Brentyn J. Ramm[1]

Dimensions of Reliability in Phenomenal Judgment

Abstract: *Eric Schwitzgebel (2011) argues that phenomenal judgments are in general less reliable than perceptual judgments. This paper distinguishes two versions of this unreliability thesis. The process unreliability thesis says that unreliability in phenomenal judgments is due to faulty domain-specific mechanisms involved in producing these judgments, whereas the statistical unreliability thesis says that it is simply a matter of higher numbers of errors. Against the process unreliability thesis, I argue that the main errors and limitations in making phenomenal judgments can be accounted for by domain-general factors: attention, working memory limits, and conceptualization. As these factors are shared with the production of perceptual judgments, errors in phenomenal judgments are not due to faulty domain-specific processes. Furthermore, this account defends phenomenal judgments against general scepticism by providing criteria for distinguishing between reliable and unreliable phenomenal judgments.*

1. Introduction

Eric Schwitzgebel (2011) claims that introspective judgments about conscious experience are generally unreliable, and are far less reliable than perceptual judgments about the world. Some of the cases

Correspondence:
Brentyn J. Ramm, School of Philosophy, The Australian National University, Australia. *Email: brentynramm@gmail.com*

[1] Portions of this work were completed while the author was a visiting scholar at the Department of Philosophy, Ohio State University (USA).

Schwitzgebel presents against reliability include the failure to distinguish details about our own phenomenology, such as the basic details of visual imagery (Chapter 3). In Chapter 7, he highlights the uncertainty we have about the character and location of our emotions and that sometimes we may not notice our emotions at all such as feelings of grumpiness. Another case is that we tend to fail to notice broad features of our visual experience such as the fact that it is not clear all the way to the edges. Rather, there is only a small central region of clarity. Philosophers also disagree about what visual patterns manifest when one's eyes are closed (*ibid.*, Chapter 8).

Based upon the weight of such problem cases, Schwitzgebel draws the pessimistic conclusion that naïve introspection[2] is 'faulty, untrustworthy, and misleading — not just sometimes a little mistaken, but frequently and massively mistaken' (*ibid.*, p. 129). More specifically Schwitzgebel makes the comparative claim that phenomenal judgments are in general far less reliable than perceptual judgments. He holds, 'Descartes, I think, had it quite backward when he said the mind — including especially current conscious experience — was better known than the outside world' (*ibid.*, p. 136) and 'Our judgements about the world tend to drive our judgements about our experience. Properly so, since the former are the more secure' (*ibid.*, p. 137). It is this comparative claim that I will be focusing on here.

What does Schwitzgebel's comparative pessimistic thesis amount to? We can distinguish between two pessimistic theses:

> The Process Unreliability Thesis: Errors in the formation of phenomenal judgments are due to factors specific to the formation of phenomenal judgments.
>
> The Statistical Unreliability Thesis: Phenomenal judgments are less reliable than perceptual judgments overall in terms of proportion of errors.

Here I will argue against the process unreliability interpretation of introspective pessimism by showing how a domain-general framework can account for the main introspective errors and limitations. My alternative claim will be:

---

2 Naïve introspection involves phenomenal judgments which have not been supplemented by training or first-person methods. These are the type of phenomenal judgments I will focus on in this paper. In what follows I drop the term 'naïve' and simply refer to phenomenal judgments.

> The Domain-General Thesis*:* Errors and limitations in the formation of phenomenal judgments are due to factors that are domain-general in the sense that they are shared with the formation of perceptual judgments.

The domain-general thesis is compatible with a statistical version of the comparative pessimistic thesis. I will not be arguing against the statistical unreliability thesis here.

It is not clear in Schwitzgebel (2008; 2011) which thesis the problem cases are meant to provide evidence for. That it is the *statistical unreliability thesis* is implicit in Schwitzgebel (2012) where he argues for a domain-general account of introspection. The distinction between process unreliability and statistical unreliability needs to be made explicit, as this has important consequences for questions about the reliability of phenomenal judgments. In particular, if the *process unreliability thesis* is true, then the errors are likely to infect introspection alone. If the alternative *domain-general thesis* is true, then perceptual judgments will be prone to the very same errors as phenomenal judgments.

I propose that the production of phenomenal judgments involves a number of domain-general factors such as attention, working memory, and conceptualization. In the domain-general framework developed here attention selects experienced features or objects,[3] and activates/forms concepts in working memory which produce (or perhaps partly constitute) judgments about the experience.

Previous authors have suggested that attentional and conceptual processes can account for many introspective processes (Bayne and Spener, 2010; Block, 2007; Carruthers, 2011; Engelburt and Carruthers, 2010; Hill, 2011; Prinz, 2004; Schwitzgebel, 2012; Watzl and Wu, 2012), with a central role for domain-general processes (Carruthers, 2011; Prinz, 2004; Schwitzgebel, 2012). It remains unclear, however, whether such explanations are capable of accounting for most errors and limitations in making phenomenal judgments.

The best defence of the reliability of introspection may be to give up on a separate introspective process altogether, and rather just talk of phenomenal judgments that involve the same domain-general

[3] I am neutral here as to whether experience is transparent such that the features and objects which engage attention are only ever external (perhaps represented) properties and objects (Harman, 1990; Tye, 1995; 2000), or whether properties of experience can (at least sometimes) be directly attended (Block, 1996; Kind, 2003).

processes as perceptual and rational judgments. One may well think that a domain-general account puts pressure on the need to posit a special domain-specific self-monitoring mechanism (Armstrong, 1968; Goldman, 2006; Lycan, 1996; Nichols and Stich, 2003). Previous authors use a domain-general account to motivate scepticism about the existence of a special introspective process (Carruthers, 2011; Prinz, 2004; Schwitzgebel, 2012). However, I do not pursue this question here. The aim of the paper will not be to assess arguments for and against particular philosophical theories of introspection. Rather, here I focus upon discussing empirical evidence for and against the *domain-general thesis*. Furthermore, a domain-general account of errors is consistent with the existence of domain-specific introspective processes, as long as these are not strongly modular and hence preclude a role of domain-general processes in producing phenomenal judgments.

Here I attempt to give a scientifically informed account of domain-general processes that could plausibly account for the major classes of introspective errors and limitations. Accounting for all errors is obviously beyond the scope of a single paper. The aim here is rather to develop a general framework which could in principle account for most of the introspective errors and limitations. This project should also be of general interest to cognitive scientists as the assumption is that the same processes are also involved in making conscious judgments in general, and will hence be involved in most psychological tasks ranging from perceptual judgment to moral judgment, to mathematical calculation.

As an empirical hypothesis I cannot claim to establish the necessary truth of the *domain-general thesis*. Evidence which undermines the hypothesis can be uncovered at any time. Rather, my claim is that there is currently enough supporting evidence to make the thesis more sufficiently probable than the *process unreliability thesis*.

The aim of the paper will not be to argue for an exhaustive theoretical account of every problem case considered here, thus deciding whether it is definitely an error or not. Rather, the goal is to argue for the disjunction: problem cases for phenomenal judgments either do not count as errors or if they do count as errors then they can be accounted for by domain-general factors.

