



Deposited via The University of York.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/196761/>

Version: Published Version

---

**Article:**

Fischer, Eugen, Allen, Keith Malcolm and Engelhardt, Paul (2023) Fragmented and conflicted: folk beliefs about vision. *Synthese*. 84. ISSN: 0023-7857

<https://doi.org/10.1007/s11229-023-04066-w>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



# Fragmented and conflicted: folk beliefs about vision

Eugen Fischer<sup>1</sup> · Keith Allen<sup>2</sup> · Paul E. Engelhardt<sup>3</sup>

Received: 6 May 2022 / Accepted: 25 January 2023  
© The Author(s) 2023

## Abstract

Many philosophical debates take for granted that there is such a thing as ‘the’ common-sense conception of the phenomenon of interest. Debates about the nature of perception tend to take for granted that there is a single, coherent common-sense conception of vision, consistent with Direct Realism. This conception is often accorded an epistemic default status. We draw on philosophical and psychological literature on naïve theories and belief fragmentation to motivate the hypothesis that untutored common sense encompasses conflicting Direct Realist and Indirect Realist conceptions: there is no such thing as ‘the’ common-sense conception of vision that could enjoy epistemic default status. To examine this hypothesis, a survey administered an agreement rating task with verbal and pictorial stimuli to lay participants. We found many laypeople simultaneously hold conflicting Direct Realist and Indirect Realist beliefs about vision. Against common assumptions, Direct Realist beliefs are not clearly dominant, and consistent adherence to Direct Realism is not the norm, but the exception. These findings are consistent with recent accounts of belief fragmentation. They forcefully challenge common methodology in philosophical debates about the nature of perception and beyond.

**Keywords** Experimental philosophy · Naïve theories of vision · Direct vs Indirect Realism · Problem of perception · Belief fragmentation · Philosophical method

## 1 Introduction

Philosophical debates frequently take for granted that laypeople share unified and coherent sets of pre-scientific ‘common-sense’ beliefs about phenomena of interest

---

✉ Eugen Fischer  
e.fischer@uea.ac.uk

<sup>1</sup> School of Politics, Philosophy, Language and Communication Studies, University of East Anglia, Norwich, UK

<sup>2</sup> Department of Philosophy, University of York, York, UK

<sup>3</sup> School of Psychology, University of East Anglia, Norwich, UK

(e.g., Daly, 2010; Jackson, 1998). This assumption is common, for example, in debates about free will (e.g., O'Connor & Franklin, 2021), consciousness (e.g., Chalmers, 1996), folk psychology (e.g. Fodor, 1987), material objects (e.g., Scholl, 2007), time (e.g., Callender, 2017), colour (e.g., Allen, 2016), and the nature of visual experience (e.g., Martin, 2002). These common-sense conceptions are philosophically significant because they are often thought to enjoy an *epistemic default status*: we should accept these conceptions, absent good reasons to the contrary; or else, philosophical positions at odds with them should provide an error theory explaining how common sense could go amiss.

However, recent psychological and philosophical work on belief (review: Porot & Mandelbaum, 2021) and naïve theories (review: Shtulman, 2017) suggests we cannot simply take for granted the existence of unified and coherent common-sense conceptions of familiar phenomena like vision: Folk beliefs are often 'fragmented' and conflicted. Different cognitive processes, operating under different conditions, generate conflicting beliefs, which are never systematically screened for coherence; they are stored at different locations in long-term memory, in different 'belief fragments', which are internally coherent, but may conflict with one another (Bendaña & Mandelbaum, 2021; Leiser, 2001). As a result, different lay conceptions of the same phenomenon may be held not only by members of different communities (as Berniūnas et al., 2021, suggest for free will) and different members of the same community (as Latham et al., 2021, and Lee et al., 2022, suggest for time), but even by the same individual (as Adams & Hansen, 2020, argue in the case of colour).

This paper investigates common-sense views of vision in light of this 'fragmentation challenge' and provides empirical evidence of inter- and intrapersonal conflict between pre-scientific beliefs about vision. Evidential experimental philosophy standardly employs surveys and experiments to examine the evidential value of philosophically relevant intuitions about hypothetical cases (reviews: Machery, 2017; Stich & Tobia, 2016). In a natural extension of this research, the present paper examines whether folk *beliefs* deserve the epistemic default status they are often accorded: The paper deploys findings about belief fragmentation to question this common methodological practice, and empirically investigates fragmentation among beliefs central to current philosophical debates. A survey examines folk theories (Direct Realist and Indirect Realist conceptions of vision) that are directly and patently in conflict with each other. The surprising finding that both theories are widely held, often by the same individuals, provides further evidence of belief fragmentation and suggests that there is no such thing as 'the' common-sense conception of vision that could enjoy epistemic default status simply in virtue of being endorsed by the folk. The paper thus challenges the epistemic default status commonly accorded to Direct Realism in debates about the nature of perceptual experience and calls into question, more generally, the methodological practice of according folk beliefs, *qua* 'common sense', an epistemic default status.

Section 1.1 introduces the 'common-sense' beliefs of interest as well as common philosophical assumptions about them. To motivate the critical investigation of these common assumptions, Sect. 1.2 reviews extant evidence of belief fragmentation. To render the issue empirically tractable, Sect. 1.3 motivates empirical hypotheses that challenge the targeted philosophical assumptions.

## 1.1 Direct Realism and common sense

Debates about the nature of perceptual experience standardly take for granted three assumptions about common sense: (1) There is a single, coherent, common-sense conception of vision; (2) this conception is captured by Direct Realism and challenged by Indirect Realism; (3) the common-sense conception enjoys ‘epistemic default status’.

According to *Direct Realism*, there are physical objects, like tables and trees, and their properties, including shape, size, and colour (‘realism’), and these objects and properties are ‘directly’ present in perception. This ‘directness claim’ has been cashed out in potentially complementary ways (review: Lyons, 2017). Two prominent interpretations include the metaphysical claim (‘perceptual directness’) that we do *not* see physical objects or their properties *by* or *in virtue of* perceiving (or being aware of) things that are distinct from them, such as mental images, ‘ideas’, or sense-data, and the epistemological claim (‘phenomenal directness’) that the presence of physical objects and their properties is *not inferred* from the presence of anything else (mental images, etc.). These claims contrast with *Indirect Realism*, which maintains that people see physical objects *by* or *in virtue of* perceiving, or being aware of, mental images, ‘ideas’, or sense data that the objects we look at cause in us (e.g., Jackson, 1977; Price, 1932; Robinson, 1994). Some versions of Indirect Realism also maintain the presence of physical objects and their properties is inferred from the presence of sense-data (e.g., Russell, 1912).

Direct Realism is standardly taken to capture our common-sense beliefs about perception. In arguing *against* Direct Realism, Hume, for instance, famously claimed that when people follow the ‘blind and powerful instinct of nature, they always suppose the very images, presented by the senses, to be the external objects’ (1777/1975, p. 151). In defending common sense against Hume, Reid agrees at least that the view that we are directly aware of ideas or images in our own minds ‘is directly contrary to the universal sense of men who have not been instructed in philosophy. When we see the sun or moon, we have no doubt that the very objects which we immediately see, are very far distant from us, and from one another’ (Reid, 1785/1969, p. 212). The still dominant view on the relationship between Direct and Indirect Realism is that ‘[we] all start out being Direct Realists. If an argument shakes our faith in this position, our initial reaction is to cling fast to the Realism, but to conclude that we are not directly aware of physical objects in the way we thought’ (Smith, 2002, p. 16).

Accordingly, the assumption that Direct Realism represents pre-reflective common sense is essential to ‘the problem of perception’, which is the focus of many debates about the nature of perceptual experience. As an authoritative statement puts it, ‘arguments at the heart of the problem of perception challenge th[e] *Direct Realist* perspective on perceptual experience. But since this perspective is embedded within our *ordinary conception* of perceptual experience, the problem gets to the heart of our ordinary ways of thinking’ (Crane & French, 2021, §1). The problem is generated by ‘arguments from illusion’ and ‘from hallucination’, which seek to show that the mere possibility of illusion and hallucination are inconsistent with Direct Realism. The resulting problem is therefore standardly understood as the challenge ‘that if illusions and hallucinations are possible, then perception, *as we ordinarily understand it*,

is impossible' (Crane & French, 2021, §2, emphasis added; cf. Brewer, 2011; Martin, 2002; Smith, 2002).

Common-sense beliefs about perception are (often implicitly) equated with the stable beliefs of scientifically untrained adults (e.g., Snowdon, 1981, p. 176), and taken to be widely, or even universally, shared (e.g., Hume, 1777; Lewis, 1997, p. 325; Strawson, 1959, p. 10). Some, like Hume, argue that the common-sense belief in Direct Realism should be rejected on the basis of philosophical reflection (e.g., Hume, 1777; Russell, 1912) or in light of our scientific understanding of perception (Russell, 1962, p. 13). Many others think that in virtue of capturing common sense, Direct Realism—not to be confused with the more specific position of Naïve Realism—should be accorded some form of epistemic default status: we should accept Direct Realism unless we have good reason to believe otherwise (e.g., Allen, 2020; Fischer et al., 2021; Genone, 2016; Martin, 2002; Reid, 1785/1969; Searle, 2015; Strawson, 1979).

This assumption of default status is common to proponents of different forms of Direct Realism, including both contemporary forms of Naïve Realism (e.g. Allen, 2020; Genone, 2016; Martin, 2002) and intentionalist or representationalist theories of perception (e.g. Searle, 2015; Strawson, 1979). The assumption can be motivated in different ways. Common-sense beliefs about vision are sometimes regarded as part of the 'massive central core of human thinking which has no history' (Strawson, 1959, p. 10), and therefore as so fundamental to how we conceive of ourselves and our relationship to the world that they are practically impossible to relinquish (e.g., Allen, 2020; Strawson, 1979). These common-sense beliefs are also taken to reflect the phenomenology of visual experience, and the phenomenology is (at least implicitly) assumed to represent a form of observational evidence (e.g., Genone, 2016; Martin, 2002; Nudds, 2009; Tye, 2000; for discussion, see Raineri, 2021). Indeed, these two motivations often go together, linked by the common assumption that the observational evidence is readily apparent and helps explain our common-sense beliefs.

Even philosophers of perception who propose theories that appear inconsistent with common sense, like Indirect Realist sense-datum theories of perception, often acknowledge the epistemic default status of common sense. This is either by trying to show that the inconsistency is merely apparent (e.g., Ayer, 1973), by accepting the need for an error theory that explains why common sense goes wrong (e.g., Boghossian & Velleman, 1989; Russell, 1912), or by defending their theories through reference to supposedly key insights of common sense to which they hold on. One example of this is the 'Phenomenal Principle': 'If there sensibly appears to a subject to be something which possesses a particular sensible quality, then there is something of which the subject is aware which does possess that sensible quality' (Robinson, 1994, p. 32). The Phenomenal Principle is central to many presentations of the argument from illusion, and is often motivated on the grounds that it is common sense (e.g., Moore, 1957, p. 134; Price, 1932, p. 63; Robinson, 1994, pp. 31–58).

*Prima facie*, there is a tension between the fact that philosophers of perception have confidently attributed to common sense both Direct Realism and the Phenomenal Principle often used to support Indirect Realism. This raises at least an initial doubt about the common philosophical assumption that there is a single, coherent common-sense conception of vision, consistent with Direct Realism, and deserving epistemic default

status. Findings about belief fragmentation provide a forceful empirical motivation for further examining this assumption.

## 1.2 Belief fragmentation

According to explanatorily ambitious ‘fragmentation’ accounts of belief storage, belief representations are stored in a large number of distinct data structures (‘fragments’) that are causally isolated: these structures are independently formed (through different processes or in different contexts), independently accessed (activated by different stimuli or the same stimuli in different contexts), and independently updated in the light of new information; fragments are internally coherent, but often conflict with each other (Bendaña & Mandelbaum, 2021; Borgoni et al., 2021). As a result, people hold many inconsistent beliefs. Such ambitious accounts of belief fragmentation rely on psychological hypotheses about cognitive architecture, and so differ from ‘minimalist’ views which conceptualise fragmentation in terms of consistency relations between beliefs (see, e.g., Schwitzgebel, 2021, p. 367).

