Psychophysical Methods and the Evasion of Introspection

M. Chirimuuta*†

In this paper I argue that certain methodological conventions within psychophysics reflect a continued uncertainty over the reliability of subjects' introspective judgements. Firstly, different psychophysical methods do not rely equally on the introspective capabilities of experimental subjects. I contrast "minimally introspective" tasks with "introspection-reliant" ones. It is only in the latter that introspection can be said to have a nontrivial role in the subjects' performance. Secondly, my distinction maps onto a number of important methodological divisions in vision science. The introspection-reliant categorization captures many of the tasks typically considered less able to yield useful information regarding the processes underlying visual sensation.

1. Introduction. Recent work on introspection in psychology has been careful to separate the specific commitments of Titchener's analytical school from the discussion of introspection in psychology more generally. In a paper that makes a significant contribution to this debate, Hatfield (2005, 260) writes that "introspection maintains an ineliminable role in psychology itself, as a source of evidence. This is especially apparent in perceptual psychology." This claim prompts two inquiries that I address in this article. Firstly, it is important to ask how introspection should be defined in the context of perceptual psychology. Secondly, it is worth considering whether Hatfield's assessment of the role of introspection is correct. I challenge the appropriateness of Hatfield's definition and offer an alternative characterization. This

*To contact the author, please write to: Department of History and Philosophy of Science, 1017 Cathedral of Learning, 4200 Fifth Avenue, University of Pittsburgh, Pittsburgh, PA 15260; e-mail: mac289@pitt.edu.

† I would like to thank Uljana Feest, Gary Hatfield, and Fred Kingdom for many helpful comments on an earlier version of this paper. Thanks also to the other panelists at the PSA 2012 Symposium, Anthony Jack and Eric Schwitzgebel, and to the audience members for questions and discussion.

Philosophy of Science, 81 (December 2014) pp. 914–926. 0031-8248/2014/8105-0017$10.00

leads to a different view of the role of introspection in the branch of perceptual psychology known as psychophysics.[1]

Hatfield (2005, 260) broadly characterizes introspection as "a mental state or activity in or through which persons are aware of properties or aspects of their own conscious experience." He later defines introspection as "deliberate and immediate attention to certain aspects of phenomenal experience" (279). This then raises the question of what it means to be aware of one's conscious perceptual experience. Is it just the same thing as looking, hearing, and so on? In certain kinds of psychophysical experiments, such as a shape constancy study, it is easy to formulate a response (Hatfield 2009, 2014). In such experiments the genuinely perceptual experience of the distorted shape of an object viewed obliquely can be separated from one's belief that the shape is just a regular circle or rectangle. Thus, such experiments can elicit two classes of responses—sensory and cognitive ones—and introspection can be characterized straightforwardly as attending to sensory experience over cognitive judgement.

Hatfield (2005, 278) also discusses the psychophysical task of metameric color matching as one example of a perceptual experiment reliant on introspection. In this experiment subjects are required to adjust the intensities of red and green lights so that the resulting composite color matches a yellow target (see fig. 1). In this case there is an important exegetical question over how to understand Hatfield's claim that introspection is an awareness or attention to aspects of conscious experience.

On one interpretation, Hatfield's definition may be quite a liberal or unrestrictive one. Namely, it is the claim that contemporary perceptual psychology relies on introspective evidence because it assumes that experimental subjects have conscious perceptual experiences and that they are capable of giving verbal or motor responses that reliably indicate the presence or absence of particular features of those conscious experiences. For example, a psychophysical experiment that measures the absolute detection threshold for a dim spot of light is said to be reliant on the subject's capacity to introspect in the sense that her subjective awareness of the spot is a crucial data point that the experimenter has access to because of the subject's capacity to introspect. And thus it must be assumed that the subject can faithfully indicate those times in which the spot enters her conscious field of view.

Yet a problem with this account is that it is not clear how it can be employed to distinguish introspection from ordinary perception, for doesn't the

1. Psychophysics is defined by Gescheider (1997) as "the scientific study of the relation between stimulus and sensation." The disciplinary demarcation between psychophysics and perceptual psychology more generally has become somewhat blurry in recent years, with many experiments that are classified as psychophysical dealing with complex perceptual states, not just simple sensation.