The plan for the paper is as follows: in Section 2, I discuss the motivation for using a domain-general approach as a response to introspective scepticism. In Section 3, I characterize the difference between phenomenal and perceptual judgments. In Sections 4, 5, and

6, I investigate in detail how domain-general factors (respectively, attention, working memory, and conceptualization) can account for problem cases for phenomenal judgments. I discuss cognitive dissociations and their challenge to a domain-general framework in Section 7. In Section 8, I argue against general scepticism towards phenomenal judgments. In Section 9, I conclude by summarizing the case for the *domain-general thesis*.

2. Background Motivations

That phenomenal judgments are generally untrustworthy (or at least exhibit very high levels of unreliability) are motivations for Daniel Dennett's (1991; 2001a) rejection of first-person methods and in Schwitzgebel taking a pessimistic position towards the possibility of reliable methods (Hurlbert and Schwitzgebel, 2007; Schwitzgebel, 2011, pp. 129, 167, though see Chapter 4). Philosophers, in conversation with me, have also professed pessimism about first-person methods on the basis of Schwitzgebel's problem cases, presumably because they interpret them as establishing the general untrustworthiness of introspection.

One motivation for the paper is to provide an initial defence of the possibility of reliable first-person methods as part of a science of consciousness.[4] That there can be reliable first-person methods is compatible with holding that naïve introspection is statistically less reliable than perceptual judgments. Hence the truth or falsity of the *statistical unreliability thesis* is not a concern from the point of view of a researcher in the science of consciousness. This seemed to be the position held by Titchener (Schwitzgebel, 2011, Chapter 5).

If Schwitzgebel's cases are taken as establishing that all phenomenal judgments are untrustworthy — 'introspective scepticism' (Bayne and Spener, 2010) — then this result arguably undermines the possibility of reliable first-person methods. This may be the case because phenomenal judgments stem from pervasively faulty processes. Thus no amount of training nor use of methods would be expected to improve them.

Introspective scepticism also presents a danger to the many everyday phenomenal judgments that are *prima facie* highly reliable: for

[4] For examples of previous defences of the use of introspection in science, see Goldman (2004) and Hatfield (2009, Chapter 16), and Kriegel (2013).

example, judging that I feel hungry, feel an intense toothache, foveal colour judgments, and similarity judgments such as orange seems more similar to yellow than to blue. As Bayne and Spener (2010, p. 8) point out, it seems perverse to doubt the reliability of these phenomenal judgments. No amount of philosophical argument will convince someone, for example, that they should doubt the fact that they are experiencing a severe toothache. Even Schwitzgebel (2011, p. 139) pulls back from distrusting these judgments, yet without a positive reason for trusting these judgments general introspective scepticism looms as threat.

Introspective scepticism based upon faulty processes can be undermined (or at least rendered unattractive) by showing that errors in phenomenal judgments stem from domain-general processes. However, if the number of errors was high enough then this could still provide a reason for introspective scepticism. I respond to this further threat in Section 8.

3. Phenomenal Judgments and Perceptual Judgments

The term 'introspection' originates from ordinary language, and suggests some sort of 'inner looking' as opposed to perception or 'outer looking'. However, I have doubts that there is a substantially separate faculty of introspection apart from those processes which produce perceptual and intellectual judgments. So as to remain neutral about the underlying processes, following Chalmers (1996, pp. 173–6), I generally use the term 'phenomenal judgment' rather than introspection.

For the purposes of this paper, a 'phenomenal judgment' is a judgment about one's current phenomenology, formed using attentional resources, on the basis of (or intended to be on the basis of) current relevant experience.[5] This includes judgments about thoughts, emo-

[5] The standard view of the basing relation is that it is a type of causal relation between the reason for having the belief (here the experience) and the belief (Korcz, 2010). An alternative to a causal relation is a constitutive relation, for instance if the concept of 'red' is partly constituted by a presently experienced phenomenal character of red. This direct phenomenal concept can in turn be part of a direct phenomenal belief such as 'the apple looks red' (Chalmers, 2010). Though this constitutive relation would only apply to a small set of beliefs. This is not the place to defend a theory of the basing relation. It should be sufficient for current purposes if the reader understands 'basing' as a form of

tions, pain, mental imagery, and sensory experiences (for example, how things seem, look, appear, feel). Some examples of judgments about experience are 'the rose looks red', 'the stove feels hot', 'the town looks far away'. Thus one may judge that 'the stove feels hot' on the basis of the stove feeling hot. I say 'intended to be on the basis of' because I want to allow that one could base their judgment upon the wrong experience such as the stove feeling cold. Otherwise the possibility of erroneous phenomenal judgments may be ruled out by definition. I say 'relevant experience' because that the stove feels hot should not be based upon the stove looking hot, but rather upon the stove feeling hot.

The relevant class of 'perceptual judgment' are judgments about objects, events, and properties of the world, formed using attentional resources, on the basis of (or intended to be on the basis of) current relevant conscious perception. For example, 'the rose is red', 'the stove is hot', and 'the town is far away'. Thus one may judge that 'the stove is hot' on the basis of the stove feeling hot (perhaps 'looking hot' also counts, if HOTNESS is a property of visual experience). However, for the purposes of this paper judging that 'the stove is hot' on the basis of seeing smoke coming from the chimney would not count, as the conscious perception does not include the stove and HOTNESS, but rather these are derived by inference from a conscious perception.

Schwitzgebel (2012, pp. 34–5) points out that it is easy to confuse judgments about sensory experience and properties of the world. For example, in a psychophysics experiment one may make repeated judgments about colour experience (the colour things look to have). In this case it is easy to slip into making a judgment about properties of the world (e.g. about the stimuli on the screen). Thus a judgment that the rose 'looks red' and 'is red' are often interchangeable in ordinary circumstances.

Importantly for present purposes, even though confusion can happen, phenomenal and perceptual judgments are nevertheless conceptually distinct. With enough effort or when I think an illusion may be involved, 'looks red' and 'is red' are sharply distinguishable. When I look at a white rose through red tinted glasses I may judge that 'the rose looks pink' but, having seen it without the glasses, nevertheless also believe that 'the rose is actually white'.

causal relation or constitutive relation between the reason for the belief (i.e. experience) and the belief.

4. Attention

In the next three sections I discuss how errors and limitations in making phenomenal judgments can be accounted for by domain-general factors. I discuss attention in this section.

Inattention is responsible for errors in both perceptual and phenomenal judgments. The effect of failing to attend in making judgments about the world is dramatically evident in studies of inattentional blindness in which subjects fail to notice unexpected events such as a gorilla walking through a basketball game when their attention is distracted (Simons and Chabris, 1999: see also, Mack and Rock, 1998; Most *et al.*, 2005; Simons, 2000).

An introspective case that is plausibly explained by inattention is my failure to notice the degree of acuity of my visual field. Schwitzgebel (2011) and Dennett (1991, pp. 53–4, 68; 2001b; 2002) point out that many people believe that their visual field is clear all the way to the edges, whereas objects in the periphery actually appear indistinct and blurry. Furthermore there is only a small central region of clarity of two degrees, whereas most would judge it to be a large window of clarity. That subjects possess a false belief about their visual field is evidenced by the surprise people express upon being shown the low resolution for shapes and colours of peripheral vision (Dennett, 1991, p. 68; 2001b).