These fragmentation accounts are currently still in their infancy, and contain only tentative suggestions, e.g., about the principles that decide which fragments get activated when. Their broad outline, however, is motivated by two well-supported claims (Leiser, 2001): First, some structural features of belief acquisition processes and of cognitive development promote the acquisition of conflicting beliefs (review: Mandelbaum, 2014). Second, our beliefs are never systematically checked for consistency. This is mainly due to constraints on working memory (review: Oberauer et al., 2016). It is also because surprisingly few people have a pronounced preference for consistent propositional attitudes (Cialdini et al., 1995), in particular where they have little personal investment in the content of the beliefs at issue (Kruglanski et al., 2018)—humans do not have a general need for cognitive consistency (*pace* the influential cognitive consistency paradigm, *cf.* Gawronski & Strack, 2012).

Evidence that conflicting beliefs are integrated into durable distinct belief fragments that are updated independently is provided by research on naïve theories and their persistence upon acquisition of scientific theories that conflict with them (see Shtulman, 2017, for a review). *Naïve theories* are causal-explanatory structures derived from a combination of innate conceptions, first-hand experience, and social learning (Gelman & Noles, 2011; Vosniadou, 1994). Prominent examples include impetus theories of motion (McCloskey, 1983) and substance theories of heat (Reiner et al., 2000; Wiser & Carey, 1983). While they arise for different reasons and take different forms, they tend to give internally coherent, perceptually grounded, object- (rather than process-) based explanations of everyday situations that attribute context-invariant properties to objects and agents.

Naïve theories are resistant to change in the face of scientific counter-instruction (Chi, 2005; Shtulman & Valcarcel, 2012; Vosniadou, 1994): Scientific theories introduce new concepts (like new notions of ‘force’, ‘molecule’, or ‘energy’), postulate imperceptible entities, and focus on process-based explanations at higher levels of abstraction and generalization (Thagard, 2014). Speeded statement evaluation studies suggest that the acquisition of such new concepts and theories does not completely

displace naïve theories which conflict with them: Across several different domains from the natural and life sciences, scientifically instructed adults proved prone to lapse into endorsing statements in line with the naïve theory but in conflict with the scientific theory when responding under time pressure; they also required more time or effort to correctly assess statements in line with the scientific theory but at odds with the naïve theory (Masson et al., 2014; Shtulman & Valcarcel, 2012; cf. Goldberg & Thompson-Schill, 2009; Kelemen et al., 2013). The phenomenon extends beyond science to further domains involving social learning. For example, speeded statement evaluation studies provided parallel evidence that in Christian adults person-based, embodied conceptions of God coexist with theologically mandated trans-personal, disembodied conceptions (Barlev et al., 2017, 2018; cf. Barrett & Keil, 1996). These findings suggest that naïve theories remain represented in the brain ('representational co-existence'), continue to be preferentially activated by verbal stimuli, and need to be effortfully suppressed, where they conflict with subsequently acquired scientific theories. We cannot unlearn naïve theories we form in childhood or adolescence; throughout adulthood, we maintain at least *implicit* belief in them, attributed on the basis of behavioural measures, in the absence of explicit endorsement.

Indeed, evidence of continued *explicit* endorsement of competing theories is provided by research on folk explanatory frameworks for human origins, illness, and death (review: Legare et al., 2012). This research documented endorsement of competing natural and supernatural theories that was not 'purely verbal' but went with explanatory reliance on the competing theories, resulting in substantive coexistence of natural and supernatural explanations. For example, in different South African samples, 50–100% of adults endorsed both biological and bewitchment explanations for the same (hypothetical) cases of AIDS (Legare & Gelman, 2008); and in a US sample, 60% of adult respondents generated mixed natural and supernatural explanations of why certain physical or psychological traits of a person (fail to) persist after their death, resulting in more or less well-integrated explanations (Watson-Jones et al., 2017). Contextual cues influence the extent to which people provide consistently natural or supernatural explanations (Astuti & Harris, 2008; Watson-Jones et al., 2017).

We will examine folk beliefs about vision. For these beliefs, extant evidence of fragmentation is provided by intrapersonal conflicts between naïve 'extramissionist' theories (according to which visual perception involves force rays leaving our eyes) and scientifically informed intromissionist beliefs (according to which visual perception involves light rays coming into our eyes) (review: Winer et al., 2002). Many children and adults who explicitly endorse extramissionist beliefs simultaneously agree with intromissionist beliefs (Cottrell & Winer, 1994), and many adults combine explicit belief in intromissionism with implicit belief in extramissionism (Guterstam et al., 2019, 2020). Once acquired, extramissionist beliefs are recruited to explain vision—when asked, most participants endorsing them take extramission to be functional for vision (Winer et al., 1996b; cf. Guterstam et al., 2019, p. 330) and are resistant to counter-instruction (Gregg et al., 2001). They have the hallmarks of 'intuitive' naïve theories (Carey, 2009; Shtulman, 2017): They are found in young children and across different age groups (Winer et al., 1996a, 1996b), cultures (Dundes, 1981), and historical periods (Lindberg, 1976), and have been articulated already by ancient philosophers (including Empedocles, Plato, and Euclid; see Meyering, 1989).

Collectively, the findings reviewed support the ‘fragmentation hypothesis’ that internally consistent but mutually conflicting sets of beliefs are stored in different data structures that can be accessed and updated independently, and where some are more amenable to updating than others. Findings simultaneously reveal complex interactions: conflicts may arise from simultaneous activation but be resolved by suppressing one of two conflicting belief sets; apparently conflicting theories may be deployed simultaneously in explanation; and the extent to which specific theories are deployed depends upon contextual cues.

### 1.3 Hypotheses

This study extends the fragmentation hypothesis to previously unexamined beliefs about vision that would seem to be in direct conflict with each other and are at the centre of current debates in the philosophy of perception: We will examine whether laypeople hold conflicting sets of beliefs (‘folk theories’) consistent with Direct Realism (‘folk Direct Realism’) and Indirect Realism (‘folk Indirect Realism’), respectively. As we understand it, *folk Indirect Realism* maintains that when viewing, e.g., an apple, we see a mental image of the apple, caused by that apple. We hypothesise this view is held in conjunction with the Cartesian Theatre conception of the mind, according to which input from the sense-organs is processed in the brain, where it results in a conscious experience that prototypically involves seeing a mental image (in an inner ‘theatre’). This conception is widely shared by laypeople (Forstmann & Burgmer, 2022). *Folk Direct Realism* rejects folk Indirect Realism: When looking at an apple, we see only the apple and no mental or other image of it, and do not infer the apple’s presence from anything else.

The conflict between these conceptions challenges the three assumptions that shape ongoing philosophical debates about the nature of perceptual experience (1–3 in Sect. 1.1). We will put these assumptions to the test by examining two hypotheses. Against the assumption (1) that there is a single, coherent common-sense conception of vision, we will examine the hypothesis that laypeople maintain incompatible pre-scientific folk theories about vision (folk Direct Realism and folk Indirect Realism) which form two different belief fragments. Philosophical accounts of belief fragmentation suggest that different fragments are activated by different stimuli (Bendaña & Mandelbaum, 2021). Consistent with this suggestion, studies that elicited extramissionist beliefs about vision with different stimuli and tasks (agreement ratings for verbal and pictorial representations, drawing tasks, forced-choice judgment tasks) elicited different responses (review: Winer et al., 2002): Extramissionist beliefs about vision are more widely endorsed when represented pictorially, rather than verbally (Winer et al., 1996b), perhaps because they are artefacts of a readily imageable spatial model of attention (*cf.* Guterstam et al., 2019, 2020). If the same spatial model is responsible for the Cartesian Theatre conception (as suggested by Webb & Graziano, 2015) that we hypothesise underpins folk Indirect Realism, then pictorial representations of the ‘Cartesian Theatre’ will activate folk Indirect Realism more strongly than verbal statements and increase endorsement:

**H<sub>1</sub> [fragmentation hypothesis]** Many laypeople endorse both folk Direct Realism and folk Indirect Realism, (i) when these are presented in the same stimulus format (verbally) and (ii) even more so when they are presented in different stimulus formats (verbal statements vs pictures).

Assumption (2) claims that folk Direct Realism captures ‘the’ single, coherent common-sense conception of vision. However, if **H<sub>1</sub>** is correct, it is not a foregone conclusion that folk Direct Realism will be as dominant as is standardly assumed in the philosophy of perception. We will therefore assess this previously unexamined dominance assumption:

**H<sub>2</sub> [mainstream dominance hypothesis]** Folk Direct Realism is explicitly maintained (i) by a majority of laypeople and (ii) far more frequently than folk Indirect Realism (say, by a ratio of at least two to one).

Given the rather widespread endorsement that recent studies (Forstmann & Burgmer, 2022) observed for the Cartesian Theatre conception that we hypothesise underpins folk Indirect Realism, we expect **H<sub>2</sub>** will remain unsupported.

The expected findings would also challenge the remaining assumption (3) that philosophers should accept the common-sense conception as a default, absent good reasons against it: While one can attempt to justify this assumption in various different ways, it is predicated on the existence of a set of folk beliefs about visual perception that is (1) coherent and (2) (almost) universally accepted or at least clearly dominant. If (as per **H<sub>1</sub>**) many laypeople are torn between mutually incompatible (Direct Realist vs Indirect Realist) conceptions of vision (i.e., if there is considerable intrapersonal disagreement) and if (*pace* **H<sub>2</sub>**) these two conceptions are maintained to roughly the same extent (i.e., if there are high levels of interpersonal disagreement), then there is no such thing as ‘the’ common-sense conception of vision to which philosophers could appeal. Even if each conception enjoyed epistemic default status simply in virtue of being maintained by many laypeople, the intra- and inter-personal conflict with a roughly equally popular folk conception would undermine their default standing prior to any inquiry.

## 2 Survey

To assess **H<sub>1</sub>** and **H<sub>2</sub>**, we conducted a survey with lay participants who had at most minimal exposure to philosophy, psychology, or natural science.

### 2.1 Methods

#### 2.1.1 Participants

100 participants were recruited through UK-based Prolific and remunerated. Participants were between 18 and 47 years old (mean 23 years). 36 were male, 61 female, and 3 non-binary. Participants were screened according to their subject of study and restricted to participants without higher-level education (UK A-Level and equivalent or

higher) in natural sciences, psychology or philosophy. All participants self-identified as ‘fluent’ in English. For details concerning sample size calculation and participants, see Supplementary Materials, Sect. A.

### 2.1.2 Materials and procedure

Participants were presented with 30 verbal and 12 image items that articulated or illustrated putative beliefs of folk Direct Realism and folk Indirect Realism, using exemplars at low levels of abstraction (*cf.* Wilson, 2006, p. 79). Participants rated 3 verbal items in each of ten conditions. Items probe for the metaphysical and epistemological beliefs commonly associated with Direct Realism and Indirect Realism, respectively.

#### Beliefs commonly associated with Direct Realism

- (1) *Perceptual Direct Object* (‘Direct Object’, for short) about stationary objects that are in sustained focus in stereotypical situations of perception: ‘When you look at a tower, you see only the tower and not any mental or other image of the tower.’
- (2) *Phenomenal Direct Object* about objects with distinctive looks that are hard to mistake: ‘When you look at a banana, you do not consciously work out that it is a banana, you just see there is a banana.’
- (3) *Perceptual Direct Properties* about objects with stereotypical shapes, sizes, or colours: ‘When you look at a tomato, you see the colour of the tomato and not the colour of an image of the tomato.’
- (4) *Phenomenal Direct Properties* about objects with stereotypical and nameable colours or shapes: ‘When you look at a fire engine, you just see there is something red and you do not consciously work out that there is something red.’

