Accepted (Sept 8, 2025) by *Philosophy and Phenomenological Research*

SCIENTIFIC OR NAÏVE?
PERCEPTIONS OF DIRECT AND INDIRECT REALISM, AND WHY THEY MATTER

Eugen Fischer[1], Keith Allen[2], and Paul E. Engelhardt[3]

[1] School of Politics, Philosophy, and Area Studies, University of East Anglia, Norwich, UK

[2] Department of Philosophy, University of York, York, UK

[3] School of Psychology, University of East Anglia, Norwich, UK

Philosophical debates about the nature of perception are standardly informed by an empirical assumption about folk beliefs: They assume there is such a thing as 'the' common-sense conception of vision, and that this conception is captured by Direct Realism. This naïve theory is thought to compete with scientifically informed Indirect Realism. This paper discusses how to render these claims empirically tractable and reports a scientific accuracy rating study whose findings suggest instead that both Direct Realism and Indirect Realism capture pre-scientific conceptions. We discuss two philosophical consequences: 'Common sense' is too conflicted about vision to deserve epistemic default status in philosophical debates about perception, and the 'problem of perception' needs to be reconceptualised as arising not from a challenge to our ordinary understanding of vision, but from a patent conflict within this understanding. Together, these two upshots facilitate a fresh approach to the problem which can usefully be transferred to further 'aporetic' problems that arise from conflicts between beliefs to which we are pre-theoretically attracted.

Naïve and Direct Realism, Indirect Realism, problem of perception, folk conceptions of vision, experimental philosophy

Experimental philosophy seeks to make philosophical problems more tractable by bringing empirical evidence to bear on them. One of its signature innovations is the identification of previously overlooked evidentiary links between philosophically relevant claims and empirical findings. This paper will identify previously unnoticed evidentiary connections running from laypeople's perception of the scientific accuracy of competing conceptions of vision, on the one hand, to a key assumption shaping philosophical debates about the nature of perception, on the other. It will report an experiment on the perceived scientific accuracy of Direct Realism and Indirect Realism, and will deploy findings to address the longstanding 'problem of perception' (Brewer, 2011; Crane & French, 2021; Robinson, 1994; Smith, 2002).

1. *Introduction: Conflicting common-sense conceptions?*

Philosophical debates about the nature of perception typically take for granted that there is such a thing as 'the' naïve or common-sense view of perception. At any rate, it is generally assumed that there is such a thing as 'the' common-sense view of vision, the most prominently discussed of the five senses. This naïve view is taken to be consistent with *Direct Realism (DR),* which maintains that physical objects are 'directly' present in perception (Lyons, 2023): we do *not* see physical objects *in virtue of* being aware of mental images or other things that are distinct from them (*'Perceptual Directness'*); the presence of physical objects is *not inferred* from the

presence of anything else (mental images, etc.) (*'Phenomenal Directness'*); and when you look at a physical object like a tomato you typically know what it is (e.g., that it is a tomato), and not just what it looks like ('*Epistemic Directness'*).

This supposedly naïve view is generally taken to be challenged by a scientifically informed *Indirect Realism (IR)*. According to IR, we see physical objects, like tables and trees ('Realism'), but do so only indirectly, *by* or *in virtue of* being aware of things that are distinct from them, such as mental images, 'ideas', or 'sense-data' that the objects we look at cause in us ('*Perceptual Indirectness*'; e.g., Jackson, 1977; Robinson, 1994; Russell, 1912/2000; *cf.* Locke 1700/ 1975). This challenge to Perceptual Directness may be accompanied by rejections of the other two directness claims: because we perceive physical objects indirectly, by being aware of mental images, we must infer the presence of the physical object from that of the mental image ('*Phenomenal Indirectness*'), and can, strictly speaking, get to know by looking at the object only what the mental image makes available to us, namely, what the object looks like to us (*'Epistemic Indirectness*'). Some philosophical Indirect Realists have endorsed all three claims (e.g., Russell, 1912/2000), others have been more selective (e.g., Jackson, 1977; see Lyons, 2023 for discussion). Most think that their position is at least in some conflict with common sense.

In brief, philosophical debates about the nature of perception are informed by this empirical assumption:

> (A) There is such a thing as 'the' (one, coherent) common-sense conception of vision, and this conception is captured by DR and inconsistent with IR.

This assumption shapes (1) the methodology of those debates and (2) their central problem:

(1) Many contributors to these debates accord the common-sense conception of vision epistemic default status (or even stronger epistemic privilege): They think that, in virtue of capturing this conception, DR should be accepted in the absence of good reasons to the contrary (e.g., Allen, 2020; Genone, 2016; Martin, 2002; Reid, 1785/1969; Searle, 2015; Strawson, 1979), and that, since it conflicts with this conception, IR should be maintained only together with an error theory that explains how common sense could go wrong (e.g., Hume, 1777/1975; Russell, 1912/2000; Boghossian & Velleman, 1989).

(2) As standardly conceived, the 'problem of perception' is developed by arguments that proceed from the uncontroversial assumptions that illusions and hallucinations are possible, to IR conclusions that appear to rule out supposedly common-sense DR: 'the problem of perception is that if illusions and hallucinations are possible, then perception, *as we ordinarily understand it*, is impossible' (Crane & French 2021, §1, our emphasis).

This formulation of the problem presupposes that, as per (A), 'the' common-sense conception of vision is consistent with DR, so that the pertinent arguments 'from illusion' and 'from hallucination' bring out a conflict between this common-sense conception and common-sense beliefs captured by their uncontroversial premises (Brewer, 2011; Martin, 2002; Robinson, 1994; Smith, 2002). This conflict generates an 'aporetic' problem (from the Greek 'aporia' or 'impasse').

The present paper will critically examine the philosophically consequential empirical assumption (A). There is considerable evidence that laypeople are familiar with, and endorse, conflicting conceptions of vision. According to intromissionist conceptions, vision involves light rays entering into our eyes. According to extramissionist conceptions, vision involves force rays leaving our eyes. The prevalence of these potentially conflicting conceptions among

laypeople of different age groups has been studied through tasks including agreement rating and drawing tasks (review: Winer et al., 2002), and with sophisticated behavioural measures (Guterstam et al., 2019). These studies found that both conceptions are maintained by laypeople, including children and adults, and sometimes by the same individuals. These findings suggest that laypeople may not have one coherent conception of vision, but may be torn between competing conceptions.

Yet more directly relevant to (A), surveys using agreement rating tasks with verbal and pictorial stimuli have revealed that laypeople (screened for scientific innocence) are almost evenly split between endorsement of DR and of IR, with many 'torn souls' feeling compelled to endorse both (Fischer, Allen, & Engelhardt, 2023; Ms). Some of these studies employed a standard validation protocol and concluded that questionnaire responses were due to stable, inferentially deployed information-bearing states, i.e., beliefs (Fischer, Allen, & Engelhardt, Ms). Folk conceptions of perception are typically equated with the stable beliefs of scientifically untrained adults (e.g., Snowdon, 1981, p.176). These studies thus suggest that at least two conflicting conceptions of vision co-exist in the lay mind, which are consistent with DR and IR, respectively.

The present paper follows up this suggestion and examines whether it challenges (A). According to (A), there is only one 'common-sense conception' of vision. This assumption is potentially compatible with the co-existence of conflicting conceptions of vision, namely, if one of these conceptions (say, extramissionism or DR) is naïve, and the other (say, intromissionism or IR) is scientifically informed. By contrast, if both of two conflicting conceptions are not only endorsed by a critical mass of laypeople (as previous studies have shown for DR and IR) but also *have an equal claim to being 'common sense'*, (A) is false for the fundamental reason that there is no such thing as 'the' common-sense conception of vision.

To empirically evaluate assumption (A), we now explicate this assumption as an empirical hypothesis about whether DR and IR are treated as naïve and scientific, respectively, in lay cognition (Section 2). We then report an experiment that uses a scientific accuracy rating task to assess this hypothesis (Sect.3). Finally, we spell out key philosophical consequences (Sect.4): a central method in philosophy of perception is flawed; the problem of perception needs to be reconceptualized; and insight into the methodological flaw, together with the reconceptualization of the problem, jointly facilitate a fresh approach to the problem. This approach may also work for further 'aporetic' problems that arise from belief conflicts.

2. Hypotheses

In philosophical debates about the nature of perception, assumption (A) is typically part-and-parcel of the view that DR is reflective of common sense, while IR is informed by science. The assumption that DR reflects common sense is shared by proponents of 'naïve realist' (e.g. Allen, 2020; Genone, 2016; Martin, 2002) and representationalist forms of DR (e.g., Searle, 2015; Strawson, 1979), and even by those who reject DR in favour of IR (e.g., Hume, 1777/1975; Russell, 1912/2000). IR, by contrast, is often regarded as broadly reflective of a scientific understanding of the causal processes that underlie perception, which has developed since the early modern period (Jackson, 1977; Putnam, 1994; Robinson, 1994; Valberg, 1992; *cf.* Locke 1700/1975). The details of our best scientific theories have of course changed considerably since the 17th century. To what extent IR is accurate by the lights of the currently

best available science is a matter of ongoing debates.[1] But IR is often seen as scientifically informed insofar as it captures a 'causal picture' of vision, according to which perceptual experiences are distinct psychological effects of physical causes, whose nature is determined in part by facts that are internal to perceiving subjects. IR seeks to articulate this scientific understanding in an intuitive manner, conceptualising the relationship between the perceiving subjects and their perceptual experiences in familiar terms, namely, as a form of seeing.

The view that IR is informed by science does not entail that IR is unfamiliar to laypeople, and it is plausible to suppose that the 'causal picture' of vision is part of general knowledge (Raineri, 2021; Valberg, 1992; pp.9-11). Laypeople are familiar with both naïve and scientifically informed conceptions, across several domains, including illness (e.g., 'colds' caused by exposure to low temperatures or by viruses), human origins, and death (religious vs biological views) (Legare et al., 2012; Leiser, 2001; Watson-Jones et al., 2017). 'Common sense' gets continuously updated with theories that seep from academic into general discourse. These scientific theories often undergo some degree of simplification, and perhaps 'vulgarisation', in entering wider cultural consciousness. For example, although it is widely 'known' that water boils at 100°C, this is only true at sea level. Similarly, the familiar Newtonian view that white light contains all colours of light (probed by Shtulman & Valcarcel, 2012, and below) is only true of broadband-spectral white light, but not, for example, of modern LEDs. On the most plausible development of the idea that IR is scientifically informed, IR has followed a similar trajectory from science, via simplification and 'vulgarisation', to the intuitive claim that seeing physical objects involves seeing mental images they cause in us, which over 40% of lay participants endorse in recent surveys (Fischer, Allen, & Engelhardt, 2023; Ms).

