Diagnostic Experimental Philosophy

Eugen Fischer and Paul E. Engelhardt

RESUMEN

El muy debatido programa 'restriccionista' de la filosofía experimental pretende delinear la medida en la que los filósofos pueden legítimamente confiar en intuiciones sobre casos posibles. Este árticulo muestra que este programa puede ser (i) puesto al servicio de la resolución del problema diagnóstico (en la estela de J. L. Austin) y (ii) reivindicado para construir y comprobar experimentalmente las explicaciones psicolingüísticas de intuiciones que exponen su falta de valor evidencial: El artículo desarrolla una explicación psicolingüística de las intuiciones paradójicas que provocan descripciones verbales de casos, y presenta dos experimentos que apoyan la explicación. Esta explicación desmitificadora ayuda a resolver las paradojas sobre la percepción (conocidas como 'argumentos a partir de la alucinación').

PALABRAS CLAVE: *filosofía experimental, intuición, filosofía del lenguaje ordinario, J. L. Austin, el problema de la percepción.*

ABSTRACT

Experimental philosophy's much-discussed 'restrictionist' program seeks to delineate the extent to which philosophers may legitimately rely on intuitions about possible cases. The present paper shows that this program can be (i) put to the service of diagnostic problem-resolution (in the wake of J.L. Austin) and (ii) pursued by constructing and experimentally testing psycholinguistic explanations of intuitions which expose their lack of evidentiary value: The paper develops a psycholinguistic explanation of paradoxical intuitions that are prompted by verbal case-descriptions, and presents two experiments that support the explanation. This debunking explanation helps resolve philosophical paradoxes about perception (known as 'arguments from hallucination').

KEYWORDS: *Experiemntal Philosophy, Intuition, Ordinary Language Philosophy, J. L. Austin, Problem of Perception.*

INTRODUCTION

The present paper explores a new approach for use in experimental philosophy's much discussed 'Warrant Project'. This project deploys em-

pirical methods and findings to assess intuitions' evidentiary value and philosophers' warrant for accepting them [review: Stich and Tobia (2016)]. Most contributors to it are 'restrictionists' who seek to expose reliance on unsound intuitions, so as to restrict philosophical use to intuitions which have evidentiary value [review: Alexander (2012), pp.70-89]. Restrictionists have typically focused on intuitions adduced as evidence for philosophical theories. They have employed surveys and experiments mainly to study the sensitivity of particular intuitions to apparently otiose parameters like cultural background and socio-economic status, gender, age, personality type, order of presentation and framing effects. From sensitivity to such parameters, restrictionists infer lack of evidentiary value. This approach has attracted methodological objections [review: Fischer and Collins (2015)]. In particular, studies concerning demographic parameters (like gender or cultural background) have also encountered replication issues [see, e.g., Lam (2010); Nagel et al. (2013), Seyedsayamdost (2015a, 2015b)]. It is therefore timely to explore an alternative approach, known as the 'Sources Project' [Pust (2012)] or 'cognitive epistemology' [Fischer (2014)].

This emerging research program develops and tests psychological explanations of intuitions, which help assess their evidentiary value. It traces intuitions back to *automatic inferences*, i.e., to largely automatic cognitive processes which duplicate inferences with normative *or* heuristic rules. Like cognitive psychology, cognitive epistemology conceptualises *intuitions* as judgments based on automatic inferences [Kahneman and Frederick (2005)], which are accompanied by high levels of subjective confidence [Thompson et al. (2011)], regardless of whether or not the thinker endorses them upon reflection. The best developed approach in cognitive epistemology constructs 'GRECI explanations' [Fischer et al. (2015)] which trace intuitions back to cognitive processes which are **g**enerally **re**liable but predictably engender **c**ognitive **i**llusions, under specific circumstances. E.g., a prominent research-line traces intuitive knowledge-attributions to a 'mind-reading' competency, which is argued to be generally reliable [Boyd and Nagel (2014)] though subject to specific biases, including an egocentrism bias [Alexander et al. (2015)], focal bias [Gerken and Beebe (2016)], and 'anti-inferential bias' [Turri (2015)]. The ultimate aim is to develop 'epistemic profiles' of cognitive processes that tell us under what conditions we may (not) trust the intuitions they generate [Weinberg (2015)]. This approach has been used to defend (restricted) use of intuitions as evidence for philosophical theories [e.g., Nagel (2012)].

We seek to extend these efforts in two crucial ways [see also Fischer et al. (2015); Fischer and Engelhardt (2016)]. First, we shift the focus

from domain-specific processes (like 'mind-reading') to domain-general processes which potentially shape philosophical thought on *any* topic, but have so far been neglected by experimental philosophers: We examine routine language processes which go on automatically whenever we produce or understand texts or utterances – crucially, e.g., when thinkers formulate or read verbal descriptions of possible cases, in philosophical thought experiments or argument. Second, we explore a fresh philosophical use of the new approach, beyond philosophical theory construction: We focus on the conditions under which generally reliable cognitive processes generate cognitive illusions, and employ these debunking explanations to develop new diagnostic solutions to characteristic philosophical problems – viz., problems which ultimately arise from clashes between (unsound) paradoxical intuitions and background beliefs.

This paper considers a paradox at the root of the classical 'problem of perception' [Smith (2002)], which has again become a focus of debate [review: Crane and French (2015)]. We analyse the intuitive paradox and outline an empirically grounded 'diagnostic' approach for resolving it (Section I). To implement the approach, we then develop a psycholinguistic explanation that debunks the underlying intuitions (Section II). We finally present two experiments that test the key hypotheses on which our explanation relies (Section III). We thus explore the use of fresh psycholinguistic tools and the pursuit of diagnostic aims in experimental philosophy – adding a new 'diagnostic' string to its bow.