#### Beliefs commonly associated with Indirect Realism

- (5) *Generic Indirect* about stationary objects that are in sustained focus in stereotypical situations of perception: ‘When you look at a table, you see an image of the table, rather than just the table.’
- (6) *Causal Indirect* about stationary objects that are stereotypically ‘shiny’ and reflect or emit light: ‘When you look at a star in the night sky, you see an image that is caused by the star.’
- (7) *Phenomenal Size/Shape* about familiar situations of non-veridical perception, where things appear to have nameable shapes or sizes: ‘When you look at a red car from the top of a tower, the car looks small and you see a red patch that really is small.’
- (8) *Phenomenal Colour* about familiar situations of non-veridical perception, where things appear to have nameable colours: ‘When you look at a sheet of paper under intense red light, the paper looks red and you see a rectangular patch that really is red.’

## Beliefs commonly associated with defences of Direct Realism against Indirect Realism

- (9) *Counter-phenomenal Size/Shape* about familiar situations of non-veridical perception, where things appear to have nameable sizes or shapes: ‘When you look at a large ship anchored out at sea, the ship looks small, but you don’t see anything that really is small.’
- (10) *Counter-phenomenal Colour* about familiar situations of non-veridical perception, where things appear to have nameable colours: ‘When you look at a white lab coat under intense red lighting, the coat looks red, but you don’t see anything that really is red.’

Conditions (7) and (8) articulate the Phenomenal Principle invoked by the key philosophical arguments for Indirect Realism (see Sect. 1.1), for exemplary ‘primary qualities’ (size and shape) and ‘secondary qualities’ (colour), respectively. (9) and (10) articulate the denial of the Phenomenal Principle, associated with defences of Direct Realism against Indirect Realism. A list of all 30 verbal items is provided in the Supplementary Materials (Sect. B).

The 12 image items are also displayed in the Supplementary Materials (Sect. B). Items I1–I3 illustrate the Cartesian Theatre conception that we hypothesise underpins folk Indirect Realism and takes the viewer to see an image of the object of sight, in their mind or head (Fig. 1). Dashed lines from eyes to object were intended to represent the direction of gaze and attention. Solid lines from object to eyes were intended to represent causal impact on sense-organs. Further images represent a (putatively Direct Realist) view which posits pictorial representations in the mind or head but does not place the viewer into a (quasi-) perceptual relationship to them (I4–I6), a version of Direct Realism which posits non-pictorial (verbal) representations (I7–I9), and a ‘naïve realist’ version of Direct Realism (I10–I12).

A *norming study* was conducted to ensure that verbal items were intelligible and that key image items (including I3) were interpreted as intended (see Suppl. Materials, Sect. C).

In the *main study*, verbal items were presented in random order. Participants used a 7-point Likert scale to rate how much they agreed with each item. To guard against acquiescence bias (Jackman, 1973), two versions of our questionnaire presented response options in different order: Half the participants received “strongly disagree” as first response option for all items, the other half were given “strongly agree” as first option. The rating task was followed by two open-text questions designed to probe participants’ interpretation of phenomenal items (7–10).

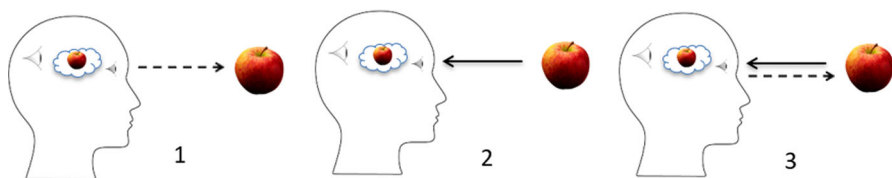


Fig. 1 Image items I1–I3

After that, 12 image items were presented in a fixed order designed to draw attention to differences between images. Participants first viewed all 12 images on the same screen, simultaneously, then each individually. Upon individual presentation, participants used a 7-point Likert scale to rate how much they agreed that the image ‘correctly represents what happens when we see an apple’. As before, response options were presented in different orders (‘strongly disagree’ first vs ‘strongly agree’ first) to different participants. Finally, participants were shown all 12 images again, asked ‘Which image best represents what happens when we see?’, and the open text question ‘Please explain your previous response: What do you think happens when someone sees something?’ Responses to these two questions allowed us to assess how participants interpreted images.

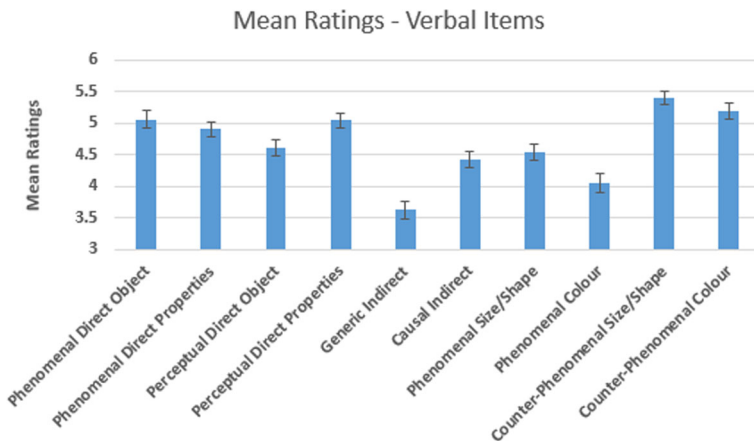
### 2.1.3 Analysis

We scrutinized participants’ responses for evidence of non-engagement (e.g., sustained series of identical responses to different items) but found none. We scrutinized verbal items by examining the correlations of item ratings in each condition. In each condition, correlations were significant ( $p$ ’s < .035, most  $p$ ’s < .01) and at least of medium size ( $\geq .3$ , except in Counter-phenomenal-size/shape, where they were between .2 and .3).

To assess our hypotheses, we sought to determine the proportions of participants who hold folk Direct Realism, folk Indirect Realism, both, and neither. To do so, we considered responses to verbal items and performed two series of cluster analyses. We first used responses to the core claims Direct Object (1 above) and Causal Indirect (6 above), respectively, as indicator of adherence to folk Direct Realism and folk Indirect Realism, respectively. To be able to use responses to a wider range of items as indicator, we then conducted an exploratory factor analysis to determine which of the beliefs examined tend to be held together with the core beliefs Direct Object and Causal Indirect, and can be regarded as part of the same folk theory. After using each indicator of adherence to folk Direct Realism and folk Indirect Realism (henceforward ‘Direct Theory’ and ‘Indirect Theory’, for brevity), we employed three different criteria (set out below) for cluster analyses, to individuate groups of interest. We do not think that any one of these analyses (or any further analysis) offers a uniquely ‘best’ classification, but rather that a good understanding of how widely the relevant folk theories are held can be obtained by considering how these classifications affect the assessment of our hypotheses. Finally, we considered how ratings of verbal and image items align.

## 2.2 Results

Mean ratings for each condition (Fig. 2) informed subsequent analyses and the interpretation of results.



**Fig. 2** Mean ratings for verbal items on 7-point scale (whole sample). Error bars show standard error of the mean

### 2.2.1 Verbal item ratings: core beliefs

We first considered responses to the core claims Direct Object and Causal Indirect. Direct Object is the metaphysical core claim of Direct Realism distinguishing it from Indirect Realism. Causal Indirect articulates the prototypical form of Indirect Realism, as evidenced by the fact that it received higher agreement ratings in our study than Generic Indirect (4.42 vs 3.62;  $t(99) = 5.56, p < .001$ ), which it entails (*cf.* Kahneman & Frederick, 2002; Tversky & Kahneman, 1982). This suggests using ratings in these conditions as simple indicator of adherence to folk Direct Realism and folk Indirect Realism, respectively.

**K-means cluster analysis (K-means1)** We first performed a K-means cluster analysis (in SPSS). This automatic analysis identifies groups of participants in a data-driven manner, without any preconceptions about what the groups should look like. Roughly speaking, it identifies maximally homogeneous groups of participants, who are as similar to each other as possible in the mean ratings they give to the items of interest. The analysis uses vector quantization to partition the individuals into the desired number of clusters, so that each individual is assigned to the cluster with the centroid nearest to its mean ratings (Ding & He, 2004; Wu, 2012). Each centroid thus constitutes the prototype for a cluster, whose members are more similar to it than to any other prototype. The analysis software is provided with the number of groups to be identified and no further information (e.g., about the scales employed) and attempts to reduce the squared Euclidean distances between individuals and cluster centroids. The mainstream hypothesis ( $H_2$ ) predicts there will be at least two groups: ‘direct theorists’ who accept Direct Object and reject Causal Indirect, and ‘indirect theorists’ who accept Causal Indirect and reject Direct Object. Our fragmentation hypothesis ( $H_1$ ) predicts there will be a third group, which accepts both Direct Object and Causal

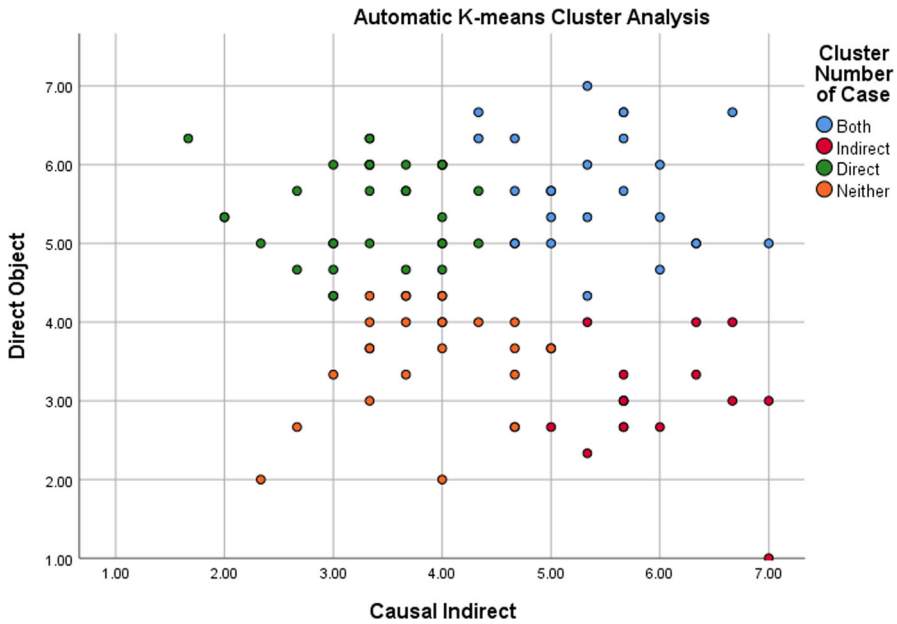


Fig. 3 Scatterplots of the four groups based on K-means cluster analysis (K-means1). Dots conflate participants with identical mean ratings; not all participants show up as a distinct dot

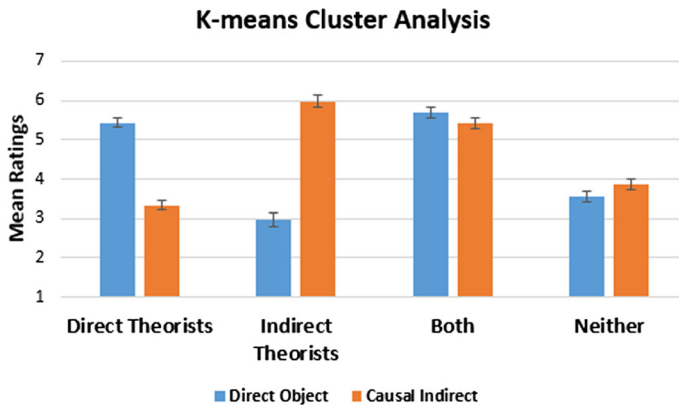
Table 1 K-means cluster analysis: locations of centroids of clusters 1–4

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
X-axis: causal indirect	5.52	5.75	3.39	3.67
Y-axis: direct object	5.62	3.02	5.63	3.77

Indirect. A fourth group might fail to accept either, e.g., because they do not hold any beliefs about the matter. To test these predictions, we instructed the programme to identify four groups and assessed whether the groups identified correspond to the four logically possible positions.