While these ideas about the naïve or scientific status of DR and IR, respectively, are part of the common ground of philosophical debates about the nature of perception, they raise the largely ignored question of how we should decide which, if any, of two conflicting conceptions that widely co-exist in laypeople's minds should be deemed 'common sense'. Since common sense gets continuously updated, we do not think that historical provenance provides a useful criterion. We propose to consider instead how conceptions are deployed in lay cognition.

Psychologists have extensively examined the co-existence of exemplary naïve and scientific theories in laypeople's minds (review: Shtulman, 2017). Key studies have employed speeded verification tasks (Barlev, Mermelstein, & German, 2018; Shtulman & Legare, 2020; Shtulman & Valcarcel, 2012): They asked lay participants to make, as quickly as possible, truth-value judgments about statements that are either consistent with both a naïve and a scientific theory, or inconsistent with both, or consistent with one, but not the other. Several studies revealed increased response times and lower proportions of correct responses in the conditions where naïve and scientific theory deliver conflicting verdicts. As commonly interpreted, these findings show that representations of both theories are stored in long-term memory (Barlev, Mermelstein, & German, 2018) and get initially activated upon engagement with the verification task; hence the naïve theory must be effortfully suppressed, to facilitate the scientifically informed response (Goldberg &Thompson-Schill, 2009; Kelemen & Rosset, 2009; Shtulman & Legare, 2020; Shtulman & Valcarcel, 2012).

This common interpretation allows us to distinguish naïve from other theories that are co-represented in a person's long-term memory, by their cognitive role:

[1] These include, for example, 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).

> One of competing co-represented theories *T functions as a naïve theory*, in people's judgment and reasoning, if and only if they effortfully suppress T in favour of an alternative theory T_S they deem more scientifically accurate but in conflict with T, in verification and reasoning tasks.

According to the criterion we propose, a conception of a phenomenon forms part of a society's 'common sense', if it functions in this way as a naïve theory in lay (non-expert) cognition.[2]

We therefore explicate assumption (A) as follows: It implies that only one of the two conceptions of vision it refers to, DR and IR, functions as a naïve theory in lay cognition, and that DR functions thus. As we have just seen, proponents of (A) expect that IR is among the more scientific theories the naïve theory conflicts with, and that the gist of this scientific theory (the 'causal picture' of vision), like Newtonian physics, has seeped into public awareness, at any rate in developed ('WEIRD'; Henrich et al., 2010) countries. Taking these expectations into account, our explication of (A) yields the following empirical hypothesis about current residents of developed countries, who lack advanced (university-level) formal training in natural sciences, psychology, or neuroscience:

> **H_0** Laypeople regard DR as less scientifically accurate than IR, regard the two as conflicting, and suppress the former in favour of the latter, in judgment and reasoning tasks.

Three prior studies have provided evidence that laypeople regard DR and IR as conflicting (Fischer, Allen & Engelhardt, Ms, Studies 1-3).

Our present study will examine the decisive first part of **H_0**:

> **H** Laypeople regard DR statements as less scientifically accurate than IR statements.

This part of **H_0** is decisive, because suppression efforts can occur due to a wide variety of cognitive and motivational factors (Anderson & Green, 2001; Gernsbacher & Faust, 1995). Only the difference in perceived scientific accuracy posited by **H** turns efforts to suppress DR into evidence that DR functions as a naïve theory, in lay cognition. Indeed, if DR functions as a naïve theory and IR as a scientific theory in whose favour DR is suppressed, we would assume that IR is deemed not only more scientifically accurate than DR (as per **H**) but scientifically accurate, *tout court*, whereas DR is not. The cautious component hypothesis **H** is not only decisive, but also *prima facie* plausible: As we have just seen, IR is generally assumed to have seeped from early modern scientific discourse into wider cultural consciousness, and such provenance would predict that laypeople deem this view scientifically accurate, irrespective of whether or not it is borne out by more recondite current science that laypeople do not know.

To evaluate the philosophically consequential assumption (A), we render it empirically tractable as hypothesis **H_0** and assess the decisive and *a priori* plausible part **H** of this hypothesis. More specifically, we will assess **H** against an alternative hypothesis. Dennett (1991) famously argued that scientists and laypeople alike cannot help feeling intuitively attracted to the 'persuasive' imagery of the 'Cartesian Theatre', a location in the brain where information from the different senses is brought together into a single conscious experience. The appeal of 'Cartesian Theatre' conceptions of the mind has been documented through

---

[2] This criterion does not require conscious insight into the origin or acquisition of T: People need not be able to give an informative explanation of why they deem T_S more scientific. We disregard features of the *content* of naïve theories, which are irrelevant for present purposes – e.g., that they give perceptually grounded, object-based explanations of everyday situations that attribute context-invariant properties to objects (Gelman & Noles, 2011; Vosniadou, 1994).

agreement rating and drawing tasks (Forstmann & Burgmer, 2022), and can potentially be explained by reference to an implicit model of attention that models attention in spatial terms (Graziano, 2022). On their most intuitive elaboration, such conceptions posit the existence of an 'inner eye' peering into an inner visual space ('theatre'), and thus render IR intuitive. This suggests the alternative hypothesis that lay belief in IR is not so much a matter of scientific theories seeping into cultural consciousness, but is an artefact of the implicit spatial model of attention (which may of course also influence scientific theorising). We tentatively infer that the conflict between IR and DR might be not one between scientifically informed and naïve theories of vision, but a conflict between two pre-scientific conceptions (whose precise nature we will discuss in Sect.3.4).

Pre-scientific conceptions can be deemed consistent or inconsistent with successful scientific theories. IR statements are suggestive of a 'Cartesian Theatre' conception that posits the existence of inner visual spaces and inner eyes. While this conception is seductive, such statements might strike laypeople as unscientific, upon even brief reflection. By contrast, nothing about DR precludes it *a priori* from being consistent with scientific theories. We therefore suggest this alternative hypothesis:

H* Laypeople regard neither IR nor DR statements as scientifically accurate; if the two differ in perceived scientific accuracy, IR statements will be deemed even less scientifically accurate than DR statements.

This hypothesis implies that the philosophically consequential assumption (A), explicated as empirical hypothesis $\mathbf{H_0}$, is false.

3. Experiment

To evaluate assumption (A), we assessed hypothesis **H** against the alternative hypothesis **H***. To do so, we employed a scientific accuracy rating task. **H** predicts that mean scientific accuracy ratings for IR will be significantly higher than mean ratings for DR. **H*** predicts that mean ratings for neither DR nor IR will be significantly above the neutral scale midpoint, and that mean ratings will be higher for DR than for IR, if the two differ at all. Finally, we can take a more fine-grained approach that allows for differences between individuals, and determine the proportion of participants who regard IR, but not DR, as scientifically accurate, i.e., provide mean ratings for IR statements that are significantly above mid-point and mean ratings for DR statements that are not. This will provide us with an upper bound of participants of whom $\mathbf{H_0}$ could potentially be true.

3.1 Methods

Participants: 200 participants were recruited and remunerated through the online platform *Prolific*. They were aged 18-80 (mean 29 years) and included 86 men, 108 women, 5 non-binary and 1 who 'preferred not to say'. All were native speakers of English, resident in the UK. All had at least some secondary education (GCSE or equivalent), but none had a higher than undergraduate degree. None had studied a natural science, psychology, or philosophy at university level.

Materials: These laypeople were given 24 verbal statements to read and 3 animations to watch. 8 verbal statements were taken from our *Direct/Indirect Realist Belief Inventory (DIRBI)* (Fischer, Allen & Engelhardt, Ms) and used concrete examples (bus, tomato, etc.) to express the key claims of DR outlined at the beginning of this paper (*Perceptual Direct* with a

generic version [PD below] and a causal version [PDC], *Phenomenal Direct* [PHD], and *Epistemic Direct* [ED]), or the IR counterclaim (PI, etc.):

PD When you look at a bus, you see just the bus and not a mental image of the bus.
PI When you look at a tomato, you see a mental image of the tomato and not just the tomato.
PCD When you look at a fork, you see just the fork and not a mental image caused by the fork.
PCI When you look at a strawberry, you see a mental image caused by the strawberry and not just the strawberry.
PHD When you look at a table, you just see there is a table, and do not consciously work out that it is a table.
PHI When you look at a pen, you consciously work out that it is a pen, and do not just see there is a pen.
ED When you see a potato, you know that it is a potato and not just that it looks like a potato.
EI When you see a horse, you know just that it looks like a horse and not that it is a horse.

Exploratory Factor Analysis conducted in validating the DIRBI inventory showed that all DR items load positively on precisely one factor, whereas all IR items positively load on another. For analysis, we can therefore combine them in a DR and an IR condition.

Scientific accuracy ratings for these items can be driven either by knowledge of relevant scientific facts and theories, or by heuristic cues. The most important of these cues is whether statements 'sound scientific'. Since exemplary scientific statements do not refer to buses and tomatoes, but use more abstract terms, we added 8 abstract reformulations of the above items, as statements about 'physical objects', such as:

PD-A When you look at a physical object, you see just the object and not a mental image of the object.

Reliance on the heuristic cue will show up in higher accuracy ratings for abstract than concrete items.

8 fillers were included to support the interpretation of findings: In their seminal speeded verification study on the persistence of naïve theories upon scientific instruction, Shtulman and Valcarcel (2012) had observed verification responses to sets of statements that are either consistent or inconsistent with both the naïve and the scientific theory of a domain, or consistent with only one, but not the other. We chose four pairs of such 'partly consistent' items, which most participants in that study had correctly judged true [T] or false [F] in accordance with the scientific theory, despite conflict with the naïve theory:

[Germs-T] Heat kills germs.
[Germs-F] Water kills germs.
[Light-T] White light contains all colours of light.
[Light-F] Black objects reflect all colours of light.
[Heat-T] Pressure produces heat.
[Heat-F] Coats produce heat.
[Seasons-T] The earth's tilt causes the seasons.
[Seasons-F] The earth's distance from the sun causes the seasons.

In the present study, a combination of high scientific accuracy ratings for the T-items and low scientific accuracy ratings for the F-items will indicate that participants are familiar with the relevant scientific theories and manage to deploy this knowledge to resist the pull of the competing naïve theory. We can then see whether a similar pattern obtains for critical items (as suggested by the rationale informing **H**).

Fischer and colleagues (2023) found that laypeople are more inclined to agree with IR when this is represented by an image of a 'Cartesian Theatre' rather than a verbal statement. We therefore also included an animation illustrating a 'Cartesian Theatre' account of vision and two prior animations that highlighted different components of the image ('inner eye' looking

at an inner screen, and moving cogs) and were intended as conveying statements PI and PHI (for further details, see Appendix A).

To inform future research, we finally sought to assess whether perceived scientific accuracy requires the kind of dedicated study undertaken here: If it is very important to people that their beliefs about vision are scientifically accurate, their agreement with statements will be based on the perceived scientific accuracy of the statements, and agreement ratings, as elicited by belief inventories, will correlate highly with scientific accuracy ratings. If so, these need not require separate study. To assess this suggestion, we finally asked participants: 'How important is it to you that your beliefs are scientifically accurate?'