I. A Philosophical Paradox and a Diagnostic Approach

Many characteristically philosophical problems arise from philosophical paradoxes: Persuasive arguments lead from intuitive or uncontroversial premises to a conclusion q apparently at odds with a common-sense conviction p. This creates the impression of a difficulty about the very possibility of p, and has us wonder, 'How is it possible that p (given that q)?' [Fischer (2011), Papineau (2009)]. E.g.: Arguments 'from illusion' and 'from hallucination' both lead to the conclusion that when people use their five senses, they are (directly) aware only of subjective perceptions or sense-data. This conclusion apparently clashes with common sense and thus motivates the question: 'How is it possible for us to see and otherwise perceive physical objects (given that we are directly aware only of sense-data)?' – the 'problem of perception' [Smith

(2002)]. J. L. Austin famously argued that we should not try to answer this question, but come to see through it, by 'unpicking', in the arguments that motivate it, 'a mass of seductive (mainly verbal) fallacies' [Austin (1962), pp. 4-5]. Regardless of whether they hunt for *verbal* fallacies, *diagnostic approaches* (in our sense) try to resolve problems of this kind by showing that we lack warrant for accepting the conclusions q that engender them, so that the impression that a problem about p arises from q is unwarranted, and the question, 'How is it possible that *p*, *given that q*?', is ill-motivated.

Where the arguments for *q* rely on intuitions that are themselves accepted without any (further) evidence or argument, cognitive epistemology can provide the critical input for such a diagnostic approach, and provide it with empirical foundations: We are justified in accepting claims at odds with common sense only if we have positive reasons in their support [Williams (2001), p. 25].[1] The paradoxical argument suggests that some of the underlying intuitions are at odds with common sense. Where thinkers accept these intuitions in the absence of supporting evidence or argument, they can only point to the mere fact that they have these intuitions, to justify accepting them. Thinkers' warrant for accepting them then depends upon the intuitions' *evidentiary value*: on whether or not the mere fact the thinkers at issue have them, as and when they do, speaks for the intuitions' truth (and thus constitutes a positive reason in their support).

If available, a debunking explanation then shows us that some of the relevant intuitions are generated by a process that predictably misfires under conditions which prevail (also) in the formulation of the paradoxical argument. Such a finding reveals that these intuitions have no evidentiary value, and that the targeted thinkers lack warrant for accepting these intuitions – and hence for accepting the argument that leads from them to the paradoxical conclusion which engenders their problem [Fischer et al. (2015)]. Elsewhere, we have developed and tested such an explanation of the intuitions that underlie 'arguments from illusion' [Fischer and Engelhardt (2016)]. We now do the same for 'arguments from hallucination'. Together, our – debunking – explanations facilitate a fresh diagnostic resolution of the 'problem of perception'.

Lacking empirical knowledge about the phenomenon of hallucination, analytic proponents of the 'argument from hallucination' have typically assumed merely that it *is possible* that 'one 'perceives' a physical object which is not there at all' [Ayer (1956/1990), p. 90], and proceeded from brief descriptions of *hypothetical* cases, mainly of visual hallucination

[also e.g. Jackson (1977), Smith (2002)]. They typically distinguish between different senses of perception-verbs like "see": The relevant case-descriptions, they explain, use these verbs in a purely *phenomenal sense* which serves merely to describe people's experiences and which lacks the existential and spatial implications that "see" ordinarily has. Here is a classic statement:

> Let us take as an example Macbeth's visionary dagger: since we are concerned only with what is possible, the fact that this episode may be fictitious does not matter. There is an obvious [ordinary] sense in which Macbeth did not see the dagger; he did not see the dagger for the sufficient reason that there was no dagger there for him to see. There is another [viz., phenomenal] sense, however, in which it may quite properly be said that he did see a dagger; to say that he saw a dagger is quite a natural way of describing his experience. But still not a real dagger; not a physical object; not even the look of a physical object, if looks are open to all to see. If we are to say that he saw anything, it must have been something that was accessible to him alone, something that existed only so long as this experience lasted; in short a sense-datum [Ayer (1956/1990), p. 90].

The argument's second half then postulates that in this possible scenario the subject's experience is qualitatively indistinguishable from the experience of seeing a physical object; it assumes that qualitatively indistinguishable experiences involve awareness of the same kind of object; and it concludes that all cases of perception involve awareness of sense-data. This second half has attracted considerable criticism [review: Smith (2002)].

Already the argument's first half, however, involves a curious leap of thought: The verb "to see" is explicitly used here in a phenomenal sense, to 'describ[e] his experience' [Ayer (1956/1990), p. 90]: "He saw a dagger" is to mean 'He had an experience as of / like that of / seeing a dagger'. An experience (event) is being described by comparing it to that of seeing a certain physical object, and said to be (in some respects) similar. This does not require that the object 'seen' be around, when the subject has the experience thus described; as Ayer stresses [ibid.], "see" carries no existential or spatial implications when used in this sense. What it does require is that the experience be in some respects similar to that of seeing a solid, physical dagger (rather than, say, that of seeing the shadow of a dagger). But the case-description explicitly postulates that 'an experience of this sort is like the experience of seeing a real physical object' [ibid.], rather than like that of seeing a shadow or other non-physical object, and the second half of

the argument even assumes that the two are *exactly* alike. In the phenomenal sense, Macbeth can therefore be said to see 'a real dagger', but cannot be said to see a non-physical object. What is 'special' is the sense in which the argument uses the verb "see" (viz. the phenomenal sense), not the object 'seen' in this sense.

As typically stated, the argument from hallucination does not itself provide any further argument to show that having a hallucination and, e.g., undergoing an experience like that of seeing a dagger, involves seeing some non-physical object. Rather, the argument's protagonists simply leap from the premise that Macbeth 'did see a dagger' to the conclusion that there is a non-physical object of sight [e.g., Price (1932), pp. 28-9; Ayer (1956/1990), p. 90; Jackson (1977), pp. 50-51; Smith (2002), pp. 194-5)]. Why did they do this? And why did they do this even after stressing that the use of "see" and its cognates in this argument is special (viz., phenomenal)?