This case seems to be at least in part a failure in attentional orienting, in particular a failure to attend to that location while keeping their eyes fixated (Schwitzgebel, 2011; Engleburt and Carruthers, 2010; Hill, 2011; Waltz and Wu, 2012). It is has also been previously argued that subjects' belief in a large window of clarity is due to them making a judgment about dynamic vision in contrast to gaze-fixed vision (see Hill, 2011, p. 27; Engelburt and Carruthers, 2010, p. 251). In normal vision, the eyes perform a saccade a few times a second (e.g. Land, 1999), which the visual system combines into a single visual scene. We would not expect subjects to be aware of information at such short durations any more than we would expect them to see the images of a film as static rather than moving, or a spinning flame on the end of a pole as a point of light rather than a circle of flame. Dennett's and Schwitzgebel's subjects were only mistaken, then, if their beliefs were about fixed-gaze vision. If their belief in a large window of clarity referred to typical, dynamic vision, then they (and most of us) were correct after all.

Interestingly, both Schwitzgebel and Dennett provide first-person methods for keeping one's eyes fixated while orienting attention to the periphery. An example would be focusing on one's thumbnail held at arm's length while attending to peripheral vision (Hill, 2011, p. 27). Upon doing this, the lack of acuity becomes highly evident. This shows that, while naïve first-person observations often fail, more sophisticated first-person techniques can succeed.

Another challenge to the reliability of first-person judgments is the apparent fact that I often fail to notice some of my experiences, such as being angry, sad, depressed, or anxious. Schwitzgebel (2011, pp. 122–3) gives the case of my partner mentioning that I seem to be grumpy about doing the dishes. I carefully reflect on this and deny that I am feeling grumpy. But from the look on my face, and the way I bang the dishes about, it is evident that I am grumpy. Perhaps upon further reflection I do actually detect feelings of grumpiness. It seems that my initial judgment was mistaken. Even though I attempt to carefully reflect, it is likely that the anger will cause me to be disposed to reject the accusation that I am grumpy without attending sufficiently carefully. If I am repressing the emotion then I may be disposed (have an attentional bias) to attend briefly to the emotion and then rapidly disengage attention, hence inhibiting explicit recognition of its presence (Derakshan, Eysenck and Myers, 2007).

Alternatively no error may have occurred at all. Other possibilities are that the emotion did not reach consciousness, or was merely dispositional, or perhaps I did not really feel grumpy at T1, but only felt so after the fact when my partner suggested that I am grumpy. Also even if I retrospectively recall that I did seem to feel grumpy at the time, my initial judgment may still be correct, as this may be a false memory created by the suggestion that I am grumpy.

Even if the above explanations of errors are correct, it does not follow that the attention used in making phenomenal judgments is a domain-general faculty. Perhaps there is a special faculty of 'introspective attention' that is responsible for these errors. For examples of the use of this term see Bayne and Spener (2010, p. 12), Hohwy (2011, p. 270), Schwitzgebel (2011, pp. 126, 175).

A reason for rejecting this hypothesis is the observation that attending to visual images, pains, thoughts, and emotions can distract someone from making correct perceptual judgments and vice versa. If there is a separate introspective attention then I should be able to fully attend to my thoughts and the world simultaneously without any interference. For example, on this view it is difficult to account for why I

am more likely to walk into street poles when I am absorbed in my thoughts. This strongly suggests that in attending to my thoughts I was withdrawing attention from the world.

There is also evidence for a single attentional mechanism in the domains of emotion and pain. For example, it has been found that felt pain is reduced when attention is distracted (e.g. Buhle and Wager, 2010; Legrain *et al.*, 2009) and high working memory load reduces negative affective responses to negative images (Schmeichel, Volokhov and Demaree, 2008; Van Dillen and Koole, 2007) and reduces attentional capture by pain (Legrain *et al.*, 2011). The best explanation of these findings seems to be that judgments about emotions and pains (so-called introspections) utilize the same attentional system.

Of course, there are exceptions when I do not seem to be interfered with by attending to multiple domains, for instance when attending to my thoughts or feelings and the traffic as I am driving. Musing that I feel happy today does not always seem to interfere with my ability to drive efficiently. However, such cases do not necessitate that there are domain-specific attentional mechanisms, as they can be explained by rapid switches of attention between the two activities, or by one of the activities continuing 'on autopilot' while the other consumes my attention.

Furthermore, on a resources theory, attention is a single limited resource that can be spread between multiple domains. On this view significant interference only occurs when attentional resources are exhausted. For example, it has been found that talking on a hands-free cell phone while driving significantly reduces driving performance even though the two activities rely upon different sensory modalities (Strayer and Johnston, 2001). Finally, while there is empirical evidence in favour of a single domain-general selection process, I am unaware of any empirical research which supports the existence of a special introspective attentional mechanism.

5. Working Memory Limitations

What about unreliability in judging aspects of experience that I am in fact currently attending to, such as thoughts and emotions? My claim here is that the limitations of working memory explain the difficulty of judging features of complex and dynamic events such as emotions and much of our mental imagery.

Working memory is usually thought to have a central capacity-limited component or an upper limit in processing resources (e.g. Baddeley, 2000; 2003; Cowan, 2001; Just and Carpenter, 1992; Halford, Wilson and Phillips, 1998), and time-limited short-term memory stores (e.g. phonological, visual-spatial) which are subject to decay and interference. Some theorists include a short-term store for affective information (Mikels *et al.*, 2008). I suggest that limits in judging complex phenomena stem from an upper limit to the number of concepts that can be formed/activated at any one time in working memory (Block, 2007, pp. 487–9).[6] Furthermore, the unreliability in judging dynamic phenomena stems from the failure to maintain concepts active in working memory.

5.1. Complexity and the Capacity Limits of Working Memory

Unsurprisingly in vision and other modalities, when phenomena are highly complex, it is more difficult to categorize all of their properties. For instance, it is difficult to report all of the items in a complex scene. This can be explained by the limits of working memory. Cowan (2001) presents a large amount of experimental evidence that there is a capacity limit to working memory of approximately four items (chunks). For example, there is a limit of approximately four items in visual working memory (Luck and Vogel, 1997), recall of verbal material (Broadbent, 1975), in the discrimination of complex aromas (Laing and Francis, 1989), and in recalling items from visual iconic memory (Sperling, 1960). That a capacity limit of four holds between so many domains, including at least some paradigmatically introspective processes, suggests a single central capacity limit.

An alternative interpretation is to posit that there are modality-specific resources such as separate capacities for visual and auditory information rather than a central attentional resource (Wickens, 1984; 2002). However, importantly Saults and Cowan (2007) found that for a simultaneously presented visual and auditory array recall was limited to 3–4 items, thus providing evidence for a single capacity limit which goes across sensory modalities. This finding is difficult to account for by a multiple capacities view.

