce the concept of mental imagery, understood as early perceptual processing that is not directly triggered by sensory input (and attempt to dispel some common misconceptions about the term). In sections V and VI, I argue that mental imagery can fill the functional role of biasing representation much better than its alternatives and it can combine the strengths of the associationist and of the propositional approaches without inheriting their weaknesses. Finally, in section VII, I give yet another strong reason in favor of the view that the biasing representation of implicit bias is mental imagery, which comes from some of the most efficient ways of counteracting implicit bias (namely, procedures involving mental imagery).

[1] What Jan De Houwer means by proposition is very different from what philosophers mean by propositional attitudes. Proposition, for De Houwer, means any representation that represents a relation (De Houwer 2009, 2011). Needless to say, many kinds of representations would qualify according to this definition, besides propositional attitudes. Some perceptual states would count as would mental imagery.

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

III. Empirical Findings

As noted above, my aim in this section is to review the most important empirical arguments against the association view and against the propositional view. I start with the arguments against the association view because they are more developed and more influential (see especially Mandelbaum 2013 and Mandelbaum 2016, from which the next four examples are taken). Many of these examples are not about implicit bias concerning social groups, but about cognitive biases in general. My assumption is that considering implicit bias about social groups against the backdrop of the wider category of cognitive biases and drawing on the larger set of examples can help us to pinpoint some important features of implicit bias.

III.a. Sugar

Subjects are shown two bottles. Both of them are filled from a bag that contains sugar. Then they are supposed to affix a label on each bottle. One says 'Sucrose'. The other says 'Sodium cyanide'. It is the subjects themselves who do this. Later they nonetheless prefer the beverage that was sweetened with the sugar from the bottle that says 'Sucrose' (Rozin, Markwith, and Ross 1990; see also Mandelbaum 2013).

This is clearly biased behavior. The subjects firmly believe that both bottles contain the same substance. When asked, they say exactly this. But their behavior goes against these firmly held beliefs.

How could the association view explain this? There is, presumably, an association between the label 'Sodium cyanide' and being dangerous as sodium cyanide is a poison. But this association does not itself explain the behavior. The behavior is about two cups of sugary beverages, and the experimenter poured sugar into these cups from bottles that had the different labels. The association is between the label and danger. The label then is attached to the bottle, the bottle is filled with sugar and the sugar is poured into the cup. But the cup somehow inherits this association to danger after all these moves. The associationist has no clear explanation of how this happens.

The propositionalist would argue that the sugar that used to be in the bottle with the 'Sodium cyanide' label inherits the property of being dangerous from the words on the label. So the property of being dangerous is predicated of the content of the bottle with the label 'Sodium cyanide'. And then given that the content of the bottle is poured in the cup, it is predicated of the cup. And this is what explains the biased behavior (involving the cup).

III.b. Celebrity Contagion

Subjects could bid for pieces of clothing worn by celebrities. People generally pay more money for clothes worn by celebrities than for identical clothes that are fresh off the shelf. But what this experiment showed was that they pay a little less if these clothes were washed since the celebrity wore them (Newman, Diesendruck, and Bloom 2011).

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

Again, this is biased behavior, and the propositionalist claims that it can be explained only if we take the biasing representation to be a propositional attitude. We have two representations here. The first one (A) is the biasing representation about the connection between the clothes and the fact that they were worn by a celebrity. And this explains why we pay a lot of money for clothes worn by celebrities. But we also have another representation (B) about the connection between clothes and washing. And our biased behavior of paying less money for washed celebrity clothes than for unwashed celebrity clothes is explained by the interaction between the biasing representation (A) and the representation about washing (B). But, the argument goes, mere associations do not combine with other representations in a content-sensitive manner. So the biasing representation is a propositional attitude.

III.c. Enemy of My Enemy

Subjects are made to dislike person A (by implicit clues). Then they are told that person A dislikes person B. The surprising result is that subjects then tend to like person B. The enemy of my enemy is my friend (Gawronski, Walther, and Blank 2005, see also Gawronski et al. 2006).

Subjects show biased behavior inasmuch as just because person A dislikes someone, this gives no reason to like or dislike this person. But the biasing representation, according to the propositionalist, must be a propositional attitude as it is only in terms of propositional attitudes that we can derive a positive from two negatives.

III.d. Peer Pressure

Your knowledge of other people's views influences implicit bias. The task is merely to sit down next to an outgroup person in an experimental setting. But if the people around you disagree with your high-prejudice view, you still tend to sit closer to this outgroup person (where both the disagreeing peers and the outgroup people are paid participants acting on the experimenter's instructions) (Sechrist and Stangor 2001).

Here, again, your biasing representation interacts with other mental states (this time, about your beliefs about what people around you think) to produce the biased behavior. But, as the propositionalist is quick to point out, associations are just not the kind of mental states that could interact with other mental states in a content-preserving manner.

III.e. Magic Bullet

Here is another experimental finding that the propositionalists could use. (A version of this argument in Mandelbaum [2016], although it makes the converse claim: if biasing representations were associations, conditioning could counteract them, but conditioning does not actually counteract bias, hence biasing representations are not associations.) It has been shown that one single presentation of an explanation can sometimes change implicit bias (Cone and Ferguson 2015).

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

If the biasing representation were an association, then this would not be possible. Associations can be counteracted only by repeated exposure (the way they were formed). One single presentation of an explanation would not make much difference to the association. In contrast, if the biasing representation were a propositional attitude, then one single explanation could change this propositional attitude. So the biasing representation must be a propositional attitude.

III.f. Logical Form

So far, I have focused on arguments against the associationist view. But there are also important considerations against propositionalism (see Madva 2016 for a summary). The most decisive of these is that implicit bias is not sensitive to logical form. For example, exposure to the sentence 'Old people drive badly' strengthens one's biasing representation about old people's bad driving. Nothing surprising so far. But even exposure to sentences that explicitly deny this, for example, the sentence 'It is not true that old people are bad drivers', also strengthens this implicit bias (Gawronski et al. 2008; cf. Deutsch et al. 2009; Deutsch and Strack 2010).