The explanation we will develop and experimentally test below rests on the hypothesis that competent speakers build up habits of spontaneous inference from familiar words. These habits are shaped by the words' dominant use. Under certain circumstances, speakers/hearers cannot help making inferences licensed by a dominant ordinary use or sense of a word, even in contexts where they use the word in a special sense that does not licence these inferences. When used in the phenomenal sense, 'Macbeth sees a dagger' states that Macbeth has an experience similar to that of seeing a dagger, but does not imply that there is anything he sees (in *any* sense of the verb). When "see" is used in its dominant visual sense, however, there is something the subject sees, namely, before his eyes, in his vicinity. Thus, a contextually inappropriate spatial inference leads from the premise that, in some (viz. the phenomenal) sense,

(1) 'Macbeth sees a dagger' to the intuition

(2) There is a dagger before Macbeth's eyes, in his vicinity.

But, by assumption, there currently is no physical object answering to the description of 'dagger' before Macbeth's eyes (though perhaps elsewhere). I.e.:

(3) There is no physical dagger before Macbeth's eyes.

Proponents of the argument therefore infer from (2) and (3) that

(4) There is a non-physical dagger before Macbeth's eyes.

Reflective integration with intuitive introspective conceptions of the mind suggests this non-physical dagger is before Macbeth's inner eye, in his mind (the most intimate vicinity).

Though very basic, the inference to (4) – a kind of disjunctive syllogism – is beyond the capabilities of automatic cognition [Reverberi et al. (2012)]. This move is controlled. The one automatic inference is to (2). We submit that the intuitive argument ultimately rests on this intuition, and (2), in turn, on contextually inappropriate spatial inferences from the phenomenal use of "see". We will now develop and test a psycholinguistic GRECI explanation (see above) that debunks this intuition, and shows that proponents of the argument lack warrant for accepting it.

II. A PSYCHOLINGUISTIC EXPLANATION

Stereotypes are sets of properties which come to mind first, and are easiest to process, when we hear those expressions. They are associated with both nouns [Hare et al. (2009)] and verbs [Ferretti et al. (2001), Harmon-Vukic et al. (2009)], and can be elicited, e.g., through sentence-completion tasks:

Joe is so easily manipulated. He is so_____.

Jack is good at manipulating people. He is so _____.

Verbs can be associated with typical features of events, agents, and 'patients', which jointly form structured stereotypes (a.k.a. 'generalized situation schemas'). E.g., manipulators are typically cunning and shrewd, and their victims naïve and gullible, so the verb "manipulate" is associated with 'cunning' and 'shrewdness' as agent-properties, and 'naiveté' and 'gullibility' as patient-properties. Such associations guide spontaneous inferences from those expressions.

Stereotypical inferences are supported by automatic activation processes in semantic memory [McRae and Jones (2013)]. According to the well-supported *graded salience hypothesis* [Fein et al. (2015), Giora (2003)], a linguistic stimulus activates – i.e. makes more readily available for use by cognitive processes from word recognition to inferencing – all semantic and stereotypical features associated with the expression, in any of its uses or senses. Crucially, the speed and strength of initial activation depends

upon the 'salience' of the sense or use. Such *salience* is a function of familiarity, conventionality, frequency, and prototypicality [Giora (2003)], i.e., of how familiar the sense or use is to the subject, of how conventional it is, of how frequently the expression is employed in this sense (rather than another), and (where appropriate) of how good examples of the relevant category it stands for in that sense. Features associated with the expression's most salient use are activated regardless of context. E.g., the ambiguous stimulus "mint" activates the probe 'candy' rapidly and strongly, even where it is clearly used in a less frequent sense (prime: 'All buildings collapsed except the mint') [Simpson and Burgess (1985), Till et al. (1988)].

Such context-insensitive activation processes which are 'stimulus-driven' initially run in parallel with context-sensitive processes which are 'expectation-driven'. Their outputs are subsequently integrated: Processes including reinforcement and decay [Oden and Spira (1983)], and suppression [Faust and Gernsbacher (1996)] may mitigate initial activation, namely, in the light of contextual cues, explicit indications of deviation from relevant stereotypes, and explicit marking of less salient uses [Givoni et al. (2013)]. Within a neo-Gricean framework, the interplay of these processes is captured by the

> *I-heuristic*: In the absence of explicit indications to the contrary, infer that situations talked about conform to stereotypes associated with the nouns and verbs used. [Levinson (2000); cp. Garrett & Harnish (2007)].

The processes indicated occur in both comprehension and production of speech and text [Levelt (1989), Pickering and Garrod (2013), Stephens et al. (2010)]. They are hence set to duplicate inferences in line with the I-heuristic not only in interpersonal communication but also in the sort of sub-vocalised speech characteristic of philosophical thought.

The strength of stereotypical association is largely a function of observed co-occurrence in the subject's physical and linguistic environment [Lund and Burgess (1996)]. In the absence of biasing media (e.g., British tabloids continually vilifying the EU), stereotypes therefore encode reasonably accurate information about the co-occurrence of features and relations in the subject's environment [Bullinaria and Levy (2007)]. In particular, when the activation of stereotypical associates interacts with context-sensitive processes, the resulting inferences in line with the I-heuristic tend to lead to reasonably accurate conclusions [Fischer et al. (2015)]. Under certain conditions, however, the underlying activation processes are liable to misfire and generate cognitive illusions. One such

condition is *extensive disparity of salience between different uses of an expression*: Where one use of an expression is very much more salient than all others, and the expression enjoys strong stereotypical association with certain features, in this most salient use, speakers/hearers are liable to infer the presence of these features also from a less salient use of the expression, regardless of contextual propriety – and even in the face of explicit indications that the expression is used in a less salient sense [Giora (2003), Simpson and Burgess (1985)]. In this case, even competent speakers/hearers will unwittingly violate the I-heuristic and make inappropriate stereotypical inferences.

