The dominance of the visual[1]

Dustin Stokes (University of Utah)
Stephen Biggs (Iowa State University)

1. Introduction

Perception-studies has long been vision-centric in at least two ways. First, vision has been studied far more often than other sense modalities. Second, ignoring other senses has been excused partly on the ground that one can learn everything that can be learned about perception-in-general by studying vision. Thus, vision is taken to be the paradigmatic perceptual sense, and its structure is then imputed to the other sense modalities in a way that is insensitive to differences among modalities.

Recent theorists reject this vision-centrism. O'Callaghan (2011), for example, argues that studying audition reveals, among other things, both a diversity of perceptual objects and a banality of multimodal experiences that "an exclusively visual perspective" wouldn't admit. He also suggests that these diverse perceptual objects are unified in an interesting way, which implies that studying only vision leads to a false account of the property *being a perceptual object*. In a similar spirit Batty suggests that:

> It is commonplace to hold that visual experience represents objects. After all, visual experience displays a rich form of perceptual organization that allows us to think and speak of individual apples and oranges, particular cats and dogs. Olfactory experience lacks such organization. If we take this organization as necessary for the representation of objects, we are led to the conclusion that olfactory experience does not represent objects. But we need not, and ought not, accept this conclusion. Olfactory experience does indeed represent objects—just not in a way that easily read off the dominant visual case (Batty 2011: 162).

Others suggest that vision-centrism biases theorizing toward direct realism with regard to material objects. Lycan (2000), for example, suggests that direct realism would seem far less plausible if theorizing focused first on smell, since smells strike us, phenomenologically, as providing only mediated access to the ordinary material objects that emit them; McLaughlin (1989) claims the same for audition.

We sympathize with these concerns, broadly construed: theorizing about perception *should* be based on extensive studies of each sense, not exclusively, or even primarily, on studies of vision. We wonder, nonetheless, why such theorizing has been vision-centric. Is vision special? Even those who appeal for diversity often answer "yes".

---

[1] This work was thoroughly collaborative and the paper thoroughly co-authored--the order of authors was chosen randomly.

O'Callaghan calls the focus on vision "understandable" since we "undeniably are visual creatures" (p. 143). Batty agrees that "[w]e are visual creatures", suggesting that a focus on vision "should come as no real surprise" since it "reflects human preferences in general" (p. 162). But what does it mean to say that we are visual creatures, or that we prefer vision? What makes vision special?

One can approach this question in many ways. One might imagine losing senses, or investigate actual losses, considering which loss would-be, or is, most destructive—although people may say they care most about vision, losing proprioception would surely be more destructive. One might imagine building a creature from the ground up, asking which sense would be needed first in order for the others to develop qua senses—Martin (1992) uses this strategy to argue that touch is especially important. One might consider whether vision has intrinsic properties that no other sense has, and whether any such properties are significant—Strawson (1959) uses this strategy to argue that vision is special because it alone is intrinsically spatial.[2]

We pursue a different, complementary path, starting from an extrinsic property of *the visual*: its *dominance* of other senses.[3] Quite roughly, the visual dominates another sense *S***,** say audition**,** with respect to property *P* if the visual asymmetrically affects how auditory stimuli that are relevant to identifying *P* are processed, where the effect is asymmetric in that the auditory has no comparable effect on how visual stimuli that are relevant to identifying *P* are processed. We find that the visual dominates with respect to a wide range of properties in psychologically and epistemically significant ways, such that the dominance of the visual partly explains why we can rightly say that vision is special.

We first identify three levels at which a sense can be dominant, the levels of *experience*, *experience-based judgment*, and *all-things-considered judgment* (§2). Then, taking touch as our test case, we argue that vision exercises two kinds of dominance, *perception-perception* dominance, in which visual perception affects how we interpret

[2] More specifically, Strawson claims that, because only vision is intrinsically spatial, only vision allows one to think of particular objects as distinct from oneself (making vision psychologically significant), and thus, only vision allows one to know about the world apart from oneself (making vision epistemically significant). That said, Strawson suggests that touch coupled with kinesthesia might give rise to spatial concepts, and ultimately, knowledge of objective particulars. See Strawson (1959), especially pp. 64-66 and subsequent discussion.

[3] For reasons that will become clear below, we use 'the visual' as a general term to denote both visual perceptual experience and visual imagery.

non-visual stimuli (§3), and *imagery-perception* dominance, in which visual imagery affects how we interpret non-visual stimuli (§4). We then consider why vision exercises these kinds of dominance over touch (§5), and how this makes vision both psychologically and epistemically special (§6). This allows speculation about the conditions in which vision dominates touch (§7). We close with a rough generalization to the relation between vision and other senses (§8).

To be clear, we don't aim to legitimize vision-centrism in perception-studies. We aim, instead, to identify (some) relations among the senses that partly explain extant vision-centrism. In fact, our line of thought has the implication that modalities should be analysed differently, given important phenomenological and informational aspects of vision by contrast with non-visual modalities. By this same token, a methodological prescription of our analysis is that we should avoid an overcorrection in which perception-studies would treat vision as just another sense.