This is easy to explain if we presuppose an associationist picture: the sentence 'It is not true that old people are bad drivers' strengthens our association between old people and bad driving because the two co-occur in this sentence (even though the sentence explicitly states that being old and bad driving do not go together). The propositionalist, in contrast, is in trouble: if the biasing representation were a propositional attitude (presumably a propositional attitude about old people driving badly), then exposure to a sentence that denies the content of this propositional attitude should not strengthen the implicit bias (but see complications with the logical form experiments in Johnson et al. 2013 and the helpful discussion of the propositionalist's options in Sullivan-Bissett [2019]). But, as the experiment shows, it does.

III.g. Neither Proposition nor Association

The evidence presented in the first five subsections above show that implicit bias is sensitive to the semantic content of our representations. It is sensitive to the content of representations other than the biasing representation. The content of the biasing representation is combined with the content of other representations, which then produces the biased behavior. The evidence presented in the last subsection shows that implicit bias is not sensitive to logical form.

In light of these findings, both of the classic accounts of implicit bias are problematic. The biasing representation would need to be both sensitive to semantic content and insensitive to logical form. But associations are not sensitive to semantic content. And propositional attitudes are not insensitive to logical form. In the next section, I argue that mental imagery is much better suited to fulfill the theoretical role of the biasing representation: it is sensitive to semantic content (unlike associations) and insensitive to logical form (unlike propositions).

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

IV. Mental Imagery

By *mental imagery* I mean perceptual processing that is not directly triggered by sensory input. This is the sense in which psychologists and neuroscientists use this term (although philosophers often use it in a much narrower sense). Here is the definition from a recent review article on mental imagery in the leading cognitive science journal, *Trends in Cognitive Science*: 'We use the term 'mental imagery' to refer to representations . . . of sensory information without a direct external stimulus' (Pearson et al. 2015: 590; see also Nanay 2018 for a philosophical summary).

When you close your eyes and visualize an apple, you have perceptual processing that is not directly triggered by sensory input: we know that there is activation as early as in the primary visual cortex and this activation is clearly not directly triggered by sensory input (that is, retinal stimulation) because your eyes are closed.

Closing your eyes and visualizing an apple is a voluntary exercise of mental imagery. But mental imagery can be, and often is, triggered involuntarily. Flashbacks of unpleasant scenes would count as mental imagery, but they are not voluntary. The same goes for earworms, a paradigm example of auditory mental imagery. Both of these phenomena consist of perceptual processing that is not directly triggered by sensory input (Nanay 2015a, 2017a, 2017b, forthcoming).

Mental imagery is a perceptual process. The only difference between sensory stimulation-driven perception and mental imagery is that the former, but not the latter, is directly triggered by sensory input. This means that mental imagery represents the world the way perception does: imagistically or, as it is often put, nonpropositionally. Without going into the details of the debate about the format of perception (and of mental imagery), I assume that a representation has imagistic content if it represents magnitudes by means of magnitudes (Peacocke 2019: 52–59; see also Nanay 2015b).

Given that inference is a relation between propositional attitudes and mental imagery is not a propositional attitude (as it has imagistic content), mental imagery does not enter into inferences. But not all content-sensitive transitions between mental states are inferences and there can be content-sensitive transitions between mental states with imagistic content.

Mental imagery can lead to, and even justify, various other mental processes in a content-sensitive manner. Here are two examples. First, you are trying to wrap a box into a gift wrap. You look at the box, you look at the gift wrap and how big a piece you tear off depends (in a content-sensitive manner) on your exercise of mental imagery. Second, you are playing snooker or billiards and you need to sink a ball at the other end of the table, with many other balls in the way, but you figure out a way of making the cue ball ricochet twice before hitting the ball exactly from the right angle. In both of these two examples, there are content-sensitive transitions between mental imagery and other mental processes, but neither of them is an inference: these transitions are not mediated by beliefs or other propositional attitudes. And neither of them is an association (cf. Mandelbaum and Quilty-Dunn 2020). (I say more about content-sensitive transitions between mental imagery and other mental processes in section V, below.)

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

This definition of mental imagery is silent about a number of features of mental imagery. Mental imagery may be voluntary (when you close your eyes and visualizing an apple) or involuntary (when you have flashbacks or an earworm). It may also be conscious or unconscious. It may be emotionally charged (but it does not have to be). And it may be action-oriented (but it does not have to be). The kind of mental imagery that plays an important role in implicit bias is affectively charged and action-oriented involuntary mental imagery. Whether it is conscious or unconscious is subject to some high-profile empirical debates in the implicit bias literature, but both positions would be consistent with my view as mental imagery may be conscious or unconscious.

This will raise some philosophical eyebrows. Philosophers sometimes take mental imagery to be necessarily conscious, but there are three kinds of reasons to think that mental imagery can be conscious or unconscious: conceptual, methodological, and empirical reasons.

First, the conceptual reason. Perception can be conscious or unconscious. If the stimulus is masked or presented for a very short period of time, the subject still perceives it, but has no conscious experience of it (see, for example, Kentridge, Heywood, and Weiskrantz 1999; Kouider and Dehaene 2007). But if perception per se can be unconscious, it would be completely ad hoc to postulate that mental imagery cannot be. Remember that mental imagery is a form of perceptual processing: perceptual processing that is not directly triggered by sensory input. If perceptual processing that *is* directly triggered by sensory input can be unconscious, it is difficult to see why perceptual processing that is *not* directly triggered by sensory input (that is, mental imagery) would have to be conscious.