We hypothesize that this happens in the argument from hallucination.[2] The verb "see" is used in at least 12 different senses (see below). These include a literal visual sense ('Bob saw the builders fixing the road'), a metaphorical epistemic sense ('Jane saw Joe's point', i.e. she knew what his point was), and an extended phenomenal sense ('Hitting his head, Jack saw stars'). According to our first hypothesis,

> H$_1$ The visual sense of "see" is much more salient than its other (e.g. epistemic or phenomenal) senses.

The phenomenal sense is employed in the case-descriptions that serve as initial premises of the argument from hallucination ('Macbeth saw a dagger'). We hypothesize further that proponents of the argument then make contextually inappropriate stereotype-driven inferences from those premises: We assume that the most salient visual sense of "S sees X" has a strong stereotypical association with spatial patient-properties, namely, with 'X is in front of S' and 'X is before S's eyes' (*directional implications*) and with 'X is around (S) to be seen' and 'X is within S's range of vision' (*proximity implications*). On this basis we hypothesize that

> H$_2$ Competent speakers/hearers infer spatial patient-properties stereotypically associated with the visual sense of "see" also from markedly less salient uses, regardless of contextual propriety.

Such a stereotypical inference, we submit, leads proponents of the argument from hallucination from non-salient uses like 'When hallucinating, Macbeth sees a dagger' to the intuitive judgment that there is a dagger before Macbeth's eyes (outer or inner). H$_1$ and H$_2$ could jointly explain the intuition we identified as the intuitive source of the argument from hallucination. But both hypotheses may well seem incredible: Various

linguists we consulted were sceptical that the visual sense of "see" should be markedly more salient than, say, the epistemic sense; and some philosophers found the suggestion that competent speakers might make spatial inferences even from sentences like 'Jane saw Joe's point' outright preposterous. We conducted three studies to test our hypotheses.

III. Experiments

In following up H_1, we bear in mind that salience is a function of familiarity, conventionality, frequency, and prototypicality. 'I see your point' does not strike us as less familiar or conventional than 'I see a knife'. We suggest that salience differences are due to differences in frequency and prototypicality. We conducted a corpus analysis to establish frequency differences, and a production experiment to examine prototypicality differences.

We analysed a sample of 1000 sentences using "see" drawn at random from the *British National Corpus* (BNC). We took a list of senses of "see" with dictionary explanations and examples from the *Macmillan Dictionary* (MEDAL), and assessed in which of these senses sentences employed the verb. We found that "see" was used far more frequently in a visual sense (68%) than in an epistemic (12.4%) or doxastic sense (9.7%), and far more frequently in either of these than in the phenomenal sense (1.1%). The remainder (8.8%) was made up of a variety of further uses, all individually below 5%. Chi-square tests showed these frequency differences were significantly different from chance ($\chi^2(3) = 1162.1$, $p < .001$). Paired comparisons (e.g. visual vs. epistemic, epistemic vs. phenomenal) showed that all differences were significant ($p < .01$) (except for epistemic vs. doxastic, $p > .05$). In particular, visual uses were statistically more frequent than epistemic and doxastic uses which, in turn, were statistically more frequent than phenomenal uses. Would they also stand out in terms of prototypicality?

Experiment 1

Approach and Predictions: Prototypicality is a matter of what kind of examples or cases of, e.g., 'seeing' come to mind first and foremost, and are deemed the best examples of, say, seeing [Smith et al. (1974)]. It is usually assessed through listing, sentence-completion, or typicality-rating tasks [Battig and Montague (1969), Chang (1986)]. We employed a sen-

tence-completion task to garner evidence for prototypicality differences between different senses of "see". In line with H_1, we predicted that, in completing sentence stems like "Jane sees___", participants will come up with significantly more visual than epistemic or phenomenal examples of seeing, i.e., produce more sentences that employ the verb in a visual than in an epistemic or phenomenal sense.

Methods: 21 undergraduate psychology students from the University of East Anglia participated for course credit. We administered a paper-and-pencil questionnaire asking them to provide up to 10 written completions of 4 sentence stems. Since philosophical statements of arguments from hallucination often employ "is aware of" alongside "see", and the former may make subjects more likely to think of non-visual objects (and hence, subsequently, of non-visual senses of "see"), we presented stems with both verbs, viz. Jane sees____, Cathy was aware of ____, Bob saw____, Harry is aware of____, in this order. To discourage mono-topical response strategies, participants were instructed to 'try to give varied responses'. Two judges were given dictionary explanations and examples from MEDAL (as above) and asked which senses were used by the "see"-sentences participants produced. They also assessed whether "aware"-sentences employed this verb in a perceptual, or non-perceptual manner, guided by the question: Is the protagonist seeing, hearing, or feeling whatever it is s/he is said to be aware of, in the situation talked about? Ambiguous sentences were not classified.

Results: For "see" stems, we could classify 308 out of 312 completions. Of these, 269 (87%) clearly employed MEDAL's sense 1 ('notice with eyes') and 14 (4.5%) clearly sense 2 ('meet/visit someone'). 7 (2.2%) employed sense 4 (know or understand something), 3 (1%) sense 5 ('consider in a particular way'), 3 (1%) sense 6 ('imagine someone/something'), 2 (0.6%) sense 7 ('find something out'), and 5 (1.6%) the phenomenal sense. For further analysis, we grouped together *visual senses* 1 and 2, *doxastic senses* 5 and 6, *epistemic senses* 4 and 7, and compared them with the *phenomenal sense*: 288 (93.5%) of classified completions employed a visual sense,[3] 9 (2.9%) an epistemic sense, and 6 (1.9%) a doxastic sense. A chi-square test revealed that the observed frequency differences between visual, epistemic, doxastic, and phenomenal senses were significantly different from chance ($\chi^2(3) = 771.0$, $p < .0001$). A comparison of visual versus epistemic uses showed that the observed frequencies were also significantly different from chance ($\chi^2(1) = 262.1$, $p < .001$). The comparison of epistemic and doxastic uses revealed no significant difference, as did their comparison

with the phenomenal use ($p > .20$). The latter is arguably due to the small sample size and the huge preponderance of visual uses. The visual use of "see" is much more prototypical than any of the other uses.