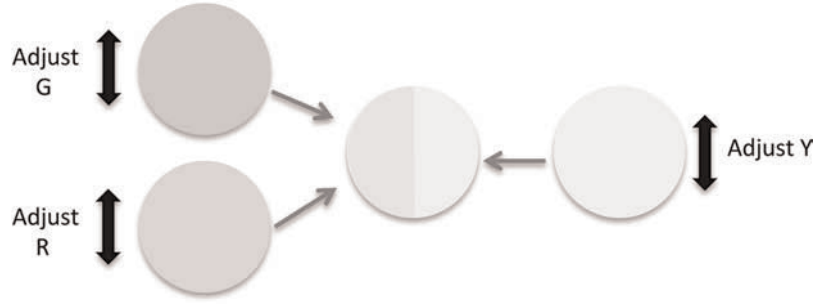

Figure 1. Metameric match experiment. The subject is asked to adjust the intensities of R, G, and Y monochromatic lights so that the yellows are indistinguishable. Color version available as an online enhancement.

subject's activity in the detection task just boil down to her looking for a dim spot of light? This worry could prompt us to take Hatfield as endorsing a more restrictive definition. For Hatfield (2005, 279) also suggests that what characterizes introspection over ordinary perception is that one attends to one's experience of an object, not just the object itself: "immediate attention to . . . phenomenal experience." This makes introspection importantly different from perception because, as many would have it, perception is "transparent"; our perceptual encounter with the world is not interrupted with moments of attention to experience itself.[2] The difficulty with this reading is that it then becomes unclear how the more restrictive definition of introspection could apply to many of the psychophysical tasks that Hatfield wants it to apply to, such as stimulus detection and the metameric matching experiment. Subjects perform such tasks by directing their attention to external stimuli, namely, the colored lights, and need not attend to their own phenomenal experience, qua experience. Nor do they need to consider their experience in a more fine-grained or detailed way than in ordinary perception.

In short, the problem is that while Hatfield's restrictive definition has the virtue of allowing one to demarcate introspection from perception, it cannot reasonably be applied to the range of psychophysical tasks that Hatfield claims it does. Yet the more liberal definition makes all perceptual activity concurrently introspective in a somewhat trivial sense. It strikes me that a different approach to defining introspection—in the context of psychophysics—is needed, one that does not characterize introspection in terms of an object of attention or focus of awareness. In this paper I propose that the

2. Here I invoke a weak notion of transparency. I do not commit myself to the strong representationalist thesis of Tye, Byrne, and others, i.e., that there is no subjective aspect to perceptual experience.

tasks that can be said to involve introspection are the ones that rely on experimental subjects' capacity to analyze and compare sensory experiences that bear nonobvious relationships of similarity and difference to each other. Thus, on my account introspection can be part of the process of perceiving and attending to an external object and need not be overtly directed at phenomenal experience. The subject may interpret her task to be simply that of attending to the external stimulus, but she can be reporting on aspects of her phenomenal experience nonetheless. It is also a feature of my view that the extent to which tasks rely on introspection is a matter of degree. In the next section I give a set of examples of common psychophysical tasks that are either "introspection reliant" or "minimally introspective." In the third section I describe how the cluster of introspection-reliant tasks—though not described in this way by scientists themselves—has commonly attracted suspicion from psychophysicists as being less likely to produce data that are "objective" and informative about neural mechanisms. I ask whether this is a mere coincidence, or whether the methodological norms of psychophysics reflect a certain wariness toward introspection.

2. Introspection in Psychophysics as Controlled Comparison

2.1. Examples of Introspectively Demanding and Undemanding Psychophysical Tasks. The *metameric match paradigm*, illustrated in figure 1, has been used to diagnose specific types of color vision deficiency since the late nineteenth century. Differences in the number of retinal cone types an individual has, and the spectral sensitivities of those cone classes, lead to measurable differences in the proportion of red to green in a composite light that he or she judges to look identical to a yellow monochromatic standard. Note that in this task the only perceptual judgment that the subject need make is over whether the composite light and monochromatic light are visually indistinguishable. If the lights are presented as abutting (as in fig. 1), then the subject simply has to judge whether or not the color field is homogeneous. No attention to the specific qualities of the perceived color is required.