Procedure: The study was delivered online via *Qualtrics*. Participants received this initial instruction:

> This study examines what kind of claims people regard as scientifically accurate and which claims they regard as naïve or unscientific. You will read below a number of statements. All of them are phrased in plain English and do not use any scientific terms.
>
> None of the statements are likely to appear in a science textbook. But some may be more scientifically accurate than others.
>
> Please tell us to what extent you think these statements are scientifically accurate. Please rate them on a scale from "I am highly confident that this statement is scientifically inaccurate" to "I am highly confident that this statement is scientifically accurate". The mid-value is "I am not sure if this statement is scientifically accurate or inaccurate".

Verbal statements were then presented in random order, followed by the animations in a fixed order designed to render salient relevant features (Appendix A). Animations were followed by interpretation questions and rating prompts. For all ratings, participants used a 7-point scale anchored as explained. Finally, they responded on a 7-point scale to the importance question. Verbal items included an attention check.

Design and analysis: We manipulated *Realism* (Direct vs Indirect) and *Abstraction* (Concrete vs Abstract) in a 2 × 2 within-subject design. We excluded 4 participants who failed the attention check. There was no evidence of strategic responding. There were no outliers ≥ 3 SDs from any condition mean. Therefore no further exclusions were made. We conducted an omnibus test (repeated-measures ANOVA) and paired comparisons.

3.2 Results

The omnibus test revealed an interaction that was significant ($F(1,195)= 37.9$, $p<.001$) and large ($\eta^2 =.16$) (Figure 1). We observed a small main effect of Realism ($F(1,195)= 4.93$, $p=.028$, $\eta^2 =.03$), but no abstract-concrete effect ($p=.85$). As Figure 1 shows, in the abstract condition mean scientific accuracy ratings for DR and IR are virtually identical and at mid-point: Collectively, participants are entirely unsure whether or not these statements are scientifically accurate. In the concrete condition, participants grow more confident: They are confident (if ever so slightly) that DR statements are scientifically accurate and that IR statements are scientifically inaccurate (whence the interaction). This impression from visual inspection is confirmed by t-tests following up the ANOVAs (Table 1).

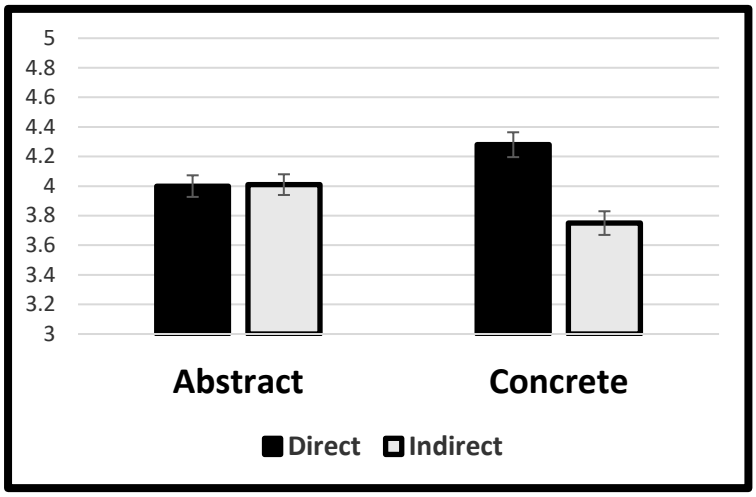

Figure 1. Mean scientific accuracy ratings by condition on 7-point scale. Error bars show the standard error of the mean.

Table 1. Paired comparisons between conditions (paired-samples t-tests) and against neutral midpoint 4 (one-sample t-tests). All df=195.

	T	*p* (two-tailed)	Cohen's *d*
DR vs IR	4.02	<.001	.29
DR-A vs IR-A	-.099	.92	-.07
DR vs DR-A	3.96	<.001	.28
IR vs IR-A	-5.04	<.001	-.36
DR vs midpoint	3.23	<.001	.23
IR vs midpoint	-3.17	<.001	-.23
DR-A vs midpoint	.00	.50	.00
IR-A vs midpoint	.16	.44	.01

To be able to assess hypotheses **H** and **H*** in more detail, we next considered mean ratings and their difference from midpoint, for each of the 16 critical items (Table 2). The effect sizes of the comparisons to mid-point were all negligible (Cohen's d<.2) or small (d<.5). The former indicate lack of collective confidence, the latter a small degree of confidence. In both concrete and abstract conditions, the sample collectively remains unsure whether the metaphysical statements of DR (PD and PCD) are scientifically accurate but attains slight confidence that the causal statement (PCI, PCI-A) of IR is scientifically *in*accurate. In the concrete condition, the sample is slightly confident that the epistemic statement (ED) of DR is scientifically accurate and its IR counterpart (EI) is inaccurate. This difference drives the more global DR-IR difference observed above for the concrete condition: The ED-EI comparison is the only comparison between Direct and Indirect Realist items in the concrete condition to attain a medium effect size (p<.001, d=.50, whereas all other d's ≤.2). Further item-level results revealed differences in the abstract condition obscured by the condition-level analysis: The difference between ED and EI in the concrete condition is here mirrored by a slightly smaller difference between ED-A and EI-A (p<.001, d=.39). But between PHD-A and PHI-A there is a

difference in the opposite direction, as the IR item is deemed more scientifically accurate than the DR item (p<.001, d=.44). This merits further study (though the correlational findings below suggest these findings should be disregarded in the present context).

Table 2. Mean scientific accuracy ratings per critical item with paired comparisons to neutral mid-point 4 and their effect size. Asterisks indicate significance level (two-tailed; *** p<.001, ** p<.01, * p<.05, • p<.08, NS not significant).

Concrete	**Mean (SD)**	**Sign.**	**Cohen's d**	***Abstract***	**Mean (SD)**	**Sign.**	**Cohen's d**
PD	4.22 (1.67)	•	.13	PD-A	4.16 (1.60)	NS	.10
PI	3.70 (1.67)	*	-.18	PI-A	4.06 (1.66)	NS	.03
PCD	4.11 (1.62)	NS	.07	PCD-A	3.88 (1.71)	NS	-.07
PCI	3.60 (1.63)	***	-.25	PCI-A	3.69 (1.54)	**	-.20
PHD	4.00 (1.72)	NS	.00	PHD-A	3.44 (1.66)	***	-.34
PHI	4.19 (1.72)	NS	.11	PHI-A	4.60 (1.61)	***	.37
ED	4.78 (1.64)	***	.48	ED-A	4.52 (1.44)	***	.36
EI	3.52 (1.67)	***	-.29	EI-A	3.70 (1.48)	**	-.20

To support the interpretation of these findings, we conducted further analyses. First, we considered the correlation table for statements' scientific accuracy ratings (Appendix B): Positive correlations among ratings of concrete metaphysical statements of DR and IR (i.e., PD-PCD and PI-PCI), respectively, were high (≥.5). Negative correlations with ratings for items of the opposite metaphysical persuasion (e.g., PD-PI) were of small-to-medium strength (-.24 to -.3). Among epistemological item ratings, both positive and negative correlations were weaker and some negative correlations were even insignificant (e.g., ED-A vs EI-A). Correlations between metaphysical and epistemological items of the same persuasion (e.g., PD-PHD or PD-ED) were small to medium, and weaker for IR than DR (e.g., no correlations between PI-ED or PI-EI). Correlations that crossed both DR/IR and metaphysical/ epistemological divides (e.g. PD-PHI or PD-EI) remained insignificant.

We inferred that metaphysical and epistemological statements are not perceived as parts of the same theory or conception whose scientific accuracy is then assessed as a whole. We therefore focused further analyses on the metaphysical statements which show a more robust pattern of correlations, and are also a stronger focus of philosophical debates about DR vs IR.

To support the interpretation of findings about these metaphysical statements, about which the sample as a whole seems to largely sit on the fence, we turned from means to frequencies and considered for these statements the numbers and proportions of participants who deemed them scientifically inaccurate (rating 1-3), scientifically accurate (5-7) or remained unsure (4) (Table 3).

Table 3. Numbers of participants (and percentage of whole sample) rating metaphysical statements as scientifically in/accurate.

	Inaccurate (1-3) N (%)	**Unsure (4) N (%)**	**Accurate (5-7) N (%)**
PD	70 (35.7)	35 (17.9)	91 (46.4)
PI	91 (46.4)	32 (16.3)	73 (37.2)
PCD	72 (36.7)	41 (20.9)	83 (42.3)
PCI	96 (49.0)	32 (16.3)	68 (34.7)
PD-A	69 (35.2)	48 (24.5)	79 (40.3)
PI-A	73 (37.2)	42 (21.4)	81 (41.3)
PCD-A	86 (43.9)	37 (18.9)	73 (37.2)
PCI-A	88 (44.9)	47 (24.0)	61 (31.1)

The largely neutral mean ratings for these conditions thus hide considerable individual variation. However, no metaphysical statement of either conception is judged scientifically accurate by even half the sample. Specifically, no metaphysical DR statement is deemed scientifically accurate by more than 46.4% of participants, and no metaphysical IR statement is deemed scientifically accurate by more than 41.3%. In line with the observed mean ratings (Figure 1), roughly equal numbers of participants find the IR and DR statements scientifically accurate in the abstract condition, while the concrete statement PD is deemed scientifically accurate by a higher number of participants than the corresponding IR statement PI. However, numeric differences remained shy of statistical significance for all comparisons (for all DR-IR comparisons, $\chi^2<1.98$, well shy of the critical χ^2 value (df=1) of 3.84).

Overall, these findings favour our hypothesis **H*** that neither DR nor IR are regarded as scientifically accurate over the hypothesis **H** that laypeople deem IR more scientifically accurate than DR (see Sect. 3.3).

To be able to assess more directly the hypothesis $\mathbf{H_0}$ that, in laypeople's cognition, DR functions as a naïve theory and IR as a scientific theory, we considered further proportions: In an individual's cognition, DR can only function as a naïve theory that is suppressed in favour of theories that are perceived as more scientifically accurate, if the individual deems these theories more scientifically accurate. Indeed, we would expect the individual to deem these theories scientifically accurate, *tout court*, and DR not so. According to $\mathbf{H_0}$, the relevant theories include IR. We therefore determined the proportions of participants who deem IR more scientifically accurate than DR (IR>DR), and the proportions who credit scientific accuracy to IR but not DR (IR>4 and DR≤4) (Table 4). The resulting figures mark, respectively, a maximum and a plausible upper bound for participants in whose cognition DR functions as a naïve theory in the way $\mathbf{H_0}$ suggests. We observe that the proportion of participants who meet the more stringent means criterion (31.6%) is almost identical with the proportion of participants whose mean IR ratings are at least 1 point above their mean DR ratings (30.6%) – intuitively, enough of a difference to make a difference. We infer that, for about a third of participants, DR can potentially be taken to function as a naïve theory suppressed in favour of a more scientific IR theory.