The second reason is methodological. Most behavioral or neuroimaging experiments on mental imagery—including the most famous ones—do not actually take the conscious experience of the subject into consideration. Take for example the famous mental rotation tasks, one of the most widely used paradigms in the study of mental imagery. There is a linear correspondence between the time required for deciding whether two three-dimensional shapes are the same and the degree of rotation between these two shapes (Shepard and Metzler 1971). Your task is to decide whether two complex three-dimensional shapes are the same. And you are quicker to respond (with a yes or no answer) if the two shapes are oriented in such a way that less mental rotation is required between them. Whatever these experiments say about mental imagery (and we can stay away from this question), it must be a claim that is silent about whether mental imagery is conscious, as these experiments are response time experiments and the reasons for inferring the exercise of mental imagery are not introspective ones, but come from the timing of the subjects' responses (for which they did not have to be conscious of any kind of mental imagery—although they obviously needed to be conscious of the task they were performing). The mental imagery involved in this task may or may not be conscious. Therefore, the concept of mental imagery that mental rotation experiments are concerned with should not (and cannot) have consciousness as a built-in feature.

The third reason is empirical: positing unconscious mental imagery can explain a number of empirical findings (especially about priming and about aphantasia) better

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

than not positing unconscious mental imagery. (See Nanay [2021] for a detailed argument; Koenig-Robert and Pearson [2020] for further support.). Aphantasia is the name of the condition in which subjects have no conscious mental imagery. They cannot conjure up the image of an apple and some of them do not have involuntary mental imagery either. Aphantasia is not a monolithic phenomenon—some aphantasia subjects have problems with voluntary control of mental imagery but they do have involuntary mental imagery—flashbacks and dreams, for example. Some others have the exact same behavioral patterns in mental imagery-involving tasks (like mental rotation) as normal subjects and their perceptual processing is also indistinguishable from that of normal subjects, but they have no accompanying experience. The most empirically plausible explanation of the condition of these aphantasia subjects (again, this only applies to a subset of people diagnosed with aphantasia) is that they do have mental imagery (regardless of whether we pick out this mental phenomenon behaviorally or neurally), but this mental imagery is unconscious.

This distinction between conscious and unconscious mental imagery is important in this context because of the ongoing empirical debate about whether implicit bias is really implicit and if so to what extent. It has been argued that subjects are in fact aware of their biasing representations, at least in some contexts. What is important from our point of view is that both sides of this debate could accept that the biasing representation is mental imagery as mental imagery may be conscious or unconscious.

Mental imagery may also be emotionally charged. If you visualize your long-deceased grandmother, this mental imagery can be emotionally charged, for example. But not all mental imagery is emotionally charged. When you close your eyes and visualize an apple, your mental imagery is unlikely to be emotionally charged (unless you really like—or hate—apples). Finally, mental imagery can be action guiding. When you get up in the middle of the night in your pitch-dark bedroom and switch on the light, the mental state that guides your action is your mental imagery, for example. But, as before, not all mental imagery is action-guiding. The mental imagery of the apple you visualized presumably does not guide any of your actions. The diversity of these varieties of mental imagery can also explain the diversity of implicit biases (emotional versus semantic, for example; see, for example, Holroyd and Sweetman 2016).

V. Mental Imagery as the Biasing Representation

My claim is that the biasing representation is mental imagery. This mental imagery is often emotionally charged, often action guiding and often unconscious. It is always involuntary.

Taking the biasing representation to be mental imagery can explain many examples of biased behavior. Take the example I started the article with, when subjects are reluctant to drink lemonade from a sterilized bedpan. The perceptual state of seeing the yellow liquid in the bedpan-shaped drinking vessel triggers the mental imagery of urine and it is this biasing representation that explains our reluctance. The imagery itself does not have to be conscious (although it might

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

be). It is emotionally charged (presumably the emotion is disgust). And it is action guiding in the sense that it interferes with our action.

Here is another example (from Gendler 2008). The chef reorganizes her kitchen. The cleaver used to be above the dishwasher, but it is now next to the stove. She knows this—she placed the cleaver from here to there herself. But in the rush of preparing a meal, she still reaches to where the cleaver used to be—above the dishwasher. This is biased behavior: it goes against her beliefs. And we can explain this in terms of the mental imagery that she has of the cleaver above the dishwasher. Again, this mental imagery does not have to be conscious. But, in this example, it is very much action-guiding mental imagery: it guides the chef's action the same way as the mental imagery of the light switch guides your action in your pitch dark bedroom.

Could we explain the chef's behavior in terms of an unconscious belief? No. We know from a vast number of studies in the neuroscience of 'attentional templates' that visual actions of the kind performed by the chef amount to early cortical processing in the visual system that is not directly triggered by sensory input—in short, to mental imagery (Keogh and Pearson 2021). The belief view does not explain this. The mental imagery view does.

And the same explanatory scheme also applies to biased behavior concerning other racial and gender groups. Subjects who are not Black are more likely to misperceive a phone as a gun if a Black person holds it than if a white person does so (Payne 2001; see also Siegel 2020). Here, again, the perceptual state of a Black person holding a phone gives rise to the mental imagery of a Black person holding a gun. Again, this mental imagery does not have to be conscious—and when white people rate Black people as more dangerous, it is possible that the mental imagery that grounds these judgments is not conscious. The same is true of the biased behavior of standing farther away from some people than others in the elevator.

It is easy to see that mental imagery fits the profile of biasing representation we identified in the last section. We have seen that the experimental results about implicit bias show that the biasing representation is sensitive to semantic content but insensitive to logical form. We have also seen that neither associations nor propositions can satisfy both of these requirements. But mental imagery does. It is very much sensitive to semantic content and especially top-down influences. If I visualize a cat, the features of the cat I visualize very much depend on what kinds of cats I have seen in my life and also on my background beliefs and knowledge about cats. The chef's mental imagery of the cleaver depends on previously stored memories. And we have plenty of empirical evidence about how mental imagery can be influenced in a top-down manner.

Mental imagery can also be combined with other mental representations. Suppose that I put an apple in front of you on the desk. You are looking at this apple, but then I ask you to close your eyes and grab this apple. You can do this reliably. We know from a vast amount of neuroscientific evidence on attentional templates that you use (potentially unconscious, but normally very much conscious) mental imagery to do so, see the discussion three of paragraphs earlier. Now I repeat the procedure, but when your eyes are closed, I move the apple to your left by fifty centimeters and I

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

tell you I did so. You can still pick up the apple with your eyes closed and do so reliably and, again, by means of utilizing your mental imagery. But the mental state that guides your action is the result of the combination of the content of your mental imagery of the apple (where it was when you closed your eyes) and the verbal information I gave you about how I moved it. I suggest that the content of the biasing representation combines with the content of other mental states in producing implicit bias in the same way as the content of this mental imagery of the apple on the desk is combined with the information I give you.