[6] Ramm and Halford (2012) provide evidence that conceptual combination draws upon capacity-limited processing resources, and thus that new concepts are formed in working memory.

As discussed above, research in emotion and pain provides some evidence for the domain-general hypothesis. There has been far more extensive research in visual imagery and iconic memory, and hence more evidence supporting the hypothesis that these draw upon the same capacity-limited resources as visual perception. For example, it has been found that mental imagery is associated with many of the same brain areas as high-level visual perception (Ishai and Sagi, 1995; Kosslyn, Ganis and Thompson, 2001; Kosslyn, Thompson and Alpert, 1997), and so it is plausible that mental imagery relies upon the same capacity limits as visual perception. It has also been found that it takes longer to generate a more complex imagined letter or shape than a simple one (Bruyer and Scailquin, 1998; Kossyln, 1980; Dror and Kosslyn, 1994). Furthermore, the generation and rotation of images is interfered with more by random letter generation (a task that requires capacity-limited resources) than articulatory suppression (which is considered to be a relatively automatic process) (Bruyer and Scailquin, 1998).

The capacity limits in extracting information from experiences also accounts for the classic problem case of why I do not know how many spots are on an imagined speckled hen. Studies on subitizing (automatic recognition of number) show a limit of approximately four in subitizing items in briefly viewed displays (Mandler and Shebo, 1982; Trick and Pylyshyn, 1994) and after-images (Atkinson, Campbell and Francis, 1976; Simon and Vaisnavi, 1996). Again, whilst such evidence shows that knowledge of experience is not infallible, neither in such cases is there a reason to believe that it is more fallible than perceptual knowledge in similar cases. Judging the number of spots on a briefly perceived speckled hen is just as difficult as judging the number of spots on an imagined hen (although the instability of mental images probably contributes an additional source of errors/limits).

The evidence suggests, then, a capacity limit to making phenomenal judgments. Importantly, this limit applies to both phenomenal and perceptual judgments (see also Hill, 2011, pp. 28–31). This hypothesis also partly explains the variability in subjects' reports regarding the patterns seen when one's eyes are closed. As these visual patterns are both complex and dynamic, inaccuracy and variability would be expected in reports. This is no different from the fact that we would not expect subjects to give a perfectly accurate account of the colours and patterns of a fireworks display. When the 'inner light show' and 'outer light show' share the same degree of complexity one would

expect both to lead to similar levels of inaccuracy in reports. One may also expect more individual variability in eyes-closed patterns than fireworks, due to differences in lighting conditions, and perhaps even differences in visual systems.

Of course, this cannot be the whole story. It is odd for instance that historically Goethe, Purkinje, Muller, and Helmholtz all report 'wandering cloudy stripes', but such descriptions disappear after the early nineteenth century (Schwitzgebel, 2011, pp. 142–9). It seems like theory must have been driving these descriptions, in particular influence from reading previous authors such as Purkinje.

5.2. Dynamicity and the Failure to Maintain Activation of Concepts

If a phenomenon is rapidly changing over time this adds further difficulties to making a phenomenal judgment. In fact, emotions, thoughts, and pains seem to be more like sounds in their dynamically changing character. Visual imagery also seems to be unstable, which can be considered a form of dynamicity. Schwitzgebel asks, what are the gross and fine features of emotion and pain? We seem to be reliable at identifying broad features of emotions, pleasures, and pains such as intensity, pleasantness, unpleasantness, and often use dynamic terms such as surging, crescendo, fading, etc. (Lambie and Marcel, 2002, p. 230).

I propose here that limitations in making phenomenal judgments about dynamic experiences are due to failures in retaining concepts active in short-term memory. Suppose that there are rapidly flashing coloured lights (green, red, yellow, blue, green, yellow, red). I am attending to them and they are not so brief that I am unaware of them.[7] Yet I fail to judge that 'there was a blue light'. One possibility is that BLUE was not encoded in short-term memory and thus I was unable to report it. Another possibility is that BLUE was encoded in short-term memory, but it was interfered with by preceding or subsequent concepts such as RED. Or finally BLUE was encoded but the information was lost due to decay. In all cases, it is a failure to maintain a concept active which explains the oversight. The same limitations also would presumably apply to comparable perceptual judgments about the

[7] At very short durations the phenomena may be masked and not even reach consciousness at all: Breitmeyer and Ogmen (2006); Kouider and Dehaene (2007).

dynamically changing shape of ripples on a lake, the topology of flame shape, and the number of forks on flashes of lightning.[8] Overall, then, errors and limitations involving complexity and dynamicity do not seem to derive from domain-specific introspective processes.

6. Conceptual Errors

I have proposed that working memory places a limit on how many concepts can be formed/activated at a time, and thus how much information I can extract from an experience, and how long concepts remain active. Two other ways in which conceptual errors/limitations may occur include: 1. missing concepts, 2. uncertainty in using concepts.[9]

6.1. Missing Concepts

As Schwitzgebel points out there are many aspects of the character of emotions that I cannot decide upon. Even holding fixed working memory limitations there are many details of my emotion and pain experiences that I simply cannot categorize. The main reason for this is that I usually lack the appropriate concepts for delineating the details of these experiences. For example, I may not possess adequate concepts for distinguishing between annoyance and jealousy, and thus think that I am merely feeling annoyed rather than jealous. It's not a matter of just possessing the words. It may come as a surprise to me when I realize that I feel jealous. I may have an incomplete concept of JEALOUSY, for example I know how people act when jealous, but presumably I do not know what it's like to feel jealous until I actually experience it (and realize what it is I'm feeling). The same applies to feelings such as love, grief, awe, aesthetic pleasure, etc. I can possess the words for these without possessing a phenomenal concept for these feelings.

Does the difficulty in acquiring concepts of certain emotions suggest that there is a domain-specific limitation here? In a sense it does, but this does not entail the existence of a faulty or more limited

[8] Watzl and Wu (2012) also emphasize this point.

[9] A third type of conceptual error which I do not discuss here is the activation of incorrect concepts. For example, suppose I am exposed to something cold such as ice, but I have been primed to expect a sensation of heat and hence I briefly mistakenly judge that I am feeling heat. Again this type of error affects both phenomenal and perceptual judgments.

domain-specific process. As to why we may possess fewer concepts in one domain than another it is helpful to consider how we often acquire phenomenal concepts. An example is colour experience. I can easily acquire the phenomenal concept of turquoise because someone can point to something that is turquoise, such as a turquoise opal. By doing this I am provided with a phenomenal sample of turquoise. I learn what it is like to experience it, how it relates to other colours (e.g. that it is a greenish-blue) and I can then identify it again on other occasions.

Emotions on the other hand are more difficult to learn presumably because it is more difficult to provide a phenomenal sample. Someone cannot simply flick a switch and turn on my feeling of tenderness. They may set up the appropriate situation or show a scene from a movie where characters are portraying this feeling, but it is not guaranteed that I will thereby feel tenderness, and certainly not as reliably as I can be made to experience turquoise. Thus it seems that I will be unlikely to acquire concepts for the myriad affective subtleties, like anger 22 versus anger 23, as I can potentially do for colour experience. This does not show that there is a unique process for making judgments about emotions that is less reliable or more limited than for making judgments about colours. Rather, there is a non-introspective limitation here. If colours were as difficult to manipulate as emotions, and as unstable, then they would be just as difficult to learn.