Contrast this task with an *asymmetric match paradigm* (fig. 2). In this case the two central stimuli are not matched for luminosity, but the subject must say whether or not they match in hue regardless of their visible difference in brightness. This requires that the subject analyze her experience of the two colors in terms of separate dimensions of hue and brightness and make a judgment as to the identity of just one of these dimensions, disregarding the difference in the other. Thus, the subject must make a series of comparisons between pairs of stimuli in order to find the pair that holds the unique but nonobvious relationship of sameness of hue. This relationship is nonobvious in that it is not marked by a simple defining characteristic like a homogenous spatial profile.

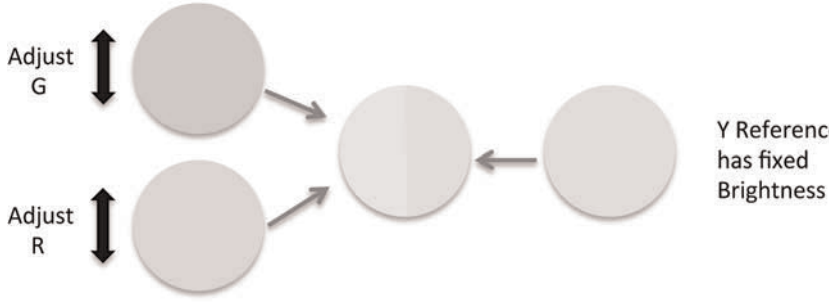

Figure 2. Asymmetric match experiment. The subject is asked to adjust the proportions of R and G monochromatic lights so that the yellows match in hue. The intensity of the Y light is fixed, so the yellows cannot be matched for brightness and are therefore distinguishable even when hue is judged to be equivalent. Color version available as an online enhancement.

It should be fairly intuitive that this task is "introspection reliant" in a way that the metameric matching task is not. The contrast between these two tasks points us to the central intuition behind my new characterization of introspection. The idea is that the metameric matching task is "minimally introspective" because it can be performed without any careful comparison of the phenomenal qualities one experiences on presentation of the two stimuli. The metameric paradigm relies on introspection only in the minimal sense that it assumes that the subject can know and reliably report when her conscious visual field is homogeneous with respect to color.[3] The asymmetric matching task, on the other hand, is "introspection reliant" because it does require this careful comparison of sensory experiences that bear nonobvious relationships of similarity and difference to each other.

Asymmetric matching paradigms have been used to study achromatic perception of lightness and darkness (fig. 3*a*; see Gilchrist 2004) and to study color constancy (Foster 2011). Figure 3*b* gives an example of an asymmetric task in which the observer views a scene under two different lighting conditions. She is instructed to adjust the color of the central patch in one image until it looks as if made from the same paper as the central patch in the other. Importantly, even when the patches are matched, there will still be a visible difference in color between them, and the experiment relies on the subject having a clear sense of what sameness of material would look like in spite of these differences. Again, the task is introspection reliant in comparison to a task in which the subject just has to report on the absolute identity or dis-

3. That is, it relies on introspection defined in the first, permissive sense. To reiterate the discussion of sec. 1, the problem with the minimal notion of introspection is that it cannot distinguish introspection from ordinary perception.