It is important to note that this dependence of mental imagery on top-down information does not mean that the content of mental imagery is conceptual content or that a concept is attached to the mental imagery. Mental imagery is perceptual processing, which may or may not be influenced in a top-down manner. But even if it is, this does not entail that its content is conceptual. To take the example of the bedpan, the mental imagery of the yellow liquid that is triggered by the perception of the bedpan does not have to have concepts like 'urine' attached to it. It is enough if the valence of the mental imagery is influenced in a top-down manner (by its association with urine). Mental imagery is influenced by our concepts—it does not have to be conceptual.

Finally, a somewhat odd feature of the implicit bias literature in philosophy is that it often focuses on the relation between concepts: the concept of being a homemaker and the concept of being a woman, for example. But this focus is in tension with the vast majority of empirical findings about implicit bias, where the trigger of biased behavior is a perceptual state. And the biasing representation is the representation that mediates between the perceptual trigger and the biased behavior. My account has a wider range of explanations of the perceptual (or quasi-perceptual, if you wish) nature of this mediating biasing representation than the classic views. The classic views would be committed to saying that the perceptual state triggers a concept and that concept is either associated with or is propositionally related to another concept and this other concept is responsible for the behavior. We have both empirical and conceptual reasons to think that perception is not always linked to actions by means of concepts (see Nanay 2013 for a summary). If so, then my account can explain the mediation between perception and action without any direct appeal to concepts.

One might wonder whether the mental imagery account is a genuine alternative to the associationist and the propositionalist account. After all, the trigger somehow leads to mental imagery and the mental imagery somehow leads to action. So one could argue that even if we accept the mental imagery account, the nature of these transitions is still an open question and they can be explained either in an associationist or in a propositionalist manner. This is a fair point, but it should also be pointed out that the question about how a perceptual state gives rise to mental imagery and then how mental imagery turns into behavior is a very different question from the one about what kind of connection between concepts leads to biased behavior. So even if the associationism versus propositionalism debate is not put to bed entirely, we have made some progress.

Take microbehavior, for example. One striking phenomenon often discussed under the heading of implicit bias is how our biasing representation influences our

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

microbehavior (Bessenoff and Sherman 2000; Chen and Bargh, 1997; McConnell and Leibold, 2001; Wilson, Lindsay, and Schooler 2000; see also Levy 2015 for a summary). Not what answer we choose in a questionnaire, but, for example, how much we look in the eyes of people with different skin color, how far away we stand from them in an elevator and so on. And here the mental imagery account has a real advantage. If the biasing representation is (action-guiding) mental imagery, then we have a direct and straightforward way of explaining how little differences in one's (action-guiding) mental imagery, which are not sensitive to one's explicitly held beliefs) are responsible for little differences in one's behavior.

VI. Best of Both Worlds

My aim in this section is to go through the most important empirical arguments against the association view and against the propositional view (discussed above in section III), and to show that none of them apply to the mental imagery view. More generally, the mental imagery view combines the strengths of these two views inasmuch as mental imagery, like associations, but unlike propositions, is insensitive to logical form. And it is also, like propositions, but unlike associations, sensitive to semantic content.

The first argument was about sugar. The dangerousness signaled by the label that said 'Sodium cyanide' was inherited first by the bottle that this sign was affixed to (by the subjects themselves) and then to the liquid in a cup that the sugar was poured into from this bottle. This might rule out pure associationism given that the association is supposed to stand between the word 'Sodium cyanide' and some kind of negative valence. The associationist would need to explain why this negative valence is inherited by the bottle and then by the liquid and it is not clear that dumb, content-insensitive, automatic associations can do this job. But if we take the biasing representation to be mental imagery, then no parallel problem arises. If you imagine the poisoned apple the Evil Stepmother prepares for Snow White to be juiced and then made ice cubes from and you imagine these ice cubes put in Snow White's drink, the original negative valence of the mental imagery of the apple is inherited by the mental imagery of the apple juice and of the ice cubes. In the light of what we know about mental imagery, we should expect subjects' reaction toward the two cups of sweetened liquids at the end of the experiment to be different as they are colored by differently valenced mental imagery.

The second argument was about celebrity contagion. This argument relied on the implicit premise that associations cannot be combined with other contentful representations. In the experiment, the biasing representation is combined with another representation about the effects of washing on clothing. And while this argument may work against associationism, it does not work against the view that the biasing representation is mental imagery. Mental imagery, as we have seen in the last section, can be combined with other contentful mental representations. Just as the mental imagery of the apple in front of me can be combined with the information I got from you about how you moved the apple, the (positively valenced) mental imagery I have when I look at the tuxedo worn by Sean Connery can also be combined with the contentful representation about the effects of dry

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

cleaning on tuxedos. I am looking at the tuxedo, this is the triggering state, I form a positively valenced mental imagery. This is the biasing representation, and that is what makes me offer more money for it. This mental imagery may, again, not be conscious. And this mental imagery combines with my background knowledge about the effects of dry cleaning, and this combination of these two mental states explains my biased behavior of offering less money for the dry-cleaned tuxedo.