The ratings of interest place participants into a two-dimensional coordinate space, with mean ratings for Causal Indirect items on the x-axis and mean ratings for Direct Object items on the y-axis (see Fig. 3). The centroids of the four clusters identified by our analysis fall into the four quadrants of this space, to either side of the ‘neutral’ mid-point values of ‘4’ (see Table 1).

Each centroid thus represents one of the four predicted positions and indicates their ‘prototypical’ mean ratings. The centroids facilitate a similarity-based classification of participants, which classifies members of each cluster as adherents to the position to whose prototypical ratings their scores are most similar. We thus classify participants as members of one of the four predicted groups: ‘direct theorists’ (whose prototypical



**Fig. 4** Mean ratings for the four groups based on K-means cluster analysis (K-means1). Error bars show the standard error of the mean

position is to agree with Direct Object, but disagree with Causal Indirect), ‘indirect theorists’ (prototype: agree with Causal Indirect but disagree with Direct Object), ‘torn souls’ (prototype: agree with both claims) and ‘objectors’ (prototype: agree with neither). In our sample of 100, this approach identifies 33 direct theorists, 16 indirect theorists, 25 torn souls (‘both’) and 26 objectors (‘neither’) (Fig. 3).

This data-driven, similarity-based classification of participants identifies four groups whose mean responses to the items of interest (Direct Object and Causal Indirect) are consistent with the four positions of interest (Fig. 4). However, it does not align perfectly with intuitively meaningful classifications. For instance, seven participants get classified as objectors (‘Neither’) despite agreeing with Causal Indirect. We therefore complemented this automatic classification with manual classifications employing two more intuitive criteria.

**Manual cluster analyses: Threshold1 and Discrepancy1** The most intuitive criterion is a threshold criterion. According to *Threshold1*, a participant qualifies as ‘direct theorist’ if and only if she agrees with Direct Object, i.e., gives Direct Object items a mean rating numerically above mid-point ‘4’, but does not agree with Causal Indirect, i.e., gives Causal Indirect items a mean rating  $\leq 4$ . The converse holds for ‘indirect theorists’. ‘Torn souls’ give both Direct Object and Causal Indirect mean ratings numerically above 4; ‘objectors’ give neither such ratings. In our sample of 100, this criterion recognizes 36 direct theorists, 24 indirect theorists, 27 torn souls (‘both’), and 13 objectors (‘neither’). This classification agrees with previous K-means classification in 85% of cases. It also succeeds in identifying four groups whose mean responses to the items of interest (Direct Object and Causal Indirect) are consistent with the four positions of interest. Details and full results are reported in the Supplementary Materials (Sect. D).

The threshold criterion is sensitive only to whether a participant agrees with Direct Object or Causal Indirect, and not to the strength of the agreement. But if a participant has a clear preference for one claim over the other, the clearly preferred claim may

guide her thinking about vision at the expense of the competing claim, even if she ‘agrees’ with both, in the sense of rating both numerically above mid-point ‘4’ when presented with them. The threshold criterion may thus increase the number of ‘both’ classifications (which is critical for our  $H_1$ ) beyond that of genuinely torn souls.

We therefore applied, next, an empirically derived discrepancy criterion: According to criterion *Discrepancy1*, a participant qualifies as ‘direct theorist’ if and only if their mean rating for Direct Object items is at least one point (on the 7-point Likert scale) higher than their mean rating for Causal Indirect items, and is not below mid-point ‘4’; similarly for ‘indirect theorist’. This criterion recognizes 39 direct theorists, 28 indirect theorists, 19 torn souls (‘both’), and 14 objectors (‘neither’). The analysis results in 72% agreement with the K-means analysis. As before, the mean ratings of the four groups for Direct Object and Causal Indirect are consistent with the four positions of interest. Details and full results are reported in the Supplementary Materials (Sect. D). Headline results of all analyses are summarised in Table 4.

### 2.2.2 Verbal item ratings: sets of beliefs

The first three analyses used attitudes towards a single core claim—Direct Object and Causal Indirect, respectively—as simple indicator of adherence to folk theories. To be able to use responses to a wider range of items as indicator, we needed to determine which of the beliefs examined tend to be held together with those core beliefs and can be regarded as part of the same folk theory. To do so, we conducted an exploratory factor analysis, before conducting three cluster analyses analogous to the above.

**Factor analysis** Exploratory factor analysis (using principal components extraction and without rotation) revealed four factors with Eigenvalues over 1. These were retained in line with the Scree test (Cattell, 1966) (see Supplementary Materials, Sect. D). We interpreted factor loadings of .512 or larger as significant (Stevens, 2012). The four factors explained approximately 71% of the variance in verbal item ratings (see Table 2)—well above the 60% threshold commonly employed in the social sciences (Hair et al., 2014).

High positive loadings on the same factor mean that ratings for items in different categories move together and correlate in the same direction. Negative loadings on the same factor mean that ratings move together but in the opposite direction. The factors obtained are therefore indicative of which of the beliefs tapped by our different item categories commonly go together, in our lay sample. Loadings on Factor 1 reveal that belief in Direct Object typically goes with those other five of the ten beliefs we examined that are commonly associated with the philosophical theory of Direct Realism. We infer that these six beliefs are part of a folk version of Direct Realism. Loadings on Factor 3 reveal that belief in the prototypical version (Causal Indirect) and the generic version (Generic Indirect) of the key claim of Indirect Realism go together and form part of a folk version of Indirect Realism. Interestingly, the observed loadings suggest that none of the other beliefs examined form part of this folk theory—including beliefs associated with philosophical theories of Indirect Realism, such as the Phenomenal Principle (Phenomenal Size/Shape and Phenomenal Colour, whose ratings load on Factor 2). Items with significant loadings on Factors 1 to 3, respectively, displayed

**Table 2** Results of the factor analysis on the ten verbal item categories

	Factor 1	Factor 2	Factor 3	Factor 4
1. Causal Indirect	-.26	.16	.66	.25
2. Generic Indirect-	.14	.05	.80	-.14
3. Direct Object	.69	.23	-.24	.32
4. Direct Property	.56	.32	-.22	.40
5. Phen. Direct Object	.70	.31	.06	-.48
6. Phen. Direct Property	.73	.20	.10	-.49
7. Counter Size/shape	.59	-.27	.41	.33
8. Counter Colour	.71	-.27	.35	.17
9. Phen. Size/shape	-.21	.81	.16	-.03
10. Phen. Colour	-.16	.79	.03	.21
Percentage of Variance	28.23	17.36	15.18	10.04
Eigenvalues	2.823	1.736	1.518	1.004
Cronbach's $\alpha$	.843	.787	.693 (.749 with one 'bad' item deleted)	

*Note:* Significant loadings highlighted in grey

reasonable-to-good internal consistency (see Cronbach's  $\alpha$ 's in Table 2), further suggesting they tapped into the same knowledge structures (which we interpret as folk Direct Realism, the Phenomenal Principle, and folk Indirect Realism, respectively).

**Cluster analyses: K-means2, Threshold2, and Discrepancy2** To determine how many of our participants held one of these folk theories, or both, or neither, we first conducted a k-means cluster analysis ('K-means2') based on the two relevant factors 1 ('Direct Factor') and 3 ('Indirect Factor'). To do so, we used the saved factor scores generated by the factor analysis and instructed the analysis software to identify four groups. As in the initial k-means cluster analysis, the centroids fell into the four quadrants of the relevant coordinate space and permitted similarity-based classification of participants, namely, as adherents to the position to whose prototypical ratings their scores are most similar. As before, we followed up with two cluster analyses, which used a threshold and a discrepancy criterion, respectively. The manual cluster analysis 'Threshold2' classified a participant as 'direct theorist' if and only if their mean rating for items articulating the six examined component beliefs of folk Direct Realism was numerically above mid-point '4' and their mean rating for items articulating the two examined component beliefs of folk Indirect Realism was  $\leq 4$ . The converse holds for 'indirect theorists'. For both item classes, 'torn souls' had means numerically above 4, and 'objectors' means  $\leq 4$ . The cluster analysis 'Discrepancy2' employed a discrepancy criterion proportionally similar to the previous discrepancy criterion.

The discrepancy analysis yielded high agreement with the k-means analysis (98% 'hit rate' across the two key categories of Direct and Indirect Theorists), whereas agreement between threshold analysis and k-means analysis was comparatively low (66%). The mean ratings of the four groups patterned roughly as before in all analyses, but the torn souls ('both') identified by K-means2 and Discrepancy2 gave considerably higher mean ratings to Direct Factor items than to Indirect Factor items, and in Discrepancy2 the same went for objectors ('neither'). Threshold2 classified more participants as Direct Theorists than any other classification. The details of these analyses,

and their full results, are reported in the Supplementary Materials (Sect. D). Headline results are reported below, in Table 4.

### 2.2.3 Image ratings

We finally examined agreement ratings for the key images (I3, I6, I9, and I12, see Supplementary Materials, Sect. B), which participants gave in response to the question, ‘Does this image correctly represent what happens when we see an apple?’ To assess whether participants interpreted these images as intended, we first examined mean ratings for these images as well as responses to the subsequent tasks, where participants indicated which image best represents what happens when we see, and then explained their answer.

We considered the mean ratings by the four different groups. Intended interpretations predict certain patterns of agreement: For example, the intended Indirect Realist interpretation of I3 predicts higher mean ratings by adherents of folk Indirect Realism than folk Direct Realism. We first scrutinized the mean image ratings for the predicted patterns (see Supplementary Materials, Sect. E). Results suggested that only I3 had been interpreted as intended, while I6 (similar to I3, but without the inner eye) had been given the same interpretation as I3. Analysis of image preferences and their open-text explanations further confirmed that participants placed the same interpretation on I3 and I6 and took both to illustrate Indirect Realist ideas (see Supplementary Materials, Sect. E). The interpretation of the remaining images remains unclear. We can hence regard only ratings for I3 and I6 as indicative of either Direct or Indirect Realist beliefs about vision, with high ratings indicative of Indirect Realist beliefs.

To facilitate assessment of our hypotheses, we considered how many members of each of our four groups agreed that these images correctly represent what happens when we see something (as indicated by ratings ‘5’–‘7’). We considered ratings for I3 and I6 by group, for each of our four groups, as classified by the two automatic K-means cluster analyses and the two manual analyses that best agreed with them (see Table 3).

**Table 3** Percentages (and proportions) agreeing with image (rating 5–7) by group, on four different classifications

	Direct	Indirect	Both	Neither
<b>Image 3 (70%)</b>				
K-means1	60.6 (20/33)	81.4 (13/16)	68.0 (17/25)	76.9 (20/26)
Threshold1	61.2 (22/36)	79.2 (19/24)	70.3 (19/27)	76.9 (10/13)
K-means2	64.3 (18/28)	78.8 (11/14)	71.9 (23/32)	69.2 (18/26)
Discrepancy2	66.7 (20/30)	78.6 (22/28)	68.2 (15/22)	65.0 (13/20)
<b>Image 6 (69%)</b>				
K-means1	60.6 (20/33)	62.5 (10/16)	76.0 (19/25)	76.9 (20/26)
Threshold1	58.3 (21/36)	70.8 (17/24)	74.1 (20/27)	84.6 (11/13)
K-means2	67.8 (19/28)	78.6 (11/14)	65.6 (21/32)	69.2 (18/26)
Discrepancy2	66.7 (20/30)	78.6 (22/28)	63.6 (14/22)	65.0 (13/20)

Percentages for whole sample (N = 100) in bold

## 2.3 Discussion

We used six different classifications to determine what proportion of our lay sample hold folk theories of vision in line with Direct Realism and Indirect Realism, respectively. Table 4 summarises the findings from verbal items.