https://doi.org/10.1086/677890 Published online by Cambridge University Press

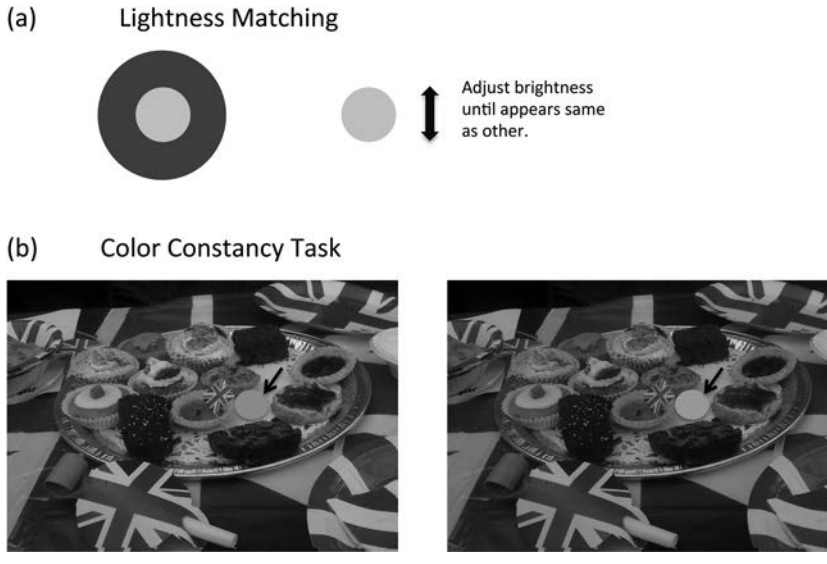

Figure 3. (*a*) Achromatic asymmetric match experiment where black annulus influences perceived brightness of one of the circles. The subject is asked to determine the point of subjective equality of the brightness of the two circles. (*b*) Asymmetric color constancy experiment. The subject is asked to adjust the color of one of the patches (marked with an arrow) until it looks as if it is made from the same paper as the other. Color version available as an online enhancement.

tinguishability of two stimuli. In particular, it relies on the subject's ability to make a "judgment call" on the one best match, given a range of close contenders that vary along a number of different dimensions. I describe the introspection-reliant tasks as requiring controlled comparison because the demand placed on the subject is to perform some kind of analysis and comparison, but within parameters that are prespecified by the experimenter.

Another kind of paradigm that intuitively fits the idea of controlled comparison is a rating scale task. In a series of experiments published recently (To et al. 2008, 2010; Tolhurst et al. 2010; see fig. 4*a*) subjects were presented with nearly 300 pairs of photographs—an original and a modified version—and were asked to rate how similar the pairs were on a scale from 0 (completely identical) to any arbitrarily high value. In one of these publications (Tolhurst et al. 2010), we present results of a simple two-alternative-forced-choice (2-AFC) contrast discrimination experiment in which subjects just had to report which of a pair of otherwise identical photographs contained a small, high-contrast central patch (see fig. 4*b*). We then apply a model of contrast discrimination to the rating scale data. The rating scale task falls under my introspection-reliant category, while the contrast discrimina-