In the third argument, the subject was made to dislike Person A and then she was told that Person A dislikes Person B. This disposes the subject positively toward Person B. Thinking about Person A involves the triggering of negatively valenced mental imagery. When we then think of Person B and we are told that Person A dislikes Person B, the mental imagery that we form of Person B is influenced by the negatively valenced mental imagery we formed of Person A in a content-sensitive manner. This process of forming mental imagery of Person B is sensitive to the information about Person A's dislike as well as the negatively valenced mental imagery of Person A. Hence, we should expect positively valenced mental imagery of Person B. Here is an analogy. Imagine your least favorite politician playing golf. That is probably a negatively valenced mental imagery. Now image this politician falling over mid-strike and landing with their face in the mud. This is a negatively valenced extra information, but the end-result is very much positively valenced. In short, we do not need propositions and inferences in order to get positive from two negatives. Content-sensitive transitions between mental imagery can do the trick just as well. If this example was not convincing enough, there is plenty of empirical evidence that the brain can connect up non-propositional magnitude representations in such content-sensitive manner (See, for example, Komer et al. [2019] for an account of such content-sensitive transitions between continuous spatial representations; see Peacocke [2019: 52–59] for a philosophical overview of how logical operations over magnitude representations are implemented in the brain.)

The fourth argument was about peer pressure. Implicit bias depends on what we know about the people around us. We behave in a more biased manner if we believe that the people around us are biased. This, again, should not come as a surprise if the biasing representation is mental imagery as our mental imagery is very much influenced by our other mental states, including our beliefs about what people around us believe or expect us to do. This experiment shows that biased behavior depends on our beliefs. But given that mental imagery also depends on our beliefs, this should not come as a surprise (there is also evidence from neuroscience about the involvement of early perceptual processes in handling peer disagreement; see Berns et al. 2005).

Finally, the argument according to which one single presentation of an explanation can change implicit bias could be seen as powerful against associationism, but it has no traction against the mental imagery view. If the biasing representation is mental imagery, then there is no reason to think that one single explanation could not change this mental imagery (for example, by showing, in some visual form, that what we imagined is not possible or not probable).

The pattern is clear. All these arguments rely on the assumption that associations cannot be combined with other mental states, whereas propositions can be. As implicit bias clearly depends on our background beliefs and other mental states,

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

this rules out associationism. But it does not rule out the view according to which the biasing representation is mental imagery as mental imagery can also be combined with beliefs.

How about the anti-propositionalist argument? Sentences like "It is not true that old people are bad drivers" strengthens biasing representations about old people's bad driving (Gawronski et al. 2008; see also Deutsch and Strack 2010; Madva 2016). But if propositionalism were true, then sentences like this would weaken the biasing representation. Note that if the biasing representation is mental imagery, then sentences of this kind would evoke the mental imagery of old people driving badly. The negation operation does not influence the evoked mental imagery (for arguments along similar lines, see Currie and Ichino 2012; Sullivan-Bissett 2019).

Again, there are systematic reasons for this. Propositions are sensitive to logical form. Associations are not. And mental imagery is not either. Given that implicit bias is insensitive to logical form, we need biasing representations that are also insensitive. And mental imagery is a good candidate for this as well.

I gave a classic best-of-both worlds explanation: if we consider biasing representations to be mental imagery, then we can preserve the advantages of the propositionalist view without its disadvantages and we can also preserve the advantages of the associationalist view without its disadvantages. Because there are structurally similar accounts in the implicit bias literature, I need to contrast my account with them. To do so, I focus on two of these: Neil Levy's (2015) patchy endorsement account and Ema Sullivan-Bissett's (2019) unconscious imagination account.

The general moral is that both Levy's concept of patchy endorsement and Sullivan-Bissett's concept of unconscious imagination are far in the propositionalist side of the divide. Levy explicitly claims that patchy endorsements 'have some propositional structure' and that they 'feature in some inferences' (2015: 816). Mental imagery does not have propositional structure (not even some!) and it does not feature in inferences. And given Mandelbaum's insistence on the fragmentation of belief (Bendana and Mandelbaum forthcoming), Levy's view may not be too far removed from Mandelbaum's own propositionalism.

Similarly, Sullivan-Bissett's (2019) view is more permissive toward the propositionalist than is mine: she would allow at least some of the unconscious imaginative episodes to be propositional attitudes. Imagination is not the same as mental imagery—imagination is a very specific exercise of mental imagery. And while there are good empirical and theoretical reasons to think that mental imagery can be, and often is, unconscious, it is much more controversial to say that imagination can be unconscious (for a classic argument against this view, see Kind 2001).

Nevertheless, Levy's view could be used as a structural parallel to the mental imagery account of implicit bias that could be helpful to highlight how content-sensitive transitions work in the case of mental imagery. For Levy, patchy endorsements enter into fewer inferential relations than beliefs. That is what makes them patchy. But they still enter into some inferential relations. That is what makes them different from associations. The mental imagery account gives a

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

structurally similar picture, where the non-inferential content-sensitive transitions between mental imagery and other mental processes are more sparse than (in principle infinite) inferential relations but more ubiquitous than (the very few and automatic) associations.

VII. Argument from Counteracting Implicit Bias

The final argument for the mental imagery view comes from the success of counteracting implicit bias with the help of the manipulation of the subjects' mental imagery. Among the most efficient ways of manipulating implicit bias is with the help of manipulating mental imagery. For example, visualizing or putting ourselves imaginatively in the shoes of a member of another racial or gender group can reduce implicit bias significantly. And, crucially, the extent of this reduction correlates with the details and vividness of the imagery involved (Lai et al. 2014; Blair, Ma, and Lenton 2001; Blair 2002; see also Peck et al. 2013 for further relevant findings).

It is important that this procedure rely fully on the manipulation of the conscious and voluntary mental imagery of the subjects: they ask the patients to visualize certain faces or imagine themselves to be certain people consciously and voluntarily. But as we have seen, this is just one way of triggering mental imagery and one that is dependent on a lot of factors that might prevent the subject from succeeding in visualizing what she is asked to visualize. It is also difficult to control whether the subject does in fact visualize the outgroup face she is asked to visualize. Finally, such visual imagery is difficult to maintain for longer than a couple of seconds. Nonetheless, in spite of all these practical problems, the imagery-involving procedure is among the most efficient ways of reducing implicit bias.

But why would evoking mental imagery (in this unreliable way that is difficult to control and maintain) be an efficient way of counteracting implicit bias? It is difficult for the propositionalist to explain this: if the biasing representation is a propositional attitude, then mental imagery—a perceptual process—should have no direct impact on it. And it should not have any more impact on it than other perceptual processes (like actual stimulation-driven perception). But it does.