2. Dominance at Different Levels

The McGurk effect (McGurk and MacDonald 1976) is perhaps the most familiar example of vision influencing another sense. In one version of the effect (Rosenblum et al. 1997), subjects see a video of a person saying "va", and hear an audio-track of a person saying "ba". Saying "ba" requires compressing the lips together but saying "va" doesn't. So, the video and audio-track are in conflict.[4] Subjects report *hearing* "va".

Exploring this effect unearths a few levels of dominance. Subjects encountering McGurk stimuli report hearing "va" *even after* learning that the audio-track plays "ba", which implies that the phenomenal character of their auditory experience is as of "va", not as of "ba".[5] In the McGurk effect, then, subjects receive a visual stimulus appropriate to one phoneme, an auditory stimulus appropriate to another phoneme altogether, and have an *auditory* experience appropriate to the *visual* stimulus, not the auditory

[4] How does the video conflict with the audio? The stimulus received by the eyes suggests that the face has the property *looks-like-produces-"va"*, the stimulus received by the ears suggests that the face has the property *sounds-like-produces-"ba"*, the probability that the face has both properties simultaneously nears 0, but the product of the probability that the face has each separately doesn't near 0. The former nears 0 because, Michael Winslow aside, people rarely mouth one phoneme and produce another. Of course, mismatch is more likely between video and audio tracks, but a mismatch in which the visual and auditory stimuli are so temporally similar still nears 0.

[5] Here and throughout we presume that subjects' reports about the phenomenal character of their experiences are typically reliable such that if subjects report that their auditory experience is as of "va", then typically it is. If this presumption is false, our claims about experience can be rephrased as claims about *judgments about experience*. Given this rephrasing, inferring that vision dominates audition at the level of experience from subjects' reports that they hear, say, "va" only requires presuming that subjects are reporting their experiences *honestly*, not that they are reporting them *accurately*.

stimulus—presumably because the stimulus received by the eyes changes the phenomenal character of the auditory experience in a way that the stimulus received by the ears doesn't change the phenomenal character of the visual experience. So, presuming that audition has no comparable effect on vision, vision dominates audition with respect to phoneme *at the level of the phenomenal character of experience*.[6] (Henceforth, unless context suggests otherwise, 'experience' is elliptical for 'the phenomenal character of experience', and, if it is clear which property is at issue, 'with respect to ...' is presumed.)

Dominance at the level of experience implies a further kind of dominance. Perceptual experience often guides judgment: Ethan has a visual experience as of a pine tree, and judges that there is a pine tree. Sometimes, however, one judges against perceptual experience: Lindsay has a visual experience as of a ghost, but judges that there isn't a ghost, perhaps because of her background beliefs that she's exhausted and ghosts don't exist. Even when one judges against perceptual experience, we can talk about the judgment *one would make if one were judging on the basis of perceptual experience alone*, as Lindsay would judge that there *is* a ghost if not for her background beliefs. These are *experience-based judgments*.

Experience-based judgments come in two kinds. First, for any subject *S* and sense modality *M*, *narrow* experience-based judgments are those *S* would make on the basis of experience in *M* alone, independently of evidence from background beliefs, other senses, etc. Second, for any subject *S* and sense modalities M_1 ... M_n, *broad* experience-based judgments are those *S* would make on the basis of experiences in M_1 ... M_n alone, independently of evidence from background beliefs, other senses, etc. To illustrate the difference, consider an unusual subject, Cathy, who has a visual experience as of a speaker mouthing "va" and an auditory experience as of the speaker uttering "ba" in response to McGurk stimuli. Cathy's narrow experience-based judgment for vision is "va", and for audition is "ba". Her broad experience-based judgment for the pair <vision, audition> could be "va", "ba", or neither, if she withholds judgment, depending on various factors (e.g., whether she's more inclined to trust vision, audition, or neither). We

[6] Does audition ever have a comparable effect on vision? One might think that it does, noting that subjects report hearing "da" when the auditory stimulus is "da" and the visual stimulus is "va", and often report hearing the fusion "da" when the auditory stimulus is "ba" and the visual stimulus is "ga". These results, however, don't provide evidence that audition has a comparable effect on vision because they don't show that subjects have *visual* experiences that are appropriate to the auditory stimuli. At any rate, we use the McGurk effect primarily to illustrate different levels at which one sense might dominate another, not to argue that the McGurk effect is a case visual dominance—though, we think it is.

can speak of broad experience-based judgments *simpliciter* when $M_1 \ldots M_n$ includes all sense modalities.

When one sense dominates at the level of experience, it also dominates at the level of narrow experience-based judgment, and, if the modalities at issue are restricted to just these two modalities, at the level of broad experience-based judgment. For ordinary subjects encountering McGurk stimuli, for example, since vision dominates audition at the level of experience, it also dominates audition at the level of narrow experience-based judgment, and, if only vision and audition are at issue, at the level of broad experience-based judgment. (Since dominance at the level of narrow experience-based judgment implies dominance at the level of broad experience-based judgment when the sensory modalities are appropriately restricted, we often presume appropriate restrictions, and speak simply of dominance at the level of experience-based judgment.)