For "aware" stems, we could classify 298 out of 324 sentence completions. Of those 160 were *non-perceptual* and 138 were *perceptual*. A chi-square test revealed that the observed frequencies were not significantly different from chance ($\chi^2(1) = 1.62$, $p > .20$). Thus, there was no significant preference to apply "S is aware of X" to either perceptual or non-perceptual cases. Hence "aware" may indeed make people think of non-visual objects and senses of "see", and we investigated whether the number of non-visual completions for "see" increased from the first "see"-stem (presented before the first "aware"-stem) and the second (presented after the "aware"-stem). We found an equal number (10) of non-visual uses in completions of both; in the second "see" stem, they occurred a little earlier than in the first (mean positions: 4.6 vs. 3.2), but an independent samples t-test revealed this difference in ordinal position was not statistically significant ($t(18) = 1.37$, $p = .19$).

Discussion: This means that competent speakers/hearers think first and foremost of visual cases of seeing – even when they have been primed to think of non-perceptual cases (through completions of an "aware" stem). The fact that 94% of completions made visual use of "see" and only 3% and well under 2%, respectively, used the epistemic or phenomenal sense, strongly suggests that the disparity in prototypicality between these uses is even larger than the already extensive disparity in frequency that emerged from our corpus analysis. Since salience is largely a function of frequency and prototypicality, our findings are fully consistent with our hypothesis H_1 and establish the kind of extensive disparity in salience between different uses of "see" that may, according to our hypothesis H_2, lead competent speakers/hearers to infer features that are stereotypically associated with the most salient use, also from less salient uses that are devoid of those associations.

Experiment 2

Approach and Predictions: In examining H_2, we focused on spatial (directional and proximity) inferences from epistemic uses of "see". These uses are no less (and probably more) prototypical than the phenomenal uses of philosophical interest (Exp.1), are demonstrably more frequent (Corpus Analysis), and probably more familiar and conventional. In a nutshell, they are more salient than phenomenal uses. Demonstrating

that competent speakers make spatial inferences from epistemic uses hence warrants the stronger conclusion that such inferences will be made not only from the uses examined but also from other uses of comparable and lower salience – including doxastic and phenomenal uses.

We used a forced-choice plausibility-ranking task:[4] Participants have to indicate which of two similar sentences (a or b) strikes them as more plausible. Judgments must be made as quickly as possible and even in the absence of a clear-cut preference. Sentences differ only in a word or phrase ('minimal pairs'), e.g.:

44a. Mona is aware of the drivers ahead of her in the queue.

44b. Mona sees the drivers ahead of her in the queue.

56a. Laura sees the drivers behind her in the queue.

56b. Laura is aware of the drivers behind her in the queue.

The sequel or context is either consistent (as in 44) or inconsistent (as in 56) with the hypothesised stereotypical inference from the verb of interest (e.g. directional inference from "see" to 'The drivers she sees are in front of her'), i.e., *'s-consistent'* or *'s-inconsistent'*.

Suppose that directional inferences are made from "S sees X" to 'X is in front of S'. Then their clash with s-inconsistent sequels as in 56 ('behind her') will reduce the subjective plausibility of the overall sentence. But no such inferences will be made from the spatially neutral "is aware of". So if participants make directional inferences from "see", their preference for "see" – over "aware"–sentences will be significantly lower in s-inconsistent items like 56, than with s-consistent items like 44. Our hypothesis H_2 then translates into the bold prediction that this will hold true even where "see" is used in a metaphorical, epistemic sense clearly devoid of spatial implications, as in:

13a. Joe sees the problems that lie ahead.

13b. Joe is aware of the problems that lie ahead.

99a. Jack sees the problems he left behind.

99b. Jack is aware of the problems he left behind.

I.e.: Even in such metaphorical contexts, participants will prefer 'see'-sentences significantly less often in s-inconsistent items (like 99) than in s-consistent items (like 13).[5]

Methods: 49 undergraduate psychology students from the University of East Anglia participated for course credit. We administered a paper-and-pencil questionnaire with 120 minimal pairs. 24 critical items contrasted "see" and "aware", to test for directional inferences, and 24 tested for proximity inferences by contrasting "see" and "think of" (where it is evident that you can just as well think of things that are not around). Critical items consisted of pairs of s-consistent and s-inconsistent items, for visual and epistemic contexts, respectively, with 6 items in each of the following categories:

Visual, s-consistent

29a/b. Hillary spent the week at home. She saw/thought of her parents a lot.

Visual, s-inconsistent

36a/b. Shirley spent the week all alone. She saw/thought of her parents a lot.

Epistemic, s-inconsistent

50a/b. Bill thinks of/sees points to be made in favour of the motion. There are none at hand.

Epistemic, s-consistent

55a/b. Will thinks of/sees points to be made in favour of the motion. There are a couple at hand.

Epistemic contexts were created by combining "see" with objects that can be known or understood but not literally 'seen' (like problems or points in favour). The design of the study was 2 × 2 (visual/epistemic × s-consistent/s-inconsistent). Separate repeated-measures ANOVAs (with follow-up t-tests) were run on see/aware and see/think items separately.