Table 4. Numbers of participants (and percentage of whole sample) with higher scientific accuracy for metaphysical IR than DR statements, by different criteria. "Mean IR" and "Mean DR" are the means across the four metaphysical items in each condition.

Lax criteria	**N (% of sample)**	**Stringent criteria**	**N (% of sample)**
PI>PD	65 (33.2)	PI>4, PD≤4	48 (24.5)
PCI>PCD	56 (28.6)	PCI>4, PCD≤4	48 (24.5)
PI-A>PD-A	72 (36.7)	PI-A>4, PD-A≤4	59 (30.1)
PCI-A>PCD-A	67 (34.2)	PCI-A>4, PCD-A≤4	45 (23.0))
Mean IR > Mean DR	80 (40.8)	*Mean IR>4, Mean DR ≤4*	62 (31.6)

To further support the interpretation of findings, we finally considered mean ratings and their difference from midpoint, for our 8 filler items (Table 5). These articulated statements that are true (T) or false (F) according to scientific theories that have percolated from science into secondary schools (or yet further 'down'), and are simultaneously inconsistent (T) or consistent (F) with naïve theories. We found a clear pattern, broken only by the one item ('Seasons-F') that had attracted the lowest proportion of correct verification responses (49%) in Shtulman and Valcarcel's (2012) study: All statements consistent with the relevant scientific theory were deemed scientifically accurate (mean ratings above midpoint), all statements inconsistent with it were deemed scientifically inaccurate (mean ratings below midpoint), and participants were moderately or highly confident about this (as indicated by the medium ($d \geq .5$) or large ($d \geq 8$) effect size of most comparisons to neutral midpoint). These findings cohere well with the high proportions of correct verification responses observed in that earlier study (added for convenience to Table 5) and suggest that (all but one of) the relevant scientific theories have become familiar to laypeople (first-year psychology students, in the case of Shtulman & Valcarcel's sample). Taken together, present and previous findings suggest that where laypeople are familiar with scientific theories, they mostly manage to suppress conflicting naïve theories, at any rate in a context which prizes scientific accuracy, like a scientific survey.

Table 5. Mean scientific accuracy ratings per filler item with paired comparisons to neutral mid-point 4 and their effect size. Asterisks indicate significance level (two-tailed).

	Mean (SD)	**Sign.**	**Cohen's *d***	% correct verification responses (Shtulman & Valcarcel, 2012, SM)
Germs-T	5.45 (1.49)	***	.97	89
Germs-F	2.31 (1.57)	***	-1.08	83
Light-T	5.06 (1.77)	***	.60	75
Light-F	3.19 (1.99)	***	-.41	67
Heat-T	4.85 (1.62)	***	.53	73
Heat-F	2.72 (1.94)	***	-.66	71
Seasons-T	4.59 (1.80)	***	.33	75
Seasons-F	3.90 (1.67)	NS	-.06	49

Scientific accuracy ratings for animations were largely aligned with ratings for verbal items: Those participants who interpreted the animations as intended (125 for 'Screen', 138 for 'Cogs'), gave virtually the same ratings for 'Screen' (3.73, SD 1.35) and 'Cogs' (4.07, SD 1.37) as the whole sample gave for the statements PI and PHI they were respectively intended to illustrate (see Table 2). But participants who correctly interpreted the 'Cartesian' image as conveying either PI or the conjunction of PI and PHI, gave it a higher mean rating (4.58, SD 1.35) than any verbal IR statement, which was significantly above scale midpoint $t(79)=3.815$, $p<.001$, $d=.43$. However, the low number of 'correct interpreters' of this animation (N=80) does not support further conclusions (see Appendix A for further animation results and discussion).

Overall, participants reported it was slightly important to them that their beliefs are scientifically accurate, as indicated by a mean importance rating of 2.85 (SD 1.24), where 1 = 'very important' and 3 = 'slightly important', significantly different from the neutral midpoint 4 ('neither important nor unimportant') $t(195)=-13.03$, $p<.001$, $d=-.93$.

This suggests that participants will be sensitive to mismatches between their credences and the perceived scientific accuracy of beliefs, and may lower credence where they perceive scientific inaccuracy. However, the comparatively low importance reported suggests that laypeople do not base their agreement with DR and IR statements exclusively or even primarily on the perceived scientific accuracy of these statements. We infer that, for laypeople, scientific accuracy ratings and agreement ratings will correlate, but not strongly enough to obviate the need for both measures. In other words, perceived scientific accuracy cannot be simply inferred from agreement (or *vice versa*); the present study was needed to complement extant agreement rating studies (Fischer, Allen, & Engelhardt, 2023; Ms).

3.3 Discussion

These findings allow us to adjudicate between hypotheses **H** and **H***. **H** predicted that mean scientific accuracy ratings would be higher for IR than for DR. We found the almost exact opposite: While ratings were the same, and at midpoint, for both IR and DR in the abstract condition, the mean rating in the concrete condition was *below* midpoint for IR, and *above* midpoint for DR (Figure 1). These findings are considerably more congenial to **H***: **H*** predicted that (i) mean ratings for neither IR nor DR would be above midpoint, and that (ii) if there was any difference between the two, mean ratings would be even lower for IR than for DR. The former prediction was borne out in the abstract condition, the latter in the concrete condition. Contrary to **H***(i), though, mean ratings for DR were significantly above midpoint in the concrete condition. However, the effect size of the comparison was driven by the single, epistemological, item (ED) and remained small ($d=.23$), markedly smaller than the mostly medium or large effects of midpoint comparisons observed for the 'T-filler' statements that are consistent with scientific theories which have seeped into general knowledge (Table 5).

For the metaphysical statements that form the core of DR and IR (at any rate in philosophical debates), **H*** is fully borne out: No metaphysical statement received mean ratings above midpoint. Mean ratings were no different from mid-point for metaphysical DR statements and just below midpoint for most metaphysical IR statements (all d's $\leq .25$) (Table 2). These findings from means cohere well with findings from frequencies (Table 3): The metaphysics of DR and IR are both deemed scientifically accurate by less than half the sample, and neither attracts scientific credit from significantly more participants than the other.

An empirical rationale for restricting attention to the metaphysical statements that form the core or DR and IR is that correlations between scientific accuracy ratings for metaphysical

and epistemological statements are rather weak (above and Appendix B), in particular for IR (where all *r*'s<.32, all but one $r \leq .29$, and many metaphysical-epistemological correlations fail to reach significance).This suggests that DR and especially IR are not regarded as coherent theories of vision that are assessed as a whole for scientific accuracy.

This conclusion, in turn, speaks against a natural worry: People may well treat scientific accuracy as a property of whole accounts, and take the accuracy of a statement to increase the more true and relevant information it contains, where information will be deemed relevant when it is regarded as part of the same account. Hence, the worry goes, folk Direct Realists and folk Indirect Realists will deem, e.g., metaphysical DR and IR statements, respectively, more scientifically accurate when these are conjoined with epistemological statements of the same persuasion, so that Indirect Realists will deem the conjunction 'PI & PHI' more scientifically accurate than each individual conjunct. Hence the low scientific accuracy ratings we observed might be due to presenting the information from DR and IR accounts of vision in too small chunks.

The present conclusion blocks this worry: Since metaphysical and epistemological statements, especially of IR, are not regarded as part of the same coherent account of vision, the addition of an epistemological to a metaphysical statement (as in 'PI & PHI') will not be deemed to add *relevant* information, so that accuracy ratings for our IR (and perhaps DR) statements will not increase through mere conjunction. In addition, mean accuracy ratings for the more informative causal metaphysical statements PCD, PCD-A, PCI, and PCI-A were all numerically *lower* than ratings for the less informative statements PD, PD-A, PI, and PI-A, respectively (Table 2), further disarming the present worry. If the metaphysical and epistemological statements of interest are not regarded as part of comprehensive accounts that are assessed as a whole for scientific accuracy, we should restrict our attention to the metaphysical statements, whose ratings are embedded in stronger positive and negative correlations with each other. As noted above, ratings for these statements fully bear out **H***.

Findings thus support **H*** and refute **H**. **H** was the decisive part of the key hypothesis $\mathbf{H_0}$ that, in lay cognition, DR functions as a naïve theory and IR as a scientific theory. Present findings imply this cannot generally be the case. We assessed $\mathbf{H_0}$ more directly by identifying a plausible upper bound of participants of whom this could potentially be true: Less than one third of participants deemed IR, but not DR scientifically accurate (Table 4).

Finally, we compared mean scientific accuracy ratings for the statements of interest to mean accuracy ratings for statements about whose truth-value there is disagreement between naïve theories and scientific theories with which laypeople are generally familiar (Table 5). If, in accordance with $\mathbf{H_0}$, IR was a familiar scientific theory and DR a naïve theory, we would expect the same pattern of responses: namely, as for the 'T-fillers', high scientific accuracy ratings for statements that are true according to IR (but false according to naive DR) and, as for the 'F-fillers', low scientific accuracy ratings for statements that are false according to IR (but true according to naïve DR). But this is not what we find, least of all for the key metaphysical statements (Table 2). The difference of participants' attitude to these IR and DR statements, on the one hand, and to familiar scientific and naïve theories in mutual conflict, on the other, also shows up in frequency findings: Filler items 'Light-T' and 'Light-F' articulated statements that are true and false, respectively, on the Newtonian theory of light. 62% of participants deemed Light-T more scientifically accurate than Light-F, and 46% of participants simultaneously found Light-T scientifically accurate (rating >4), but not Light-F (rating ≤4). These figures are almost twice as high as the analogous figures for the critical metaphysical

items PD and PI (Table 4). The observed responses to critical items clearly do not reflect widespread knowledge of both a naïve and a conflicting scientific theory.

So what drives the responses to these items, instead? The lack of a concrete-abstract effect and the observed interaction between our two independent variables speak against the hypothesis that our participants were simply basing their scientific accuracy ratings on the heuristic cue of how scientific the items sounded: As explained above (Sect.3.1), such reliance would have uniformly increased the ratings for abstract ('physical object') over concrete (e.g., 'bus') versions of our items. But we observed an increase only for IR items, while ratings decreased for DR items built from the same linguistic stock (Figure 1, moving right-to-left). We infer that the observed responses do tap into knowledge structures (in the epistemologically undemanding sense of cognitive science), but that these knowledge structures do not include knowledge of a pair of familiar theories of vision, one scientific and one naïve.

3.4 Interpretation

Where does this conclusion leave us? Which knowledge structures did responses tap? We will consider two possibilities, informed by the rationale informing **H** and **H***, respectively, and bring out the consequences of each for the philosophically influential assumption (A) that there is such a thing as 'the' common-sense conception of vision, which is consistent with DR and inconsistent with IR.