Of course, one might fight one's disposition to trust one or another sense in cases of conflict, or to trust the senses taken together when they converge; a subject encountering McGurk stimuli for the first time who has been told that the audio-track plays "ba" might report that it plays "ba", whatever her senses say. Experience-based judgments ignore such background beliefs. Since these *all-things-considered judgments* often stem from factors that are independent of online perception, they are relevant to the present project just in case those independent factors aren't especially salient, in which case all-things-considered judgments follow broad experience-based judgment.

3. Perception-Perception Dominance

We now explore vision's perception-perception dominance over other senses by focussing on visual perception's asymmetric effect on tactile perception.[7] After running through several experiments, we sketch a picture of visual dominance over touch. To prefigure, the picture implies neither that the visual invariably controls judgment when it conflicts with the tactile, nor that the tactile never controls judgment when it conflicts with the visual, but does imply that the visual affects how tactile stimuli are processed in cases of conflict in a way that the tactile doesn't affect how visual stimuli are processed.

[7] Though haptic and passive touch (i.e., touching and being touched) may be distinct senses, we simply use 'touch'. See Fulkerson (this volume).

Rock and Victor (1964) placed a small plastic square (25mm x 25 mm; 1mm thick) in a box. Some subjects saw the square through an eyepiece that used a transparent optical element to distort the image along its horizontal axis such that the square's apparent width was roughly halved. Other subjects touched the square by reaching around to the back of the box, and then through a black, silk cloth. Still other subjects (those in the experimental group) did both simultaneously. (The cloth prevented subjects in the experimental group from seeing a distorted image of their own hand, which otherwise would have allowed them to infer that the viewer distorts what they see.) After five seconds of exposure to the square, all subjects were asked to draw what they sensed. Subjects who touched-but-didn't-see the square drew squares. Subjects who saw-but-didn't-touch the square drew rectangles that were about twice as tall as wide. Subjects who touched-and-saw the square *also* drew rectangles that were about twice as long as wide, behaving just like subjects who saw-but-didn't-touch the square.

Rock and Victor repeated the experiment with a different measure of perceived shape: rather than drawing the target object, subjects chose from an array of candidate "matches". Even when the matches were selected by touch alone, subjects in the vision-only and vision-plus-touch conditions picked roughly the same *rectangular* "matches" for the original square. Subjects in the touch-only condition, unsurprisingly, picked squares. (See Miller 1972, and Welsh and Warren 1980 for similar results.)

In these experiments, subjects' eyes received a stimulus appropriate to a particular rectangle, their skin received a stimulus appropriate to a particular square, and their judgments followed vision. These results suggest, then, that vision affects how we interpret tactile stimuli with respect to shape in a way that touch doesn't affect how we interpret visual stimuli, i.e., that, with respect to shape, vision exercises *perception-perception* dominance over touch.

At what levels does this effect occur? Since subjects' choice of shape reflects their online perceiving, not background considerations, the effect occurs at least at the level of broad experience-based judgment. Does the effect extend to the levels of narrow experience-based judgment, and ultimately, experience? It would be bizarre if subjects who experienced tactile sensations as of square drew/selected-as-matches the very same rectangles as subjects who merely saw the object—and did so without protesting (e.g.,

reporting a conflict between their senses). Plausibly, then, the effect originates at the level of experience, with the visual affecting the phenomenal character of tactile experience. A supplementary experiment supports this claim (Rock and Victor 1964). The experiment used a viewing apparatus that made objects appear smaller than they were. Subjects in the see-and-touch condition were asked to open and close their eyes, reporting whether the object felt different with their eyes closed. This is akin to asking subjects to open and shut their eyes during a McGurk-presentation to test whether what they hear changes. Most subjects (23 of 38) reported that the object felt larger when they closed their eyes, then smaller when they opened them. This suggests both that the effect occurs at the level of experience, and that the effect is at least partly isolated, since subjects presumably didn't believe that the object was changing size as their eyes open and shut. Subjects' surprise also reinforces the (antecedently plausible) claim that the effect extends to all-things-considered judgments—subjects would not be surprised by the conflict if they already believed that vision was misleading. Since the effect occurs at the levels of experience, experience-based judgment, and all-things-considered judgment, we can say that vision dominates touch at these levels.[8]

This effect is amazingly resilient. Using a Rock-Victor-style task, Power and Graham found that the effect persists even when subjects are told to "really *feel* the object very thoroughly" (1976, p. 164). They also found that it holds even for those who have special facility with touch (expert potters). Others have found that repeated exposure doesn't undermine naivety: without receiving feedback pointing to inconsistency, subjects don't notice it spontaneously, even after repeated trials (see, respectively, Over 1966 and Power 1980). Still others have found that the effect persists even when the visual field is dimmed or objects are blurred (respectively, Warren 1979; Fishkin et al. 1975), suggesting that even degraded visual information (known to be degraded) often dominates at the level of experience.