Results confirmed our predictions: In visual cases, "see" was preferred over "aware" 96% of the time in contexts consistent with *directional inferences* (like item 44), but only 5% of the time in contexts inconsistent with them (like item 56). This drop was significant (comparison s-consistent/s-inconsistent: $t(48) = 45.35$, $p < .001$). Crucially, a

similar drop, if from lower heights, could be observed in epistemic cases. In these cases, metaphorical "see" and "aware" were felt to be equally good choices in s-consistent contexts (like item 13): metaphorical "see" was preferred 44% and "aware" 56% of the time. The slight preference for "aware" in this epistemic/s-consistent condition was not significantly different from chance ($t(48) = 1.90$, $p = .06$). But, crucially, metaphorical "see" was preferred over "aware" only 8% of the time when the context was inconsistent with directional inferences (as in item 99) – a significant drop (comparison s-consistent/s-inconsistent: $t(48) = 9.26$, $p < .001$). See Figure 1.
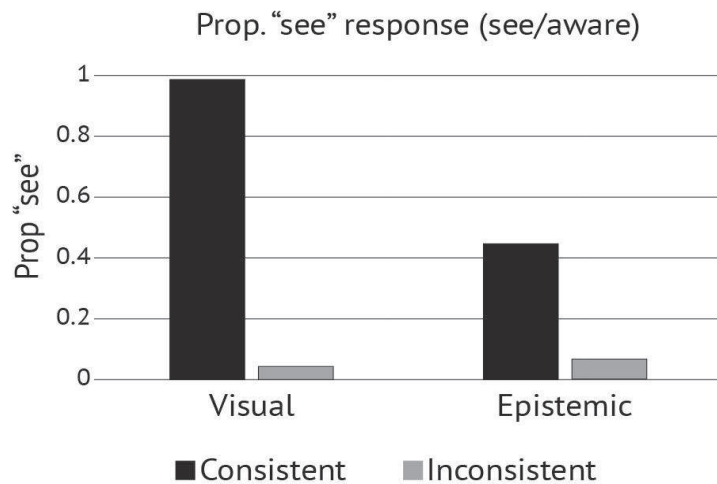

Figure 1 – Exposing directional inferences: Results showing proportion of preferences for "see" over "aware" in visual and epistemic conditions, for s-consistent and s-inconsistent contexts.

A similar pattern emerged in testing for *proximity inferences*: In visual cases, "see" was preferred over "think of" 95% of the time in contexts consistent with such inferences (like item 29), but only 2% of the time in contexts inconsistent with them (like item 36). This drop was significant (comparison s-consistent/s-inconsistent: $t(48) = 43.83$, $p < .001$). Again, there was a similar drop, if from lower heights, in epistemic cases: In these cases, metaphorical "see" was preferred over "think of" 58% of the time, in s-consistent contexts (like item 55) – a marginal preference

still shy of statistical significance ($t(48) = 1.99$, $p = .053$). But, crucially, metaphorical "see" was preferred over "think of" only 5% of the time when the context was inconsistent with proximity inferences (as in item 50) – another significant drop (comparison s-consistent/s-inconsistent: $t(48) = 12.75$, $p < .001$). See Figure 2.
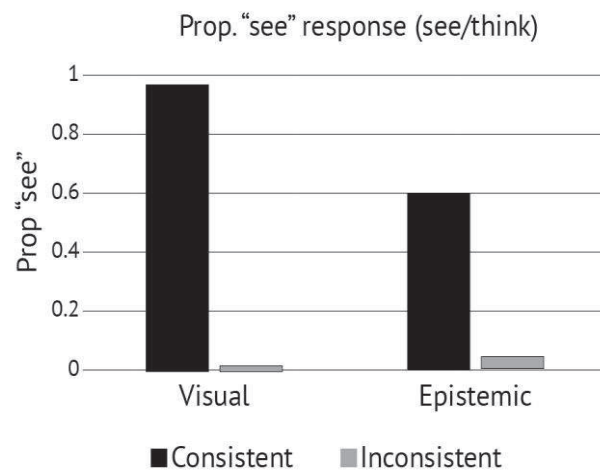

Figure 2 – Exposing proximity inferences: Results showing proportion of preferences for "see" over "think of" in visual and epistemic conditions, for s-consistent and s-inconsistent contexts.

Discussion: Our hypothesis H₂ implies that where "see" is used in a less salient metaphorical sense in which it is clearly devoid of spatial implications, competent speakers will nonetheless reject "see" in favour of verbs that work equally well in the relevant contexts but lack spatial associations. This is precisely what we found: As long as items were s-consistent, participants had no pronounced preference for "aware" or "think of" over the evidently metaphorical epistemic use of "see" – these verbs were deemed to work roughly equally well in epistemic/doxastic contexts. By contrast, participants had a very pronounced preference for "aware" and "think of" in such contexts when items were inconsistent with the spatial implications of "see" – despite their evident irrelevance. This suggests that participants made spatial (directional and proximity) inferences from the verb even in these inappropriate contexts: These inferences resulted in

perceived inconsistencies with the context, which in turn lowered the subjective plausibility of the 'see'-sentences judged in s-inconsistent contexts.

IV. CONCLUSION

These experiments and previous corpus analysis support hypotheses H_1 and H_2 (Section III). These hypotheses provide a psycholinguistic explanation (Section II) of the intuitions we have identified as the root of philosophical arguments from hallucination (Section I). This explanation debunks these intuitions: Arguments from hallucination employ a non-salient (phenomenal) use of a verb ("see") that simultaneously has a use with very high salience. In case of such extensive disparity in salience, stereotype-driven automatic inferences are liable to take us from true premises which correctly make a non-salient use of the verb ('Macbeth sees a dagger'), to intuitive judgments that are contextually inappropriate ('There is a dagger before Macbeth's eyes', inner or outer), whose fluency endows them with high subjective plausibility [Alter and Oppenheimer (2009), Thompson et al. (2011)], even so. The finding that a case-description employs non-salient uses of words which have highly salient different uses in ordinary discourse thus provides an 'undermining defeater' [Pollock (1984)] for any stereotype-driven intuitions prompted by the description: The mere fact that competent speakers/hearers have this intuition in response no longer *eo ipso* speaks for the intuition's truth, and gives us no warrant to believe that something of the sort will be true in or of the case described. As we have seen at the outset (Section I), such debunking explanations of argumentatively unsupported intuitions facilitate a fresh diagnostic resolution of the targeted paradoxes.[6]