The different classifications appear to yield markedly different results, and contain an outlier (Threshold2). To understand and assess these differences, we need to refer back to the different criteria the classifications employ. The first three classifications take into account only the ratings for the metaphysical key claims of Direct and Indirect Realism, namely Direct Object and Causal Indirect, respectively. The second three classifications are based on the ratings also for further claims which typically coincide with those key claims: For folk Direct Realism, these are five claims that all have higher

**Table 4** Number of participants (N = 100) classified as adhering to Direct Realist and Indirect Realist folk theories of vision, on six different classifications

Classification	Direct	Indirect	Both	Neither	Direct + Both	Indirect + Both
K-means1	33	16	25	26	58	41
Threshold1	36	24	27	13	63	51
Discrepancy1	39	28	19	14	58	47
K-means2	28	14	32	26	60	46
Threshold2	47	10	34	9	81	44
Discrepancy2	30	28	22	20	52	50

The two columns furthest to the right add up figures from previous columns, and represent the total number of all participants who accept either Direct or Indirect Realism (regardless of whether they *also* accept the conflicting folk theory)

mean ratings than Direct Object; for folk Indirect Realism this is a more generic claim with lower mean ratings than Causal Indirect (Fig. 2). Accordingly, several participants who do not agree with Direct Object get pushed above the mid-point threshold by the other claims of folk Direct Realism they agree with, while several participants who endorse Causal Indirect get dragged below the threshold by the generic claim they do not agree with. This leads to the higher number of folk Direct Realists and lower number of folk Indirect Realists in the second threshold classification (Threshold2). In what follows, we will set aside Threshold2 in assessing our hypotheses. This is because we would hesitate to recognize someone who fails to endorse Direct Object as holding a folk theory that is Direct Realist in a philosophically meaningful sense; similarly, we would hesitate to deny that proponents of Causal Indirect hold a folk theory that is Indirectly Realist, just because they fail to agree with a more generic claim they could (but don't) deduce from it. (For simplicity, we also do not report further analyses for Discrepancy1, which pattern in all relevant respects with Threshold1).

As already noted, we do not think that any one of the remaining analyses (or any further analysis) offers a uniquely 'best' classification; rather, a good understanding of how widely and strongly the relevant folk theories are held can be obtained by considering how these classifications affect the assessment of our hypotheses (repeated for convenience).

**H<sub>1</sub> [fragmentation hypothesis]** Many laypeople endorse both naïve theories of vision that can be regarded as folk versions of Direct Realism and Indirect Realism, respectively, (i) when these are presented in the same stimulus format (verbally) and (ii) even more so when they are presented in different stimulus formats (verbal statements vs pictures).

Part (i) is borne out by the observation that, depending upon the classification used, 19–32% of our participants endorsed both competing theories, when claims were presented verbally (Table 4, 'Both' column). A *fifth of participants* endorsed both key claims (19%, Discrepancy1) or both theories (22%, Discrepancy2) without a clear preference for one over the other, and stand to be equally strongly influenced by either in their thinking about vision. A *quarter of participants* (27%) agreed with both the directly conflicting key claims of folk Direct Realism and folk Indirect Realism (Threshold1). A *third of participants* (32%) endorsed both theories, on an automatic classification that took into account consistency and magnitude of agreement to a range of component claims of these theories (K-means2).

Part (ii) of **H<sub>1</sub>** is supported by observing that image ratings reveal yet more participants hold both folk theories: *two thirds* of the 'Direct Theorists' who endorsed verbal statements of folk Direct Realism (but not of folk Indirect Realism) agreed that images representing folk Indirect Realism represent correctly what is going on when people see a mundane object (I3: 64–67%; I6: 67–68%; see Table 3, K-means2 and Discrepancy2). Strikingly, 61% of 'Direct Theorists' who endorsed the metaphysical core claim of folk Direct Realism, Direct Object, but not the prototypical claim of folk Indirect Realism, Causal Indirect, agreed with the Indirect Realist image (I3) that directly contradicts Direct Realism's metaphysical core claim; and their agreement was merely about 10% lower than that of participants who endorsed verbal statements

**Table 5** Conflict vs. consistency: number of participants (N = 100) whose verbal and image ratings provide evidence of agreement with both folk Direct Realism and folk Indirect Realism ('Conflicted') and of agreement only with folk Direct Realism ('Direct Realism Only'), on four different classifications

	Conflicted	Direct Realism only	Ratio
K-means1	45	13	3.46:1
Threshold1	49	14	3.50:1
K-means2	50	10	5.0:1
Discrepancy2	42	10	4.2:1

of both theories, and ca. 20% lower than that of pure 'Indirect Theorists' (see Table 3, I3, Threshold1, cf. K-means1).

To comprehensively assess  $H_1$ , we add those 'Conflicted Direct Theorists' (Direct Theorists who agree with image I3) to Torn Souls (who agree with verbal statements of both folk theories). This reveals that about half of our participants held conflicting folk theories of vision (see left-hand column in Table 5). To gauge to what extent conflicted or consistent belief in folk Direct Realism is the norm, we finally compare the number of 'conflicted' and 'consistent' adherents to folk Direct Realism: The 'conflicted' combine verbal agreement with folk Direct Realism with agreement to verbal or pictorial representations of folk Indirect Realism, whereas the 'consistent' agree with verbal statements of folk Direct Realism, but agree with no statement (verbal or other) of folk Indirect Realism. We find that the conflicted outnumber the consistent folk Direct Realists by ratios ranging from 3.5:1 (on the narrow measure that takes into account only ratings of key claims) to 5:1 (see Table 5).

In summary, holding conflicting folk theories of vision is considerably more common than consistent adherence to folk Direct Realism and no conflicting such theory. While the proportion of consistent adherents to folk Indirect Realism remains to be established, and the beliefs of Objectors (who hold neither theory, 13–26% of our sample) remain to be explored, present findings suggest fragmented and conflicted beliefs about vision are the norm, not the exception.

We now turn to the mainstream dominance hypothesis we did not expect to stand up to scrutiny:

**$H_2$  [mainstream dominance hypothesis]** Folk Direct Realism is explicitly maintained (i) by a majority of laypeople and (ii) far more frequently than folk Indirect Realism (say, by a ratio of at least two to one).

This is typically asserted on the assumption that laypeople will hold only one of these two incompatible folk theories. However, we have seen that a non-negligible number of our participants agreed with both. We therefore assess part (i) of  $H_2$  by considering the figures both for 'Direct Theorists' and for 'Direct Theorists + Both' in Table 4. We see that a less than crushing, but clear majority of participants (52–63%, setting aside Threshold2 for the reasons explained) endorse folk Direct Realism, either on its own, or along with folk Indirect Realism. However, only a minority of participants (28–39%) agree with verbal statements expressing just folk Direct Realism, as generally assumed by proponents of  $H_2$ . We also see that only a small minority (10–14%) remain not

**Table 6** Folk Direct Realism vs folk Indirect Realism: ratios on four different classifications

	Direct theorists + torn souls: indirect theorists + torn souls
K-means1	1.4:1
Threshold1	1.2:1
K-means2	1.3:1
Discrepancy2	1.04:1

merely verbally loyal to folk Direct Realism but also desist from a simultaneous pictorial affair with folk Indirect Realism ('Direct Realism Only' in Table 5). True fidelity to folk Direct Realism is achieved by the few, not the many.

To assess part (ii) of  $H_2$ , we consider the ratio of participants endorsing folk Direct Realism (Direct Theorists + Torn Souls) to participants endorsing folk Indirect Realism (Indirect Theorists + Torn Souls), all based on ratings of verbal items (see Table 6; calculated from the numbers in Table 4). We find that the core claim of folk Direct Realism, Direct Object, is slightly more popular than the core claim of folk Indirect Realism, Causal Indirect (K-means1 and Threshold1), and the same goes for the examined larger components of the two theories (K-means2). But even this slight difference seems to vanish the moment we take measures that ensure participants will qualify as Direct Theorists only if they have a clear preference for the whole of folk Direct Realism over folk Indirect Realism (Discrepancy2).

This negative assessment of the mainstream dominance hypothesis still allows that folk Direct Realism might enjoy a more subtle dominance: Laypeople who endorse folk Direct Realism might endorse this Direct Realist view more strongly than laypeople who endorse folk Indirect Realism endorse this Indirect Realist view. Further analyses ruled out this suggestion (see Supplementary Materials, Sect. F).

To facilitate interpretation of our findings (and in response to a reviewer query), we finally conducted a follow-up study to examine whether laypeople regard folk Direct Realism and folk Indirect Realism as incompatible—or interpret them as making compatible claims, e.g., using the verb “see” in different senses: Laypeople recruited in the same way as in the main study provided compatibility ratings for items that paired statements using “see” in the same sense and in different senses, and for items pairing statements of Direct Object and Causal Indirect beliefs (see Supplementary Materials, Sect. G). Findings suggest laypeople do not place a ‘different-sense’ interpretation on the main study’s key items that would render them compatible and that insight into the resulting conflict is not universal, but widespread.

### 3 General discussion

#### 3.1 Main findings and empirical discussion

Our survey with lay participants examined folk beliefs about vision, with a focus on two sets of conflicting beliefs, consistent with Direct Realism ('folk Direct Realism')

and Indirect Realism ('folk Indirect Realism'), respectively. We found substantial variation between individuals and substantial conflict within individuals. Conflicted belief about vision was widespread, and folk Direct Realism was endorsed neither much more widely nor much more strongly than folk Indirect Realism.

First, *many laypeople's beliefs about vision are conflicted*: About half of our participants endorsed both folk Direct Realism and folk Indirect Realism when these were presented in different stimulus formats (by verbal statements vs pictures). A fifth to a third of participants endorsed both positions even when these were presented in the same stimulus format (namely, verbally). Conflicted beliefs about vision proved to be the norm, not the exception: Conflicted Direct Realists, who endorse both folk Direct Realism and folk Indirect Realism, when given verbal or pictorial representations, outnumbered consistent Direct Realists, who endorse the former theory, but not the latter, by ratios ranging from 3.5:1 to 5:1, depending upon the criteria used to classify participants as Direct Realists.

Second, *Folk Direct Realism is prominent, but not clearly dominant*: A clear, but not crushing majority (52–63%) of laypeople endorsed folk Direct Realism. But only a small minority (10–14%) maintained *only* folk Direct Realism, without simultaneously holding conflicting beliefs in folk Indirect Realism that could be activated by verbal stimuli or images. Most folk Direct Realists held their Direct Realism as part and parcel of an incoherent store of conflicting beliefs about vision. Overall ('pure' plus 'conflicted') endorsement of verbal statements of folk Direct Realism was only slightly more common than analogous endorsement of folk Indirect Realism, with ratios ranging from 1.4:1 to 1.04:1. The co-existence of conflicting beliefs thus undermined the dominance of folk Direct Realism.

Fragmentation accounts of belief storage suggest a straightforward interpretation of these findings: folk Direct Realism and folk Indirect Realism are contents of different belief fragments. In support of this interpretation, we now consider to what extent our findings are consistent with the key assumptions of such accounts (see Sect. 1.2) or are open to alternative explanations.

First, fragmentation accounts assume *conflicts between belief fragments*. Present findings provide evidence of relevant intrapersonal belief conflicts, if agreement with conflicting statements (and with images) is interpreted as expression of conflicting beliefs. An alternative interpretation suggests that our lay participants lack (some of) the beliefs of interest and (sometimes) decide on the spot, driven, for example, by acquiescence bias (Jackman, 1973). If this were correct, then—absent prior beliefs and sustained reflection during responding—participants' ratings should be influenced by the order in which response options are presented. In verbal item ratings, we indeed observed a medium-sized order effect for Generic Indirect, and a small effect for Causal Indirect, though no significant effect for other item categories (Supplementary Materials, Sect. F). However, we observed no order effects for image items illustrating Causal Indirect (*ibid.*), and image ratings cohered well with subsequent open-text responses explaining 'what happens when we see an apple' (Supplementary Materials, Sect. E). These free-text responses were all consistent with shallow understanding, but such understanding of complex causal systems (like vision) can translate into comparatively stable beliefs that are held until forcefully challenged (Rozenblit & Keil, 2002). Together with verbal ratings of Direct Realist items (which displayed no

order effects), image ratings provide evidence that about half our participants hold conflicting beliefs about vision.

