Synthese (2023) 201:166
https://doi.org/10.1007/s11229-023-04149-8

ORIGINAL RESEARCH

X-Phi and the challenge from ad hoc concepts

Michelle Liu[1,2]

Received: 28 August 2022 / Accepted: 3 April 2023

Abstract
Ad hoc concepts feature prominently in lexical pragmatics. A speaker can use a word or phrase to communicate an ad hoc concept that is different from the lexically encoded concept and the hearer can construct the intended ad hoc concept pragmatically during utterance comprehension. I argue that some philosophical concepts have origins as ad hoc concepts, and such concepts pose a challenge for experimental philosophy regarding these concepts. To illustrate this, I consider philosophers' 'what-it's-like'-concepts and experimental philosophy of consciousness.

Keywords Ad hoc concepts · What it's like · Lexical pragmatics · Consciousness · Experimental philosophy

1 Introduction

In this paper, I defend three claims:

(i) There are ad hoc concepts.
(ii) Some philosophical concepts find their origins as ad hoc concepts.
(iii) Philosophical concepts which find their origins as ad hoc concepts pose a challenge for experimental philosophy regarding these concepts.

Ad hoc concepts are concepts communicated by speakers that are different from lexically encoded concepts and constructed pragmatically by hearers during utterance comprehension. They feature prominently in relevance theory, a post-Gricean inferential approach to pragmatics (e.g. Sperber & Wilson, 1998; Wilson & Sperber, 2002), especially in the work of Robyn Carston (2002, 2010a, 2010b, 2012, 2019, 2021; also Wilson & Carston, 2007).

Ad hoc concept constructions are not only commonplace in comprehending ordinary utterances, but also occur in philosophical discussions. Some philosophical

✉ Michelle Liu
michelle.liu1@monash.edu

1 University of Hertfordshire, Hatfield, UK

2 Monash University, Melbourne, Australia

Published online: 03 May 2023

concepts, as I argue, find their origins as ad hoc concepts, i.e. (ii). I illustrate this using 'what it's like'-concepts, which are commonly used in philosophical discussion on consciousness.

Philosophical concepts which find their origins as ad hoc concepts pose a challenge for corresponding experimental philosophy regarding these concepts, i.e. (iii). Focusing on the X-Phi literature on consciousness as a case study, I problematise the commonly-made moves from certain empirical data concerning laypeople's judgements to conclusions that laypeople do not have or grasp relevant philosophical concepts, do not have philosophers' intuitions featuring these concepts, or that philosophical ideas relying on these concepts are problematic. While the paper focuses on the case of consciousness, the argument here generalises. Such moves can be problematic when the relevant philosophical concept finds its origin as an ad hoc concept, or when having the relevant intuition presupposes the acquisition of an ad hoc concept. The challenge raised here is in line with the kind of 'challenge from ordinary language' that drawing conclusions about philosophically significant concepts solely on the basis of judgements elicited in experimental settings can easily be unjustified (see Hansen, 2020; Baz, 2012, 2014, 2015).

The structure of the paper is as follows. Section 2 focuses on (i) and clarifies what ad hoc concepts are. Section 3 argues for (ii) by drawing on recent work on the semantics of 'what it's like', a phrase that philosophers routinely use to refer to phenomenal properties of conscious states. Section 4 addresses (iii) and illustrates it with reference to experimental philosophy of consciousness. Section 5 concludes the paper.

2 Ad hoc concepts

Carston (2002, p. 322) introduces ad hoc concepts as follows:

> This term is used to refer to concepts that are constructed pragmatically by a hearer in the process of utterance comprehension. The idea is that speakers can use a lexically encoded concept to communicate a distinct non-lexicalized (atomic[1]) concept, which resembles the encoded one in that it shares elements of its logical and encyclopaedic entries, and that hearers can pragmatically infer the intended concept on the basis of the encoded one. The description of such concepts as 'ad hoc' reflects the fact that they are not linguistically given, but are constructed online (on the fly) in response to specific expectations of relevance raised in specific contexts.

The passage points out two important features of ad hoc concepts corresponding to the two sides of communication. First, on the speaker's side, a speaker can use a word or phrase to communicate an ad hoc concept that is different from the lexically encoded concept. Second, on the audience's side, a hearer can construct the intended ad hoc concept through pragmatic inference. I shall elaborate on these two features

[1] Contra Carston, one need not commit to the idea that ad hoc concepts are atomic. Allott and Textor (2012) argue that ad hoc concepts are clusters of information, which are distinct from concepts proper. For them, the latter are memory addresses that give access to a body of information associated with an individual or kind.

in turn. The following exposition on ad hoc concepts shall remain brief since detailed discussion can be found in the literature (see especially Carston, 2002; Allott & Textor, 2012).

2.1 Speaker-side

To illustrate the first feature, consider the following scenario:

(1) Meera and Luke are watching John skilfully swinging on the monkey bars. Meera, pointing at John, says to Luke: 'John is such a *gibbon*.'

The word 'gibbon', according to its encoded meaning, denotes a type of non-human primate. In uttering 'John is such a gibbon', Meera uses the lexically encoded concept GIBBON to communicate a non-lexicalised ad hoc concept GIBBON*,[2] conveying the meaning that John is good at performing the locomotion of brachiating like a gibbon. GIBBON* is a non-lexicalised concept that is distinct from the lexicalised concept GIBBON.[3]

According to relevance theory, ad hoc concepts contribute to the truth-conditional contents of utterances which express such concepts (see Wilson & Carston, 2007). An ad hoc concept conveys a sense that is often either broader or narrower than the encoded one. The former is known as 'semantic broadening' and the latter 'semantic narrowing'. In the case of semantic broadening, the extensions denoted by the encoded sense form a subset of those denoted by the ad hoc concept. In the case of (1), the extensions of GIBBON* form a category of entities—including both gibbons and humans—that is broader than the category denoted by the lexically encoded meaning of 'gibbon', which would only include gibbons. Consider also (2):

(2) The bathwater is *boiling*.

Used as a hyperbole, (2) expresses the ad hoc concept BOILING*, denoting the property of *being very hot*, which is more general than the property of *boiling*. In the case of semantic narrowing, the ad hoc concept picks out a subset of the extensions denoted by the encoded sense. Consider (3):

(3) He is more of a *man* than you will ever be.

In (3), the extensions denoted by the ad hoc concept MAN* forms a subset of the extensions denoted by the encoded sense. The ad hoc concept refers to a specific type of men, likely the type that embodies stereotypical features such as bravery and honour.

[2] Throughout the paper, I follow the convention of using '*' to label ad hoc concepts.

[3] Meera's utterance in (1) is an instance of metaphor. While relevance theorists appeal to ad hoc concepts to explain metaphor, other theorists who appeal to ad hoc concepts for explanatory purposes (e.g. explaining loose use, polysemy, etc.) might not explain metaphor this way. Here, I am not committed to the relevance theorists' way of explaining metaphor. (1) is merely used as an example to make sense of the notion of ad hoc concepts.

2.2 Audience-side

Let us now turn to the audience-side of communication. A hearer can latch onto an ad hoc concept in a conversational context through the process of pragmatic inference. This process draws on the hearer's knowledge of the encoded meaning of the word, and its logical and encyclopaedic entries. The logical entry of a concept refers to 'a set of inference rules, or "meaning postulates", which capture certain analytic implications of the concept' (Carston, 2002, p. 321). For instance, the logical entry of the lexically encoded concept GIBBON contains an inference rule that renders the output 'a type of non-human primate'. The encyclopaedic entry of a concept contains background knowledge associated with the concept, including commonsense knowledge, scientific information, cultural beliefs and idiosyncratic experiences (op. cit.). The encyclopaedic entry of GIBBON may contain the scientific information that gibbons are endangered and the commonsense knowledge that they have long arms, live in rainforests, and are good at brachiating.

Consider scenario (1) again:

(1) Meera and Luke are watching John skilfully swinging on the monkey bars. Meera, pointing to John, says to Luke: 'John is such a *gibbon*.'

The mention of 'gibbon' activates the logical entry of the word (e.g. a gibbon is a type of non-human primate), as well as its encyclopaedic entry, which contains a number of features, including:

(4) *Encyclopaedic entry of gibbon*

a. have long arms
b. very agile
c. good at swinging (from branch to branch)
d. live in rainforests in Southeast Asia.

(4) contains stereotypical properties associated with gibbons, e.g. (a), (b), and (c). Such features are likely to be easily accessible when the word 'gibbon' is mentioned. In the context of (1), the speaker (Meera) directs the attention of the hearer (Luke) to the way John skilfully goes on the monkey bars. Assuming John does not have particularly long arms, stereotypical features of gibbons pertaining to agility and arm-based locomotion, e.g. (b), (c), are likely to receive high activation and thus be easily accessible. Given the context, Meera's utterance is likely to generate the following contextual implications for Luke:

(5) a. John is very agile.
b. John is good at swinging (from bar to bar).