The rationale for **H** assumed that IR seeped into cultural consciousness from scientific discourse. Our findings suggest that (*pace* **H**) only a minority of laypeople perceive IR as scientifically accurate. But it remains possible that IR seeped into cultural consciousness from scientific discourse in early modernity and has, since then, come to be perceived as unscientific. It could be that while Newtonian physics, including Newtonian optics, continue to be taught in school, IR features, if at all, only in philosophy curricula (as currently in the UK, from where our sample is drawn).[3] If this is correct, IR may not, or no longer, be recognized as scientific in nature by most laypeople, and DR will not be suppressed in favour of IR *as the more scientific theory*. Rather (*pace* $\mathbf{H_0}$,), *both* DR and IR will be suppressed by most laypeople in favour of theories that are perceived as (more) scientific. That is: In most laypeople's cognition, both DR and IR will function as naïve theories, so that both will qualify as common-sense conceptions. Interpreting present findings in line with the rationale informing **H** thus leads to the rejection of (A): *Pace* (A), common sense will contain not one coherent, but at least two conflicting conceptions of vision.

The rationale for **H*** suggested that IR and DR both entered common sense from independent but equally pre-scientific sources: We suggested that IR beliefs are an artefact of an implicit model of attention, and are as pre-scientific as DR (Sect.2). We now unpack this suggestion to argue that it takes us from present findings to the same ultimate conclusion, i.e., the rejection of (A).

According to 'attention schema theory' from cognitive neuroscience (Graziano, 2022; Graziano & Kastner, 2011; Webb & Graziano, 2015), implicit perceptual models of attention help people track others' focus of attention, support their own endogenous control of selective attention, and are partly accessible to higher cognition, giving rise to intuitions and beliefs. Implicit models prioritize usefulness and economy over accuracy, and provide simple useful caricatures, rather than accurate representations of the process they depict. Thus, our implicit

[3] E.g., IR is covered by the epistemology module of the current AQA philosophy A-level syllabus, but does not appear in the corresponding psychology syllabus. Last accessed 11/11/2024.

model of other people's focus of visual attention portrays people as emitting force rays from their eyes towards the objects of their attention: The perceptual model represents direction of attention by a subthreshold motion signal that falsely indicates a flow from the viewer's eyes to the object. This signal activates motion processing centres (Guterstam et al., 2020) and influences judgments about the physical behaviour of objects of sight: when people see an observer looking at an object (rather than elsewhere, or being blindfolded), they judge that the object can tilt over in the observer's direction without falling over, as if it was held up by a force emanating from the observer's eyes (Guterstam et al., 2019). If higher cognition has some access to this implicit model, it could account for extramissionist beliefs about vision.

We suggest a similar explanation for IR beliefs about vision. According to attention schema theory, the processes of social cognition employed to track others' attention are also used to control our own (Graziano & Kastner, 2011; Graziano, 2022): The perceptual model they compute for the purpose consists of a set of attributes assigned to spatial locations. The model ties an attention signal to a place (a subject of attention) and directs it towards a spatially located object (of attention). Where the model constructed is of one's own, rather than another's attention, it ties the attention signal to a place in one's own head. Since attention can be directed not only towards external objects but also towards one's own thoughts, beliefs, feelings, sensations, etc., the model represents our attention as directed towards an object spatially located in our head (in the case of thoughts or beliefs) or body (for corporeal sensations). If higher cognition has some access to this implicit model, it could account for intuitions and beliefs people have about themselves. Graziano (2022) seeks to explain common intuitions and beliefs about consciousness. We suggest that a perceptual model of attention that ties attention signals to locations in the head and directs them towards objects also placed in the head is well placed to explain belief in the Cartesian Theatre conception and IR beliefs about vision embedded in that conception. On this account, IR beliefs have a pre-scientific source and are typically formed prior to and independently from scientific instruction.

We now turn to DR. Its positive claims (e.g., 'When you look at a physical object, you see just the object…') derive their content largely from the contrast with opposing IR beliefs ('…and not a mental image of the object'). This marked difference to naïve causal-explanatory theories suggests that DR beliefs are not part of a naïve *theory* in the psychological sense of a *causal-explanatory framework* acquired prior to scientific instruction (*cf.* Gelman & Noles, 2011; Vosniadou, 1994), but are formed in response to pre-scientific IR beliefs on the basis of experiential knowledge. The most relevant knowledge structure, we suggest, are situation schemas that organise experiential event knowledge, namely, about *seeing*-events, and are automatically activated by verbal stimuli (e.g., the verb 'to see') and perceptual cues (McRae & Jones, 2013; Rumelhart, 1978). These schemas contain information that is acquired through observation of patterns in the physical environment and in linguistic discourse; this information is about the typical and diagnostic features of agents (viewers), patients (objects of sight) and the relations between them (Ferretti et al., 2001; Hare et al., 2009; McRae et al., 1997). Dependency networks in complex schemas can encode causal, functional, and nomological information (Sloman et al., 1998).

We submit the situation schema associated with the verb 'to see' is automatically activated the moment beliefs about vision are articulated or considered. If this schema does not include information about mental images that causally mediate between viewers and objects of sight, it will support Direct Realist responses the moment IR beliefs are considered. Validation studies for the DIRBI inventory provide evidence that the relevant schema indeed does not

contain information about mental images (Fischer, Allen & Engelhardt, Ms., Study 3): A classification task made participants decide whether cases of hallucination qualify as cases of 'seeing'. Such similarity-based classification judgments seize on situation schemas (review: Hampton, 2015). Regardless of whether participants held IR or DR beliefs about vision (or both), they frequently mentioned the subject's having mental images among the salient features of cases of hallucination. According to IR, this feature makes these cases more similar to vision; according to DR, it makes them more different. Even so, previously assessed DR/IR beliefs made no difference to participants' classifications. We infer that the situation schema deployed for these judgments is the same for people with DR and IR beliefs, and that this schema contains no information about mental images being causally or otherwise involved in seeing. Hence the schema will support DR responses to IR beliefs.

While these explanations of IR and DR beliefs call for further development and empirical assessment, they may help us to understand present key findings: According to the proposed explanations, both DR and IR beliefs have pre-scientific cognitive sources. As laypeople know, naïve conceptions may be proven factually inaccurate by science. The proposed accounts would therefore explain why the metaphysical core beliefs of neither DR nor IR were regarded as scientifically accurate by even half of our lay sample, and why DR/IR beliefs in general attracted scientific accuracy ratings that patterned so differently from those for pairings of naïve and scientific theories. This suggests, again, that (*pace* $\mathbf{H_0}$) *both* DR and IR will be suppressed by most laypeople in favour of theories that are perceived as (more) scientific. The explanations informed by the rationale for **H*** thus lead to the same conclusion as the explanation informed by the rationale for **H**: DR and IR both function as naïve theories in lay cognition, and thus qualify as part of our society's 'common sense'.

For now, we stress the philosophically relevant upshot of present findings: Previous studies found that both DR and IR conceptions of vision are prevalent in the general population, with each being endorsed by at least 40% of each sample (Fischer, Allen, & Engelhardt, 2023; Ms). Present findings suggest that both conceptions are 'common sense'. *Pace* (A), there hence is no such things as 'the' common-sense conception of vision. Rather, at least two conflicting conceptions, consistent with DR and IR, respectively, compete within common sense.

4. Philosophical Discussion and Future Directions

This empirical finding has important consequences for philosophical debates about the nature of perception: It forcefully challenges a dominant methodology (Sect. 4.1), motivates the reconceptualization of their central 'problem of perception' (Sect. 4.2), and supports a fresh approach to the problem that may carry over to further aporetic problems (Sect. 4.3).

4.1 Against default epistemic status

The 'default-and-challenge' model of justification (Brandom, 1996; Williams, 2001; *cf.* Austin, 1957/1979) accords common-sense beliefs an epistemic default status: It suggests these beliefs should be accepted in the absence of good reasons to the contrary. While this suggestion is highly contested in epistemology, it is built into standard philosophical methodologies including the method of cases and the quest for reflective equilibrium (*cf.* Daly, 2010; Machery, 2017). The more specific assumption about perception certainly is widely accepted in the philosophy of perception:

(A*) Common-sense beliefs about perception should be accepted as a default (i.e., in the absence of good reasons to the contrary).

Accordingly, a key criterion used to assess philosophical theories of perception is how well they capture or respect supposedly pre-theoretical common-sense thinking about perception. Thus, Naïve Realist theories of perception are often recommended on the grounds that they best capture this common-sense thinking (for discussion: Fish, 2009, pp.18-19; Raineri, 2021). Representationalists often argue that their theories are able to accommodate the supposedly common-sense belief that perceptual experiences are 'direct' (e.g. Strawson, 1979; Searle, 2015). And many Indirect Realists (including sense-datum theorists) accept they have to provide an error theory that explains why supposedly Direct Realist common sense got it wrong (e.g., Russell, 1912/2000).

As commonly interpreted, (A*) relies on the empirical assumption (A) that has now been refuted. This refutation renders (A*) untenable: If two conceptions of vision that directly contradict each other have equal claim to being 'common sense', then some common-sense beliefs about vision are mutually inconsistent, and this inconsistency already constitutes a good reason to give up at any rate some of these beliefs. Not all of them can serve as initial common ground. The consequences for philosophical methodology in debates about the nature of perception are considerable: at first blush, it would seem that appeal to common sense needs to be replaced by appeal to phenomenological introspection (phenomenology; *cf*. Raineri, 2021) or to science (methodological naturalism; *cf.* Drayson, 2021).

Both moves would mark a radical change. The move to reliance on phenomenology may seem unproblematic. Indeed, reliance on common sense and on phenomenology are often at least implicitly conflated, on the assumption that the deliverances of introspection are 'pre-theoretically' evident and inform widely shared common-sense beliefs about vision (e.g., Crane & French, 2021). Previous studies have shown that laypeople are collectively and sometimes individually torn between DR and IR beliefs about vision (Fischer, Allen & Engelhardt, 2023; Ms) – as they are torn about further topics (Knobe, in press). If common-sense beliefs about vision are reflective of visual phenomenology, then this phenomenology will have to differ between individuals and even for the same individual, perhaps across different situations – even if Direct Realists appealing to phenomenology would then appear to have missed this variation, suggesting a lack of adequate methods. Alternatively, if the phenomenology of visual experience is the same for all or most humans, at least some of the conflicting common-sense beliefs about vision will to fail to reflect it. In this case, phenomenology and common-sense beliefs need to be treated as independent, calling for different methods (Merleau-Ponty, 1945/2013; *cf*. Allen, 2019). Either way, the change from reliance on common sense to reliance on phenomenology would imply a departure from common methods that may seem no less radical than the naturalistic change to reliance on science.