[8] Vision dominates touch at the level of all-things-considered judgment in only 80% of subjects. Of the remaining 20%, some discovered a conflict between their senses when they (failing to follow instructions) looked away from the viewing apparatus, thereby encountering a tactile-only stimulus of the square; others spontaneously reported that they were suspicious of the viewing apparatus. Such subjects were more likely to resolve the tension "more in the direction of touch" than were naive subjects (p. 596). Since Rock and Victor say little about this result, it's not clear whether all such subjects drew/selected-as-matches squares, all such subjects drew/selected-as-matches rectangles that were closer to square than those drawn/selected by naïve subjects, or only some such subjects did one of these.

Power (1980) explored the limits of our tolerance for inconsistency. In two Rock-Victor-style tasks, he varied the visually apparent length and then angles of a square. The apparent length varied such that the square appeared to be a rectangle with two sides being 50% or 80% longer than the others. The apparent angles varied such that the square appeared to be either a diamond (with acute angles of 75 degrees) or a parallelogram (with acute angles of 70 degrees). Even in the cases of most extreme distortion, subjects' judgments followed vision.[9]

These results suggest that, when visual and tactile stimuli conflict, vision *often* dominates touch with respect to size and shape at the levels of experience, experience-based judgment, and all-things-considered judgment. But vision doesn't *always* dominate. One wonders, then, what modulates this dominance.

Heller (1983) identifies one factor. He presented subjects with clear acrylic, three-dimensional shapes, inscribing conflicting two-dimensional coloured shapes (e.g., an octagonal clear piece of acrylic might inscribe a blue square). Subjects felt the shapes with their hands while viewing the shapes through stained glass, which prevented them from seeing the acrylic, but allowed them to see the inscribed two-dimensional coloured shape, although it was noticeably blurred. Most subjects (11 of 12) chose matches for the target object that were consistent with the *tactile* stimulus, not the visual stimulus. Heller concludes that vision doesn't dominate if two conditions are met: first, visual stimuli are sufficiently blurry; second, the discrepancy between visual and tactile stimuli is extreme.

Here, touch affects how subjects interpret visual stimuli at the level of all-things-considered judgment. Does it therefore affect vision in just the way that vision affects touch in the earlier experiments? Many subjects in Heller's experiment claimed they couldn't see adequately, some asking whether they should choose what they saw or what they touched. This contrasts sharply with results from Rock and Victor, in which subjects claim to feel, say, a rectangle even while touching a square, even though vision and touch are both *clear*, where vision is clear in that, subjectively, visual experience seems as if it is entirely clear, rather than blurry, or otherwise misleading. Faced with conflicting clear

[9] These subjects wore a thin glove to prevent them from seeing their distorted hand. When subjects wore no glove, their judgments typically compromised between the stimulus to their eyes and hands. This suggests that either the tactile stimulus played an increased role in determining judgment, or subjects relied on inference from their belief about the distorting effect of the lens (acquired from seeing their distorted hand). All subjects (gloved or not) judged the square to be a diamond or parallelogram, suggesting that even if one knows that vision is quite distorted, it is not fully dominated by un-distorted touch at any level.

vision and degraded touch, one surely would follow vision. So, although Heller identifies a limit on vision's dominance, he hasn't found a case in which touch affects vision in a way that challenges visual dominance.

Over (1966) identifies another factor that modulates visual dominance of touch. He found that when visual and tactile stimuli are wildly discrepant, subjects judge that an angle is the *mean* of the visually and tactically presented angles. Whether these judgments reflect compromise at the level of experience, or conflict resolution at the level of broad experience-based judgment is unclear.

Lederman and Abbot (1981) and Heller (1982) identify yet another factor that modulates visual dominance. Each found that judgment might follow touch more closely than vision when assessing *texture*. Perhaps, then, touch dominates vision with respect to texture, at least at the level of all-things-considered judgment. At the very least, this result suggests that vision is less likely to dominate touch when touch's proper sensible is at issue than when a common sensible is. One might be tempted to infer that vision doesn't dominate other senses when the other senses' proper sensible is at issue, but the McGurk effect blocks this inference.

This begins to suggest a picture of visual *perception-perception* dominance of touch: given conflicting stimuli that are not wildly discrepant, and an appropriate target property *P* (e.g., shape, size), vision is likely to dominate touch with respect to *P* at each of the three levels, unless visual experience is much more degraded than tactile experience. This speculative picture relies on the vague properties *being an appropriate target property*, *being wildly discrepant*, and *being much more degraded than*. An ideal account would specify these properties—which some of our subsequent discussion does (see §5-7). Even as it stands, nonetheless, this picture strikes us as plausible, informative, and well suited to the data.[10]

[10] To be clear, this picture doesn't imply that visual dominance is "all or none", such that the visual dominates *completely* when certain conditions are met, but otherwise doesn't dominate at all. The conditions can be taken, accordingly, to mark the limit of a continuum ranging from the visual completely controlling judgment to the tactile completely controlling judgment. What the picture requires is that, given where this point lies, the visual dominates touch. An example illustrates the broader point. Suppose that in a Rock-and-Victor-style experiment subjects are asked to judge the shape of a four-inch by four-inch square. Suppose that the viewing apparatus initially presents a clear image of a two-inch by eight-inch rectangle. Suppose that subjects in this condition judge that the shape is a rectangle with its long sides four times longer than its short sides. Suppose that, as the viewing apparatus becomes increasingly blurry subjects judge that the shape is a rectangle with its long sides three times longer than its short sides, and so on, until, with enough blurring of vision, subjects judge that the object is square. The proposed picture of visual dominance is compatible with this possibility.