The derivation of (5) relies on (4b) and (4c). Note that the logical entry of 'gibbon' encodes the entailment that gibbons are a type of non-human primate. But John is a human. Contextual implications in (5) are only derivable if the concept GIBBON is adjusted to the ad hoc concept GIBBON* both to include certain stereotypical features of gibbons and be broadened to refer to creatures with these features. In this way, our hearer, Luke, arrives at the truth-conditional content of Meera's utterance, i.e. JOHN IS SUCH A GIBBON*, and thereby successfully constructs the ad hoc concept GIBBON*.

So, the hearer starts with the encoded meaning of the speaker's utterance 'John is such a gibbon'. Drawing on stereotypical features in the encyclopaedic entry of the word 'gibbon', which are made accessible by the conversational context, the hearer arrives at an interpretation of the speaker's utterance which includes contextual implications like (5) and constructs the corresponding ad hoc concept GIBBON*.

It is worth noting that once an ad hoc concept is conventionalised, processing utterances involving it does not require the abovesaid pragmatic adjustment. Consider (6):

(6) a. The policies are *written in stone*.
b. Starmer was an *evangelist* for the EU.

Ad hoc concepts, when conventionalised, give rise to polysemy or idioms (see Carston, 2019, 2021). In (6a), the idiom 'written in stone' means 'difficult to alter'. This idiomatic meaning, which is conventionalised, has been broadened from the constructed literal meaning. In (6b), the dominant and original sense of the word 'evangelist' denotes a person who preaches to others in the hope of converting them to the Christian faith. The word has another sense, as used in (6b), denoting someone who is a firm advocate of something. This sense of the word is conventionalised. In comprehending (6b), a hearer is likely to access this encoded sense directly without pragmatic inference. However, it is important to note that a concept may be conventionalised for some language users while remaining ad hoc for others. For instance, the relevant sense of 'evangelist' in (6b) may be familiar for some speakers, but for others only the dominant Christian sense alone may be familiar. In the latter case, in comprehending (6b), the hearer would still need to latch onto the ad hoc concept—EVANGELIST*—through pragmatic inference from the dominant, lexically encoded sense of the word.

3 Consciousness and the semantics of 'what it's like'

On the assumption that ad hoc concepts are common, as suggested by Carston and others (e.g. Bezuidenhout, 2001; Recanati, 2004, p. 75),[4] I shall explore what follows for philosophical concepts and X-Phi. In this section, I argue that some philosophical concepts also find their origins as ad hoc concepts. This is the second thesis that I wish to put forward, i.e. (ii). In supporting (ii), I show that philosophers' 'what it's like'-concepts, which are routinely used in philosophical discussion on conscious experiences, are such concepts.

Phenomenal consciousness (which I shall refer to using the term 'consciousness') is standardly introduced and defined using the Nagelian 'what it's like'-locution ('WIL-locution' for short) and the related 'there's something it's like'-locution ('TSIL-locution' for short): a mental state is *conscious* if *there is something it is like* to have that state (Nagel, 1974). Appealing to the recent literature on the semantics of WIL-sentences (e.g. Stoljar, 2016; D'Ambrosio & Stoljar, 2022), in this section I show that philosophers' WIL-concepts, expressed by corresponding WIL-sentences in the

[4] In explaining a wide range of literal and figurative uses of language, Carston (2002) appeals to ad hoc concepts in her general theory of lexical pragmatics (see also Wilson & Carston, 2007).

relevant philosophical discussion on consciousness, find their origins as ad hoc concepts and are narrower than the encoded WIL-concepts used in ordinary English. Call philosophers' WIL-concepts 'WIL$_P$-concepts'. There are many WIL$_p$-concepts, e.g. the concept of *what it is like to see something red* in Jackson's (1982) thought-experiment about the super-scientist Mary; the concept of *what it is like to be a bat from the inside* in Nagel's (1974) seminal paper. WIL$_P$-concepts are specific concepts that fall under the general philosophical concept CONSCIOUSNESS. What it is like to see something red and what it is like to be a bat from the inside are different ways of being conscious. Here I shall primarily focus on WIL$_P$-concepts. I shall make brief remarks about the general concept CONSCIOUSNESS towards the end of the section.

The structure of this section is as follows. Section 3.1 makes preliminary clarification on WIL-sentences as used in ordinary English. Section 3.2 explains Stoljar's recent affective account of the semantics of WIL-sentences. Section 3.3 clarifies philosophers' WIL$_P$-concepts. Section 3.4 argues that WIL$_P$-concepts are ad hoc concepts. Section 3.5 turns to the related TSIL-locution, and the general concept CONSCIOUSNESS.

3.1 Preliminary clarification

In this section, I make five preliminary points about ordinary WIL-sentences.

First, WIL-sentences are commonly used in ordinary English and are non-technical (Farrell, 2016; Hellie, 2004; Stoljar, 2016). Consider (7)-sentences:

(7) a. John knows what it is like to live in the forest.
b. The Chancellor has no idea what it is like being poor.
c. Juan tells the children what it was like when there were no computers.
d. What will it be like in the UK in the future?

The phrase 'what it is like' can be followed by an infinitive clause, e.g. (7a); a gerund, e.g. (7b); a subordinate clause, e.g. (7c); or be embedded in a question, e.g. (7d). Given WIL-sentences used in philosophical discussion on consciousness usually take the syntactic form 'what it is like to φ' (e.g. 'what it is like to see something red') where 'φ' refers to a conscious state (e.g. seeing something red),[5] we will focus on constructions like (7a) where 'what it is like' is followed by an infinitive clause referring to an event or a state.

Examples from (7) bring out a second point about WIL-sentences, namely, that they can have *experiential* uses and also *non-experiential* uses. Some WIL-sentences, given appropriate contexts, focus on the experiential or psychological effects that an event has on an individual. For instance, in a context of talking about the tranquillity of being in nature, (7a) may be understood to be about what the *experience* of living in the forest is like; and in a context of discussing poverty-induced stress and anxiety, (7b) alludes to what being poor *feels* like. In other cases, WIL-sentences are not understood to be about experiences. Consider (7c). One might imagine Juan, in telling the children about what the world was like before there were computers, to mention that writers

[5] Note that Nagel (1974) was concerned with what it is like to be a certain organism rather than what it is like to have a certain conscious state.

used typewriters and that it was common to write letters. (7d), in the context of Brexit, might elicit answers about the impact of Brexit on the economy, migration, and so on. These are non-experiential uses of WIL-sentences.

The third and fourth points are about the meanings of two words in the phrase 'what it's like'. It is usually thought that 'like' in 'what it's like' does not mean 'resemble' (Nagel, 1974: fn6; Lewis, 1988, p. 282; Hellie, 2004; Stoljar, 2016).[6] OED's entry for the word 'like' lists one of its meanings as '[u]sed in questions asked about the characteristics or nature of someone or something' as in (8a). The same sense is also used in declarative sentences like (8b).

(8) a. What is Shanghai *like*?
b. You know what these politicians are *like*.

(8a) is intuitively interpreted as asking about salient properties or facts about Shanghai rather than what Shanghai resembles (see Hellie, 2004). In answering (8a), one would most likely give a list of words denoting various properties that one might associate with the city, e.g. 'modern', 'busy', 'flashy'. One can of course compare Shanghai with some other cities. For instance, one might answer that 'it's like Hong Kong', meaning that Shanghai is in some way similar to Hong Kong. But even in this case, the comparative construction is used to bring out salient features of Shanghai, i.e. the properties that it has in common with Hong Kong. Similarly, in hearing (8b), one might expect the speaker to have in mind some specific properties of the politicians being referred to, e.g. manipulative, mendacious, corrupt.

Fourth, the word 'it' in 'what it's like' is an expletive pronoun and behaves as it does in sentences like 'it is awful to hear the news'. The latter is equivalent to 'To hear the news is awful', or more naturally 'Hearing the news is awful'. Similarly, (7a)—'John knows what it is like to live in the forest'—is equivalent to 'John knows what living in the forest is like'.

Finally, WIL-sentences like (7a) involve covert subjects. We can distinguish two kinds of subjects—the subject of the 'what it's like'-clause (call it 'x') and the subject of the action of φ-ing (call it 'y'). In (7a), x is identical to y and the sentence is understood to mean that 'John knows what it is like *for a person for that person* to live in the forest'. But the two subjects can come apart. Imagine that John has a son who recently decided to turn away from civilisation and to live in the forest where he can forage for food. In this case, we can ask what it is like *for* John *for* his son to live in the forest, where John is the subject of the 'what it's like'-clause but not the infinitive clause 'to live in the forest'. Nevertheless, when the subjects are not explicitly stated, as in (7a), they are understood as coincident. Below I follow Stoljar (2016) and focus on the construction 'what it is like *to* x *for* y', using two different prepositions to clearly distinguish the two covert subjects.