By undermining the methodological assumption (A*), the refutation of the empirical assumption (A) thus presents a major methodological challenge to extant philosophical debates. In response to this challenge, one could try to save (A) by rejecting the straightforward idea that the relevant 'common sense' includes all conceptions of vision that function as naïve theories, in lay cognition (as per our criterion from Sect.2). Instead, one might restrict the scope of what counts as 'common sense', which arguably is a partially normative notion. The hope would be that suitable restrictions on the scope of 'common sense' will yield a coherent set of beliefs about vision whose epistemic default status can be plausibly motivated. We now consider two versions of this response that are anchored in the extant literature, and discuss how present findings and future research would help assess them empirically.

The *'core response'*, restricts what counts as 'common sense' about vision to those topical beliefs that belong to the 'massive central core of human thinking which has no history', and do not change over time or place (Strawson, 1959, p.10): only these enjoy epistemic default status. One challenge to this *a priori* plausible view arises from evidence of diachronic variation in related lay conceptions that may also have been thought to exhibit stability given underlying psychological similarities, like lay conceptions of colour (Adams & Hansen, 2020). The current study suggests a different criticism: Even if there are cross-culturally stable common-sense beliefs about vision, this 'central core' will include conflicting beliefs. This response is suggested by our preferred explanation of present findings (Sect. 3.4): On this account, the metaphysical beliefs of IR and of DR are both generated by a process and a knowledge structure that may plausibly be assumed to be well-nigh universal, namely, by an implicit model of attention (for IR) and by a situation schema capturing experiential event knowledge about seeing-events (for DR).[4]

These competing hypotheses require further investigation. A natural starting point would be to engage in cross-cultural validation of the new Direct/Indirect Realism Belief Inventory (DIRBI) and use the measure to study whether agreement ratings from participants in different cultural contexts reveal similar patterns of intra- and inter-personal conflict as have been found in the case of Anglophone lay participants in the UK (Fischer, Allen, & Engelhardt, 2023; Ms).

Another 'restrictive' strategy to save assumption (A) would restrict the scope of relevant 'common sense' to reflectively endorsed beliefs, in the hope that only a coherent set of beliefs about vision would end up being endorsed upon reflection – or that reflection would at least decrease the level of inter- and intrapersonal disagreement, in favour of DR over IR, or vice versa. This *'reflection response'* has recent precedent: Early findings from experimental philosophy provided evidence that philosophically relevant case judgments ('intuitions') are subject to epistemically otiose presentation effects and demographic variability that seem to deprive them of evidentiary value (review: Machery, 2017). In response, some philosophers suggested that only reflectively endorsed case judgments are philosophically relevant, and hypothesised that these judgments would display less, or no, undue variability (e.g., Ludwig, 2007; see Hannon, 2018, for discussion).

Empirical examination of this hypothesis found against it: a variety of reflection prompts (analytic priming, response justification pressure, financial incentives, forced response delay) did not change a range of philosophically relevant case judgments laypeople made, which also did not differ between participants differing in reflectiveness (Kneer et al., 2022; *cf.* Gerken & Beebe, 2016; Roberts, Allen, & Schmidtke, 2020). In a similar way, the second 'restrictive' response can be empirically assessed by administering the DIRBI inventory in conjunction with familiar reflection prompts or measures of reflectiveness: the restrictive response would predict that reflection prompts greatly reduce the level of inter- and intrapersonal disagreement and produce responses that are more clearly aligned with either DR or IR, and that more reflective participants will be, collectively and individually, less torn between these two views than less reflective participants. If these predictions were confirmed, and the reflection response adopted, common sense might be saved as a source of philosophically relevant evidence at the price of changing the character of appeal to common sense: Common sense

[4] Since situation schemas are built up from observed statistical regularities in both the physical environment and the discourse environment (McRae & Jones, 2013), cross-linguistic differences in how people talk about vision will affect the schema. We therefore do not suggest that the content of the schema will be exactly the same across languages and cultures, only that (almost) all individuals' schemas will support DR responses to IR beliefs, in virtue of the regularities observed (by non-blind individuals) in the physical environment.

could only be appealed to at the end of a reflection process and could not serve as the initial common ground proponents of epistemic default assumption (A*) would typically like it to be.

4.2 Reconceptualising the problem of perception

The methodological assumption (A*) that shapes philosophical debates about the nature of perception relies on the empirical assumption that there is such as a thing as 'the' common sense conception of vision. So does the central problem those debates address: As standardly conceived, the 'problem of perception' is the problem of how perception, *as we ordinarily understand it,* is possible, namely, in the face of persuasive arguments for IR known as arguments 'from illusion' and 'from hallucination' (Crane & French, 2021; Robinson, 1994; Smith, 2002). Proponents of the problem thus take for granted that (A) there is one ordinary understanding of vision and that this understanding is captured by DR, so that arguments for IR challenge this ordinary understanding. Present findings, and our proposed explanation of them, speak against (A) and thus motivate a threefold reconceptualization of the problem.

First, the 'problem of perception' has been seen as, in effect, one venue of the wider clash between what Sellars (1963) calls the 'scientific' and the 'manifest image' of the world (Drayson, 2019; Putnam, 1994, pp.465-487; Valberg, 1992). This is true insofar as IR is typically seen as developing a scientifically informed 'causal picture of perception', and DR is taken to capture a more naïve conception that is reflective of the phenomenological character of the experience through which the world manifests itself to us. But present findings and their proposed explanation suggest this assessment cannot be right.

Present findings suggest that neither DR nor IR function as scientific theories in lay cognition. Both compete *within* common sense, at any rate prior to restrictions designed to render common sense coherent (Sect. 4.1). If IR seeped into common sense from early modern science, and if it still deserves to be regarded as scientifically informed, this part of the conflict between the scientific and the manifest image of the world will play out within today's common sense. If, by contrast, our preferred explanation of IR (Sect. 3.4) is correct, IR can also be said to articulate how the world manifests itself to us: Partial awareness of the implicit model of other's attention has been suggested to generate extramissionist beliefs that are deployed in making gut physical judgments (e.g., of permissible tilt) (Guterstam et al., 2019), as well as the occasional feeling of being stared at from behind (Graziano & Kastner, 2011) – which arguably are part of the way the world (deceptively) presents itself to us. If, similarly, IR beliefs and the phenomenology of 'seeing images' in the head are generated by partial access to an implicit model of self-directed attention, these are part of the way a specific part of the world, namely, the subject of perception, naturally presents itself to itself. If, as we further suggested, DR is reflective of the experiential world knowledge about *seeing*-events, DR also articulates part of the manifest image of the world. The conflict between DR and IR then reflects a tension within the manifest image.

Second, it is common to regard the problem of perception as arising from a *latent* or hidden tension within common-sense beliefs, which is exposed by philosophical reflection that may appeal to science (e.g., Robinson, 1994; Valberg, 1992): As standardly conceived, the problem is developed by arguments 'from illusion' and 'from hallucination' that proceed from uncontroversial assumptions (e.g., that things sometimes appear a different size, shape, or colour than they actually are) to metaphysical conclusions captured by IR, which are incompatible with the alleged DR of common sense about ordinary cases of vision. Philosophical reflection thus brings out a previously unrecognized tension between hitherto

uncontroversial folk beliefs about related, but distinct phenomena: Upon (possibly scientifically informed) reflection, the folk understanding of ordinary cases of vision appears inconsistent with the folk understanding of illusion and hallucination.

Present findings and their proposed explanation (Sect. 3.4), by contrast, suggest that the problem of perception arises, more fundamentally, from a *patent* conflict between different folk conceptions of *one and the same* thing: Laypeople have competing DR and IR conceptions of ordinary cases of vision. The conflict between these two is obvious prior to philosophical reflection or any appeal to science. It is due to cognitive sources that precede philosophical reflection and scientific investigation, namely, an implicit model of attention and experiential event knowledge about *seeing*. We suggest that philosophical reflection resulting in the development of the arguments from illusion and from hallucination is prompted by this patent conflict: These arguments are formulated in philosophical attempts to adjudicate between the patently conflicting folk conceptions. The arguments do not develop the core problem, which consists in the conflict between folk conceptions consistent with DR and IR, respectively. Rather, these arguments 'merely' add the complication of apparent further conflicts, namely, between the DR conception of vision and folk conceptions of illusion and hallucination.

Third, and most importantly, present findings complete the case for reformulating the question that articulates the problem. As previous studies have shown, laypeople are, individually and collectively, torn between DR and IR conceptions of vision. In other words, both conceptions are 'ordinary' in the sense of being widely shared by laypeople. The present study has shown that both conceptions are 'ordinary' also in the sense of being non- or pre-scientific in cognitive role or function (see Sect.2). But if both DR and IR are part of our widely shared, non-scientific 'ordinary understanding' of vision, we cannot sensibly ask,

> 'How is vision, as we ordinarily understand it, possible (viz., in the face of persuasive arguments for IR)?'

The most similar question we can sensibly ask is:

> 'How can one of two conflicting folk conceptions of vision (viz., folk DR and folk IR) be correct or warranted in the light of conflicts with each other and with further folk beliefs about vision?'

Relevant further beliefs include the beliefs in the occurrence or possibility of illusions and perfect hallucinations that the eponymous philosophical arguments appear to show incompatible with DR. The problem articulated by this new question has been dubbed 'the true problem of perception' (Fischer, 2023). We now finally consider how present findings help us address this problem – with an approach that might profitably be tried out on similar problems.

4.3 Resolving the problem of perception

The methodological assumption that common-sense judgments and beliefs about familiar phenomena should be accepted in the absence of good reasons to the contrary is built into influential philosophical methods (Sect.4.1). An important strand of experimental philosophy that seeks to implement these methods empirically has relied on this methodological assumption and studied folk conceptions of philosophically relevant phenomena (like agency and free will), to find out which of competing conceptions (e.g., compatibilism or incompatibilism) is the majority view of adult laypeople and thus qualifies as 'common sense'; these studies aimed to settle questions of burden of proof (for a review, see Sommers, 2010). By contrast, we will explore how the domain-specific *refutation* of the influential

methodological assumption, such as the present refutation of (A*), can help us to directly address one important kind of philosophical problem.

Plato regarded a sense of wonder in the face of the familiar as the starting point of all philosophising (n.d./2004, 155b–d). He had in mind puzzlement about *the very possibility* of familiar facts, rather than curiosity about their causes. Such puzzlement is articulated by the question commonly used to formulate the problem of perception: 'How is perception, as ordinarily understood, even possible?' This distinctive puzzlement is the hallmark of *one* kind of characteristically philosophical problem, which is also exemplified by sceptical problems and (key versions of) the problems of mental causation, free will, or induction. Like the problem of perception, these are aporetic problems: Common sense or brief intuitive reasoning command assent to a set of individually plausible but mutually inconsistent propositions, which form an 'aporetic cluster' (Rescher, 2001).