Similarly, according to associationism, the best way to unlearn an association is extinction and extinction is achieved by repeated exposure to perceptual stimuli that goes against the association (if the association is between A and B, then the extinction would involve exposure to A and non-B or non-A and B). But the research I cite shows that mental imagery is a more efficient way of counteracting implicit bias than extinction and it is unclear how the associationist could explain this.

The mental imagery view is obviously well suited to explain this: if the biasing representation is mental imagery, then it should not come as a surprise that the bias can be recovered by manipulating the subject's mental imagery.

The procedure to counteract implicit bias is in some sense very similar to the procedures used in psychiatry to treat involuntary mental imagery of depression or posttraumatic stress disorder patients. Here, too, the most efficient way of treating these symptoms is by manipulating the patient's mental imagery (by, say, imagery

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

rescripting, a method of changing various features, including the valence, of mental imagery episodes) (see Holmes et al. 2010; James et al. 2015).

VIII. Conclusion

There are many unresolved issues even if we accept the mental imagery view of implicit bias. Assuming that the biasing representation is mental imagery, there are still unanswered questions about how this mental imagery is triggered by the perceptual stimulation. A lot of recent empirical research has focused on top-down influences on perceptual processing and, given that mental imagery is a perceptual process, this would be an example of such top-down influence, but then the question is about where these top-down influences come from.

And there are further questions about how this mental imagery triggers the biased behavior. There have been advances in this questions as well, as we now understand the role of mental imagery in action initiation and action guidance better (see Jeannerod 1997 for pioneering work on this), but there are still unresolved issues, for example, about how sensory stimulation-driven perceptual information fails to overwrite the biasing mental imagery in certain high-stake situations (for example, in the case of quick actions by police officers toward people from certain racial groups).

Finally, if it is true that the biasing representation is mental imagery, then we can make the procedures of counteracting implicit bias even more efficient. As this biasing representation is involuntarily triggered, given that we know that involuntary mental imagery can be modified more quickly with involuntary mental imagery (Murphy et al. 2015; Slofstra et al. 2016; Clark et al. 2016), the mental imagery-involving procedures for counteracting implicit bias would work even more efficiently if they used involuntarily triggered mental imagery (and not just asking the subject to visualize an outgroup face, which she might or might not do and who knows for how long).

BENCE NANAY
UNIVERSITY OF ANTWERP
bence.nanay@uantwerpen.be

References

Bendana, Joseph, and Eric Mandelbaum. (forthcoming) 'The Fragmentation of Belief'. In Cristina Borgoni, Dirk Kindermann, and Andrea Onofri (eds.), *The Fragmented Mind* (Oxford University Press).

Berger, Jacob. (2020) 'Implicit Attitudes and Awareness'. *Synthese*, 197, 1291–1320.

Berns, Gregory S., Jonathan Chappelow, Caroline F. Zink, Giuseppe Pagnoni, Megan E. Martin-Skurski, and Jim Richards. (2005) 'Neurobiological Correlates of Social Conformity and Independence during Mental Rotation'. *Biological Psychiatry,* 58, 245–53.

Bessenoff, Gayle R., and Jeffrey W. Sherman. (2000) 'Automatic and Controlled Components of Prejudice toward Fat People: Evaluation versus Stereotype Activation'. *Social Cognition*, 18, 329–53.

Blair, Irene V. (2002) 'The Malleability of Automatic Stereotypes and Prejudice'. *Personality and Social Psychology Review,* 6, 242–61.

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

Blair, Irene V., Jennifer E. Ma, and Alison P. Lenton. (2001) 'Imagining Stereotypes Away: The Moderation of Implicit Stereotypes through Mental Imagery'. *Journal of Personality and Social Psychology*, 81, 828–41.

Brownstein, M. (2016) 'Attributionism and moral responsibility for implicit bias'. *Review of Philosophy and Psychology*, 7, 765–86.

Chen, Mark, and John A. Bargh. (1997) 'Nonconscious Behavioral Confirmation Processes: The Self-Fulfilling Consequences of Automatic Stereotype Activation'. *Journal of Experimental Social Psychology*, 33, 541–60.

Clark, I. A., Emily A. Holmes, M. W. Woolrich, and C. E. Mackay. (2016) 'Intrusive Memories to Traumatic footage: The Neural Basis of Their Encoding and Involuntary Recall'. *Psychological Medicine,* 46, 505–18.

Cone, Jeremy, and Melissa J. Ferguson. (2015) 'He Did *What*? The Role of Diagnosticity in Revising Implicit Evaluations'. *Journal of Personality and Social Psychology,* 108, 37–57.

Currie, Greg, and Anna Ichino. (2012) 'Aliefs Don't Exist, though Some of Their Relatives Do'. *Analysis*, 72, 788–98.

Del Pinal, G., and S. Spaulding. (2018) 'Conceptual centrality and implicit bias'. *Mind & Language*, 33, 95–111.

Deutsch, Roland, Robert Kordts-Freudinger, Bertram Gawronski, and Fritz Strack. (2009) 'Fast and Fragile: A New Look at the Automaticity of Negation Processing'. *Experimental Psychology,* 56, 434–46.

Deutsch, Roland, and Fritz Strack. (2010) 'Building Blocks of Social Behavior: Reflective and Impulsive Processes'. In Bertram Gawronski and B. Keith Payne (eds.), *Handbook of Implicit Social Cognition: Measurement, Theory, and Applications* (New York: Guilford Press), 62–79.

De Houwer, Jan. (2009) 'The Propositional Approach to Associative Learning as an Alternative for Association Formation Models'. *Learning and Behavior,* 37, 1–20.

De Houwer, Jan. (2011) 'Evaluative Conditioning: A Review of Functional Knowledge and Mental Process Theories'. In Todd Schachtman and Steve Reilly (eds.), *Associative Learning and Conditioning Theory: Human and Non-human Applications* (New York: Oxford University Press), pp. 399–416.