6.2. Uncertainty in Using Concepts

On other occasions, I possess the appropriate concept, and it is currently activated, but I am uncertain as to how to employ it. An example, given by Schwitzgebel, is in making judgments about visual imagery. Schwitzgebel asks if you can give medium-level details of your mental image of a house (Schwitzgebel, 2011, Chapter 3). How clear is it? How detailed is it? How stable is it? Your uncertainty in answering these questions suggests that you are an unreliable introspector.[10]

[10] Contrary to Schwitzgebel's claim, Pearson, Rademaker and Tong (2011) used an ingenious binocular rivalry task to show evidence that subjects are reliable at judging the vividness of their mental images. Using the same paradigm, Rademaker and Pearson (2012) found that subjects can improve this metacognitive ability with training. My thanks to Helen Yetter-Chappell for pointing me towards these studies.

Even though we presumably do get such questions wrong on occasion, there are analogous examples of perceptual judgments where there are similar amounts of uncertainty. For example: 'Where exactly does that haystack begin and end?' 'How large is that bus?' 'How bushy is that tree?' These are difficult questions to answer. Yet this does not show that my perception of haystacks, buses, or trees is unreliable, only that I am not always sure how to apply these categories. Of course, the instability and low resolution of mental images most likely makes answering many questions about them more difficult than many visual perceptions.

7. Cognitive Dissociations

Cognitive dissociations present a challenge to a domain-general theory as they provide examples of deficits in making judgments within specific modalities, while leaving general cognitive abilities intact. For example, visual associative agnosia is the inability to identify visually perceived ordinary objects such as cups despite intact visual perception (Farah, 2004). Subjects with this condition can also recognize a familiar object as a cup if they are allowed to touch it, thus showing that the conceptual system is intact, and that the deficit is specific to visual recognition. Such neurological conditions suggest that there are domain-specific informational links between different sensory modalities and the activation of concepts in working memory.

One particularly relevant example for the present enquiry is alexithymia, which is a clinical condition which literally means 'no words for feelings' (Samur *et al.*, 2013).[11] It is characterized by a subject's difficulty in describing emotions. For example, sufferers may not be able distinguish whether they are feeling angry, sad, or anxious, or distinguish whether they are feeling a bodily sensation or an emotional state, and they also may not be able to identify others' emotions. Alexithymia is a challenge for a domain-general cognitive theory because sufferers have difficulties describing their emotions even though their attentional and working memory systems and other conceptual abilities remain intact, thereby suggesting the existence of a domain-specific process.

In fact, two types of alexithymia have been distinguished (Larsen *et al.*, 2003). Type I is associated with a lack of emotional experience.

[11] Thank you to an anonymous reviewer for pressing me on this case.

Since sufferers feel a blunting or an absence of emotions, this has a follow-on effect to the cognitive processing of emotions in general. An analogy here is colour-blindness (Lane *et al.*, 1996). A subject that is colour-blind has problems identifying colour, but this is due to not experiencing them in the first place, not due to a general attentional or conceptual malfunction. Hence this form of alexithymia is not problematic for the current hypothesis.

Type II alexithymia appears to be analogous to visual associative agnosia. It is characterized by no deficit in feeling emotions, but predominantly in identifying and describing them. This form of alexithymia is associated with damage to the person's corpus callosum and hence a disorder in interhemispheric communication. This suggests that there can be a domain-specific disruption between the experiencing of an emotion and the activation of a concept for the emotion in working memory.

This is compatible with the present domain-general framework for three reasons: firstly, it does not entail that identifying emotions relies upon a special introspective attentional mechanism, working memory resource, or conceptual system, only that there is an additional domain-specific informational link (particularly in interhemispheric communication).

Secondly, the existence of domain-specific processes does not invalidate the present framework because the hypothesis was that *most* errors can be accounted for by domain-general processes (not that there are no domain-specific processes which contribute to phenomenal judgments). This is consistent with domain-specific processes also contributing to the judgment as long as their role in producing errors in neurally typical individuals is relatively small. While I cannot conclusively prove that this is in fact the case, I think that there can be indirect evidence for this hypothesis. In particular, in accounting for the unreliability of different phenomenal judgments I have drawn upon attentional, working memory, and conceptual factors, as well as properties of the target phenomenal states such as complexity and dynamicity. There has so far been no need to posit any other significant sources of error in accounting for problem cases in typical individuals.

Thirdly, alexithymia is also associated with a difficulty in identifying others' emotional states (a judgment about the world) as well as one's own emotional states (a phenomenal judgment). Lane *et al.* (1996) present evidence that individuals with alexithymia have difficulty in recognizing the emotions displayed by faces. That is, it is not

just a deficit unique to identifying one's own emotional states, but in emotion processing in general. Goldman (2006) also reviews evidence that a deficit in feeling a specific emotion is paired with a deficit in recognizing the same specific emotion in others. This suggests that making judgments about one's own emotions and those of others share underlying processes, which is a version of the domain-general thesis. Thus, even if evidence was found that these informational links were significantly unreliable in neurally typical subjects, alexithymia still does not provide evidence for a separate introspective process apart from processes involved in making judgments about the world. Rather it is consistent with the domain-general thesis.

8. Against Introspective Scepticism

In Sections 4–6, I argued that most errors and limitations of phenomenal judgments can in principle be accounted for by the domain-general dimensions of attention, working memory limitations, and conceptualization. If true then this account undermines the *process unreliability thesis.* I provide reasons in support of the *domain-general thesis* in the next section.

What about the statistical unreliability thesis? It can still be held that phenomenology is in general more complex, dynamic, and elusive, and thus we will always be less reliable at judging it than facts about the world. The *statistical unreliability thesis* is compatible with the domain-general account. According to the current framework, the greater unreliability of phenomenal judgments (if they are indeed less reliable) is typically due to properties of the target of the judgment such as complexity and dynamicity, or other factors such as a lack of expertise in making phenomenal judgments, rather than a fault in specific processes which produce the judgment. This can lead to statistical differences in the reliability of phenomenal judgments, but this uniqueness does not imply that the errors stem from separate domain-specific processes.

Perhaps a type of scepticism can be based upon statistical unreliability. In this section I consider some motivations for holding the statistical unreliability thesis. I then present an argument for blocking a move from statistical unreliability to introspective scepticism.

In my view, Schwitzgebel's (2011, p. 136) claim that experience is almost always 'gelatinous, disjointed, swift, shy, changeable' is incorrect, especially when it comes to perceptual experience. For

example, my visual experience of desks, the sky, and trees are very stable. In general, however, the dynamicity of thoughts, emotions, and mental imagery is a point in favour of comparative statistical unreliability.