https://doi.org/10.1086/677890 Published online by Cambridge University Press

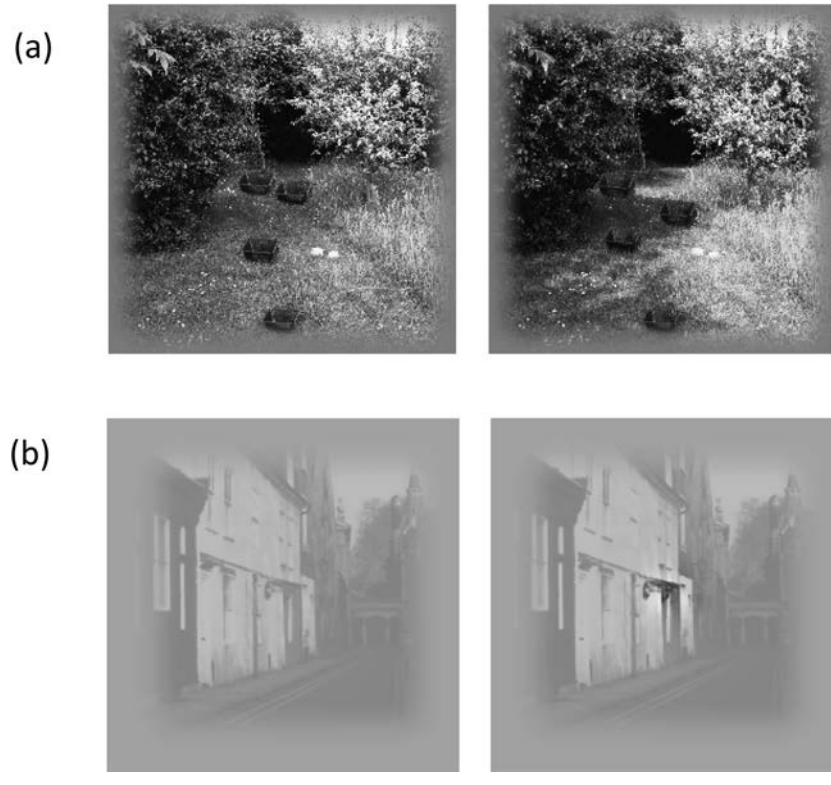

Figure 4. Examples of stimuli of the sort used by Tolhurst et al. (2010). (*a*) Rating scale task. For each of 294 image pairs, subjects were asked to rate how similar or different they appeared on a numerical scale of their own devising. (*b*) 2-AFC contrast discrimination. Subjects had to report whether the high-contrast central patch appeared in the first or second stimulus.

tion task is minimally introspective. In the former, the subject must make a judgement as to the relative similarity of a large number of pairs of stimuli, which differ in different ways, whereas in the latter task she detects the presence or absence of a high-contrast patch in a rather automatic fashion. Figure 4 illustrates how similar stimuli can be used in these two very different experiments, so it is not complexity of stimulus per se that determines how introspectively demanding the task is. Rather, the determining factor is the nature of the response that the subject must make to the stimulus, that is, whether the response is simply a choice between saying the high-contrast patch appeared first or second out of two stimuli, or whether it calls for a more careful examination of the perceived properties of the stimuli.

Before moving on, I would like to emphasize that my two categories are intended to reflect a qualitative difference in how introspectively demanding these tasks are, and that I say nothing in this paper about how to quantify this difference and how it is that introspective demands admit of degree. For example, the question of whether or not metameric matching is even less introspectively demanding than contrast discrimination will be left unanswered. It seems plausible that introspective demands, like attentional demands, come in degrees, but I offer no suggestions of how one might measure this. It is also plausible that there will be some tasks that occupy middle ground between my categories and are hard to classify either way. I do not deal with such cases here. My aim in presenting a set of tasks that are intuitively more reliant on introspection than the others has been to highlight one way that introspection may be said to play a role in perceptual psychology, and to this end I have focused on the most clear-cut cases.

2.2. Other Classifications of Psychophysical Tasks. One of the attractive things about psychophysics as a subject for philosophy of science is the fact that throughout its short history methodological questions about the best way to measure sensory responses have been debated in a perspicacious way by leading protagonists.[4] Moreover, such controversies still resonate in the living memory of the discipline and are recounted even in the most recent textbooks. One way in which methodological debates commonly unfold is with a distinction first being drawn between two broad classes of psychophysical techniques, and the relative merits of the two classes are subsequently discussed.

In their textbook Kingdom and Prins (2009) devote a chapter to the "dichotomies" that have been most significant to psychophysicists past and present. The first of these, Brindley's (1970) distinction between Class A and Class B observations, is particularly relevant to my account of introspection. Brindley characterized Class A observations as any tasks in which the observer just had to report on the absolute similarity or dissimilarity in the appearance of a pair of stimuli. For example, the measurement of the detection threshold for a spot of light is Class A because the subject need only indicate whether the trial in which the spot is present is distinguishable or not from the reference stimulus in which the spot is absent. Likewise, the measurement of the discrimination threshold for the brightness of the spot is Class A, as it just requires the subject to report whether the trial in which the luminosity of the spot is increased looks different from the trial in which the luminosity remained at baseline. In contrast, Brindley (1970, 133) categorized as Class B "any observation that cannot be expressed as the identity

4. Feest (2014) presents a wonderful example of this in the debate between Gestaltists and atomists over how to measure sensory experience.

or non-identity of two sensations," for example, "all those [observations] in which the subject must describe the quality or intensity of his sensations, or abstract from two different sensations some aspect in which they are alike."