De Houwer, Jan. (2019) 'Implicit Bias Is Behavior: A Functional-Cognitive Perspective on Implicit Bias'. *Perspectives on Psychological Science,* 14, 835–40.

De Houwer, Jan, Sarah Thomas, and Frank Baeyens. (2001) 'Association Learning of Likes and Dislikes: A Review of 25 Years of Research on Human Evaluative Conditioning'. *Psychological Bulletin*, 127, 853–69.

Dunham, Yarrow, Andrew S. Baron, and Mahzarin R. Banaji. (2008) 'The Development of Implicit Intergroup Cognition'. *Trends in Cognitive Sciences,* 12, 248–53.

Forscher, Patrick S., Calvin K. Lai, Jordan R. Axt, Charles R. Ebersole, Michelle Herman, Patricia G. Devine, and Brian A. Nosek. (2019) 'A Meta-analysis of Procedures to Change Implicit Measures'. *Journal of Personality and Social Psychology,* 117, 522–59.

Gawronski, Bertram, Eva Walther, and Hartmut Blank. (2005) 'Cognitive Consistency and the Formation of Interpersonal Attitudes: Cognitive Balance Affects the Encoding of Social Information'. *Journal of Experimental Social Psychology,* 41, 618–26.

Gawronski, Bertram, and Galen V. Bodenhausen. (2006) 'Associative and Propositional Processes in Evaluation: An Integrative Review of Implicit and Explicit Attitude Change'. *Psychological Bulletin,* 132, 692–731.

Gawronski, Bertram, and Etienne P. LeBel. (2008) 'Understanding Patterns of Attitude Change: When Implicit Measures Show Change, but Explicit Measures Do Not'. *Journal of Experimental Social Psychology*, 44, 1355–61.

Gendler, Tamar Szabó. (2008) 'Alief in Action (and Reaction)'. *Mind & Language,* 23, 552–85.

Greenwald, Anthony G., Debbie E. McGhee, and Jordan L. K. Schwartz. (1998) 'Measuring Individual Differences in Implicit Cognition: The Implicit Association Test'. *Journal of Personality and Social Psychology,* 74, 1464–80.

Greenwald, Anthony G., and Brian A. Nosek. (2009) 'Attitudinal Dissociation: What Does It Mean?' In Richard E. Petty, Russell H. Fazio, and Pablo Brinol (eds.), *Attitudes: Insights from the New Implicit Measures* (New York: Psychology Press), 65–82.

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

Hahn, Adam, Charles M. Judd, Holen K. Hirsh, and Irene V. Blair. (2014) 'Awareness of Implicit Attitudes'. *Journal of Experimental Psychology: General,* 143, 1369–92.

Holmes, Emily A., Ella L. James, Emma J. Kilford, and Catherine Deeprose. (2010) 'Key Steps in Developing a Cognitive Vaccine against Traumatic Flashbacks: Visuospatial Tetris versus Verbal Pub Quiz'. *PLoS ONE,* 5, e13706.

Holroyd, Jules. (2016) What Do We Want from a Model of Implicit Cognition?' *Proceedings of the Aristotelian Society*, 116, 153–79.

Holroyd, Jules, and Joseph Sweetman. (2016) 'The Heterogeneity of Implicit Bias'. In Michael Brownstein and Jennifer Saul (eds.), *Metaphysics and Epistemology,* Vol. 1 of *Implicit Bias and Philosophy* (Oxford: Oxford University Press), 80–113.

James, Ella L., Michael B. Bonsall, Laura Hoppitt, Elizabeth M. Tunbridge, John R. Geddes, Amy L. Milton, and Emily A. Holmes. (2015) 'Computer Game Play Reduces Intrusive Memories of Experimental Trauma via Reconsolidation Update Mechanisms'. *Psychological Science,* 26, 1201–15.

Jeannerod, Marc. (1997) *The Cognitive Neuroscience of Action*. Oxford: Blackwell.

Johnson, Gabbrielle M. (2019) 'The Structure of Bias'. *Mind,* 516, 1193–1236.

Johnson, Dan R., Daniel M. Jasper, Sallie Griffin, and Brandie L. Huffman. (2013) 'Reading Narrative Fiction Reduces Arab-Muslim Prejudice and Offers a Safe Haven from Intergroup Anxiety'. *Social Cognition*, 31, 578–98.

Kentridge, R. W., C. A. Heywood, and L. Weiskrantz, L. (1999) 'Attention without Awareness in Blindsight'. *Proceedings of the Royal. Society London. B*, 266: 1805–11

Keogh, Rebecca, and Joel Pearson. (2021) 'Attention Driven Phantom Vision: Measuring the Sensory Strength of Attentional Templates'. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 376. https://doi.org/10.1098/rstb.2019.0688.

Kind, Amy. (2001) 'Putting the Image Back in Imagination'. *Philosophy and Phenomenological Research,* 62, 85–109.

Komer, Brent, Terrence C. Stewart, Aaron R Voelker, and Chris Eliasmith. (2019) 'A Neural Representation of Continuous Space Using Fractional Binding'. In Ashok Goel, Colleen Seifert, and Christian Freksa (eds.), *41st Annual Meeting of the Cognitive Science Society (CogSci 2019): Creativity + Cognition + Computation*, vol. 3 (New York: Curran), 2038–43.

Koenig-Robert, Roger, and Joel Pearson. (2020) 'Decoding Non-conscious Thought Representations during Successful thought Suppression'. *Journal of Cognitive Neuroscience,* 32, 2272–84.

Kouider, Sid, and Stanislaus Dehaene. (2007) 'Levels of Processing during Non-conscious Perception: A Critical Review of Visual Masking'. *Philosophical Transactions of the Royal Society B,* 362, 857–75.

Lai, Calvin K., Maddalena Marini, Steven A. Lehr, Carlo Cerruti, Jiyun-Elizabeth L. Shin, Jennifer A. Joy-Gaba, Arnold K. Ho *et al.* (2014) 'Reducing Implicit Racial Preferences: I. A Comparative Investigation of 17 Interventions'. *Journal of Experimental Psychology: General,* 143, 1765–85.