Second, fragmentation accounts posit that mutually inconsistent belief fragments are nevertheless *internally consistent*. Our factor analysis provided initial evidence of internal coherence (see Table 2): Wherever sets of items contradict each other (Causal or Generic Indirect vs. Direct Object or Property; Phenomenal Size/shape or Colour vs. Counter-phenomenal Size/shape or Colour), and one category loads significantly and positively on one factor, the other loads negatively on that factor, indicating that ratings for these items move in the opposite direction. The three factors with significant loadings, crucially including factors 1 (Direct Realism) and 3 (Indirect Realism) are indicative of internally consistent sets of beliefs.

Third, fragmentation accounts assume *independent access*: Consistent with this assumption, agreement with folk Indirect Realism proved sensitive to the stimulus format (verbal statements vs images); images elicited agreement with folk Indirect Realism from over 60% of participants who endorsed verbal statements of folk Direct Realism only (Table 3). Similar to other studies that elicited explicit agreement with conflicting beliefs (see Sect. 1.2), even different verbal statements presented in the same context (the same questionnaire, for the same task) managed to elicit agreement with both directly conflicting folk theories.

Finally, fragmentation accounts propose that conflicting belief fragments are *independently formed* and *independently updated*. Our findings do not address the questions of how beliefs in folk Direct Realism and folk Indirect Realism, respectively, are formed or updated. It is, however, hard to see how internally coherent sets of directly contradictory beliefs could be formed and persistently maintained over time, unless they were formed and updated independently. While further work, beyond the remit of this paper, is required to establish that belief fragmentation provides the best available explanation of present findings (see below, Sect. 3.3), we therefore tentatively conclude that folk Direct Realism and folk Indirect Realism are constitutive of two distinct belief fragments.

### 3.2 Philosophical consequences: the ‘fragmentation challenge’ and potential responses

It is standardly assumed in the philosophy of perception that there is a single, coherent common-sense conception of vision, and that this conception is captured by Direct Realism. It is also widely assumed that this conception has some form of epistemic default status, and that we should accept the conception, unless we have good reasons not to. Accordingly, consistency with common sense is treated as evidence in support of a philosophical theory.

Findings of conflict and fragmentation challenge this widespread practice of privileging common sense. Despite being perceived as mutually incompatible, folk Direct Realism and folk Indirect Realism are maintained by roughly similar numbers of laypeople (with at most a mild preponderance of folk Direct Realists), and about half of our participants were torn between the two conceptions. If, in line with common

philosophical practice, ‘common sense’ is equated with the stable beliefs of scientifically untrained adults (see Sect. 1.1), present findings suggest there is no such thing as ‘the’ (single, coherent) common-sense conception of the phenomenon to which one could appeal. Findings of conflict and fragmentation thus pose the *fragmentation challenge* to philosophical appeals to common sense: Folk beliefs do not deserve epistemic default status simply in virtue of being ‘common-sense’ in the sense of folk or pre-scientific beliefs; consistency with these beliefs does not speak for a philosophical position, and conflict does not speak against it. Debates about the nature of perception should not appeal to ‘common sense’. We now develop this challenge by examining what we take to be the most straightforward, and most likely, responses.

To preserve the ability to appeal to common sense, philosophers might seek to recover, from among competing belief fragments, a nugget of coherent ‘true’ common sense, restricted to pre-scientific beliefs, while setting aside folk beliefs ‘tainted’ by science (e.g., Snowdon, 1990, pp. 121–122). This *restrictionist response* could be motivated by the thought that the common-sense beliefs enjoying epistemic default status are part of the ahistorical ‘central core of human thinking’ (Strawson, 1959, p. 10) which it is therefore theoretically undesirable, or perhaps even practically impossible, to relinquish (e.g., Allen, 2020; Brewer, 2011; Strawson, 1979). The response would also make sense, for example, where folk beliefs have assimilated scientific concepts or findings only partially or incorrectly, and thus offend against both ‘true’ common sense and proper science, or where folk beliefs have assimilated concepts or ideas of an immature or otherwise defective science (like Freudian psychoanalysis). Standard views of Direct and Indirect Realism invite this response: Philosophers of perception typically regard Direct Realism as the untutored folk conception of vision (Sect. 1.1), and Indirect Realism as a position acquired through philosophical reflection or exposure to science (Brewer, 2011; Chalmers, 2006; Crane & French, 2021; Smith, 2002). According to a response along these lines, Direct Realism captures ‘true’ common sense, whereas Indirect Realism captures folk beliefs that half-assimilate science and true common sense may well deserve epistemic default status, even when held simultaneously with folk beliefs which do not.

Against this first response, we suggest that both folk Direct Realism *and* folk Indirect Realism are formed prior to philosophical reflection or scientific instruction. We observe that the Cartesian Theatre conception is built around a pre-scientific theory of sense-perception: Input from the sense-organs is processed in the brain, where it results in a conscious experience that prototypically involves seeing a mental image (in an inner ‘theatre’). This act of inner perception, and only this act, requires input from the sense-organs to come together in a single location, behind the eyes, and requires the conscious experience to be the effect of earlier neural processing (Dennett, 1991). Folk Indirect Realism elaborates this pre-scientific theory of sense-perception. But it does so without recourse to new, scientific concepts—it rather assimilates the relationship between the subject and the output of cerebral processing to the familiar relationship between a viewer and an object of sight. Folk Indirect Realism is no more scientifically informed than the Cartesian Theatre conception.

This conception is widely held by laypeople: Studies using drawing tasks and verbal agreement ratings have found evidence of widespread acceptance of the conception’s ‘materialist substrate’: Laypeople locate consciousness (but not unrelated neurological

processes or unconscious thinking) in a single, confined area in the prefrontal cortex (behind the eyes), and they take conscious experience to occur only after (and arguably, as a result of) complete neural processing of sensory input (Forstmann & Burgmer, 2022; cf. Bertossa et al., 2008, for evidence from interviews). One study, for instance, probed for beliefs about where and when ‘the brain creates the subjective, conscious experience of seeing a tree’ by combining verbal statements with illustrations (tree with blurry edges) that suggested this experience involved seeing a (mental) image (Forstmann & Burgmer, 2022, Study 4). The widespread acceptance of ‘Cartesian’ beliefs observed in that study is consistent with the widespread agreement our study found for ‘Cartesian’ images (I1–I3).

Webb and Graziano (2015) suggested the Cartesian Theatre conception is an artefact of a grossly simplified implicit model of attention that is used to track others’ and own focus of visual attention (cf.). Alternatively, the conception could be culturally learned prior to exposure to science, as a critical review of extant findings has suggested for dualist beliefs (Barlev & Shtulman, 2021). Either way, the Cartesian Theatre conception is a naïve theory that predates the acquisition of topically relevant scientific information and may influence its assimilation. As noted, folk Indirect Realism elaborates this naïve theory by recourse to familiar, rather than scientific concepts. We conclude that both folk Direct Realism *and* folk Indirect Realism are sets of genuinely pre-scientific beliefs, organised in different belief fragments that are internally coherent but mutually incompatible. The restrictionist response seems unable to cope with the fragmentation challenge.

Many, perhaps most, philosophers of perception who appeal to common sense will (also) give another, ‘phenomenological’, response to the challenge. These philosophers tend to recognise two distinct but related sources of ‘common sense’ about vision, namely folk beliefs about vision and the phenomenological character of visual experiences (what it is like to perceive things visually from a first-person perspective). The latter is taken to inform those beliefs. For example, Strawson famously moves from the phenomenological claim that ‘mature sensible experience (in general) presents itself as ... an *immediate* consciousness of the existence of things outside us’ to a conclusion about ‘the common realist conception of the world’, namely, that it ‘does not have the character of a ‘theory’... [that ‘goes beyond’] the ‘data of sense’” (1979, p. 97). A similar ‘oscillation’ between claims about common-sense beliefs and the phenomenology of ordinary visual experience pervades discussions of contemporary Naïve Realism (Raineri, 2021): This theory is sometimes presented as capturing folk beliefs by providing ‘the best philosophical articulation of what we all pre-theoretically accept concerning the nature of our sense-experience’ (Martin, 2006, p. 404; cf. Crane, 2006, p. 133; Genone, 2016). Simultaneously, it is also presented as the view that ‘best articulates how sensory experience seems to us to be just through reflection’ (Martin, 2006, p. 354), namely, ‘first-personal reflection on its character’ (Crane & French, 2021; §1.1). The implicit assumption is that our pre-scientific folk beliefs about perception are shaped by how sensory experience appears to us from a first-person perspective.

The *phenomenological response* to the fragmentation challenge severs this link between folk beliefs about vision and the phenomenology of vision, in the hope of

thereby being able to retain epistemic default status for some aspect of how our experiences ‘strike’ us. This response accepts that because of belief fragmentation, folk beliefs deserve no epistemic default status. But it insists that phenomenology retains this status. For this to be possible, noticing how sensory experience appears may require careful attention, or in addition the use of ‘conceptual tweezers aided and abetted by argumentation’ (O’Shaughnessy, 2000, p. 452). Indeed, folk beliefs may even fundamentally misrepresent how sensory experience seems to us, as is sometimes claimed in the phenomenological tradition (e.g., Merleau-Ponty, 1945). According to the phenomenological response, the ultimate evidence for philosophical theories of perception is careful, possibly expert, reflection on the phenomenological character of visual experience (*cf.* Fish, 2009, pp. 18–23).

The present study thus moves the debate to a *purely* phenomenological perspective that severs the commonly assumed link between visual phenomenology and folk beliefs about vision. But the study does not speak to this phenomenological response: Responses to our verbal items and images need not reflect judgments about the phenomenological character of experience: Participants in our study were not primed to attend to the phenomenological character of visual experience, nor prompted to answer questions in light of how their experiences seem to them, as opposed to on the basis of other beliefs or sources of information about experience. However, the existence of disagreements in introspective judgements, even on the basis of careful reflection (e.g., Allen, 2016, Ch. 7; Schwitzgebel, 2019), suggests it is a live possibility that the fragmentation challenge will arise also for beliefs formed on the basis of introspection and reflection on its deliverances. The extent to which this phenomenological basis can provide outputs with default epistemic status requires empirical investigation into phenomenological judgments about experience (*cf.* Allen et al., 2022). The success or otherwise of the phenomenological response is a matter for further empirical investigation, but this new topic lies beyond the scope of the present paper.

A more radical response to the fragmentation challenge is the *naturalistic response* which maintains that appeals to common-sense beliefs have no role to play in motivating and adjudicating between competing philosophical theories of perception. Methodological naturalists in philosophy have long suggested that philosophical theories of natural phenomena like vision should be built on the best available scientific theories, which will help us understand, e.g., how vision works, what it is, and why visual experience is the way it is (e.g., Burge, 2010; Drayson, 2021; reviews: Kornblith, 2016; Logue & Richardson, 2021). A common criticism charges that naturalistic philosophical theories fail to account for some aspect of our common-sense conception of the world (in different contexts, see, e.g., Allen (2016) on colour, Chalmers (1996) on consciousness, and Kim (1988) on naturalized epistemology). Findings of fragmentation in folk belief provide a new and principled reason for naturalistic approaches to ignore common-sense beliefs in philosophical theorising: there is no such thing as ‘the’ common-sense conception of the phenomena of interest, and fragmented and conflicted beliefs about it deserve no epistemic privilege.