[6] However, Gaskin (2019) argues for a resemblance account of WIL-sentences. D'Ambrasio and Stoljar (forthcoming) distinguish two versions of the resemblance account—a relational version and a notional version and contend that arguments against the resemblance account problematise the relational version not the notional version. The latter version is compatible with other accounts in the literature, e.g. the affective account (Stoljar, 2016), the property account (Hellie, 2004).

3.2 The affective account

Stoljar's (2016) recently proposed affective account of WIL-sentences is said to accommodate the aforementioned features mentioned in Sect. 3.1 (see also D'Ambrosio & Stoljar, 2022).[7] According to this account, the locution 'what it's like to φ' expresses an *affective* relation between an individual and the *event* of φ-ing, and relatedly, the locution of 'what it's like to x for y to φ' expresses an *affective* relation between individual x—the subject of the 'what it's like' clause—and the event of y's φ-ing. It states a particular kind of *way* y's φ-ing affects x, where a *way* 'is a certain sort of property something has' (Stoljar, 2016, p. 1170, fn14). So, on this view:

(9) 'S knows what it is like to x for y to φ' is true in a context c if and only if there is, in c, some way W such that S knows that y's φ-ing affects x in way W.

Consider (10):

(10) Sally knows what it is like to healthcare workers for hospitals to be inundated with patients.

On this account, (10) is analysed as the following:

(11) 'Sally knows what it is like to healthcare workers for hospitals to be inundated with patients' is true in a context c if and only if there is, in c, some way W such that Sally knows that hospitals' being inundated with patients affects healthcare workers in way W.

Stoljar points out that WIL-sentences are context-sensitive, signalling a contextually salient way that the event of φ-ing can affect an individual. Consider (7a)—'John knows what it is like to live in the forest'. In the context of discussing the benefits of the natural environment, the sentence is naturally interpreted as expressing positive effects that living in the forest has. In the context of discussing a housing crisis, the effects at issue are likely to be negative.

Stoljar (2016, p. 1176) also points out that in *stereotypical* cases, the speaker has in mind a particular kind of *experiential* way that the individual at issue is being affected by an event (see also D'Ambrosio & Stoljar, 2021, forthcoming[8]):

(12) There are stereotypical contexts c such that 'S knows what it is like to x for y to φ' is true in a context c if and only if there is, in c, some *experiential* way W* such that S knows that y's φ-ing affects x in way W*.

[7] For other accounts of the semantics of WIL-sentences, see e.g. Snowdon (2010), Janzen (2011), and Gaskin (2019). For criticism of these other accounts, see Stoljar (2016) and D'Ambrosio and Stoljar (2022).

[8] D'Ambrosio and Stoljar (2021, forthcoming) note the pragmatic connection between 'imagine what an object is like' and 'imagine what it looks like' where the latter specifies the sensory modality in which one imagines the object. They (forthcoming) point out that such pragmatic connections, which are common, form part of Horn's (1984) 'division of pragmatic labour', where a general term contained in an utterance, which can be specified in various ways, is standardly understood as conveying a stereotypical reading, e.g. 'milk' in 'I am drinking a glass of milk' is standardly understood to mean cow's milk given the latter's prevalence. In a way, D'Ambrosio and Stoljar's discussion serves as a precursor to the pragmatic account advanced here with respect to the connection between WIL-concepts as encoded by ordinary WIL-sentences and philosophers' WIL_p-concepts which are tied to phenomenal consciousness.

An event affects an individual S in an *experiential* way just in case there is some way S *feels* as a result of the event, where the notion of 'feeling' is understood in a broad sense and is not restricted to bodily sensations.[9] For instance, we would usually expect the utterer of (10) to have in mind some particular way that healthcare workers experience hospitals' being inundated with patients or how they feel as a result of the latter event, e.g. exhausted, anxious. While (12) states that WIL-sentences are associated with certain stereotypical contexts where the affective relation between the individual and the event is an experiential one, it does not rule out contexts where the relevant affective relation is non-experiential. For instance, the sentence—'The leave campaigners had no idea what it would be like for the UK to leave the EU'—expresses some non-experiential way that the event of UK leaving the EU affects the country.

3.3 WIL$_P$-concepts

Given the affective account of WIL-sentences, let us turn to philosophers' WIL$_P$-concepts, which occur in the discussion on consciousness. To bring out philosophers' WIL$_P$-concepts, I shall focus on Frank Jackson's (1982) thought-experiment about Mary.

Consider the super-scientist Mary, who knows all there is to know about vision science in her black-and-white room. Jackson contends that when Mary leaves her room and sees something red for the first time, it is intuitive to think that Mary learns something new which she did not know in her black-and-white room, namely:

(13) Mary learns what it is like to see something red.

(13) involves a WIL$_P$-concept, i.e. WHAT IT'S LIKE TO SEE SOMETHING RED. Let 'S' be the subject of the conscious state of seeing something red. (13) is equivalent to (14):

(14) Mary learns what it is like to S for S to see something red.

If we apply the affective account of WIL-sentences, we have the following analysis of (14):

(15) 'Mary learns what it is like to S for S to see something red' is true in context c if and only if there is, in c, some way W such that Mary learns that S's seeing something red affects S in way W.

Note that there are multiple different ways that seeing something red could affect the subject. One might be affected in a neurological way, having a certain pattern of firing in the brain. One might be affected in an idiosyncratic way such as having thoughts about revolutionary politics. As it stands, (15) does not specify what kind of way S's seeing something red affects S.

Note also that Mary already knew about the neurological way seeing something red can affect a subject in her black-and-white room. In the particular context involving Mary, (14) is supposed to pick out an *experiential* way that seeing something red

[9] For instance, in the broad sense of 'feel', we can say that Balzac's *Lost Illusions* describes how it feels to be a young provincial poet in early 19th-century France.

affects the subject. Put differently, (14) draws on the aforementioned 'stereotypical contexts' where the relevant affective relation between the individual and the event is an *experiential* one (see Stoljar, 2016, p. 8119). On this interpretation, (14) is analysed as (16):

(16) 'Mary learns what it is like to S for S to see something red' is true in context c if and only if there is, in c, some *experiential* way W* such that Mary learns that S's seeing something red affects S in way W*.

Nevertheless, it is not just any experiential way that S's seeing something red affects S. For instance, one could be feeling slightly enraged as a result of seeing something red. But such a *contingent* experiential way that seeing something red can affect the subject is not the intended way that Mary supposedly learns from seeing something red. The relevant experiential way is *essential* or *constitutive* of the conscious state of seeing something red (Stoljar, 2016, p. 1190). So, we arrive at (17):

(17) 'Mary learns what it is like to S for S to see something red' is true in context c if and only if there is, in c, some experiential way W** such that Mary learns that S's seeing something red affects S in way W** and W** is constitutive of the conscious state of seeing something red.

This constitutive, experiential way W**—call it 'the phenomenal way'—is what Mary learns by being S, the subject of the conscious state of seeing something red. It is what defines and individuates the type of conscious state of seeing something red and is what makes the latter a *conscious* state. When philosophers use the locution 'what it's like (to S for S) to φ' in the context of discussing consciousness, they refer to this particular *phenomenal way* that S is affected by φ-ing.

3.4 WIL$_P$-concepts as ad hoc concepts

I now argue that philosophers' WIL$_P$-concepts, i.e. concepts used in philosophical discussion on consciousness using WIL-sentences, are introduced as ad hoc concepts.[10] I illustrate this using the WIL$_P$-concept WHAT IT'S LIKE TO SEE SOMETHING RED.

Recall from Sect. 2 the two crucial features of ad hoc concepts associated with two sides of communication. Consider the first feature, i.e. a speaker can use a word or phrase to communicate an ad hoc concept that is different from the lexically encoded concept. As we saw, philosophers' WIL$_P$-concepts are distinct from WIL-concepts encoded in ordinary WIL-sentences. Correspondingly, the encoded meanings of ordinary WIL-sentences (see (15)) are distinct from the intended meanings of philosophers' WIL-sentences (see (17)). Philosophers' WIL$_P$-concepts fall under the category of semantic narrowing. The WIL$_P$-concept WHAT IT'S LIKE TO SEE SOMETHING RED conveys a narrower sense than the encoded concept. We saw that according to the affective account, the encoded meaning of a WIL-sentence expresses some contextually salient way—experientially or non-experientially—that an individual is affected by an event. (13), in the relevant philosophical context, expresses

[10] To argue for this, I again rely on Stoljar's affective account of WIL-sentences. While I think Stoljar's account fares better than alternative accounts, the general point being made here, namely, that WIL$_P$-concepts are ad hoc concepts, does not depend on Stoljar's specific semantic account of WIL-sentences.

a specific way that an individual is *experientially* affected by having a phenomenal state where such a way is also *constitutive* of the state. This *phenomenal* way is what philosophers mean by WIL-sentences in the discussion on consciousness. So, in the latter context, philosophers use a lexicalised encoded WIL-concept to convey an ad hoc, non-lexicalised WIL$_P$-concept.