Levy, Neil. (2015) 'Neither Fish nor Fowl: Implicit Attitudes as Patchy Endorsements'. *Noûs*, 49, 800–23.

Machery, Edouard. (2016) 'De-Freuding Implicit Attitudes'. In Michael Brownstein and Jennifer Saul (eds.), *Metaphysics and Epistemology*, Vol 1 of *Implicit Bias and Philosophy* (Oxford: Oxford University Press), 104–29.

Madva, Alex. (2016) 'Why Implicit Attitudes Are (Probably) Not Beliefs'. *Synthese,* 193, 2659–84.

Mandelbaum, Eric. (2013) 'Against Alief'. *Philosophical Studies,* 165, 197–211.

Mandelbaum, Eric. (2016) 'Attitude, Inference, Association'. *Nous* 50, 629–58.

Mandelbaum, Eric, and Jake Quilty-Dunn. (2020) 'Non-inferential Transitions: Imagery and Association'. In Anders Nes and Timothy Chan (eds.), *Inference and Consciousness* (London: Routledge), 151–71.

McConnell, Allen R., and Jill M. Leibold. (2001) 'Relations among the Implicit Association Test, Discriminatory Behavior, and Explicit Measures of Racial Attitudes'. *Journal of Experimental Social Psychology,* 37, 435–42.

Murphy, Susannah E., M. Clare O'Donoghue, Erin H. S. Drazich, Simon E. Blackwell, Anna Cristina Nobre, and Emily A. Holmes. (2015) 'Imagining a Brighter Future: The Effects of

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press

Positive Imagery Training on Mood, Prospective Mental Imagery and Emotional Bias in Older Adults'. *Psychiatry Research,* 230, 36–43.
Nanay, Bence. (2013) *Between Perception and Action*. Oxford: Oxford University Press.
Nanay, Bence. (2015a) 'Perceptual Content and the Content of Mental Imagery'. *Philosophical Studies,* 172, 1723–36.
Nanay, Bence. (2015b) 'Perceptual Content'. In Mohan Matthen (ed.), *Oxford Handbook of the Philosophy of Perception* (Oxford: Oxford University Press), 153–67.
Nanay, Bence. (2017a) 'Pain and Mental Imagery'. *Monist*, 100, 485–500.
Nanay, Bence. (2017b) 'Sensory Substitution and Multimodal Mental Imagery'. *Perception,* 46, 1014–26.
Nanay, Bence. (2018) 'Multimodal Mental Imagery'. *Cortex,* 105, 125–34.
Nanay, Bence. (2021) 'Unconscious Mental Imagery. *Philosophical Transactions of the Royal Society B*, 376. https://doi.org/10.1098/rstb.2019.0689.
Nanay, Bence. (forthcoming) *Mental Imagery*. Oxford: Oxford University Press.
Newman, George E., Gil Diesendruck, and Paul Bloom. (2011) 'Celebrity Contagion and the Value of Objects'. *Journal of Consumer Research*, 38, 215–28.
Nier, Jason A. (2001) 'How Dissociated Are Implicit and Explicit Racial Attitudes? A Bogus Pipeline Approach'. *Group Process & Intergroup Relations,* 8, 39–52.
Payne, B. Keith. (2001) 'Prejudice and Perception: The Role of Automatic and Controlled Processes in Misperceiving a Weapon'. *Journal of Personality and Social Psychology,* 81, 181–92.
Peacocke, Christopher. (2019) *The Primacy of Metaphysics*. Oxford: Oxford University Press.
Pearson, Joel, Thomas Naselaris, Emily A. Holmes, and Stephen M. Kosslyn. (2015) 'Mental Imagery: Functional Mechanisms and Clinical Applications'. *Trends in Cognitive Sciences*, 19, 590–602.
Peck, Tabitha C., Sofia Seinfeld, Salvatore M. Aglioti, and Mel Slater. (2013) 'Putting Yourself in the Skin of a Black Avatar Reduces Implicit Racial Bias. *Consciousness and Cognition*, 22, 779–87.
Ranganath, Kate A., Colin Tucker Smith, and Brian A. Nosek. (2008) 'Distinguishing Automatic and Controlled Components of Attitudes from Direct and Indirect Measurement Methods'. *Journal of Experimental Social Psychology,* 44, 386–96.
Rozin, Paul, Maureen Markwith, and Bonnie Ross. (1990) 'The Sympathetic Magical Law of Similarity, Nominal Realism, and Neglect of Negatives in Response to Negative Labels'. *Psychological Science*, 1, 383–84.
Sechrist, Gretchen B., and Charles Stangor. (2001) 'Perceived Consensus Influences Intergroup Behavior and Stereotype Accessibility'. *Journal of Personality and Social Psychology,* 80, 645–54.
Shepard, Roger N., and Jacqueline Metzler. (1971) 'Mental Rotation of Three-Dimensional Objects'. *Science,* 171, 701–3.
Siegel, Susanna. (2020) 'Bias and Perception'. In Erin Beeghly and Alex Madva (eds.), *An Introduction to Implicit Bias: Knowledge, Justice, and the Social Mind* (London: Routledge), 99–115.
Slofstra, Christien, Maaike H. Nauta, Emily A. Holmes, and Claudi L. H. Bockting. (2016) 'Imagery Rescripting: The Impact of Conceptual and Perceptual Changes on Aversive Autobiographical Memories'. *PloS One* 11, e0160235.
Sullivan-Bissett, Ema. (2019) 'Biased by Our Imaginings'. *Mind & Language,* 34, 627–47.
Toribio, Josefa. (2018) 'Implicit Bias: From Social Structure to Representational Format'. *Theoria*, 33, 41–60.
Wilson, Timothy D., Samuel Lindsay, and Tonya Y. Schooler. (2000) 'A Model of Dual Attitudes'. *Psychological Review,* 107, 101–26.

https://doi.org/10.1017/apa.2020.29 Published online by Cambridge University Press