However, the fragmentation challenge does not merely support naturalistic approaches over competitors, it also has implications for naturalistic theorising itself. Naturalistic theories of perception cannot simply be ‘read off’ from our best scientific

theories, but require interpretation, analysis, and argument. The fragmentation challenge constrains how naturalistic theories can be developed. In particular, it rules out appeals to common sense in deciding between interpretations of scientific theories. To illustrate, consider the debate about whether contemporary predictive coding accounts of perception are forms of Direct or Indirect Realism (e.g., Clark, 2013; Drayson, 2018; Hohwy, 2007). According to predictive coding accounts, the brain makes predictions about the sensory information it receives on the basis of prior knowledge about the world. Hohwy argues that the predictive coding account entails the ‘unfashionable’ view that perception is indirect, because what we perceive is just ‘the brain’s best hypothesis’ (2007, p. 322). Clark, by contrast, argues that the predictive account is at least ‘not-indirect perception’ (2013, p. 493). But why should we prefer one interpretation of the psychological theory over the other? According to Clark, the predictive account ‘delivers a genuine form—perhaps the only genuine form that is naturally possible—of “openness to the world”’ (2013, p. 492). This assumes that delivering a form of ‘openness to the world’ is theoretically desirable. However, the fragmentation challenge implies that this assumption cannot be motivated by reference to common sense—one of the most common motivations for it. Common sense does not provide a reason to prefer a Direct (or not-indirect) over an Indirect Realist interpretation of the predictive coding account. Similar points apply to discussions of naturalistic representationalist theories of perception, particularly where the view that perceptual experiences represent objects in our environment is motivated on the grounds that representationalism better accords with intuitive beliefs about vision than Indirect Realism (sense-datum and qualia theories) (e.g., Harman, 1990, p. 39; Tye, 2000, pp. 46–47).

In summary, if common-sense beliefs are fragmented, then naturalistically inclined philosophers of perception cannot appeal to tropes from traditional forms of philosophical theorising about which views are more intuitive or better align with common sense in developing and defending their theories. As such, the fragmentation challenge presents a much deeper challenge to methodology in philosophical discussions of perception than it might at first appear.

### 3.3 Limitations and future directions

In our study, an agreement rating task with verbal and pictorial stimuli provided initial evidence of a conflict between directly opposing Direct Realist and Indirect Realist beliefs about vision. Further evidence can be garnered with other tasks (*cf.* Sect. 1.2). Further research is required to definitely exclude the alternative hypothesis that laypeople simply do not have any relevant beliefs about vision, and so gave ad hoc responses that are mere task artefacts. Some present and previous findings speak against this alternative hypothesis: In the present study manipulation of the order of response options produced suggestive order effects only for one item category (Generic Indirect), and open-text responses were consistent with prior image ratings (see Sect. 3.1). In previous studies (Bertossa et al., 2008; Forstmann & Burgmer, 2022), focus group interviews, agreement ratings, and drawing tasks provided evidence of belief in the Cartesian Theatre conception underpinning Indirect Realism (Sect. 3.2). Even so, further research

is necessary to examine whether responses are reflective of stable beliefs and to what extent these beliefs are ‘merely verbal’ or actually guide thought, being deployed in reasoning and problem-solving tasks (*cf.* Schwitzgebel, 2021, pp. 364–365). These questions can be examined by considering the test–retest reliability of a questionnaire like ours (administering the same questions to the same participants, some weeks later, to see whether individuals give the same responses) and by investigating whether judgments about pertinent items influence, e.g., the acceptance of sceptical arguments and reasoning about cases of hallucination (where endorsement of folk Direct Realism and folk Indirect Realism predict different responses). This research is currently under way.

Further research is also required to examine the persistence of the conflict between folk Direct Realism and folk Indirect Realism, *viz.*, to what extent the beliefs examined are reflective or susceptible to change upon reflection, with and without new informational input. Moreover, the relative prevalence of the conflicting conceptions of vision may vary between cultures and remains to be examined by cross-cultural studies. For example, a more holistic thinking style, associated with East Asian cultures (Nisbett, 2003), may favour Direct Realist beliefs. Finally, future research on folk beliefs about vision will need to address for the belief fragments of interest the perhaps biggest question confronting the fragmentation hypothesis: when does which fragment get activated, when does which fragment get suppressed, and when does which fragment get deployed in cognitive tasks? How do, e.g., stimulus format, contextual cues, nature of the task, and the domain considered (e.g., the perceiver’s environment vs their psychology) influence whether Direct Realist or Indirect Realist beliefs—or neither—guide our thinking? A better understanding of belief fragmentation is not only interesting in its own right but will also let us appreciate in more detail the challenge it poses to traditional forms of philosophical theorising.

While it has limitations, the present study thus motivates an ambitious research programme. Major philosophical consequences would arise, however, even on the most plausible alternative interpretation of our findings, which traces them back to the lack of stable folk beliefs in line with Direct Realism and Indirect Realism, respectively: Even in this case, philosophical debates about the nature of perception cannot appeal to folk beliefs—for want of anything to appeal to. For the immediate philosophical application, lack of pertinent beliefs would have the same fundamental upshot as belief fragmentation.

## 4 Conclusion

Like many other beliefs, folk beliefs about vision are fragmented and conflicted. Our survey provides first evidence of a conflict between Indirect Realist beliefs and Direct Realist beliefs that reject them. These conflicting folk theories are endorsed roughly equally widely, and are endorsed not merely by different members of the same community but frequently by the same individuals who appear to be torn between them. These findings complement previous findings (about the coexistence of intromissionist and extramissionist conceptions of vision) which suggest that folk beliefs about vision are fragmented. We argued that the conflict exposed by the present study occurs at the

level of pre-scientific beliefs. On an alternative interpretation (against which we provide initial evidence), present findings are reflective of lack of pertinent folk beliefs. Either way, the findings pose a radical challenge to the traditional philosophical practice of appealing to ‘the’ common-sense conception of vision to support theorising about perception: Present findings suggest there simply is no such—one, coherent, dominant—conception philosophers could appeal to, or seek to defend or explain, or even draw on to formulate desiderata for naturalistic theories and to help decide between different philosophical interpretations of scientific theories. In other words, these findings motivate giving up the evidentiary practice common in philosophical debates about the nature of perception. Many further philosophical debates assume the existence of coherent and dominant common-sense conceptions of phenomena of interest, and credit them with epistemic default status. Belief fragmentation is therefore likely to have much wider methodological consequences in philosophy.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s11229-023-04066-w>.

**Acknowledgements** For helpful comments on previous drafts of this paper, we would like to thank two anonymous referees for this journal as well as Ruth Lee, Edouard Machery, and audiences at the 2021 conference of the European Society of Philosophy and Psychology, the 1st European Experimental Philosophy Conference (2021), and the work-in-progress workshop at the Center for Philosophy of Science, University of Pittsburgh. For assistance with the follow-up study, we are indebted to Dimitra Lazaridou-Chatzigoga.

**Funding** No external funding was received for conducting this research. The main study reported in this paper was supported by internal funding from the University of York. The follow-up study was supported by internal funding from the University of East Anglia.

## Declarations

**Conflict of interest** The authors have no competing interests to declare that are relevant to the content of this article.

**Ethical approval** The use of human research participants was approved by the Arts and Humanities Research Ethics Committee of the University of York. The research conformed to the ethical standards for conducting research as outlined by the British Psychological Society.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adams, Z., & Hansen, N. (2020). The myth of the common-sense conception of colour. In T. Marques & A. Wikforss (Eds.), *Shifting concepts: The philosophy and psychology of conceptual variability* (pp. 106–127). Oxford University Press.
- Allen, K. (2016). *A naïve realist theory of colour*. Oxford University Press.
- Allen, K. (2020). The value of perception. *Philosophy and Phenomenological Research*, 100(633), 656.
- Allen, K., Quinlan, P., Andow, J., & Fischer, E. (2022). What is it like to be colour-blind? A case study in experimental philosophy of experience. *Mind and Language*, 37, 814–839.
- Astuti, R., & Harris, P. L. (2008). Understanding mortality and the life of the ancestors in rural Madagascar. *Cognitive Science*, 32, 713–740.
- Ayer, A. J. (1973). *The central questions of philosophy*. Weidenfeld.
- Barlev, M., Mermelstein, S., & German, T. C. (2017). Core intuitions about persons coexist and interfere with acquired Christian beliefs about God. *Cognitive Science*, 41, 425–454.
- Barlev, M., Mermelstein, S., & German, T. C. (2018). Representational co-existence in the God concept. *Psychonomic Bulletin & Review*, 25, 2330–2338.
- Barlev, M., & Shtulman, A. (2021). Minds, bodies, spirits, and gods: Does widespread belief in disembodied beings imply that we are inherent dualists? *Psychological Review*, 128, 1007–1021.
- Barrett, J. L., & Keil, F. C. (1996). Conceptualizing a nonnatural entity: Anthropomorphism in God concepts. *Cognitive Psychology*, 31, 219–247.
- Bendaña, J., & Mandelbaum, E. (2021). The fragmentation of belief. In C. Borgoni, D. Kindermann, & A. Onofri (Eds.), *The fragmented mind* (pp. 78–107). Oxford University Press.
- Berņiūnas, R., Beinorius, A., Dranseika, V., Silius, V., & Rimkevičius, P. (2021). The weirdness of belief in free will. *Consciousness and Cognition*, 87, 103054.
- Bertossa, F., Besa, M., Ferrari, R., & Ferri, F. (2008). Point zero: A phenomenological inquiry into the seat of consciousness. *Perceptual and Motor Skills*, 107, 323–335.
- Boghossian, P., & Velleman, D. (1989). Color as a secondary quality. *Mind*, 98, 81–103.
- Borgoni, C., Kindermann, D., & Onofri, A. (Eds.). (2021). *The fragmented mind*. Oxford University Press.
- Brewer, B. (2011). *Perception and its objects*. Oxford University Press.
- Burge, T. (2010). *Origins of objectivity*. Clarendon.
- Callender, C. (2017). *What makes time special?* Oxford University Press.
- Carey, S. (2009). *The origin of concepts*. Oxford University Press.
- Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioral Research*, 1(2), 245–276.
- Chalmers, D. (1996). *The conscious mind*. Oxford University Press.
- Chalmers, D. (2006). Perception and the fall from Eden. In T. S. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 49–67). Oxford University Press.
- Chi, M. T. H. (2005). Common-sense conceptions of emergent processes: Why some misconceptions are robust. *Journal of the Learning Sciences*, 14, 161–199.
- Cialdini, R. B., Trost, M. R., & Newsom, J. T. (1995). Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications. *Journal of Personality and Social Psychology*, 69, 318–328.
- Clark, A. (2013). Expecting the world: Perception, prediction, and the origins of human knowledge. *Journal of Philosophy*, 110(9), 469–496.
- Cottrell, J. E., & Winer, G. A. (1994). Development in the understanding of perception: The decline of extramission perception beliefs. *Developmental Psychology*, 30, 218–228.
- Crane, T. (2006). Is there a perceptual relation? In T. S. Gendler & J. Hawthorne (Eds.), *Perceptual experiences* (pp. 126–146). Oxford University Press.
- Crane, T., & French, C. (2021). The problem of perception. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. <https://plato.stanford.edu/archives/fall2021/entries/perception-problem>
- Daly, C. (2010). *An introduction to philosophical methods*. Broadview.
- Dennett, D. C. (1991). *Consciousness explained*. Little, Brown.
- Ding, C., & He, X. (2004). Cluster structure of k-means clustering via principal component analysis. In: *Proceedings, 8th Pacific-Asia conference on advances in knowledge discovery and data mining (PAKDD 2004)* (pp. 414–418).
- Drayson, Z. (2018). Direct perception and the predictive mind. *Philosophical Studies*, 175, 3145–3164.