Now consider the second feature of ad hoc concepts, according to which a hearer can latch onto an intended ad hoc concept through pragmatic inference. In Jackson's thought-experiment about Mary, the relevant ad hoc WIL$_P$-concept, i.e. WHAT IT'S LIKE TO SEE SOMETHING RED, is constructed online in the context of engaging with the thought-experiment. Jackson's thought-experiment sets up a context contrasting Mary's before-release black-and-white room with her after-release colourful world. When confronted with the question of whether Mary learns something new when seeing something red for the first time, the addressee is likely to consider whether Mary's seeing something red for the first time affects her in a special way and reflect on the potential radical change in Mary's experience. Upon subsequently hearing (13)—'Mary learns what it is like to see something red', experiential uses of the WIL-locution are likely to receive high activation and are thus easily accessible. The addressee is then likely to infer the following contextual implications from the utterance (13):

(18) a. Mary learns an experiential way seeing something red affects the perceiver.
b. This experiential way is constitutive of the experience of seeing something red.
c. Mary did not know this specific way when she was in her black-and-white room.

(18) follows from (13) where the proposition expressed by the latter is understood to include the relevant ad hoc concept WHAT IT'S LIKE TO SEE SOMETHING RED. In this way, the addressee also constructs an interpretation of (13) akin to (17).

The point made with respect to the concept WHAT IT'S LIKE TO SEE SOMETHING RED generalises and applies to other WIL$_P$-concepts. For instance, Nagel in his 1974 paper 'What is it like to be a bat?' gives various contextual clues to prompt the reader to latch onto the relevant ad hoc concept WHAT IT'S LIKE TO BE A BAT. Immediately after defining consciousness using the WIL-locution, Nagel (1974, p. 436) points out that the 'what-it's-like' character of subjective experience is not to be analysed in functional or behavioural terms. Thus, his WIL-locution is not supposed to refer to how being a bat affects a bat in certain non-experiential ways, e.g. hanging out in the cave, hunting for insects. This narrows down the kinds of affective relations between the subject of an experience and an event of undergoing that experience. Nagel then considers a bat's experience and its subjective point of view to help the reader to latch onto the relevant sense of the intended WIL$_P$-concept.

Contextual information from imagined scenarios like Jackson's and Nagel's enable philosophers to effectively introduce WIL$_P$-concepts to readers who are unfamiliar with the relevant philosophical discussion, allowing them to latch onto the intended interpretation of the WIL-locution. In philosophical discussion on consciousness, philosophers nevertheless often directly use the WIL-locution to denote phenomenal properties of conscious states without contextual information, but rather with

the tacit assumption that the reader (most likely a philosopher) is already familiar with WIL_P-concepts. WIL_p-concepts may very well be routinised or even lexicalised for philosophers of mind such that they can, upon encountering a WIL-sentence in philosophical discussion, directly home in on the relevant sense without pragmatic adjustment.

3.5 Consciousness and 'there's something it's like'

Let us now turn to the related TSIL-locution. The latter is often used to define *conscious* states:

(19) A mental state X is a conscious state if and only if *there is something it is like* to be in X.

It is important to note that in ordinary English, TSIL-sentences are far less common than WIL-sentences. The Corpus of Contemporary American English (COCA) renders 934 results for the word-string 'what it is like'. In contrast, there are only 8 results for the word-string 'there is something it is like' and all of them are occurrences in philosophical texts on consciousness.[11] It is plausible that TSIL-sentences are predominantly used in philosophical discussion.

In the discussion on consciousness, TSIL-sentences are closely associated with corresponding WIL-sentences. Consider (14) again:

(14) Mary learns what it is like to S for S to see something red.

(14), as we saw, is analysed as (17):

(17) 'Mary learns what it is like to S for S to see something red' is true in context c if and only if there is, in c, some experiential way W** such that Mary learns that S's seeing something red affects S in way W** and W** is constitutive of the conscious state of seeing something red.

Philosophers take (14) to entail (20):

(20) There is something it is like to S for S to see something red.

Given the analysis of (14) as (17), the word 'something' in (20) quantifies over the relevant phenomenal way seeing something red affects S. (20) is thus analysed as (21):

(21) 'There is something it is like to S for S to see something red' is true in a context c if and only if there is, in c, some experiential way W** such that S's seeing something red affects S in way W** and W** is constitutive of the conscious state of seeing something red.

In latching onto the intended meaning of (20), the addressee would first need to decode the encoded meaning of the sentence on the basis of the encoded meanings of the individual words in the sentence plus rules of syntax. One then needs to narrow down the range of extensions of the first occurrence of the word 'something' and interpret it as only referring to an experiential way in which the subject is affected by

[11] Data are collected from the following site: https://www.english-corpora.org/coca/.

seeing something red and which is also constitutive of the state. Plausibly, an addressee who can easily latch onto the intended meaning of (20) in the philosophical context is already familiar with philosophers' WIL-locution and has possessed or latched onto the ad hoc WIL$_P$-concept featured in (14).

Using Stoljar's recently proposed semantic account of WIL-sentences, in this section I argued that WIL$_P$-concepts deployed in philosophical discussion about consciousness are ad hoc concepts. Now, insofar as WIL-sentences understood in the relevant senses, as well as corresponding TSIL-sentences, are standardly used to introduce and define the general philosophical concept of consciousness (i.e. phenomenal consciousness), the latter concept may well be ad hoc also, in which case it is not encoded in the conventionalised meaning of the word 'consciousness' and requires the listener to construct the ad hoc meaning online in the relevant context. Alternatively, the term 'consciousness' may be polysemous in ordinary English with the phenomenal sense being one of its stable senses. In this case, the what-it-is-like talk would be to help the addressee to grasp that it is the phenomenal sense which the speaker intends when discussing 'consciousness'. Indeed, were the notion of phenomenal consciousness *uniquely* encoded in the word 'consciousness' in ordinary parlance, philosophers would not need to routinely appeal to WIL-talk and WIL$_P$-concepts when introducing the concept of phenomenal consciousness. While the philosophical concept of consciousness may or may not be a conventionalised sense of the word 'consciousness' for ordinary English speakers, it nevertheless is most likely conventionalised and plausibly the dominant sense of the word for analytic philosophers of mind.

4 The challenge for X-Phi

In the previous two sections, I have argued for the following two theses:

(i) There are ad hoc concepts.
(ii) Some philosophical concepts find their origins as ad hoc concepts.

In this section, I argue for the third thesis I wish for put forward, i.e. (iii):

(iii) Philosophical concepts which find their origins as ad hoc concepts pose a challenge for experimental philosophy regarding these concepts.

The specific challenge ad hoc concepts pose is with respect to the moves that experimental philosophers often make from certain empirical data concerning laypeople's judgements to the following types of conclusion:

(I) Laypeople do not have or grasp philosophers' concept C.
(II) Laypeople do not have philosophers' intuition X.
(III) Philosophical ideas based on C or X are problematic.

Either (I) or (II) is used to justify (III). Using X-Phi literature on consciousness as a case study (Sect. 4.1), I argue that such moves are problematic (Sect. 4.2). The argument here generalises: inferences from empirical data to conclusions (I)–(III) can be problematic when the philosopher's concept C finds its origin as an ad hoc concept or when having the intuition of X presupposes the acquisition of an ad hoc concept. I end the section by relating the criticism raised here against experimental philosophy to the kind of challenge from ordinary language found in Hansen (2020).

4.1 X-Phi on consciousness

In relation to the hard problem of consciousness, i.e. the problem of how physical processes can give rise to consciousness, there are two strands of discussion within X-Phi literature on consciousness. The first strand concerns whether laypeople have the philosophical concept of consciousness, i.e. phenomenal consciousness (e.g. Knobe & Prinz, 2008; Sytsma & Machery, 2009, 2010; Sytsma, 2009, 2010, 2016; Sytsma & Ozedemir, 2019). Within this discussion, inferences from empirical data to (I) as well as to (III) are frequently made. It is often thought that people's judgements in experimental settings are indicative of whether they possess the concept of consciousness, and the latter is thought to have implications with respect to whether the hard problem of consciousness is a genuine problem. Concerning the move from (I) to (III) with respect to the concept of consciousness, the thought is that if the concept of consciousness turns out to be a concept only possessed and grasped by philosophers, then consciousness looks like a dubious phenomenon. The puzzle it poses for physicalism might then seem much less interesting and pressing (see Stoljar, 2016, p. 1184). As a result, the hard problem of consciousness at least appears to be less of a genuine problem.