4. Imagery-Perception Dominance

In the previous section, we find that vision exercises perception-perception dominance over touch, at least with respect to size and shape. In this section, we shift to a more surprising discovery: vision exercises *imagery-perception* dominance over touch. After explicating *mental imagery*, we consider evidence that visual imagery asymmetrically affects how we interpret certain tactile stimuli.

4.1 Mental Imagery

Mental imagery is often described as *quasi-perceptual experience*, a characterization that traces back at least to Hume (1739/2000). What makes mental imagery *perceptual*? Imagery experiences resemble perceptual experiences in some significant, modality-specific way, i.e., the *phenomenal character of visual imagery* resembles the *phenomenal character of visual perceptual experiences* in a significant way, and so on for each modality for which there can be mental imagery. What makes mental imagery merely *quasi*-perceptual? First, imagery and perceptual experiences have different kinds of causes: ordinarily, only the latter are caused by confrontation with the stimuli they represent. Second, imagery lacks the assertoric force of perceptual experiences: ordinarily, only perceptual experience, absent any additional input or belief-forming basis, leads one to judge that the world is as the experience represents it to be.[11]

This last point requires elaboration. If encountering a tree causes one to have a visual experience with a phenomenal character as of a tree, one is strongly inclined to all-things-considered judge that there is a tree. By contrast, if visually imaging a tree causes one to have an experience with a phenomenal character (at least roughly) as of a tree, one is far less inclined to all-things-considered judge that there is a tree. The spatiotemporal nature of the assertoric force is also important. When encountering a tree leads one to judge (on the basis of visual experience) that there is a tree, one typically judges that there is a tree *here* and *now*. By contrast, even if visually imaging a tree leads one to

[11] For ease of presentation, we often talk as if perceptual experiences are representational. Nothing turns on this way of speaking.

make an all-things-considered judgment about a tree (e.g., that it exists, has green leaves, etc.), the judgment is far less likely to be about what is here and now.[12]

Why these differences? Perhaps the phenomenal characters of perceptual and imagery experiences differ such that only the former produce relevant, narrow experience-based judgments, as if imagery experiences come stamped with a phenomenal marker that defeats their potential assertoric force, right at the level of experience. Perhaps, instead, the process that produces imagery produces beliefs that the experiences aren't perceptual, as if mental images get stamped with a doxastic marker that can defeat their assertoric force, right when they otherwise would affect all-things-considered judgment. While these models differ (and there may be others), they agree that *mental images come marked as non-perceptual* such that they lack the default assertoric force of perceptual experiences.

Imagery, then, differs importantly from illusory and hallucinatory experiences. One might have a defeater for illusory experiences, as when one sees the Muller-Lyer illusion and is informed of its illusory nature. One might have a defeater for hallucinatory experiences, as when one remembers ingesting hallucinogens. As a general matter, neither illusory nor hallucinatory experiences, however, *come marked as non-perceptual*: when the uninformed see the Muller-Lyer illusion, they judge that the lines have different lengths; when the forgetful ingest hallucinogens, they judge that the pink elephants are parading. Even if this difference isn't metaphysically or cognitively necessary, it's certainly characteristic of typical imagery, illusory, and hallucinatory experiences. This yields the following characterization:

> *Mental imagery:* A mental image *i* is a perceptual-experience-like mental state/process that is not caused by the appropriate external stimuli (at least not in the ordinary way) and lacks assertoric force about the here and now (at least typically)—where *i* is *perceptual-experience-like* partly in that its phenomenal character resembles that of whichever modality (or modalities) it mimics.

[12] This point also can be used to distinguish ordinary visual experience from iconic memory. Unlike ordinary visual experience, iconic memory is assertive about the past, not the present, and the assertoric force has no special tie to the immediate environment. Thanks to Mohan Matthen for pointing out the importance of these spatiotemporal differences.

In sum, although imagery experiences resemble perceptual experiences in some significant, modality-specific way, the two have different kinds of causes and different default assertoric forces. This characterization isn't offered as an ultimate definition. It ignores some features that have been thought of, perhaps rightly, as definitive of mental imagery.[13] It also leaves open, inter alia, how imagery experiences resemble perceptual experiences, and what makes their resemblance significant. We treat this characterization, then, as a "working definition", adequate for our purposes, even if incomplete.

4.2 Evidence of Imagery-Perception Dominance

With this discussion of mental imagery in place, we consider the surprising discovery that visual *imagery* strongly affects how we interpret tactile stimuli during ordinary tactile *perception*.