Fischer (2023) distinguished two conceptions of aporetic problems: according to the well-established *'reasoning conception'*, these problems are developed by philosophical arguments, like the arguments from illusion and from hallucination (perception), the argument from dreaming (scepticism), or the argument from causal-explanatory closure (mental causation). These arguments bring out conflicts between individually plausible propositions about related, but distinct phenomena, captured by their premises and conclusions, respectively. For example, arguments from hallucination bring out the clash between the claim 'Perfect hallucinations are possible' (plus further plausible assumptions) and DR claims like our PD (Macpherson, 2013; Smith, 2002). These clashes motivate those questions about the very possibility of phenomena like vision.

According to the novel *'belief conception'* of aporetic problems, by contrast, the problem ultimately arises from a conflict between competing folk conceptions of the same phenomenon; these conflicting conceptions form the nucleus of the relevant aporetic cluster. The familiar arguments are then formulated to adjudicate between the incompatible folk conceptions, pulling their plausible premises into the aporetic cluster as well. Whether the reasoning conception or the belief conception apply to a given aporetic problem is an empirical question, to be addressed by empirical study of folk conceptions. As we have seen (Sect.4.2), extant evidence suggests the problem of perception fits (what we will now call) the belief conception.

To resolve an aporetic problem, philosophers adhering to the reasoning conception seek to reconcile the parties to the apparent conflict or to revise their beliefs. In his classic treatment of the problem of perception, Austin (1962) maintained the reasoning conception but sought to '*dissolve'* the problem by exposing fallacies in the arguments that motivate it. In line with (A*), Austin (1957/1979, p.185) granted common sense epistemic default status and clearly regarded it as coherent. Only philosophical arguments, he thought, create the appearance of a conflict between common-sense conceptions of veridical perception and illusion and hallucination. These arguments hence constitute the epistemic motivation for addressing the problem of how perception, as we ordinarily understand it, is possible in the face of illusions and hallucinations. The exposure of fallacies in the arguments shows this motivation is unsound, and the question ill-motivated.

We now consider how present findings contribute to a different 'dissolution' of the aporetic problem that shows its epistemic motivation unsound. Having undermined both (A*) (Sect.4.1) and the reasoning conception of the problem (Sect. 4.2), we now consider how we can 'dissolve' the aporetic problem, given that it seems to conform to the belief conception. To do so, we proceed in two steps: First, we ask what epistemic motivation philosophers may have

to address the (true) problem of perception, and identify one of DR or IR as being correct. Second, we ask whether this epistemic motivation is sound.

Why should someone need to consider the (true) problem of perception, and attempt to reconcile, or adjudicate between, those two conflicting folk conceptions of vision? Some philosophers explicitly acknowledge non-epistemic motivations for addressing aporetic problems and seek, e.g., to dispel the 'worry' or distress the underlying belief conflicts apparently caused in them (e.g., Austin, 1962, p.5; Wittgenstein 1933/2005, p.409; *cf.* Fischer, 2011, pp.217-234). We set aside these and all other psychological motivations, and focus exclusively on epistemic motivations.[5] We ask: what do philosophers addressing the problem want to find out or understand?

The problem is central to debates about the nature of perception. Recent methodological discussion in the area suggests that current philosophical theorising in the field is typically aimed at understanding how sense-perception works, how it shapes action and belief, what perceptual experience is like, and why it is that way (Fish, 2010; Logue & Richardson, 2021) These debates have evidently moved well beyond the starting point identified by Plato, and are addressing questions similar to exemplary scientific questions. Psychologists and cognitive neuroscientists also seek to explain how sense-perception works or why perceptual experience is the way it is. But these scientists largely ignore common-sense conceptions of vision, and their conflicts. So why should philosophers not follow suit? Methodological naturalists have long advised philosophers to base philosophical theories of natural phenomena like vision on the best available scientific theories rather than on common sense, to answer questions such as these (e.g. Burge, 2010; Kornblith, 2016). If this advice is sound, debates about the nature of perception do not need to consider common sense and its conflicts, and any epistemic motivation for addressing the problem of perception (true or traditional) in these debates is bound to be unsound.

So what normative epistemic assumptions might participants in debates about the nature of perception endorse, explicitly or implicitly, that would make it right for them to reject the naturalistic advice? While methodological debates in the philosophy of perception have revealed several potential reasons (Allen, 2020; Logue & Richardson, 2021), we suggest a significant one is the assumption that common-sense conceptions of vision and related phenomena have some sort of privileged epistemic status, namely, as defeasible default (e.g., Austin, 1957/1979) or indefeasible bedrock (e.g., Strawson, 1979). These motivate philosophers of perception to treat consistency with those conceptions as a desideratum their theories need to satisfy. Hence theorists would need to resolve apparent conflicts between the conceptions, in order to work out which bits of common sense need to be honoured. (A*) and its even stronger (if less prominent) cousins thus provide a key epistemic motivation for engaging with the problem of perception.

These epistemological assumptions all rely on the empirical assumption (A), that there is such a thing as 'the' common-sense conception of vision, which is challenged by findings of conflicts between common-sense beliefs (Sect. 4.1). Such conflicts can, but need not provide a reason to strip these beliefs of any privileged status accorded to them. What kind of conflicts

[5] Some Naïve Realists (e.g., Allen, 2020) explicitly address feelings of 'alienation' arising from conflicts with experientially saturated common sense. We set aside all non-epistemic motivations, since we doubt they generalize to many thinkers: *Pace* the still popular cognitive consistency paradigm, surprisingly few people have a pronounced preference for consistent propositional attitudes (Cialdini et al., 1995), in particular where they lack personal investment in the beliefs at issue (Kruglanski et al., 2018).

provide such a reason, and which do not, is a theoretical question requiring a separate paper. For present purposes, we suggest that a condition that is met in the present case is sufficient to defeat epistemic default status: The constitutive statements of DR and IR are *about the very same phenomenon*, and *directly contradict* each other (Sect. 4.2). According these common-sense beliefs epistemic default status means maintaining both should be accepted in the absence of good reasons against them. But then the *patent conflict* between the two conceptions constitutes a good reason to reject both conceptions, prior to any further investigation. If both qualify as common sense, then by (A*) both should be accepted in the absence of good reasons to the contrary. However, if we accept DR, its patent conflict with IR provides a good reason against IR. Conversely, if we accept IR, its patent conflict with DR provides a good reason against DR. We cannot accord both common-sense conceptions epistemic default status without providing *prima facie* reasons against accepting them. In this sense, the patent conflict between them defeats the attribution (A*) of epistemic default status to common-sense beliefs about vision.[6]

The refutation of assumption (A*) thus shows that the principal epistemic motivation for engaging with the aporetic problem is unsound. In other words, it 'dissolves' the problem of perception, or at any rate makes the key contribution to such a 'dissolution'. This novel approach coheres well with methodological naturalism: The refutation of (A*) provides a principled reason for adopting this methodological stance in philosophical theorising about the nature of perception: If there is no such thing as 'the' common-sense conception of vision that could provide at least initial common ground for such theorising, this provides a decisive further reason for philosophers of perception to turn to science. And once this stance is adopted, no sound epistemic motivation remains for addressing the problem of perception – true or traditional – in such theorising. The refutation of (A*) thus facilitates a *'naturalistic dissolution'* of the problem.

5. Conclusion

Conflicts between common-sense beliefs give rise to philosophically influential aporetic problems. By and large, philosophers address such problems when they have enough respect for common sense to want to use it as initial common ground. This respect is not warranted where laypeople are torn between two (or more) patently conflicting conceptions of one and the same phenomenon, which have equal claim to being 'common sense'. This paper showed folk conceptions of vision are a case in point: Previous studies used belief inventories to show that laypeople are, collectively and sometimes individually, torn between conceptions of vision that are consistent with Direct Realism and Indirect Realism, respectively. Against philosophically influential views of DR as naive and IR as scientific, this paper used scientific accuracy ratings to reveal that both conceptions function as naïve accounts in lay cognition, and proposed explanations that trace both to pre-scientific sources. Both conceptions thus seem to qualify as 'common sense'. The pointed and patent conflict thus revealed *within* common sense about vision shows this part of common sense does not deserve the respect philosophers of perception typically accord it; they should ignore it. Arguably, this very conflict is also the ultimate source of the problem of perception. Insight into the precise nature of the belief

[6] The fact that laypeople recognize the incompatibility of DR and IR (Fischer, Allen & Engelhardt, Ms) confirms that the pertinent beliefs contradict each other, *when understood as laypeople do.*

conflict underlying the problem of perception thus provides good reason to set the problem aside.[7]

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: OUP.

Allen, K. (2019). Merleau-Ponty and Naïve Realism, *Philosophers' Imprint,* 19(2), 1-25. http://hdl.handle.net/2027/spo.3521354.0019.002

Allen, K. (2020). The value of perception. *Philosophy and Phenomenological Research*, 100, 633-656. doi: 10.1111/phpr.12574

Anderson, M.C. & Green, C. (2001). Suppressing unwanted memories by executive control. *Nature*, 410, 366–369. doi: 10.1038/35066572

Austin, J.L. (1957/1979). A plea for excuses. Repr. in his *Philosophical Papers, 3rd ed* (pp.175–204). Oxford: OUP

Austin, J.L. (1962). *Sense and Sensibilia*. Oxford: OUP.

Barlev, M., Mermelstein, S., & German, T.C. (2018). Representational co-existence in the God concept. *Psychonomic Bulletin & Review, 25*, 2330-2338. doi: 10.3758/s13423-017-1421-6

Boghossian, P., & Velleman, D. (1989). Color as a secondary quality. *Mind*, 98, 81-103. doi: 10.1093/mind/xcviii.389.81

Brandom, R. (1994). *Making it Explicit: Reasoning, Representing, and Discursive Commitment*. Cambridge, Mass.: Harvard UP.

Brewer, B. (2011). *Perception and its Objects*. Oxford: OUP.

Burge, T. (2010). *The Origins of Objectivity.* New York: Oxford.

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. doi: 10.1037/0022-3514.69.2.318

Clark, A. (2013). Expecting the world: Perception, prediction, and the origins of human knowledge, *Journal of Philosophy*, 110(9), 469-496. doi: 10.5840/jphil2013110913

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*. Buffalo: Broadview.

Dennett, D.C. (1991). *Consciousness Explained.* Boston: Little, Brown.

Drayson, Z. (2018). Direct perception and the predictive mind. *Philosophical Studies*, 175, 3145-3164. doi: 10.1007/s11098-017-0999-x

Drayson, Z. (2021). Naturalism and the metaphysics of perception. In H. Logue & L. Richardson (Eds.), *Purpose and Procedure in the Philosophy of Perception*. Oxford: OUP.