Another consideration in favour of this thesis is the possibility that we are experts when it comes to making judgments about the world as social factors and biological survival depend upon this.[12] Since we have far more practice at getting these judgments right perhaps we are more likely to attend more effectively and to acquire more fine-grained concepts like expert wine tasters (Ballester *et al.*, 2008; Solomon, 1990) than for untrained judgments about experience.

Perhaps we can also chunk items of the world more efficiently and thus can make more accurate judgments about complex situations such as being in traffic in comparison to the patterns when one's eyes are closed, just as chess experts are better at assessing and memorizing positions on a chess board than novices (Chase and Simon, 1973). Also there may be few negative consequences if we get details of our emotional phenomenology, visual acuity, visual imagery, and eyes-closed phenomenology incorrect, and so perhaps we tend to be novices particularly when it comes to these experiences.

This being said, judgments about sensory experience such as 'the rose looks red' are intimately related to perceptual judgments such as 'the rose is red'. In making both types of judgment, I attend outwardly to the rose and its perceptible properties. I do not shift my attention inwards to decide how the rose looks. Schwitzgebel (2012, pp. 34–5) makes a similar point when he discusses how we easily slip between judgments about our perceptual experience and properties of the world, and takes it as evidence for overlapping processes in making judgments. However, he does not mention the implication of this — that this is a reason for thinking that a large number of phenomenal judgments and perceptual judgments (in the veridical case) are hence equivalent in reliability. Furthermore, expertise in one would then be expected to entail an equivalent level of expertise in the other. Overall, then, the novice argument is another point in favour of statistical unreliability for some though not all phenomenal judgments. The flip-side of this argument is that it should be possible (at least in principle) to train subjects to overcome these errors.

12 Thank you to Eric Schwitzgebel for suggesting this argument to me.

Assuming that the statistical unreliability of phenomenal judgments was sufficiently high (although this seems unlikely), one might hold that this justifies a form of scepticism in which one should mistrust all of these judgments. Thus it may be that a strong form of pessimism is still viable based upon their statistical unreliability. I think that the current account blocks the threat of scepticism based upon statistical unreliability by providing criteria for distinguishing between reliable and unreliable phenomenal judgments.

The proposal is that identifying the dimensions involved in making phenomenal errors vindicates the apparently trustworthy cases of phenomenal judgments such as foveal colour judgments, knowing that I am in intense pain, etc. by showing why they are so often correct. These highly trustworthy, virtually undoubtable judgments are characterized by the stability and high intensity of the experience, and the high degree of attentiveness at the time of the judgment, as well as the possession of appropriate concepts.

This enables us to locate the reliability of a particular phenomenal judgment in reliability space (Figure 1). I will provide an example using wine tasting judgments (for the sake of simplicity I do not include all of the relevant factors in the figure).[13] For example, the judgments of a wine taster who is inattentive and with a low degree of conceptual adequacy would be located at A (e.g. an inattentive novice taster); inattentive but high conceptual adequacy at B (e.g. an inattentive expert taster); attentive but low conceptual adequacy at C (e.g. an attentive novice taster); and attentive and high conceptual adequacy at D (e.g. an attentive expert taster).[14] As there are principled criteria for distinguishing between reliable and unreliable phenomenal judgments, there is no reason to doubt judgments that reside in a superior position in reliability space, thus undermining the potential slide into introspective scepticism.

A response to this account by the pessimist is to move the pessimism to a higher level and ask: how do you know that you are in a

13 Errors due to the complexity or dynamicity of the experience can be incorporated into this simplified reliability space as these tend to cause a failure of the appropriate concepts to either be activated or maintained in working memory. A more explanatorily adequate figure would also include these factors.

14 In a study by Solomon (1990) it was found that wine experts are better at picking the odd wine out of three glasses of wine than novices, which suggests that it is the possession of concepts, rather than terminology, that differentiates experts from novices (see also Ballester *et al.*, 2008).

situation where you are being appropriately attentive and have adequate concepts, etc.? However, this question also applies to perception, reasoning, and mathematics so one cannot be a higher-order pessimist about phenomenal judgments without also generalizing this pessimism to these other domains. This argument works independently of the truth of the *domain-general thesis*. Suppose evidence arose that there are multiple domain-specific attentional mechanisms. Nevertheless, by identifying a *type* of error as attentional, this provides criteria for deciding if a phenomenal judgment is likely to be erroneous.

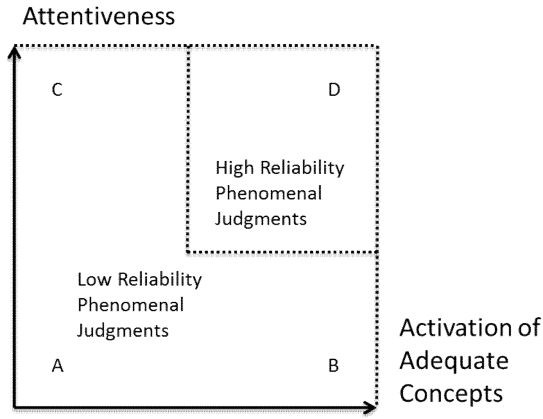

Figure 1. Reliability space for phenomenal judgments on the dimensions of attentiveness and conceptual adequacy. (A) Low attentiveness and low conceptual adequacy, (B) low attentiveness and high conceptual adequacy, (C) high attentiveness and low conceptual adequacy, (D) high attentiveness and high conceptual adequacy.

9. The Case for the Domain-General Thesis

The primary aim of the paper was to argue against the *process unreliability thesis* which holds that phenomenal judgments are unreliable due to faulty domain-specific processes. The strategy was to argue that problem cases for phenomenal judgments can either be discounted as non-errors or accounted for by domain-general factors such as attention, working memory, and conceptualization. Thus it is not that mechanisms involved in making phenomenal judgments are unreliable *per se*, but any judgment in which there are attentional and conceptual errors or working memory limits are exceeded, including

phenomenal, perceptual, mathematical, and rational judgments. If true then this account undermines the *process unreliability thesis.*

Furthermore, this would significantly limit the scope of pessimism about phenomenal judgments because perceptual judgments would also suffer from the same errors and limitations as phenomenal judgments. One outcome of the current account, if true, is that theorists who claim that phenomenal judgments are generally unreliable must mean that they are unreliable in terms of overall numbers, not that the underlying processes are unreliable. If they mean that the underlying processes themselves are generally unreliable, then the claim would overgeneralize, such that perceptual and intellectual judgments would also be generally unreliable. This is presumably an outcome that no one wants.

Should we accept the domain-general account of errors and limitations? As it is an empirical hypothesis it is always open to defeating evidence, so I cannot claim to have proved the hypothesis here. I do, however, think that it is sufficiently probable enough to warrant assenting to this thesis over a domain-specific account. The reasons in favour of the present thesis are:

(1) Attentional distraction between different modalities, suggesting that there is a single focus of attention. If there was a separate introspective attention then I should be able to attend to my thoughts and the world simultaneously without distraction.
(2) A working memory capacity-limit of approximately four chunks for many different domains (Cowan, 2001) (including stereotypically introspective cases), as well as direct empirical evidence for a single capacity across sensory modalities (Saults and Cowan, 2007).
(3) Phenomenal judgments and perceptual judgments are subject to the same types of conceptual errors.
(4) That most errors and limitations for neurally typical individuals stem from a domain-general system is compatible with cognitive dissociations such as alexithymia.
(5) The domain-general framework is more parsimonious than positing multiple domain-specific systems that are responsible for the same types of errors and limitations. Why posit special types of attention, working memory capacity, and conceptualization when a single system could perform the same functions? This arguably shifts the burden of proof onto the proponent of the *process unreliability thesis*.