In a number of places, Sytsma et al. appeal to empirical data to argue that laypeople lack the concept of phenomenal consciousness (Sytsma & Machery, 2009, 2010; Sytsma, 2009, 2010, 2016; Sytsma & Ozedemir, 2019). Notably, in 'Two Conceptions of Subjective Experience', Sytsma and Machery (2010) contend that if laypeople possessed the concept of consciousness, they would treat paradigmatic cases of phenomenal states (e.g. feeling pain, seeing red) in a similar way to philosophers in tracking phenomenal consciousness. They presented participants with a vignette describing a scenario where an agent manipulates a red object. The vignette varied along two dimensions: the agent is either a human male or a simple robot, and the scenario describes the agent as either performing a successful manipulation or getting an electric shock. Participants—philosophers and non-philosophers—were asked to rate, on a 7-point scale (where '1' means 'clearly no', '4' 'not sure' and '7' 'clearly yes'), whether the agent 'saw red' or 'felt pain' given the relevant scenario. Their results showed that philosophers treated 'feeling pain' and 'seeing red' similarly—they attributed both to the human but neither to the robot; whereas non-philosophers treated the two states differently—while they attributed both to the human, they were willing to attribute 'seeing red' to the robot but not 'feeling pain'. Sytsma and Machery take the results to show that unlike philosophers, non-philosophers do not recognise the phenomenality of prototypical phenomenal states and lack the concept of phenomenal

consciousness.[12] At this point, Sytsma and Machery (2010, p. 321) raise the question of whether the hard problem of consciousness is a genuine problem:

> At this juncture, it is important to ask why we should believe that there is really a hard problem of consciousness. That is, why should we believe that there are experiences, as defined, such that an aspect of them is expected to be left unexplained even after we have thoroughly accounted for their functional roles?

Here, we see a clear instance of reasoning from (I) to (III): that laypeople lack the concept of consciousness is used to cast doubt on the hard problem of consciousness itself.

It is worth noting that whether or not there is a hard problem does not depend on whether non-philosophers possess and utilise the concept of consciousness. Sytsma and Machery (2010, p. 323) themselves raise the possibility that people may lack the concept but be able to easily latch onto it. However, they dismiss this response on the basis of their own experience that 'ordinary people either don't understand or don't take seriously the philosophical concept of phenomenal consciousness even after a lengthy explanation' (op. cit).

It is also worth noting that one can motivate the hard problem of consciousness without relying on the general concept CONSCIOUSNESS, but only on specific WIL$_P$-concepts (see Chalmers, 2020, p. 240; Gregory et al., 2021). Just as one can have the concepts of TRIANGLE, RHOMBUS, and SQUARE without having the general concept CLOSED SHAPE, one can have WIL$_P$-concepts of specific phenomenal properties, e.g. WHAT IT IS LIKE TO FEEL PAIN, WHAT IT IS LIKE TO SEE SOMETHING RED, without possessing the general concept CONSCIOUSNESS. Indeed, the hard problem is routinely introduced using WIL$_P$-concepts and specific phenomenal states (e.g. Jackson, 1986; Chalmers, 1995; Kim, 2005, p. 13). The question then comes down to whether laypeople have or can grasp WIL$_P$-concepts.

In a recent paper, Fischer and Sytsma (2021) appeal to empirical data to suggest that laypeople lack WIL$_P$-concepts or at least have difficulties understanding the Nagelian WIL-locution. They (2021: fn3, Appendix D) asked participants to rate a list of features with respect to zombies (as standardly conceived in non-philosophical contexts) on a 7-point scale where '1' means 'very atypical', '4' means 'neutral (neither typical nor atypical)', and '7' means 'very typical', and the list included the following four items that attribute phenomenal consciousness to the beings at issue:

(A) These beings are capable of having conscious experiences.

[12] One key objection against Sytsma and Machery's conclusion is that the phrase 'see red' is ambiguous between a *non-phenomenal reading*, which one would attribute to a robot, and a *phenomenal reading*, which one would attribute to a human but not a robot (Huebner 2010, p. 137; Chalmers 2018: fn7). It may be that the phrase 'see red' tends to trigger the phenomenal reading in philosophers and the non-phenomenal reading in non-philosophers. To avoid this objection, in a subsequent study Sytsma and Ozdemir (2019) replaced the phrase 'see red' in the original vignette with 'experience red' and also 'feel pain' with 'experience pain'. Results nevertheless showed a similar pattern compared to the previous study though to a lesser extent. Laypeople attributed 'experiencing red' to the robot but not 'experiencing pain' while they attributed both to the human. However, it is worth noting that the verb 'experience' is polysemous. 'Experience red' can mean undergoing a conscious experience of seeing something red or encountering the colour red. While it is false to say that the robot 'experiences red' in the former sense, it is true to say that it 'experiences red' in the latter sense.

(B) These beings have an inner mental life, including feelings and emotions.
(C) These beings are sentient and experience their surroundings and sensations.
(D) There is something it is like to be such a being.

The mean ratings were 2.93, 2.41, 3.06, and 3.61 respectively, suggesting (A)-(D) are atypical features of zombies. However, more than half of participants (54.6%) answered '4' for (D), which was significantly higher than the corresponding proportions for the neutral rating for the other three items. Since each of (A)–(C) is supposed to capture the idea that the relevant beings have phenomenal consciousness, those who disagreed with (A)–(C) ought to have disagreed with (D) on the Nagelian interpretation. So, Fischer and Sytsma conclude that many participants did not latch onto the intended Nagelian interpretation of (D). This result, as I shall expound later, can be easily made sense of given our discussion on WIL- and TSIL-sentences in Sect. 3, and relatedly is no grounds for thinking that laypeople are unable to latch onto WIL_{P}-concepts in other contexts.

The second strand in X-Phi literature on the hard problem of consciousness concerns whether laypeople have so-called 'problem intuitions' about consciousness (Gottlieb & Lombrozo, 2018; Fischer & Sytsma, 2021; Díaz, 2021; Gregory et al.'s 2021). Problem intuitions include *explanatory* intuitions (e.g. that consciousness cannot be explained in physical terms), *metaphysical intuitions* (e.g. that consciousness is non-physical), *knowledge intuitions* (e.g. that knowledge about what it is like to see colours is not exhausted by knowledge about colour science), and *modal intuitions* (e.g. that we can conceive physical duplicates of us but lacking consciousness) (Chalmers, 2018, p. 12). Problem intuitions motivate the hard problem of consciousness and underlie corresponding arguments against physicalism. Levine's (1983) explanatory gap argument is based on an explanatory intuition. Kripke's (1980) argument against physicalism in *Naming and Necessity* is underpinned by the metaphysical intuition that the feeling of pain is distinct from C-fibre stimulation. Jackson's (1982) knowledge argument appeals to the knowledge intuition that Mary learns something new after leaving her black-and-white room. Chalmers' (1996) zombie argument is built on a modal intuition about the conceivability of philosophical zombies.

While empirical literature on these intuitions suggests a mixed picture, a number of theorists have presented studies to show that some of the problem intuitions are not widespread.[13] Such theorists then draw the conclusion that the hard problem of consciousness does not exist (Díaz, 2021; Peressini, 2014), or that related philosophical arguments based on the problem intuitions (e.g. zombie argument) do not get off the ground (Fischer & Sytsma, 2021, p. 10). So, inferences from empirical data to (II) as well as to (III) are frequently made within this discussion.

In a series of studies, Díaz (2021) tested both explanatory and metaphysical intuitions. For instance, he presented participants with versions of a vignette outlining the close connections between the feeling of pain and neural activity in the DPI (the dorsal posterior region of the insula), and then asked participants to rate two statements: 'The properties of pain are fully explained in terms of neural activity in the DPI' (which

[13] In terms of problem intuitions that are allegedly widespread, Gregory et. al.'s (2021) studies showed that a large majority of respondents reported that colour blind Mary, who knows all physical information about vision before her operation, does not know what it is like to see red, and that she will learn what it is like to see red after her operation.

tested the explanatory intuition), and 'The feeling of pain is just neural activity in the DPI' (which tested the dualistic intuition). Results suggested that both the explanatory intuition and the dualistic intuition were not widespread.[14] In light of the empirical evidence, Díaz questions whether the hard problem of consciousness exists at all.

Regarding modal intuitions, Peressini (2014, pp. 874–875) directly asked participants whether the following two imagined scenarios are possible: 'a medical procedure that would remove your inner experience without affecting your brain, so from the outside you would remain unchanged physically and behaviourally' and 'a person physically and behaviourally identical to you in all ways but who had no inner experience at all'. A majority of the participants did not judge these imagined scenarios to be possible.[15] The conceivability of philosophical zombies was empirically studied by Fischer and Sytsma (2021). They took their results to show that linguistic salience bias contributed to the conceivability of philosophical zombies when the latter are described as 'zombies', and for a majority of their participants philosophical zombies were not conceivable when described as 'duplicates'. Fischer and Sytsma took the findings to problematise the conceivability premise of the zombie argument.