Sathian et al. (1997) immobilized the right-hand index fingers of strongly right-handed subjects. They then ran small plastic domes across the immobilized index finger. The domes had gratings of varying widths cut into their faces. For some subjects, the gratings were run parallel to their immobilized finger. For others, the gratings were run perpendicular to their immobilized finger—see Figure 1. The domes were hidden from sight for all subjects. In the control condition, subjects were asked to identify the width of the gratings as either small or large. In the experimental condition, subjects were asked to identify the orientation of the gratings as either horizontal or vertical (relative to the finger). In each condition, subjects succeeded 75-80% of the time.

Subjects in the experimental condition reported using visual imagery to perform the task. Subjects in the control condition didn't. PET scans corroborated these reports, finding brain activity appropriate to visual imagery only during the experimental condition. Specifically, PET scans found increased cerebral blood flow in subjects' left parieto-occipital cortex, an area of visual cortex that activates when subjects generate visual images in response to various stimuli, e.g., verbal cues (Mellet et al. 1996; Kosslyn et al 1995). This activated region also overlaps with a large region of visual association

[13] To name two: first, empiricists can hold that imagery relies upon information acquired through previous perception in a way that perception doesn't; second, imagery is arguably under immediate voluntary control in a way that perception isn't—I can't help but see the mountain if it's in front of me, my eyes are open, and my visual system is functioning properly, I can't "get myself to see" a mountain if none is present, and I can visually image a mountain, no matter my surroundings (Reid 1764/1997; Wittgenstein 1967; McGinn 2004; Thomas 2010).

cortex in the parietal lobe, which is active during deliberate recall of visual patterns (Roland et al. 1990). These results suggest that visual imagery helps subjects perform the orientation task, even though the stimuli are only tactile.

Results from Zangaladze et al. (1999) further support this suggestion. They conducted an experiment like that described above, with one exception: they used transcranial magnetic stimulation (TMS) to disrupt activity in subjects' left parieto-occipital cortex. Success on the orientation task worsened significantly with the disruption, but success on the spacing task didn't. This further supports the claim that visual imagery plays a causal role in producing subjects' judgments about the orientation of the gratings—a role that is, moreover, beneficial, as we argue below. Self-reports fit this finding: robbed of visual imagery, many subjects said that they could feel the stimulus, but not identify its orientation. Zangaladze et al. conclude, therefore, that "visual cortical processing is *necessary* for normal tactile perception" (our italics, 1999: 588). Notice that, since subjects don't see the plastic domes, this conclusion is about the role of the neural correlate of visual *imagery* in online tactile *perception*.

Is the effect symmetric? Is haptic imagery often invoked for relevantly similar visual tasks? Zhang et al. (2004) explore this question (see also Klatzky et al. 1987). In a visual task, subjects saw images with a distinctive shape (e.g., a garlic) or texture (e.g., tree bark). In a haptic task, they touched (without seeing) 3D shapes (e.g., a heart) and heavily textured objects (e.g., sandpaper). They were asked to think of (but not report) descriptors that characterize the stimuli. They lay in an fMRI scanner throughout. Subjects regularly reported (after the task set and fMRI scan) using visual imagery for the haptic perceptual tasks, but didn't report using haptic imagery for the visual perceptual tasks. fMRI results corroborated these reports, finding strong activation of visual cortex during haptic tasks, but no significant activation of brain areas that process haptic information during visual tasks. Furthermore, the vividness of the reported imagery strongly correlated with the strength of activity in the ventral visual pathway, specifically, the lateral occipital complex, which is responsible for identifying objects by shape and high-level kind, for example 'egg' or 'dog' (Kourtzi and Kanwisher 2001). Since the ventral visual pathway is associated with conscious visual experience, this also suggests that the phenomenal character of visual imagery experiences resembles the phenomenal

character of visual perceptual experiences. This result supports the already plausible claim that the effect is asymmetric.

While we recognize that self-reports, neuroimaging, and TMS results can be problematic, we find the confluence among this evidence to be quite compelling: subjects' reports *predict* their performance in tasks, neuroimaging results, and TMS results.[14] Together, then, these experiments suggest that we often invoke visual imagery to learn about certain tactile stimuli, but we rarely if ever invoke tactile imagery to learn about visual stimuli. Recall that vision dominates another sense with respect to a certain property when the visual asymmetrically affects other senses in just this way. So, once again, vision dominates touch, now in virtue of visual *imagery* asymmetrically affecting tactile *perception*. So, the visual exercises *imagery-perception* dominance over touch.

At what level does this imagery-perception dominance occur? The visual exercises this dominance at least at the levels of all-things-considered judgment, and experience-based judgment. Whether it also exercises this dominance at the level of experience is unclear. Perhaps visual imagery tweaks tactile experience such that subjects in Sathian et al.'s experiment tactilely feel the gratings as oriented in the way that visual imagery suggests they are. Perhaps touch remains forever silent about the orientations. Future research should explore this issue. We conclude, then, only that vision exercises imagery-perception dominance of touch *at least* at the levels of all things-considered and experience-based judgment.