School of Politics, Philosophy,
Language and Commuication Studies
University of East Anglia
Norwich NR4 7TJ, UK
E-mail: E.Fischer@uea.ac.uk

School of Psychology
University of East Anglia
Norwich
NR4 7TJ, UK
E-mail: p.engelhardt@uea.ac.uk

NOTES

[1] This paragraph's argument does *not* rely on the other, more controversial half of the default-and-challenge model of justification: We do not assume that

thinkers are always justified in accepting claims endorsed by common sense, in the absence of positive reasons for doubt.

2 Elsewhere [Fischer and Engelhardt (2016)], we develop and test the hypothesis that related, if weaker, salience effects are at the root of arguments from illusion: These rely on stereotypical inferences from non-salient, but typically unmarked (phenomenal) uses of appearance verbs ('X looks F [to S]', etc.) to contextually inappropriate attributions of doxastic patient-properties ('S is inclined to think that X is F').

3 This includes 5 completions that were ambiguous between senses 1 and 2 (e.g. "Bob saw all of his friends"), not counted above.

4 For a full explanation of this powerful experimental paradigm, see Fischer and Engelhardt (2016).

5 In these items, also the spatial expressions are used metaphorically (e.g. 'lie ahead' = 'in the future'). However, if spatial inferences are made, the impression of a conflict will arise in items like 99 from the fact that salient (here: spatial) meanings are immediately activated and retained for interpretation of spatial time metaphors [Giora and Fein (1999), Fein et al. (2015), cp. Gentner et al. (2002)].

6 For help with development of materials, corpus annotation, data gathering and entry, we thank Oliver Afridijanta, Mark Curtis, Ian Hare, and Aurelie Herbelot. For helpful comments, we thank Rachel Giora and an audience at the 7th Annual Conference of the Experimental Philosophy Group UK (Reading, April 2016).

REFERENCES

ALEXANDER, J. (2012), *Experimental Philosophy*, Cambridge, Polity.

ALEXANDER, J., GONNERMAN, C., and WATERMAN, J. (2015), 'Salience, and Epistemic Egocentrism', in J. Beebe (ed.), *Advances in Experimental Epistemology*, London, Bloomsbury.

ALTER, A. L. and OPPENHEIMER, D. M. (2009), 'Uniting the Tribes of Fluency to Form a Metacognitive Nation', *Personality and Social Psychology Review*, 13, pp. 219-235.

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

AYER, A.J. (1956), *The Problem of Knowledge*, repr. 1990, London, Penguin.

BATTIG, W. F., and MONTAGUE, W. E. (1969), 'Category Norms for Verbal Items in 56 Categories: A Replication and Extension of the Connecticut Category Norms', *Journal of Experimental Psychology*, 80, pp. 1-46.

BOYD, K. and NAGEL, J. (2014), 'The Reliability of Epistemic Intuitions', in E. Machery and E. O'Neill (eds.), *Current Controversies in Experimental Philosophy* (pp. 109-127), London, Routledge.

BULLINARIA, J. A., and LEVY, J. P. (2007), 'Extracting Semantic Representations from Word Co-Occurrence Statistics: A Computational Study', *Behavior Research Methods*, 39, pp. 510-526.

CHANG, T. M. (1986), 'Semantic Memory: Facts and Models', *Psychological Bulletin*, 99, pp. 199-220.

CRANE, T. and French, C., (2015), 'The Problem of Perception', in N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, Summer 2015, http://plato. stanford.edu/entries/ perception-problem/.

FAUST, M., and GERNSBACHER, M. A. (1996), 'Cerebral Mechanisms for Suppression of Inappropriate Information During Sentence Comprehension', *Brain and Language*, 53, pp. 234-259.

FEIN, O., YEARI, M., and GIORA, R. (2015), 'On the Priority of Salience-Based Interpretations: The Case of Sarcastic Irony', *Intercultural Pragmatics*, 12, pp. 1-32.

FERRETTI, T., MCRAE, K., and HATHERELL, A. (2001), 'Integrating Verbs, Situation Schemas, and Thematic Role Concepts', *Journal of Memory and Language*, 44, pp. 516-547.

FISCHER, E. (2011), *Philosophical Delusion and its Therapy,* New York, Routledge.

— (2014), 'Philosophical Intuitions, Heuristics, and Metaphors', *Synthese*, 191. pp. 569-606.

FISCHER, E., and COLLINS, J. (2015), 'Rationalism and Naturalism in the Age of Experimental Philosophy', in their (eds.), *Experimental Philosophy, Rationalism and Naturalism* (pp. 3-33), London, Routledge.

FISCHER, E., and ENGELHARDT, P. E. (2016), 'Intuitions' Linguistic Sources'. Stereotypes, Intuitions, and Illusions', *Mind & Language*, 31, pp. 65-101.

FISCHER, E., ENGELHARDT, P. E., and HERBELOT, A. (2015), 'Intuitions and Illusions. From Explanation and Experiment to Assessment', in E. Fischer & J. Collins (eds.), *Experimental Philosophy, Rationalism and Naturalism*, (pp. 259-292), London, Routledge.

GARRETT, M., and HARNISH, R. M. (2007), 'Experimental Pragmatics: Testing for Implicatures', *Pragmatics and Cognition*, 17, pp. 245-262.

GERKEN, M., and BEEBE, J. (2016), 'Knowledge In and Out of Contrast', *Nous*, 50, pp. 133-164.