4.2 The challenge from ad hoc concepts

In the last section, we saw that empirical data regarding people's judgements concerning consciousness are used to infer (I)–(III) where 'C' stands for the concept of consciousness or WIL$_P$-concepts and 'X' stands for a problem intuition:

(I) Laypeople do not have or grasp philosophers' concept C.
(II) Laypeople do not have philosophers' intuition X.
(III) Philosophical ideas based on C or X are problematic.

However, as I shall argue in this section, such inferences from empirical data can be problematic when the concepts at issue find their origins as ad hoc concepts and that the intuitions at issue rely on having latched onto relevant ad hoc concepts. I shall primarily focus on problematising the inferences from certain empirical data to (I) and (II). Insofar as the latter are used to support (III), the inference from relevant empirical data to (III) is also problematic as a consequence of my arguments.[16]

[14] With respect to explanatory intuitions, Gottlieb and Lombrozo provided evidence 'that participants were more likely to judge scientific explanations for psychological phenomena to be impossible or uncomfortable' (2018, p. 121), when the relevant psychological phenomena have 'introspective-phenomenology', i.e. involve 'a subjective experience (a feeling of what it is like) that only the individual experiencing it can know' (2018, p. 123). In terms of other metaphysical intuitions, Liu (2021: fn5) provided preliminary evidence supporting the existence of the intuition of revelation, that the essences of phenomenal properties are revealed in phenomenal experiences. Data showed that fluent English speakers judge sentences like (a) 'I know what an itch feels like, but I don't know what the feeling of an itch really is' as strange, but not sentences like (b) 'I know what gold looks like, but I don't know what gold really is'. Liu argues, on an armchair basis, that the oddness of (a) is best explained by the hypothesis that the thesis of revelation is part of our ordinary conception of experience, which we tacitly appeal to in interpreting sentences like (a).

[15] However, it is unclear that participants made clear distinctions between different kinds of possibility, e.g. nomological, logical.

[16] There may be other reasons to reject the inference from (I) or (II) to (III) even if we grant (I) and (II). In this paper, I shall set these reasons aside.

It is important to note that the goal here is *not* to argue against (I) and (II). For instance, there may be good empirical grounds for doubting that laypeople share a particular intuition widely held by philosophers. It is also worth noting that I am not questioning *wholesale* the conclusions of the experimental studies from which the examples I mention below are drawn.[17] The goal here is to criticise certain uses of empirical data as grounds for supporting (I) and (II). In doing so, I shall primarily focus on experimental prompts that use WIL-sentences. But similar concerns may be raised regarding other wordings that are aimed at capturing the relevant ad hoc concepts.

4.2.1 Ways to fail to construct an ad hoc concept

As we saw from Sect. 2, a speaker can use a word or phrase to communicate an ad hoc concept that is different from the lexically encoded concept. The addressee can infer the intended ad hoc concept through pragmatic inference in a given conversational context. However, the addressee might fail to latch onto the speaker's intended ad hoc concept for various reasons. Below I consider four different ways that one might fail to construct the intended ad hoc concept. I illustrate them with respect to the scenario discussed in Sect. 2.1:

(1) Meera and Luke are watching John skilfully swinging on the monkey bars. Meera, pointing at John, says to Luke: 'John is such a *gibbon*.'

(a)*Ignorance on the Part of the Audience*

Suppose that Luke does not know the encoded meaning of the word 'gibbon', or possess the encyclopaedic knowledge that gibbons are agile and good at swinging from branch to branch. In this case, it is unlikely that Luke would be able to latch onto the ad hoc concept GIBBON* and infer the relevant contextual implications that John is agile and good at swinging.

(b)*Unclarity on the Part of the Speaker*

Now suppose Luke has a perfect grasp of the meaning of 'gibbon', but Meera is the one not knowing that gibbons are stereotypically associated with being good at swinging. Confused between gibbons and gorillas, Meera, instead of using the word 'gibbon', actually utters 'John is such a gorilla'. Gorillas may be good at swinging, but they are certainly not stereotypically associated with it. In this case, without further clarification on Meera's part, it would be impossible for Luke to correctly infer what Meera meant. Luke might mistakenly interpret Meera as meaning that John is very strong. In this case, our speaker is being unclear, which impedes the construction of the intended meaning of her utterance on the side of the addressee.

(c)*Lack of Contextual Information*

Suppose instead of pointing at John who is skilfully swinging on the monkey bars, we are imagining a scenario where Meera and Luke are merely talking on the phone about John. Suppose Luke asks Meera how John is. Meera then replies 'John is such a gibbon these days'. It would be almost impossible for Luke to construct the intended meaning of Meera's utterance without further contextual information to make sense of

[17] For instance, I agree with Fischer and Sytsma's (2021) conclusion that their results showed that linguistic salience bias contributed to the conceivability of philosophical zombies when the latter are described as 'zombies'.

it, e.g. John has recently developed an interest in monkey bars. In this case, Luke fails to latch onto the ad hoc concept GIBBON* due to insufficient background or contextual information.

(d) *Lack of Conversational Feedback Mechanism*

Now in scenario (c), Luke, who is talking to Meera on the phone, can easily ask Meera to clarify what she meant by her utterance. Ordinary conversational settings are endowed with feedback mechanisms between the speaker and the addressee—the addressee can ask the speaker to clarify (e.g. asking Meera what the word 'gibbon' means or what she has in mind when using the word) and the speaker can offer further information. Such feedback mechanisms facilitate successful communication (see Schober & Clack, 1989; Hansen, 2020), including helping the addressee to latch onto the intended ad hoc concept. But, suppose instead of speaking to Meera on the phone in real time, Luke receives an email from Meera which contains the sentence 'John is such a gibbon these days' without further contextual information. In this case, given that the linguistic context is an email exchange, Luke lacks an immediately available feedback mechanism, which would otherwise enable him to promptly ask Meera about what she meant in order to interpret her utterance and latch onto the ad hoc concept GIBBON*.

So, there are various ways that a hearer might fail to construct the intended meaning of a speaker's utterance and the intended ad hoc concept. Conditions like (a)–(d), which contribute to a failure in latching onto the relevant ad hoc concept, can occur in ordinary communicative contexts as well as experimental settings. These conditions could also function simultaneously. This undercuts the inferences from certain empirical data to (I) and (II), and consequently (III) insofar as (I) and (II) are used to justify (III).

4.2.2 Inferences to (I)

Recall Sytsma and Machery's testimony from their own experience that 'ordinary people either don't understand or don't take seriously the philosophical concept of phenomenal consciousness even after a lengthy explanation'. Now, if CONSCIOUSNESS and WIL$_P$-concepts are ad hoc concepts, then it is no surprise that hearers can easily fail to latch onto these concepts, just as they can fail to construct other intended ad hoc concepts. This by no means indicates that the concept of consciousness or related WIL$_P$-concepts are illegitimate, or that laypeople would fail to grasp the concepts had the latter been introduced under favourable conditions where there is an appropriate feedback mechanism in place, adequate background information, and no confusion on the part of the speaker or audience.

Consider also Fischer and Sytsma's experiment testing atypical features of zombies, which includes item (D) below:

(D) There is something it is like to be such a being.

Without adequate contextual information and conversational feedback mechanisms, the reader can easily fail to latch onto the intended Nagelian interpretation in constructing the meaning of (D). For instance, the reader might take (D) to mean that being a zombie affects the subject in some way, but unclear about which way the speaker has in mind. Alternatively, the reader might take the polysemous word 'like'

to mean 'resemble' and subsequently interpret (D) to mean 'there is something being a zombie *resembles*'. While it is true that there is something being a zombie resembles, e.g. being a normal human being, it is not true that there is something being a zombie is like in the sense that there is some experiential way that the zombie feels in virtue of being a zombie. So, (D) has many competing interpretations. It would not be surprising if many participants did not latch onto the Nagelian interpretation. After all, the TSIL-locution is not an ordinary locution and its comprehension, especially with respect to the interpretation of 'something', requires that one is already familiar with the ad hoc WIL_P-concept as used in philosophical discussion.

Since ordinary conversational contexts as well as experimental contexts can easily fail to facilitate the communication of ad hoc concepts, we need to be careful drawing conclusions concerning an ad hoc concept C from empirical data collected from ordinary conversation or experimental setting to conclusion (I), that laypeople do not have or grasp philosophers' concept C. Since C is an ad hoc concept, laypeople may not already have the concept C in the sense of routinely utilising C in thinking, but it is far from clear that they do not grasp C when appropriate conditions for facilitating the communication of C are in place.