Clearly, we don't always use visual imagery to interpret tactile stimuli. One wonders, then, what modulates visual imagery-perception dominance over touch. Some factors should be clear: we're unlikely to invoke visual imagery if visual perception is available, and we're more likely to invoke visual imagery for some tasks than for others (for example, detecting shape versus texture). That said, since relatively little research aims to explore this issue, we are left to speculate. For now, we speculate that the

[14] Consider just a few problems with neuroimaging studies. First, since PET and fMRI measurements are rather coarse-grained, treating large groups of neurons as single unit, they can find that different tasks produce the same exact activation even though not a single neuron is activated for both tasks. Second, since one judges that an area is activated for a task by contrasting activity during that task with activity during a control task, choice of control is crucial, but there are no clear methods for choice. Third, in order to identify a neural correlate of a conscious event, such as the neural correlate of visually imaging, one must have an independent measure to identify the presence of consciousness, but any such measure is controversial (cf., Chalmers 1998). Fourth, there is clear evidence that common methods for analyzing neuroimaging results can mislead us badly (cf., Bennett et al. 2010).

modulating factors here loosely resemble those for visual perception-perception dominance of touch. We return to this issue later.

5. Why the Visual Dominates Touch

Why does the visual dominate touch with respect to size and shape? What difference between vision and touch grounds visual dominance? The answer lies in the nature of *visual experience*. Some have claimed that visual experience is *uniquely spatial*. O'Shaughnessy, for example, says that "perception at a distance is uniquely visual in type", and "other varieties of perception encounter their object *without* spatial mediation" (2009: 114).[15] This claim should be rejected; after all, audition alone can place sounds at least in egocentric space, and touch alone lets us navigate in ways that require spatial information. Nonetheless, the claim gets something right: visual experience is uniquely *richly* spatial.

A kind of experience is richly spatial if and only if it allows one to identify multiple *macrospatial* properties and *allocentric* properties *all-at-once*. *Macrospatial properties* are "large-scale elements like shape, size and orientation", which contrast with "microspatial properties", which "are small-scale surface elements like texture or irregularities" (Stoesz et al. 2003, p. 41).[16] *Allocentric properties* are spatial relations among external objects that hold irrespective of any given perceiver, which contrast with egocentric properties, which are spatial relations between external objects and perceivers. Properties are identified *all-at-once* if and only if they are identified more or less immediately and more or less simultaneously.

Does visual experience allow one to identify multiple macrospatial properties and allocentric properties all-at-once? Yes. V*isual experience* presents the size, shape, and orientation of objects, as well as their relative locations, all-at-once. By contrast, unaided tactile experience, presents macrospatial properties and allocentric properties at best through a *series of experiences* that are generated by *extensive*, *slow exploration*, such that tactile experience doesn't present multiple macrospatial or allocentric properties all-at-once. At the very least, visual experience allows one to identify multiple macrospatial

[15] Others, e.g., Strawson (1959), endorse similar claims, which are rooted in Berkeley's theory of vision (1713). See Evans (1980) for related discussion, and O'Callaghan (2011) for resistance to the aspatiality of audition.

[16] Although Stoesz et al., like most who use these labels, apply this distinction only to "tactile features", macrospatial properties are common sensibles, and some microspatial properties are too.

and allocentric properties *so quickly* and *in such rapid succession* that it's natural to describe their presentation as more or less immediate and simultaneous, in a way that would be strikingly unnatural for tactile experiences.

An example reinforces this view. Suppose you enter the space that houses Kawamata's chair installation (See Figure 2). Your visual experience is immediately rich with macrospatial and allocentric information. You see the shape of the sculpture, its meandering curves, its massiveness, the chairs, their shapes, sizes, and orientations—some vertical, some horizontal. You see the chairs as tightly stacked, one on top of another, and the cavernous space surrounding the sculpture. You see all this, moreover, *all-at-once*. Plausibly, touch provides nothing like this wealth of information. That said, we only need the contrastive point that, even if tactile experiences can provide the same information, it can do so only through a *series* of experiences that result form *extensive*, *slow* exploration, i.e., "feeling around", which precludes the immediacy and simultaneity that visual experience ordinarily enjoys.[17]

Visual experience, then, is richly spatial in a way that touch isn't. This difference, we contend, provides the basis for the visual's perception-perception dominance. Can it also provide a basis for the visual's imagery-perception dominance?[18] Yes, because visual *imagery* experience significantly resembles visual *perceptual* experience (cf., §4.1). One might object that the resemblance, however significant, doesn't extend so far that visual imagery experience inherits *being richly spatial*. Perhaps one is tempted to ground this objection by observing that visual imagery doesn't seem, phenomenologically, to be richly spatial. This concern merits two quick responses. First, the results from §4.2 suggest that visual imagery experiences present the very macrospatial properties that ordinary visual experiences do, which suggests that imagery experiences are richly spatial *enough* to ground vision's imagery-perception dominance of touch. Second, when we (the authors) close our eyes and visually image Kawamata's sculpture, our imagery experiences seem to us to present many of the same macrospatial

[17] Although we talk as if visual experience represents high-level properties, such as *being a chair* and *being vertical*, we aren't wed to that view. Even if visual experience presents only colour and shape properties, visual-experience-based judgments that there are chairs oriented in various ways are immediate and simultaneous in a way that touch-experience-based judgments aren't.