Brindley's description of Class B observations is interchangeable with my characterization of introspection-reliant tasks. Indeed, the tasks that I presented as examples of my minimally introspective category—metameric matching and contrast discrimination—are Class A, whereas all kinds of asymmetric matching and rating scale tasks are Class B. In essence, both of these categorization schemes can be understood as drawing a distinction between tasks in which the experimental subject is treated somewhat like a thoughtless measuring instrument and methods that rely on the subject's status as a critical being who can attend to and reflect on her own conscious states. The point is not that the A/minimally introspective class treats the subject as if unconscious, or that it requires the subject to have sensory capacities but not cognitive ones. Rather, it is that the A/minimally introspective class makes no demands on any capacity for reflection on and comparison of occurrent sensory states, whereas tasks in the B/introspection-reliant class do.[5]

It is hard to say how influential Brindley's distinction has been. It came under immediate criticism from Boynton and Onley (1962) but was clearly accepted in some form by Marks (1978) and Teller (1984) and is discussed at length in Gescheider's (1997) psychophysics textbook. Kingdom and Prins (2009, 18) choose not to employ it as an overarching basis for classifying psychophysical experiments because of the problem that certain tasks (e.g., the discrimination of the mean angle of orientation for fields of randomly placed bars) cannot be classified as either A or B.

Kingdom and Prins's (2009, 22) preferred distinction is between tasks that measure performance and those that measure appearance, which they characterize in the following way: "If the measurement can be meaningfully considered to be better under one condition than under another, then it is a performance measure, if not it is an appearance measure." Performance tasks are any ones designed to chart perceptual "limits" (e.g., contrast discrimination, detection of a spot of light against a differently colored background). An example of an appearance task is an experiment comparing the strength of the Müller-Lyer illusion with fin angles of 45° and 60°. Even if the length

5. In support of this idea that the key distinction in play here is between subject-as-measuring-instrument and subject-as-reflective-being, it is worth noting that Brindley's one example of a psychophysical document explicitly hostile to Class B observations is the 1943 Optical Society of America (OSA) report, which, as Stevens (1951, 31) relates, "reduces psychophysics to the employment of a human observer as a null instrument under a set of strictly specified conditions." And Brindley's one example of a psychophysicist liberal with regard to Class B is Stevens (1951), who explicitly rejects the OSA definition as too narrow and restrictive (cf. Helson 1949).

of the central bars appears to be more different when the fins are 45°, there is no sense in which the subject is "better" at the task in that condition. Hence, this Class B observation can also be said to be an appearance measure. Thus, there is an overlap with my distinction: appearance tasks tend to be introspection reliant, and performance tasks tend to be minimally introspective. But it is not as well matched as is the case with Class A versus B. In particular, the metameric match task that I classify as minimally introspective turns out to be an appearance measure.[6]

3. Not All Psychophysical Methods Were Created Equal. All I have argued so far is that there is an intuitive way of differentiating psychophysical tasks that are more reliant on introspection from those that are not, and that my categorizations turn out to be roughly coextensional with categorizations of tasks developed within the psychophysical tradition. The question now is what to make of this finding. Is it just a coincidence that the distinctions coincide? It should come as no surprise to the reader that my next point will be that the categories that line up on the introspection-reliant side have tended to meet with more diffidence and suspicion from psychophysicists than those on the other side.

Brindley (1970) presents Class A observations as especially informative about the physiological mechanisms underlying perception because they can be used to test "psycho-physical linking hypotheses" that two stimuli (e.g., yellow monochromatic light and a certain mixture of red and green lights) will produce the same neural activity and hence the same sensation.

On the relative status of the two classes, he writes, "The use of Class A observations as a basis for analysing the function of the eye and visual pathway is not controversial; every writer on vision admits, at least by implication, that they can be legitimately used. On the use of the kinds of observation here called Class B, there have been differences of opinion. . . . The conservative opinion, in its most extreme form, is that only Class A observations are of any value, and in a discussion of visual mechanisms all Class B observations may be entirely disregarded" (1970, 134). Brindley himself takes this view to be too narrow, but he is critical of Stevens's (1951) "extreme liberal opinion" for failing to make the distinction. Later in the book, when discussing Hering's opponent theory of color, he writes as if it is still moot whether the kinds of phenomenological reports presented by Hering in

6. A related dichotomy is Sperling, Dosher, and Landy's (1990) Type 1 vs. Type 2 distinction. In Type 1 experiments the subject's response may be either correct or incorrect with respect to some physical dimension of the stimulus. For Type 2 the experimenter cannot classify responses as correct or incorrect. Note again that the metameric match turns out to be Type 2, even though it is Class A/minimally introspective.

support of his theory can actually be taken as evidence for a kind of color mechanism (1970, 208).