4.2.3 Inferences to (II)

Relatedly, if having the philosophical intuition X relies on grasping a certain ad hoc concept C, we need to be cautious in drawing the conclusion (II), that laypeople lack intuition X, from purported empirical data concerning laypeople's judgements collected in experimental setting. Consider the statement that Díaz (2021, p. 68) used in testing laypeople's explanatory intuition: 'Neural activity in the DPI cannot explain what [it] is like to feel pain'. The locution 'what it is like to feel pain' on its own can refer to different ways that the experience of feeling pain affects the experiencer including neurological ways.[18] So even if many participants give low ratings of agreement to the statement, it does not mean that they lack the explanatory intuition.[19] Participants might have simply constructed an ad hoc WIL-concept that is different from the relevant WIL_P-concept used in philosophical discussion. Such results thus do not demonstrate that laypeople lack the relevant problem intuition, i.e. the explanatory intuition in this case. Nor can they be used to support the claim that the existence

[18] It is worth noting that the statement that required rating was preceded by a vignette that opened with the following sentences: 'The feeling of pain is a subjective mental state; only the individual him- or herself can feel it. Neural activity is an objective physical process; outside observers can measure it.' It is unclear that this description clearly conveys the ad hoc WIL_p-concept with respect to the feeling of pain. But even if it does, it is unclear that the description, which does not contain an explicit WIL-locution, would on its own direct participants to deploy the relevant ad hoc WIL_p-concept when considering the statement that does contain the locution later on.

[19] Díaz's (2021, p. 65) primary concerns in his studies were to test laypeople's explanatory intuitions and metaphysical intuitions. For the former, he asked participants to rate the statement 'the properties of pain are fully explained in terms of neural activity in the DPI', and for the latter 'the feeling of pain is just neural activity in the DPI'. However, words like 'explain' and 'is' are polysemous or at least indeterminate in their meanings. There are different kinds of explanations—neural activity in the DPI may causally explain the presence of the feeling of pain, but it is unclear that it reductively explains why the latter arises from the former. The word 'is' is arguably ambiguous between an identity use ('The morning star is Venus') and a predication use (e.g. 'John is the embodiment of conservatism').

of the hard problem of consciousness, which can be motivated on the basis of the explanatory intuition, is dubious.

More generally, for any philosophical intuition X, in order to establish whether someone has X, we need to first establish that she has grasped the relevant concepts needed to entertain X. If the relevant concepts are ad hoc concepts, then in order to test judgements regarding X, one first needs to ensure that participants grasp the relevant ad hoc concepts as part of the experimental setting. The latter requires appropriate introduction of the concepts under circumstances which facilitate successful communication and ideally avoids features (a)-(d) discussed above, accompanied with pre- or post-testing, or both, to ensure that participants have latched onto the relevant concepts. This sets up a contrast to testing philosophical intuitions featuring only ordinary, lexicalised concepts. In the latter case, experimenters can simply assume that participants possess the relevant concepts, and that they are able to retrieve and use them in making the required judgments.

It is also worth noting why Jackson's thought-experiment is particularly adequate at providing the right conditions for the addressee to latch onto the intended WIL$_P$-concept. A recent study from Gregory et al. (2021) confirms that the knowledge intuition, as brought out by the Mary thought-experiment, is widespread. Recall Jackson's thought-experiment and consider it as a discourse between a speaker—the philosopher who puts forward the thought-experiment—and an addressee, say, a non-philosopher who engages with the thought-experiment. Our addressee is told by her philosopher interlocutor that Mary is a super-scientist who has been trapped in a black-and-white room since birth, but knows all there is to know about colour science including what is going on in people's brains when they see different colours. After having elaborated on Mary's scientific knowledge in great detail, our philosopher asks her non-philosopher interlocutor to imagine that one day Mary gets to leave her black-and-white room and see the colourful world. She then poses the following question to her addressee: 'Will Mary learn something new when she sees, say, a red tomato for the first time?' The intuition is likely to be in the affirmative (see Gregory et al., 2021), though our addressee might not put things in terms of a WIL-locution. Nevertheless, once the WIL-sentence is introduced, i.e. 'Mary learns what it is like to see something red', the addressee can immediately latch onto the intended meaning. In this case, it is clear, given the context, that 'what it is like to see something red' does not refer to physical-functional ways that the experience affects Mary, which she already knew in her black-and-white room. It is also clear that the WIL-locution does not prompt some idiosyncratic, experiential way that Mary is affected by seeing something red. The most accessible interpretation is that the locution refers to a constitutive, experiential way that seeing red affects the subject of the experience, and Mary gets to be such a subject when she sees something red for the first time.

4.2.4 The challenge

The challenge that ad hoc concepts pose to experimental philosophy can be summarised as below:

1. Experimental data on laypeople's judgements concerning certain philosophical concepts and philosophical intuitions about these concepts are often used to draw conclusions (I)–(III).
2. Some philosophical concepts find their origins as ad hoc concepts, and contexts—both ordinary and experimental—in which conditions such as (a)–(d) are present may impede successful grasp of these concepts.
3. Drawing conclusions (I)–(III) on the basis of empirical data collected from contexts in which conditions such as (a)–(d) are present is unjustified.

The challenge articulated here is in line with Hansen's 'revised challenge from ordinary language' to experimental philosophy. Hansen (2020), following Baz (2012, 2014), argues that people's responses to a philosophical question in one context do not tell us about how they would respond to it in a different context. Hansen (2020,p. 2453) points out that the application of a concept in an experimental context depends on multiple factors, including whether there are 'motivations that go beyond just wanting to perform the experimental task', whether participants are aware that 'they are taking part in an experiment', and whether the experimental task 'involves active collaboration between speakers and addressees'. The last factor is particularly pertinent to our discussion here. Active collaboration between speakers and addressees, as we saw, enables addressees to interpret speakers' utterances and latch onto relevant ad hoc concepts. While Hansen's challenge focuses on the fact that ordinary applications of concepts are context-dependent and experimental settings which strip away such pragmatic factors do not yield a clear picture of these concepts, the challenge raised here, while being the same kind, points out that some philosophical concepts are ad hoc in origin and that successful grasping of these concepts is context-dependent and requires careful setup on the part of the speaker.

5 Conclusion

In this paper, I have defended the following three claims:

(i) There are ad hoc concepts.
(ii) Some philosophical concepts find their origins as ad hoc concepts.
(iii) Philosophical concepts which find their origins as ad hoc concepts pose a challenge for experimental philosophy regarding these concepts.

With respect to (i), I showed that ad hoc concepts, which are prevalent in ordinary discourse, are concepts that speakers communicate by using encoded concepts and are constructed pragmatically by hearers from encoded concepts during utterance comprehension. With respect to (ii), I focused on the philosophical discussion on consciousness and related WIL_{P}-concepts. Appealing to recent literature on the semantics of WIL-sentences, I showed that WIL_{P}-concepts find their origins as ad hoc concepts. With respect to (iii), I problematise frequently-made moves from empirical data concerning laypeople's judgements to claims that laypeople do not have or grasp relevant philosophical concepts, do not have philosophical intuitions that feature these concepts, or that philosophical ideas relying on these concepts are problematic.

While I have focused on the case of consciousness and related X-Phi literature on consciousness, the arguments made here also apply elsewhere. Many philosophical concepts are initially introduced as ad hoc concepts and the relevant senses are eventually conventionalised for those who engage in the debate. Consider different notions of *possibility*. A philosopher may pose the question 'Is it possible to travel faster than the speed of light?' with the intention that the notion of possibility is understood in the logical rather than the nomological sense. Someone who has the nomological notion of possibility in mind may very well answer 'no' to the question the philosopher poses. But this does not reveal whether or not this person thinks that travelling faster than the speed of light is logically possible. So, simply asking participants whether a certain scenario is 'possible' (see Peressini, 2014) is inadequate for testing whether a philosophical intuition is widely shared, since participants might not latch onto the relevant sense of 'possible' that is pertinent to the philosophical intuition at issue.

Consider also different notions of *explanation*. C-fibres firing may *causally explain* the presence of the feeling of pain such that one can explain the latter by merely citing the correlation between the two. But it is far from clear that it explains the latter in a more demanding sense that some philosophers of mind are interested in, i.e. the notion of *reductive explanation* (see also fn19).[20] Someone who judges C-fibres firing to causally explain the presence of the feeling of pain does not necessarily think that it reductively explains the latter. So, simply asking participants whether certain physical features *explain* properties of pain (see Díaz's, 2021) is inadequate for testing the alleged explanatory intuition concerning phenomenal properties since there is no indication that participants would latch onto the relevant sense of 'explain'.

Successful communication of philosophical concepts that find their origins as ad hoc concepts requires careful stage setting, careful articulation on the part of the speaker, and adequate feedback mechanisms between the speaker and the addressee in order to rule out any confusion on the part of the addressee. This by no means suggests that such concepts cannot be communicated through experimental settings. But care must be taken in setting up the experiments before drawing conclusions about whether laypeople have or grasp certain philosophical concepts or have certain philosophical intuitions, and whether philosophical ideas based on the relevant concepts are dubious.