Ferretti, T.R., McRae, K., & Hatherell, A. (2001). Integrating verbs, situation schemas, and thematic role concepts. *Journal of Memory and Language, 44*, 516-547. doi: 10.1006/jmla.2000.2728

Fischer, E. (2011). *Philosophical Delusion and its Therapy*. New York and London:

[7] For comments on previous drafts, we thank an anonymous reviewer and conference audiences in Norwich (Euro X-Phi) and Warsaw (ESPP). The research reported was funded by the British Academy (SRG22\220454).

Routledge.
Fischer, E. (2023). Critical ordinary language philosophy: A new project in experimental philosophy. *Synthese*, 201, 102. doi: 10.1007/s11229-023-04068-8
Fischer, E., Allen, K. & Engelhardt, P.E. (2023). Fragmented and conflicted: folk beliefs about vision. *Synthese*, 201, 84. doi: 10.1007/s11229-023-04066-w
Fischer, E., Allen, K. & Engelhardt, P.E. (Ms). *Direct or Indirect Realism? Assessing conflicting folk conceptions of vision*. Unpublished manuscript, University of East Anglia.
Fischer, E., Engelhardt, P.E. (2020). Lingering stereotypes: Salience bias in philosophical argument. *Mind & Language,* 35, 415–439. doi: 10.1111/mila.12249
Fish, W. (2009). *Perception, Hallucination, and Illusion*. New York: OUP
Fish, W. (2010). *Philosophy of Perception: A Contemporary Introduction*. London: Routledge.
Forstmann, M., & Burgmer, P. (2022). The cartesian folk theatre: People conceptualize consciousness as a spatio-temporally localized process in the human brain. *Journal of Experimental Psychology: General, 151*(4), 781–803. doi: 10.1037/xge0001108
Gelman, S.A., & Noles, N.S. (2011). Domains and naïve theories. *WIREs Cognitive Science, 2*(5), 490–502. doi: 10.1002/wcs.124
Genone, J. (2016). Recent work on naïve realism. *American Philosophical Quarterly*, 53(1), 1–26. http://www.jstor.org/stable/44982080
Gerken, M. & Beebe, J.R. (2016). Knowledge in and out of contrast. *Noûs*, 50, 133-164. Doi¸ 10.1111/nous.12064
Gernsbacher, M.A., & Faust, M. (1995). Skilled suppression. In F.N. Dempster & C.J. Brainerd (Eds.), *Interference and inhibition in cognition* (pp. 295–327). Academic Press.
Goldberg, R.F., & Thompson-Schill, S.L. (2009). Developmental "roots" in mature biological knowledge. *Psychological Science*, 20, 480–487. doi: 10.1111/j.1467-9280.2009.02320.x
Graziano, M.S.A. (2022). A conceptual framework for consciousness. *Proceedings of the National Academy of Sciences*, 19(18), e2116933119. Doi: 10.1073/pnas.2116933119
Graziano, M.S.A. & Kastner, S. (2011). Human consciousness and its relationship to social neuroscience: A novel hypothesis, *Cognitive Neuroscience*, 2:2, 98-113. doi: 10.1080/17588928.2011.565121
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*, 116, 328–333. doi: 10.1073/pnas.181658111
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*, 117, 13162-13167. doi: 10.1073/pnas.2003110117
Hampton, J.A. (2015). Categories, prototypes and exemplars. In: N. Riemer (Ed.), *The Routledge Handbook of Semantics* (pp. 125-141). Routledge
Hannon, M. (2018). Intuitions, reflective judgments, and experimental philosophy, *Synthese*, 195, 4147–68. doi: 10.1007/s11229-017-1412-1
Hare, M., Jones, M., Thomson, C., Kelly, S., & McRae, K. (2009). Activating event knowledge. *Cognition*, *111*, 151–167. doi: 10.1016/j.cognition.2009.01.009
Henrich, J., Heine, S.J., Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, 33(2-3), 61-83. doi: 10.1017/S0140525X0999152X
Hohwy, J. (2007). Functional integration and the mind. *Synthese*, 159(3), 315-328. doi:

10.1007/s11229-007-9240-3
Hume, D. (1777/1975). *Enquiries Concerning Human Understanding and Concerning the Principles of Morals*. P. Nidditch (Ed.). Oxford: OUP.
Jackson, F. (1977). *Perception. A Representative Theory*. Cambridge: CUP.
Kelemen, D., & Rosset, E. (2009). The human function compunction: teleological explanation in adults. *Cognition*, 111(1),138-43. doi: 10.1016/j.cognition.2009.01.001
Kneer, M., Colaco, D., Alexander, J., & Machery, E. (2021). On Second Thought: Reflections on the Reflection Defense. In T. Lombrozo, J. Knobe, and S. Nichols (Eds.), *Oxford Studies in Experimental Philosophy, Volume 4*, Oxford: OUP. doi: 10.1093/oso/9780192856890.003.0010
Knobe, J. (in press). Conflicting intuitions. *Ergo*. https://philpapers.org/archive/KNOCIG-2.pdf
Kornblith, H. (2016). Philosophical naturalism. In H. Cappelen, T. Szabo Gendler, & J. Hawthorne (Eds.), *Oxford Handbook of Philosophical Methodology* (pp.147-158). Oxford: OUP. doi: 10.1093/oxfordhb/9780199668779.013.8
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. doi: 10.1080/1047840X.2018.1480619
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. doi: 10.1111/j.1467-8624.2012.01743.x
Leiser, D. (2001). Scattered naïve theories: Why the human mind is isomorphic to internet web. *New Ideas in Psychology*, 19, 175–202. doi: 10.1016/S0732-118X(01)00007-1
Locke, J. (1700/1975): *An Essay Concerning Human Understanding*, 4th edition, ed. by P. Nidditch, Oxford: Clarendon, 1975.
Logue, H., & Richardson, L. (2021). Introduction. In their (Eds.), *Purpose and Procedure in the Philosophy of Perception*. (pp. 1-22). Oxford: OUP.
Ludwig, K. (2007). The epistemology of thought experiments: First person versus third person approaches. *Midwest Studies in Philosophy, 31*(1), 128-159. doi: 10.1111/j.1475-4975.2007.00160.x
Lyons, J. (2023). Epistemological problems of perception. In E.N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy.* Summer 2023. https://plato.stanford.edu/archives/sum2023/entries/perception-episprob/
Machery, E. (2017). *Philosophy within its Proper Bounds*. Oxford: OUP.
Macpherson, F. (2013). The philosophy and psychology of hallucination. In Macpherson, F., & Platchias, D., (Eds.), *Hallucination: Philosophy and psychology* (pp. 1-38). MIT Press.
Martin, M.G.F. (2002). The transparency of experience. *Mind and Language*, 4, 376-425. doi: 10.1111/1468-0017.00205
McRae, K., Ferretti, T. R., & Amyote, I. (1997). Thematic roles as verb-specific concepts. *Language and Cognitive Processes*, *12*, 137–176. doi: 10.1080/016909697386835
McRae, K., & Jones, M. (2013). Semantic memory. In D. Reisberg (Ed.), *Oxford Handbook of Cognitive Psychology*. Oxford: OUP. doi: 10.1093/oxfordhb/9780195376746.013.0014
Merleau-Ponty, M. (1945/2013). *The Phenomenology of Perception*. Transl. D. Landes. London: Routledge.
Plato (n.d./2004). *Theaetetus*. Transl. J. Sachs. Newburyport: Focus
Putnam, H. (1994). Sense, nonsense, and the senses: An inquiry in the powers of the human mind. *Journal of Philosophy* 91, 445-517. doi: 10.2307/2940978

Raineri, C. (2021). What's so naïve about naïve realism?. *Philosophical Studies*, 178, 3637–3657. doi: 10.1007/s11098-021-01618-z

Reid, T. (1785/1969). *Essays on the Intellectual Powers of Man*. Cambridge, Mass.: MIT Press.

Rescher, N. (2001), *Paradoxes: Their Roots, Range, and Resolution*, Chicago and LaSalle.

Roberts, P., Allen, K. and Schmidtke, K. (2020). Reflective intuitions about the Causal Theory of Perception across sensory modalities, *Review of Philosophy and Psychology* 12, 257-277. doi: 10.1007/s13164-020-00478-6

Robinson, H. (1994). *Perception*. London: Routledge.

Rumelhart. D.E. (1978). Schemata: The building blocks of cognition. In R. Spiro, B. Bruce, & W. Brewer (Eds.), *Theoretical Issues in Reading Comprehension*. Erlbaum.

Russell, B. (1912/2000). *The Problems of Philosophy*, Oxford: OUP.

Searle, J. (2015). *Seeing Things As They Are: A Theory of Perception*. New York: OUP.

Sellars, W. 1963. 'Philosophy and the Scientific Image of Man' in his: *Science, Perception and Reality* (pp. 1-40). London: Routledge.

Shtulman, A. (2017). *Scienceblind. Why our intuitive theories about the world are so often wrong*. New York: Basic Books.

Shtulman, A., & Legare, C.H. (2020). Competing explanations of competing explanations: Accounting for conflict between scientific and folk explanations. *Topics in Cognitive Science, 12*(4), 1337–1362. doi: 10.1111/tops.12483

Shtulman, A., & Valcarcel, J. (2012). Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition*, 124, 209–215. doi: 10.1016/j.cognition.2012.04.005

Sloman, S.A., Love, B.C. & Ahn, W.K. (1998). Feature centrality and conceptual coherence. *Cognitive Science*, 22, 189–228. doi: 10.1207/s15516709cog2202_2

Smith, A.D. (2002). *The Problem of Perception*. Cambridge, Mass: Harvard UP.

Snowdon, P.F. (1981). Perception, vision and causation, *Proceedings of the Aristotelian Society*, 81, 175-192. doi: 10.1093/aristotelian/81.1.1

Sommers, T. (2010). Experimental philosophy and free will. *Philosophy Compass*, 5, 199-212. doi: 10.1111/j.1747-9991.2009.00273.x

Strawson, P.F. (1959). *Individuals*. London: Methuen.

Strawson, P.F. (1979). Perception and its Objects. In G. Macdonald (Ed.), *Perception and Identity*. London: Palgrave.

Valberg, J.J. (1992). *The Puzzle of Experience.* Oxford: Clarendon.

Vosniadou, S. (1994). Capturing and modelling the process of conceptual change. *Learning and Instruction*, 4, 45–69. doi: 10.1016/0959-4752(94)90018-3

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. doi: 10.1111/cogs.12430

Webb, T.W., & Graziano, M.S.A. (2015). The attention schema theory: A mechanistic account of subjective awareness. *Frontiers in Psychology*, 6, doi: 10.3389/fpsyg.2015.00500

Williams, M. (2001). *The Problem of Knowledge*. Oxford: OUP

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. doi: 10.1037/0003-066X.57.6-7.417

Wittgenstein, L. (1933/2005). *The Big Typescript TS 213*. Oxford: Blackwell.