One might think that this is all beside the point in a discussion about introspection because the reason that the value of Class B observations was held in question was not because they are introspection reliant but because of their failure to underwrite psychophysical bridge principles. But I do not think that this problem is so disconnected from the issue of introspection. It is worth noting that in order to use Class A observations as evidence for psychophysical bridge principles, Brindley (1970, 133) had to supply supporting assumptions such as the following: "whenever two stimuli cause physically indistinguishable signals to be sent from the sense organs to the brain, the sensations produced by these stimuli, as reported by the subject in words, symbols or actions, must also be indistinguishable." Now if Class B tasks were to be granted some equivalent supporting assumptions, then one could equally say that they are informative of underlying neural mechanisms. For example, in the case of the asymmetric hue matching experiment, why not assume that whenever two hue components of a stimulus cause physically indistinguishable signals to be sent to the brain, the hue sensations must also be indistinguishable? This assumption would support inferences from equality of hue sensation to the sameness of a neural signal somewhere in the brain. It would be a special case of the assumption made in support of inferences from Class A observations, quoted above.

However, Class B observations are treated differently. The reason for this difference is likely that Brindley and other theorists (e.g., Marks 1978) have been wary of attributing to subjects the kind of introspective powers that would be needed to analyze hue separately from all other sensory qualities and determine exactly the point of equivalence of hue. In other words, if these theorists had shared Titchener's faith in the analytical acumen of introspection, they would have had no reason to treat Class B observations differently from Class A.

This pattern of unequal treatment can be seen not just in the discussion of Class A and B observations but also with respect to the other dichotomies discussed by Kingdom and Prins. They note that it is fairly common for psychophysicists to refer to some tasks as more "objective" or "subjective" than others, with all the value-laden connotations of these terms. Kingdom and Prins (2009, 18–19) explain this usage in the following way:

> All psychophysical experiments are in a trivial sense subjective, because they measure what is going on inside the head, and if this is the intended meaning of the term then the distinction is redundant.[7] The dichotomy is

7. Compare the worry discussed above that all psychophysical experiments rely on introspection in a trivial or "minimal" way, and hence the distinction between introspection and perception is made redundant.

more often invoked, however, to differentiate between different types of psychophysical procedure. The distinction has been used variously to characterize Class A versus Class B observations, tasks for which there is versus tasks for which there is not a correct and an incorrect response,[8] forced-choice versus non-forced-choice procedures, and criterion-dependent versus criterion-free procedures.

The notion of subjectivity at play here lines up well with the idea that experiments are subjective if they are introspection reliant. For all the tasks on the wrong side of the subjective-objective tracks are ones that rely on the subject's judgments concerning the appearance of the stimuli, involving complex comparisons that cannot be independently verified by examining the physical properties of the stimuli themselves.[9]

There is a sense in which the title of this paper is misleading. I have not shown that psychophysicists have avoided using experimental methods more reliant on introspection, or that the use of such methods has always been questioned. Indeed, when Kingdom and Prins (2009, 26) write that "both performance-based and appearance-based experiments are important to our understanding of vision. Measures from both types of experiments are probably necessary to fully characterize the system," they are articulating a methodological pluralism that many psychophysicists would endorse. However, the crucial point is that the methods on the wrong side of the divide, those more reliant on introspection, continue to need their advocates, whereas those on the other side have been accepted without question. This is an indication of the contested status of introspection within the psychophysics tradition.