At the very least this paper blocks an argument for the *process unreliability thesis*. However, I claim that we can draw a stronger conclusion than this.

The above reasons do not prove the hypothesis, but I think that they provide sufficient support, in conjunction with being the simpler hypothesis, for provisionally accepting the *domain-general thesis* over the *process unreliability thesis* barring evidence arising to the contrary.

The present account also provides systematic criteria for distinguishing reliable from unreliable phenomenal judgments, which blocks the argument for general introspective scepticism based upon statistical unreliability. By limiting the scope of pessimism, this paper also contributes to the wider project of defending the possibility of reliable first-person methods for investigating conscious experience. The specific details of what such first-person methods might consist in is a topic for another time.

Acknowledgments

I am grateful to David Chalmers, Frank Jackson, Colin Klein, Eric Schwitzgebel, Jonathan Simon, Declan Smithies, and Daniel Stoljar for helpful comments on the manuscript.

References

Armstrong, D. (1968) *A Materialist Theory of the Mind*, London: Routledge & Kegan Paul.

Atkinson, J., Campbell, F.W. & Francis, M.R. (1976) The magic number 4±0: A new look at visual numerosity judgements, *Perception*, **5** (3), pp. 234–334.

Baddeley, A. (2000) The episodic buffer: A new component of working memory?, *Trends in Cognitive Sciences*, **4** (11), pp. 417–423.

Baddeley, A. (2003) Working memory: Looking back and looking forward, *Nature Reviews Neuroscience*, **4** (10), pp. 829–839.

Ballester, J., Patris, B., Symoneaux, R. & Valentin, D. (2008) Conceptual vs. perceptual wine spaces: Does expertise matter?, *Food Quality and Preference*, **19** (3), pp. 267–276.

Bayne, T. & Spener, M. (2010) Introspective humility, *Philosophical Issues*, **20**, pp. 1–22.

Block, N. (1996) Mental paint and mental latex, *Philosophical Issues*, **7**, pp. 19–49.

Block, N. (2007) Consciousness, accessibility, and the mesh between psychology and neuroscience, *Behavioral and Brain Sciences*, **30** (5–6), pp. 481–499.

Breitmeyer, B. & Ogmen, H. (2006) *Visual Masking: Time Slices through Conscious and Unconscious Vision*, vol. 41, Oxford: Oxford University Press.

Broadbent, D.E. (1958) *Perception and Communication*, New York: Pergamon.

Bruyer, R. & Scailquin, J. (1998) The visuospatial sketchpad for mental images: Testing the multicomponent model of working memory, *Acta Psychologica*, **98** (1), pp. 17–36.
Buhle, J. & Wager, T.D. (2010) Performance-dependent inhibition of pain by an executive working memory task, *Pain*, **149** (1), pp. 19–26.
Carruthers, P. (2011) *The Opacity of Mind: An Integrative Theory of Self-Knowledge*, Oxford: Oxford University Press.
Chalmers, D.J. (1996) *The Conscious Mind: In Search of a Fundamental Theory*, New York: Oxford University Press.
Chalmers, D.J. (2010) *The Character of Consciousness*, Oxford: Oxford University Press.
Chase, W.G. & Simon, H.A. (1973) Perception in chess, *Cognitive Psychology*, **4** (1), pp. 55–81.
Cowan, N. (2001) The magical number 4 in short-term memory: A reconsideration of mental storage capacity, *Behavioral and Brain Sciences*, **24** (1), pp. 87–185.
Dennett, D.C. (1991) *Consciousness Explained*, Boston, MA: Little, Brown and Company.
Dennett, D.C. (2001a) The fantasy of first-person science, [Online], http://ase.tufts.edu/cogstud/papers/chalmersdeb3dft.htm.
Dennett, D.C. (2001b) Surprise, surprise, *Behavioral and Brain Sciences*, **24** (5), p. 982.
Dennett, D.C. (2002) How could I be wrong? How wrong could I be?, *Journal of Consciousness Studies*, **9** (5–6), pp. 13–16.
Derakshan, N., Eysenck, M.W. & Myers, L.B. (2007) Emotional information processing in repressors: The vigilance–avoidance theory, *Cognition and Emotion*, **21** (8), pp. 1585–1614.
Dror, I.E. & Kosslyn, S.M. (1994) Mental imagery and aging, *Psychology and Aging*, **9** (1), pp. 90–102.
Engelbert, M. & Carruthers, P. (2010) Introspection, *Wiley Interdisciplinary Reviews: Cognitive Science*, **1** (2), pp. 245–253.
Farah, M.J. (2004) *Visual Agnosia*, Cambridge, MA: MIT Press.
Goldman, A. (2004) Epistemology and the evidential status of introspective reports, *Journal of Consciousness Studies*, **11** (7–8), pp. 1–16.
Goldman, A. (2006) *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*, Oxford: Oxford University Press.
Halford, G.S., Wilson, W.H. & Phillips, S. (1998) Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology, *Behavioral and Brain Sciences*, **21** (6), pp. 803–831.
Harman, G. (1990) The intrinsic quality of experience, in Tomberlin, J. (ed.) *Philosophical Perspectives 4: Action Theory and Philosophy of Mind*, Atascadero, CA: Ridgewell Publishing Company.
Hatfield, G.C. (2009) *Perception and Cognition: Essays in the Philosophy of Psychology*, Oxford: Clarendon Press.
Hill, C. (2011) How to study introspection, *Journal of Consciousness Studies*, **18** (1), pp. 21–43.
Hohwy, J. (2011) Phenomenal variability and introspective reliability, *Mind & Language*, **26** (3), pp. 261–286.
Hurlburt, R.T. & Schwitzgebel, E. (2007) *Describing Inner Experience?: Proponent Meets Skeptic*, Cambridge, MA: MIT Press.