- Drayson, Z. (2021). Naturalism and the metaphysics of perception. In H. Logue & L. Richardson (Eds.), *Purpose and procedure in the philosophy of perception*. Oxford University Press.
- Dundes, A. (1981). *The evil eye: A folklore casebook*. Garland.
- Fischer, E., Engelhardt, P. E., Horvath, J., & Ohtani, H. (2021). Experimental ordinary language philosophy. *Synthese*, 198, 1029–1070.
- Fish, W. (2009). *Perception, hallucination, and illusion*. Oxford University Press.
- Fodor, F. (1987). *Psychosemantics*. MIT Press.
- Forstmann, M., & Burgmer, P. (2022). The Cartesian folk theater: People conceptualize consciousness as a spatio-temporally localized process in the human brain. *Journal of Experimental Psychology: General*, 151(4), 781–803.
- Gawronski, B., & Strack, F. (Eds.). (2012). *Cognitive consistency: A fundamental principle in social cognition*. Guilford Press.
- Gelman, S. A., & Noles, N. S. (2011). Domains and naïve theories. *Wires Cognitive Science*, 2(5), 490–502.
- Genone, J. (2016). Recent work on naïve realism. *American Philosophical Quarterly*, 53(1), 1–26.
- Goldberg, R. F., & Thompson-Schill, S. L. (2009). Developmental “roots” in mature biological knowledge. *Psychological Science*, 20, 480–487.
- Graziano, M. S. A., Guterstam, A., Bio, B. J., & Wilterson, A. I. (2020). Toward a standard model of consciousness: Reconciling the attention schema, global workspace, higher-order thought, and illusionist theories. *Cognitive Neuropsychology*, 37(3–4), 155–172.
- Gregg, V. R., Winer, G. A., Cottrell, J. E., Hedman, K. E., & Fournier, J. S. (2001). The persistence of a misconception about vision after educational interventions. *Psychonomic Bulletin and Review*, 8, 622–626.
- Guterstam, A., Kean, H. H., Webb, T. W., Kean, F. S., & Graziano, M. S. A. (2019). Implicit model of other people’s visual attention as an invisible, force-carrying beam projecting from the eyes. *Proceedings of the National Academy of Sciences of the United States of America*, 116, 328–333.
- Guterstam, A., Wilterson, A. I., Wachtell, D., & Graziano, M. S. A. (2020). Other people’s gaze encoded as implied motion in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 117, 13162–13167.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis: Pearson new international edition*. Pearson.
- Harman, G. (1990). The intrinsic quality of experience. *Philosophical Perspectives*, 4, 31–52.
- Hohwy, J. (2007). Functional integration and the mind. *Synthese*, 159(3), 315–328.
- Hume, D. (1777/1975). *Enquiries concerning human understanding and concerning the principles of morals*. Oxford University Press.
- Jackman, M. R. (1973). Education and prejudice or education and response-set? *American Sociological Review*, 38(3), 327–339.
- Jackson, F. (1977). *Perception. A representative theory*. Cambridge University Press.
- Jackson, F. (1998). *From metaphysics to ethics: A defense of conceptual analysis*. Oxford University Press.
- Kahneman, D., & Frederick, S. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 49–81). Cambridge University Press.
- Kelemen, D., Rottman, J., & Seston, R. (2013). Professional physical scientists display tenacious teleological tendencies: Purpose-based reasoning as a cognitive default. *Journal of Experimental Psychology: General*, 142, 1074–1083.
- Kim, J. (1988). What is naturalized epistemology? *Philosophical Perspectives*, 2, 381–405.
- Kornblith, H. (2016). Philosophical naturalism. In H. Cappelen, T. Szabo Gendler, & J. Hawthorne (Eds.), *Oxford handbook of philosophical methodology* (pp. 147–158). Oxford University Press.
- Kruglanski, A. W., Jasko, K., Milyavsky, M., Chernikova, M., Webber, D., Pierro, A., & Di Santo, D. (2018). Cognitive consistency theory in social psychology: A paradigm reconsidered. *Psychological Inquiry*, 29(2), 45–59.
- Latham, A. J., Miller, K., & Norton, J. (2021). Is our naïve theory of time dynamical? *Synthese*, 198, 4251–4271.
- Lee, R., Shardlow, J., Hoerl, C., O’Connor, P. A., Fernandes, A. S., & McCormack, T. (2022). Towards an intuitive theory of time. *Cognitive Science*, 46(7), e13166. <https://doi.org/10.1111/cogs.13166>
- Legare, C. H., Evans, E. M., Rosengren, K. S., & Harris, P. L. (2012). The coexistence of natural and supernatural explanations across cultures and development. *Child Development*, 83, 779–793.

- Legare, C. H., & Gelman, S. A. (2008). Bewitchment, biology, or both: The co-existence of natural and supernatural explanatory frameworks across development. *Cognitive Science*, 32, 607–642.
- Leiser, D. (2001). Scattered naïve theories: Why the human mind is isomorphic to internet web. *New Ideas in Psychology*, 19, 175–202.
- Lewis, D. (1997). Naming the colors. *Australasian Journal of Philosophy*, 75(3), 325–342.
- Lindberg, D. C. (1976). *Theories of vision from Al-Kindi to Kepler*. University of Chicago Press.
- Logue, H., & Richardson, L. (Eds.) (2021). *Purpose and procedure in the philosophy of perception*. Oxford University Press.
- Lyons, J. (2017). Epistemological problems of perception. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Spring 2017. <https://plato.stanford.edu/archives/spr2017/entries/perception-episprob/>
- Machery, E. (2017). *Philosophy within its proper bounds*. Oxford University Press.
- Mandelbaum, E. (2014). Thinking is believing. *Inquiry*, 57, 55–96.
- Martin, M. G. F. (2002). The transparency of experience. *Mind and Language*, 4, 376–425.
- Martin, M. G. F. (2006). On being alienated. In T. S. Gendler & J. Hawthorne (Eds.), *Perceptual experience* (pp. 354–410). Clarendon Press.
- Masson, S., Potvin, P., Riopel, M., & Foisy, L. M. B. (2014). Differences in brain activation between novices and experts in science during a task involving a common misconception in electricity. *Mind, Brain, and Education*, 8, 44–55.
- McCloskey, M. (1983). Naïve theories of motion. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 299–324). Erlbaum.
- Merleau-Ponty, M. (1945). *Phenomenology of Perception*. Transl. D. Landes. London: Routledge, 2013.
- Meyering, T. C. (1989). *Historical roots of cognitive science: The rise of a cognitive theory of perception from antiquity to the nineteenth century*. Kluwer.
- Moore, G. E. (1957). Visual sense data. Reprinted in R. Swartz (Ed.), *Perceiving, sensing, and knowing*. Anchor, 1965.
- Nisbett, R. (2003). *The geography of thought: How Asians and Westerners think differently... and why*. Free Press.
- Nudds, M. (2009). Recent work in perception: Naïve realism and its opponents. *Analysis Reviews*, 69(2), 334–346.
- O'Connor, T., & Franklin, C. (2021). Free will. In E. N. Zalta (Ed.), *Stanford encyclopedia of philosophy*. Spring 2021. <https://plato.stanford.edu/archives/spr2021/entries/freewill/>
- O'Shaughnessy, B. (2000). *Consciousness and the world*. Oxford University Press.
- Oberauer, K., Farrell, S., Jarrold, C., & Lewandowsky, S. (2016). What limits working memory capacity? *Psychological Bulletin*, 142(7), 758–799.
- Porot, N., & Mandelbaum, E. (2021). The science of belief: A progress report. *WIREs Cognitive Science*, 12, e1539. <https://doi.org/10.1002/wcs.1539>
- Price, H. H. (1932). *Perception* (2nd ed.). Reprinted in 1961. Methuen.
- Raineri, C. (2021). What's so naïve about naïve realism? *Philosophical Studies*, 178, 3637–3657.
- Reid, T. (1785/1969). *Essays on the intellectual powers of man*. MIT Press.
- Reiner, M., Slotta, J. D., Chi, M. T. H., & Resnick, L. B. (2000). Naïve physics: A commitment to substance-based conceptions. *Cognition and Instruction*, 18, 1–34.
- Robinson, H. (1994). *Perception*. Routledge.
- Rozenblit, L., & Keil, F. (2002). The misunderstood limits of folk science: An illusion of explanatory depth. *Cognitive Science*, 26, 521–562.
- Russell, B. (1912/2000). *The problems of philosophy*. Oxford University Press.
- Russell, B. (1962). *An inquiry into meaning and truth*. Pelican.
- Scholl, B. (2007). Object persistence in philosophy and psychology. *Mind and Language*, 22, 563–591.
- Schwitzgebel, E. (2019). Introspection. In E. N. Zalta (Ed.), *The Stanford encyclopedia of philosophy*. Winter 2019. <https://plato.stanford.edu/archives/win2019/entries/introspection/>
- Schwitzgebel, E. (2021). The pragmatic metaphysics of belief. In C. Borgoni, D. Kindermann, & A. Onofri (Eds.), *The fragmented mind* (pp. 350–375). Oxford University Press.
- Searle, J. (2015). *Seeing things as they are: A theory of perception*. Oxford University Press.
- Shtulman, A. (2017). *Scienceblind. Why our intuitive theories about the world are so often wrong*. Basic Books.
- Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition*, 124, 209–215.
- Smith, A. D. (2002). *The problem of perception*. Harvard University Press.

- Snowdon, P. (1981). Perception, vision, and causation. *Proceedings of the Aristotelian Society*, 81, 175–192.
- Snowdon, P. F. (1990). The objects of perceptual experience. *Proceedings of the Aristotelian Society*, 64, 121–150.
- Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences* (5th ed.). Routledge.
- Stich, S., & Tobia, K. (2016). Experimental philosophy and the philosophical tradition. In J. Sytsma & W. Buckwalter (Eds.), *Blackwell companion to experimental philosophy* (pp. 5–21). Wiley Blackwell.
- Strawson, P. F. (1959). *Individuals*. Methuen.
- Strawson, P. F. (1979). Perception and its objects. In G. Macdonald (Ed.), *Perception and identity*. Palgrave.
- Thagard, P. (2014). Explanatory identities and conceptual change. *Science & Education*, 23, 1531–1548.
- Tversky, A., & Kahneman, D. (1982). Judgments of and by representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 84–98). Cambridge University Press.
- Tye, M. (2000). *Consciousness, color, and content*. MIT Press.
- Vosniadou, S. (1994). Capturing and modelling the process of conceptual change. *Learning and Instruction*, 4, 45–69.
- Watson-Jones, R. E., Busch, J. T. A., Harris, P. L., & Legare, C. H. (2017). Does the body survive death? Cultural variation in beliefs about life everlasting. *Cognitive Science*, 41, 455–476.
- Webb, T. W., & Graziano, M. S. A. (2015). The attention schema theory: A mechanistic account of subjective awareness. *Frontiers in Psychology*. <https://doi.org/10.3389/fpsyg.2015.00500>
- Wilson, M. (2006). *Wandering significance: An essay on conceptual behavior*. Oxford University Press.
- Winer, G. A., Cottrell, J. E., Gregg, V., Fournier, J. S., & Bica, L. A. (2002). Fundamentally misunderstanding visual perception. Adults' belief in visual emissions. *American Psychologist*, 57, 417–424.
- Winer, G. A., Cottrell, J. E., Karefilaki, K., & Chronister, M. (1996a). Conditions affecting beliefs about visual perception among children and adults. *Journal of Experimental Child Psychology*, 61, 93–115.
- Winer, G. A., Cottrell, J. E., Karefilaki, K., & Gregg, V. R. (1996b). Images, words and questions: Variables that influence beliefs about vision in children and adults. *Journal of Experimental Child Psychology*, 63, 499–525.
- Wiser, M. F., & Carey, S. (1983). When heat and temperature were one. In D. Gentner & A. L. Stevens (Eds.), *Mental models*. Erlbaum.
- Wu, J. (2012). *Advances in K-means clustering*. Springer.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.