Acknowledgements Thanks to three reviewers for insightful comments. Thanks also to the editors of this volume - Eugen Fischer and Nat Hansen - for helpful suggestions and support during the reviewing process. The paper was presented at the 'Philosophy's Experimental Turn and the Challenge from Ordinary Language' conference in Berlin, which was sponsored by the Fritz Thyssen Foundation. Thanks to the participants on that occasion, especially Joachim Horvath for feedback and encouragement on the project. I am also indebted to Nicholas Allott whose written comments and expertise on relevance theory have helped me greatly improve the paper.

Funding This work was funded by the Leverhulme Trust Early Career Fellowship ECF-2021-539. Open Access funding enabled and organized by CAUL and its Member Institutions.

[20] A reductive explanation involves two domains—a higher-level and a lower-level domain—and is an explanation of a higher-level phenomenon in terms of lower-level phenomena, e.g. the reductive explanation of the boiling of a certain liquid is in terms of the lower-level phenomenon, the movements of molecules which make up the liquid. The explanatory resources for a reductive explanation are drawn solely from the lower level, such that a reductive explanation of A may only refer to the lower-level phenomena B in its explanans (Kim, 2005).

Declarations

Conflict of interest The author declares that there is no conflict of interest.

References

Allott, N., & Textor, M. (2012). Lexical pragmatic adjustment and the nature of ad hoc concepts. *International Review of Pragmatics, 4*(2), 185–208.
Baz, A. (2012). Must philosophers rely on intuitions? *The Journal of Philosophy, 109*(4), 316–337.
Baz, A. (2014). Recent attempts to defend the philosophical method of cases and the linguistic (re)turn. *Philosophy and Phenomenological Research, 92*(1), 105–130.
Baz, A. (2015). Questioning the method of cases fundamentally-reply to Deutsch. *Inquiry, 58*(7–8), 895–907.
Bezuidenhout, A. (2001). Metaphor and what is said: A defense of a direct expression view of metaphor. *Midwest Studies in Philosophy, 25*(1), 156.
Carston, R. (2002). *Thoughts and utterances: The pragmatics of explicit communication*. Blackwell.
Carston, R. (2010a). Metaphor: Ad hoc concepts, literal meaning and mental images. *Proceedings of the Aristotelian Society, 110*(3), 295–321.
Carston, R. (2010b). Lexical pragmatics, ad hoc concepts and metaphor: From a relevance theory perspective. *Italian Journal of Linguistics, 22*(1), 153–180.
Carston, R. (2012). Word meaning and concept expressed. *The Linguistic Review, 29*(4), 607–623.
Carston, R. (2019). Ad hoc concepts, polysemy and the lexicon. In K. Scott, B. Clark, & R. Carston (Eds.), *Relevance, pragmatics and interpretation* (pp. 150–162). Cambridge University Press.
Carston, R. (2021). Polysemy: Pragmatics and sense conventions. *Mind and Language, 36*, 108–133.
Chalmers, D. (1995). Facing up to the problem of consciousness. *Journal of Consciousness Studies, 2*(3), 200–219.
Chalmers, D. (1996). *The conscious mind*. Oxford University Press.
Chalmers, D. (2018). The meta-problem of consciousness. *Journal of Consciousness Studies, 25*(9–10), 6–61.
Chalmers, D. (2020). Is the hard problem universal? *Journal of Consciousness Studies, 27*(5–6), 227–257.
D'Ambrosio, J., & Stoljar, D. (2021). Vendler's puzzle about imagination. *Synthese, 199*, 12923–12944.
D'Ambrosio, J., & Stoljar, D. (2022). Two notions of resemblance and the semantics of "what it's like"'. *Inquiry*. https://doi.org/10.1080/0020174X.2022.2075453
D'Ambrosio, J. & Stoljar, D. (forthcoming). *Imagination, fiction, and perspectival displacement*. Oxford Studies in Philosophy of Mind.
Díaz, R. (2021). Do people think consciousness poses a hard problem? Empirical evidence on meta-problem of consciousness. *Journal of Consciousness Studies, 28*(3–4), 55–75.
Farrell, J. (2016). "What it is like" talk is not technical talk. *Journal of Consciousness Studies, 23*(9–10), 50–65.
Fischer, E., & Sytsma, J. (2021). Zombie intuitions. *Cognition, 215*, 104807.
Gaskin, R. (2019). A defence of the resemblance meaning of 'what it's like. *Mind, 128*(511), 673–698.
Gregory, D., Hendrickx, M., & Turner, C. (2021). Who knows what Mary knew? An experimental study. *Philosophical Psychology, 35*(4), 522–545.
Gottlieb, S., & Lombrozo, T. (2018). 'Can science explain the human mind? Intuitive judgments about the limits of science. *Psychological Science, 29*, 121–130.

Hansen, N. (2020). "Nobody would really talk that way!": The critical project in contemporary ordinary language philosophy. *Synthese, 197*, 2433–2464.
Hellie, B. (2004). Inexpressible truths and the allure of the knowledge argument. In P. Ludlow, Y. Nagasawa, & D. Stoljar (Eds.), *There's something about Mary: essays on phenomenal consciousness and Frank Jackson's knowledge argument* (pp. 333–364). MIT Press.
Horn, L. (1984). Toward a new taxonomy for pragmatic inference: Q-based and R-based implicature. In D. Schiffrin (Ed.), *Meaning, form, and use in context: Linguistic applications* (pp. 11–42). Georgetown University Press.
Huebner, B. (2010). Commonsense concepts of phenomenal consciousness: Does anyone care about functional zombies? *Phenomenology and the Cognitive Sciences, 9*(1), 133–155.
Jackson, F. (1982). Epiphenomenal qualia. *Philosophical Quarterly, 32*(127), 127–136.
Janzen, G. (2011). In defence of the what-it-is-likeness of experience. *The Southern Journal of Philosophy, 49*, 271–293.
Kim, J. (2005). *Physicalism, or something near enough*. Princeton University Press.
Knobe, J., & Prinz, J. (2008). Intuitions about consciousness: Experimental studies. *Phenomenology and the Cognitive Sciences, 7*, 67–83.
Kripke, S. (1980). *Naming and necessity*. Harvard University Press.
Levine, J. (1983). Materialism and qualia: The explanatory gap. *Pacific Philosophical Quarterly, 64*, 354–361.
Lewis, D. (1988). What experience teaches. *Proceedings of the Russellian Society, 13*, 29–57.
Liu, M. (2021). Revelation and the intuition of dualism. *Synthese, 199*(3–4), 11491–11515.
Nagel, T. (1974). What is it like to be a bat? *Philosophical Review, 83*, 435–450.
Peressini, A. (2014). Blurring two conceptions of subjective experience: Folk versus philosophical phenomenality. *Philosophical Psychology, 27*(6), 862–889.
Recanati, F. (2004). *Literal meaning*. Cambridge University Press.
Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhearers. *Cognitive Psychology, 21*, 211–232.
Snowdon, P. (2010). On the what-it-is-like-ness of experience. *Southern Journal of Philosophy, 48*(1), 8–27.
Sperber, D., & Wilson, D. (1998). The mapping between the mental and the public lexicon. In P. Carruthers & J. Boucher (Eds.), *Language and thought: Interdisciplinary themes* (pp. 184–200). Cambridge University Press.
Stoljar, D. (2016). The semantics of "what it's like" and the nature of consciousness. *Mind, 125*(500), 1161–1198.
Sytsma, J. (2009). Phenomenological obviousness and the new science of consciousness. *Philosophy of Science, 76*, 958–969.
Sytsma, J. (2010). Folk psychology and phenomenal consciousness. *Philosophy Compass, 5*(8), 700–711.
Sytsma, J. (2016). Attributions of consciousness. In J. Sytsma & W. Buckwalter (Eds.), *A companion to experimental philosophy* (pp. 259–278). Blackwell.
Sytsma, J., & Machery, E. (2009). How to study folk intuitions about consciousness. *Philosophical Psychology, 22*(1), 21–35.
Sytsma, J., & Machery, E. (2010). Two conceptions of subjective experience. *Philosophical Studies, 151*(2), 299–327.
Sytsma, J., & Ozdemir, E. (2019). No problem: Evidence that the concept of phenomenal consciousness is not widespread. *Journal of Consciousness Studies, 26*(9–10), 241–256.
Wilson, D., & Carston, R. (2007). A unitary approach to lexical pragmatics: relevance, inference and ad hoc concepts. In N. Burton-Roberts (Ed.), *pragmatics*. Palgrave-Macmillan.
Wilson, D., & Sperber, D. (2002). Truthfulness and relevance. *Mind, 111*(443), 583–632.