4. Conclusions. I would like to conclude by pointing out a few advantages of my characterization of introspection. Firstly, and in agreement with Hatfield (2005), I am not saying that introspection is an inner gaze on a Cartesian theater, or that it needs its own special inner object (e.g., sense data). Furthermore, like Byrne (2005), I am able to talk about introspection without denying transparency of perceptual experience (see n. 2). My account corresponds to ways that scientists themselves classify experiments, and it also opens up the possibility that introspective activity is on some continuum with ordinary perception. One set of questions that I have not addressed here concerns the nature of the relationship between introspection and other capacities such as perception, attention, and perceptual cognition. Indeed, I have

8. In other words, performance vs. appearance or Sperling's Type 1 vs. Type 2.

9. As Feest (2014) observes, one thing that characterizes phenomenological, first-person methods is the authority of the subject. Absent external criteria for correct or incorrect responses, the experimental subject is placed in a position of authority, which some experimenters find problematic.

https://doi.org/10.1086/677890 Published online by Cambridge University Press

said nothing about the possible mechanisms of introspection. I suspect that such questions will not be answered easily and will themselves rely on a better understanding of the mechanisms behind conscious perceptual experience (which is still far beyond our current scientific grasp). In agreement with Feest (2012), I believe that an understanding of introspection can only go so far without a fuller examination of the theory of mind and awareness implicit in scientific practice. But I hope to have shown that the study of psychophysical experiments is a useful starting point in making sense of the connection between introspection and perception.

REFERENCES

Boynton, R. M., and J. W. Onley. 1962. "A Critique of the Special Status Assigned by Brindley to 'Psychophysical Linking Hypotheses' of 'Class A.'" *Vision Research* 2:383–90.

Brindley, G. S. 1970. *Physiology of the Retina and the Visual Pathway*. 2nd ed. London: Arnold.

Byrne, Alex. 2005. "Introspection." *Philosophical Topics* 33 (1): 79–104.

Feest, Uljana. 2012. "Introspection as a Method and Introspection as a Feature of Consciousness." *Inquiry* 55 (1): 1–16.

———. 2014. "Phenomenal Experiences, First-Person Methods, and the Artificiality of Experimental Data." *Philosophy of Science*, in this issue.

Foster, David H. 2011. "Color Constancy." *Vision Research* 51:674–700.

Gescheider, George A. 1997. *Psychophysics: The Fundamentals*. Mahwah, NJ: Erlbaum.

Gilchrist, Alan. 2004. *Seeing Black and White*. Oxford: Oxford University Press.

Hatfield, Gary. 2005. "Introspective Evidence in Psychology." In *Scientific Evidence: Philosophical Theories and Applications*, ed. P. Achinstein. Baltimore: Johns Hopkins University Press. Reprinted in Hatfield 2009.

———. 2009. "On Perceptual Constancy." In *Perception and Cognition*. Oxford: Oxford University Press.

———. 2014. "Psychological Experiments and Phenomenal Experience in Size and Shape Constancy." *Philosophy of Science*, in this issue.

Helson, H. 1949. "Review of 'Introduction to Color.'" *Psychological Bulletin* 46 (2): 166–69.

Kingdom, Frederick A. A., and Nicolaas Prins. 2009. *Psychophysics: A Practical Introduction*. Amsterdam: Elsevier.

Marks, L. E. 1978. *The Unity of the Senses*. New York: Academic.

Sperling, G., B. A. Dosher, and M. S. Landy. 1990. "How to Study the Kinetic Depth Experimentally." *Journal of Experimental Psychology: Human Perception and Performance* 16:445–50.

Stevens, S. S. 1951. *Handbook of Experimental Psychology*. London: Chapman & Hall.

Teller, Davida Y. 1984. "Linking Propositions." *Vision Research* 24 (10): 1233–46.

To, Michelle P. S., P. G. Lovell, Tom Troscianko, and David J. Tolhurst. 2008. "Summation of Perceptual Cues in Natural Visual Scenes." *Proceedings of the Royal Society* B 275:2299–2308.

———. 2010. "Perception of Suprathreshold Naturalistic Changes in Colored Natural Images." *Journal of Vision* 10:1–22.

Tolhurst, David J., Michelle P. S. To, M. Chirimuuta, Tom Troscianko, P.-Y. Chua, and P. G. Lovell. 2010. "Magnitude of Perceived Change in Natural Images May Be Linearly Proportional to Differences in Neuronal Firing Rates." *Seeing and Perceiving* 23:349–72.