GENTNER, D., IMAI, M., and BORODITSKY, L. (2002), 'As Time Goes By: Evidence for Two Systems in Processing Space-Time Metaphors', *Language and Cognitive Processes*, 17, pp. 537-565.

GIORA, R. (2003). *On Our Mind. Salience, Context, and Figurative Language*, Oxford, OUP.

GIORA, R., and FEIN, O. (1999), 'On Understanding Familiar and Less-Familiar Figurative Language', *Journal of Pragmatics*, 31, pp. 1601-1618.

GIVONI, S., GIORA, R., and BERGERBEST, D. (2013), 'How Speakers Alert Addressees to Multiple Meanings', *Journal of Pragmatics*, 48, pp. 29-40.

HARE, M., JONES, M., THOMSON, C., KELLY, S., and MCRAE, K. (2009), 'Activating Event Knowledge', *Cognition*, 111, pp. 151-167.

HARMON-VUKIĆ, M., GUÉRAUD, S., LASSONDE, K. A. and O'BRIEN, E. J. (2009), 'The Activation and Instantiation of Instrumental Inferences', *Discourse Processes*, 46, pp. 467-490.

JACKSON, F. (1977), *Perception. A Representative Theory*, Cambridge, CUP.

KAHNEMAN, D. and FREDERICK, S. (2005), 'A Model of Heuristic Judgment', in K. J. Holyoak and R. Morrison (eds.), *Cambridge Handbook of Thinking and Reasoning* (pp. 67-293), Cambridge, CUP.

LAM, B. (2010), 'Are Cantonese-Speakers Really Descriptivists? Revisiting Cross-Cultural Semantics', *Cognition*, 115, pp. 320-329.

LEVELT, W. J .M. (1989), *Speaking: From Intention to Articulation*, Cambridge, MA, MIT Press.

LEVINSON, S. C. (2000), *Presumptive Meanings. The Theory of Generalized Conversational Implicature*, Cambridge, Mass., MIT Press.

LUND, K., and BURGESS, C. (1996), 'Producing High-Dimensional Semantic Spaces from Lexical Co-Occurrence', *Behavior Research Methods*, *Instrucment, & Computers*, 28, pp. 203-208.

MCRAE, K., and JONES, M. (2013), 'Semantic Memory', in D. Reisberg (ed.), *Oxford Handbook of Cognitive Psychology*, Oxford, OUP, pp. 206-19.

NAGEL, J. (2012), 'Intuitions and Experiments: A Defence of the Case Method in Epistemology', *Philosophy and Phenomenological Research*, 85, pp. 495-527.

ODEN, G. G. and SPIRA, J. L. (1983), 'Influence of Context on the Activation and Selection of Ambiguous Word Senses', *Quarterly Journal of Experimental Psychology*, 35, pp. 51-64.

PAPINEAU, D. (2009). 'The Poverty of Analysis', *Aristotelian Society Supplementary Volume*, 83, pp. 1-30.

PICKERING, M. J. and GARROD, S. (2013). 'An Integrated Theory of Language Production and Comprehension', *Behavioral and Brain Sciences*, 36, pp. 329-392.

POLLOCK, J. (1984), 'Reliability and Justified Belief', *Canadian Journal of Philosophy*, 14, pp. 103-114.

PRICE, H. H. (1932), *Perception*, 2nd ed., repr. 1961, London, Methuen.

PUST, J. (2012). 'Intuition', in: E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, http://plato.stanford.edu/ archives/win2012/entries/intuition/.

REVERBERI, C., PISCHEDDA, D., BURIGO, M. and CHERUBINI, P. (2012), 'Deduction Without Awareness', *Acta Psychologica*, 139, pp. 244-253.

SEYEDSAYAMDOST, H. (2015a), 'On Normativity and Epistemic Intuitions: Failure of Replication', *Episteme*, 12, pp. 95-116.

— (2015b), 'On Gender and Philosophical Intuition: Failure of Replication and Other Negative Results', *Review of Philosophy and Psychology*, 28, pp. 642-673.

SIMPSON, G. B., and BURGESS, C. (1985), 'Activation and Selection Processes in the Recognition of Ambiguous Words', *Journal of Experimental Psychology: Human Perception and Performance*, 11, pp. 28-39.

SMITH, A. D. (2002), *The Problem of Perception*, Cambridge, Mass, Harvard U.P.

SMITH, E. E., SHOBEN, E., and RIPS, L. J. (1974), 'Structure and Process in Semantic Memory: A Featural Model for Semantic Decisions', *Psychological Review*, 81, p. 214-241.

STEPHENS, G. J., SILBERT, L. J. and HASSON, U. (2010), 'Speaker–Listener Neural Coupling Underlies Successful Communication', *PNAS*, 107, pp. 14425–14430.

STICH, S., and TOBIA, K. (2016), 'Experimental Philosophy and the Philosophical Tradition', in J. Sytsma and W. Buckwalter (eds.), *Blackwell Companion to Experimental Philosophy*, Wiley Blackwell, Malden, pp.5-21.

THOMPSON, V. A., PROWSE TURNER, J. A. and PENNYCOOK, G. (2011), 'Intuition, Reason, and Metacognition', *Cognitive Psychology*, 63, pp. 107-140.

TILL, R. E., MROSS, E. F., and KINTSCH, W. (1988), 'Time Course of Priming for Associate and Inference Words in a Discourse Context', *Journal of Verbal Learning and Verbal Behaviour*, 16, pp. 283-298.

TURRI, J. (2015), 'Skeptical Appeal: The Source-Content Bias', *Cognitive Science*, 39, pp. 307-324.

WEINBERG. J. (2015), 'Humans as Instruments, On the Inevitability of Experimental Philosophy', in: E. Fischer and J. Collins (eds.), *Method, Rationalism, and Naturalism. Rethinking Philosophical Method* (pp. 171-187), London, Routledge.

WILLIAMS, M. (2001), *Problems of Knowledge*, Oxford, OUP.