Ishai, A. & Sagi, D. (1995) Common mechanisms of visual imagery and perception, *Science*, **268**, pp. 1772–1774.
Just, M.A. & Carpenter, P.A. (1992) A capacity theory of comprehension: Individual differences in working memory, *Psychological Review*, **99** (1), pp. 122–149.
Kind, A. (2003) What's so transparent about transparency?, *Philosophical Studies*, **115**, pp. 225–244.
Korcz, K.A. (2010) The epistemic basing relation, in Zalta, E.N. (ed.) *The Stanford Encyclopedia of Philosophy*, [Online], http://plato.stanford.edu/archives/spr2010/entries/basing-epistemic/.
Kosslyn, S.M. (1980) *Image and Mind*, Cambridge, MA: Harvard University Press.
Kosslyn, S.M., Thompson, W.L. & Alpert, N.M. (1997) Neural systems shared by visual imagery and visual perception: A positron emission tomography study, *Neuroimage*, **6** (4), pp. 320–334.
Kosslyn, S.M., Ganis, G. & Thompson, W.L. (2001) Neural foundations of imagery, *Nature Reviews Neuroscience*, **2** (9), pp. 635–642.
Kouider, S. & Dehaene, S. (2007) Levels of processing during non-conscious perception: A critical review of visual masking, *Philosophical Transactions of the Royal Society B: Biological Sciences*, **362** (1481), pp. 857–875.
Kriegel, U. (2013) A hesitant defense of introspection. *Philosophical Studies*, **165** (3), pp. 1165–1176.
Laing, D.G. & Francis, G.W. (1989) The capacity of humans to identify odors in mixtures, *Physiology & Behavior*, **46** (5), pp. 809–814.
Lambie, J.A. & Marcel, A.J. (2002) Consciousness and the varieties of emotion experience: A theoretical framework, *Psychological Review*, **109** (2), pp. 219–259.
Land, M.F. (1999) Motion and vision: Why animals move their eyes, *Journal of Comparative Physiology A*, **185** (4), pp. 341–352.
Lane, R.D., Lee, S., Reidel, R., Weldon, V., Kaszniak, A. & Schwartz, G.E. (1996) Impaired verbal and nonverbal emotion recognition in alexithymia, *Psychosomatic Medicine*, **58** (3), pp. 203–210.
Larsen, J.K., Brand, N., Bermond, B. & Hijman, R. (2003) Cognitive and emotional characteristics of alexithymia: A review of neurobiological studies, *Journal of Psychosomatic Research*, **54** (6), pp. 533–541.
Legrain, V., Van Damme, S., Eccleston, C., Davis, K.D., Seminowicz, D.A. & Crombez, G. (2009) A neurocognitive model of attention to pain: Behavioral and neuroimaging evidence, *Pain*, **144** (3), pp. 230–232.
Legrain, V., Crombez, G., Verhoeven, K. & Mouraux, A. (2011) The role of working memory in the attentional control of pain, *Pain*, **152** (2), pp. 453–459.
Luck, S.J. & Vogel, E.K. (1997) The capacity of visual working memory for features and conjunctions, *Nature*, **390** (6657), pp. 279–281.
Lycan, W.G. (1996) *Consciousness and Experience*, Cambridge, MA: MIT Press.
Mack, A. & Rock, I. (1998) *Inattentional Blindness*, Cambridge, MA: MIT Press.
Mandler, G. & Shebo, B.J. (1982) Subitizing: An analysis of its component processes, *Journal of Experimental Psychology: General*, **111** (1), pp. 1–22.
Mikels, J.A., Reuter-Lorenz, P.A., Beyer, J.A. & Fredrickson, B.L. (2008) Emotion and working memory: Evidence for domain-specific processes for affective maintenance, *Emotion*, **8** (2), pp. 256–266.

Most, S.B., Scholl, B.J., Clifford, E.R. & Simons, D.J. (2005) What you see is what you set: Sustained inattentional blindness and the capture of awareness, *Psychological Review*, **112** (1), pp. 217–242.

Nichols, S. & Stich, S.P. (2003) *Mindreading: An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*, Oxford: Oxford University Press.

Pearson, J., Rademaker, R.L. & Tong, F. (2011) Evaluating the mind's eye: The metacognition of visual imagery, *Psychological Science*, **22** (12), pp. 1535–1542.

Prinz, J. (2004) The fractionation of introspection, *Journal of Consciousness Studies*, **11** (7–8), pp. 40–57.

Rademaker, R.L. & Pearson, J. (2012) Training visual imagery: Improvements of metacognition, but not imagery strength, *Frontiers in Psychology*, **3**, pp. 1–11.

Ramm, B.J. & Halford, G.S. (2012) Novelty and processing demands in conceptual combination, *Australian Journal of Psychology*, **64** (4), pp. 199–208.

Samur, D., Tops, M., Schlinkert, C., Quirin, M., Cuijpers, P. & Koole, S.L. (2013) Four decades of research on alexithymia: Moving toward clinical applications, *Frontiers in Psychology*, **4**, pp. 1–4.

Saults, J.S. & Cowan, N. (2007) A central capacity limit to the simultaneous storage of visual and auditory arrays in working memory, *Journal of Experimental Psychology: General*, **136** (4), pp. 663–684.

Schmeichel, B.J., Volokhov, R.N. & Demaree, H.A. (2008) Working memory capacity and the self-regulation of emotional expression and experience, *Journal of Personality and Social Psychology*, **95** (6), pp. 1526–1540.

Schwitzgebel, E. (2008) The unreliability of naïve introspection, *Philosophical Review*, **117** (2), pp. 245–273.

Schwitzgebel, E. (2011) *Perplexities of Consciousness*, Cambridge, MA: MIT Press.

Schwitzgebel, E. (2012) Introspection, what?, in Smithies, D. & Stoljar, D. (eds.) *Introspection and Consciousness*, Oxford: Oxford University Press.

Simon, T.J. & Vaishnavi, S. (1996) Subitizing and counting depend on different attentional mechanisms: Evidence from visual enumeration in afterimages, *Perception & Psychophysics*, **58** (6), pp. 915–926.

Simons, D.J. (2000) Attentional capture and inattentional blindness, *Trends in Cognitive Sciences*, **4** (4), pp. 147–155.

Simons, D.J. & Chabris, C.F. (1999) Gorillas in our midst: Sustained inattentional blindness for dynamic events, *Perception*, **28** (9), pp. 1059–1074.

Solomon, G.E.A. (1990) Psychology of novice and expert wine talk, *The American Journal of Psychology*, **103** (4), pp. 495–517.

Sperling, G. (1960) The information available in brief visual presentations, *Psychological Monographs: General and Applied*, **74** (11), pp. 1–29.

Strayer, D.L. & Johnston, W.A. (2001) Driven to distraction: Dual-task studies of simulated driving and conversing on a cellular telephone, *Psychological Science*, **12** (6), pp. 462–466.

Trick, L.M. & Pylyshyn, Z.W. (1994) Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision, *Psychological Review*, **101** (1), pp. 80–102.

Tye, M. (1995) *Ten Problems of Consciousness*, Cambridge, MA: MIT Press.

Tye, M. (2000) *Consciousness, Color, and Content*, Cambridge, MA: MIT Press.

Van Dillen, L.F. & Koole, S.L. (2007) Clearing the mind: A working memory model of distraction from negative mood, *Emotion*, **7** (4), pp. 715–723.
Watzl, S. & Wu, W. (2012) Perplexities of Consciousness, by Eric Schwitzgebel, *Mind*, **121** (482), pp. 524–529.
Wickens, C.D. (1984) Processing resources in attention, in Parasuraman, R. & Davies, D.R. (eds.) *Varieties of Attention*, New York: Academic Press.
Wickens, C.D. (2002) Multiple resources and performance prediction, *Theoretical Issues in Ergonomics Science*, **3** (2), pp. 159–177.

Paper received April 2015; revised September 2015.