[18] To be clear, this is a claim about the *basis* for visual dominance, not about what *constitutes* its dominance. That is, we are not claiming that the visual dominates touch in that it gives us richly spatial properties but touch doesn't. We are claiming, instead, that the visual dominates touch with respect to properties *P* that can be detected (at least in part) by either sense (e.g., shape, size, etc.), and that this dominance with respect to *P* occurs because the visual gives us richly spatial properties *Q* but touch doesn't.

and allocentric properties that the original perceptual experiences presented, and we suspect that the same holds for most people, even if visual imagery experiences are, in some sense, mere shadows of their perceptual counterparts.[19]

We think, moreover, that the visual is *uniquely* richly spatial: it is richly spatial while the auditory, olfactory, etc. are not. Brief reflection should reveal that audition is the only other relevant sense. And, although audition provides some spatial information—a sound may be heard as coming from one's right, or as being "bigger" than a second sound—it is relatively impoverished in this regard, typically providing egocentric spatial information, and little if anything by way of macro-spatial property representation. To see the point, contrast the information one can gather by listening to, say, a basketball game with the information one can gather by seeing it. We expect, then, that this case for the dominance of the visual over touch extends to other senses; the visual dominates other senses no less than it dominates touch. We revisit this claim, offering some preliminary evidence, in the final section.

6. Significance of Dominance

With this discussion of visual dominance in place, we now consider why this dominance makes vision special. We first note that this dominance is, rather trivially, *psychologically* significant. We then argue that it is also *epistemically* significant in a surprising way: visual dominance of touch is *epistemically good*. Epistemically speaking, the visual *should* dominate touch.

Vision exercises perception-perception dominance over touch with respect to macrospatial and allocentric properties. The dominance is at the levels of experience, experience-based judgment, and all-things-considered judgment. Touch doesn't dominate vision for a comparably significant range of properties. So, because vision dominates touch, it contributes disproportionally to what we experience and think. This is enough to make vision significant *psychologically*. This point extends to visual imagery-perception dominance of touch, though whether visual imagery dominates at the level of experience is less clear.

[19] Imagine a contest to quickly identify macrospatial and allocentric properties of the sculpture between two people who have seen it previously. One is free to touch, but neither see nor visually image. The other is free to visually image, but neither see nor touch.

Of more interest to philosophers, visual dominance over touch makes vision significant *epistemically*. Consider first visual perception-perception dominance. We, joining most philosophers, take perceptual experiences to provide (at least) prima facie justification for perceptual beliefs: if one has a perpetual experience as of a leafy green tree, one is prima facie justified in believing that there is a leafy green tree. So, since vision exercises perception-perception dominance over touch at the level of experience, and this dominance leads to dominance in judgment, vision contributes more to justification for perceptual beliefs (in cases of conflict, and plausibly ordinary cases) than touch does.

Although this shows that visual dominance of touch is epistemically *significant*, it doesn't show that this dominance is epistemically *good*. To support the latter claim, we consider, in broad outline, epistemic externalist and epistemic internalist views, presuming that readers can extend our discussion to their preferred versions.

Externalism about justification holds, roughly, that factors that are inaccessible to a believer can affect whether her beliefs are justified. The most prominent version of externalism is process reliabilism, which holds, roughly, that a belief is justified just in case it is produced by a reliable process (see, e.g., Goldman 1967). Accordingly, a process reliabilist will hold that beliefs produced by visual perception-perception dominance of touch are justified just in case the process by which vision exercises this dominance is reliable. We contend that it is. So, we conclude, given reliablism, visual perception-perception dominance of touch is epistemically good.

Why think that the process by which vision exercises perception-perception dominance over touch is reliable? Since visual experience is richly spatial but touch isn't, vision is better situated than touch is to inform us about various properties (e.g., size, shape, location, etc.), and thus, one's best bet for acquiring information about these properties from tactile stimuli is to use vision, until it becomes significantly degraded. This point extends to visual imagery-perception dominance: as per the results discussed in §4, visual imagery is better situated to inform us about various properties (e.g., orientation) than touch is, and thus, absent visual experience, one's best bet for acquiring relevant information from tactile stimuli is to use visual imagery—hence, the success of subjects who use visual imagery to identify the orientation of felt gratings. So, the

process that produces each kind of dominance is reliable. (This model can be buttressed with an evolutionary claim: use of the visual to form beliefs about macrospatial and allocentric properties, at least for creatures like us, affords clear adaptive advantages.)

One might object that this account of the reliability of visual dominance faces a dilemma that goes as follows. First horn: for any property *P*, if touch can provide any information about *P* that vision can, then the visual isn't better situated to provide information about (and so does not dominate with respect to) *P*. Second horn: If vision can provide information about *P* that touch can't, then tactile stimuli aren't relevant to detecting *P* and so the visual doesn't dominate with respect to *P*—since dominance would require the visual affecting how tactile stimuli that are relevant to identifying *P* are processed. We reject each horn.

Regarding the first horn: even if touch can provide any information about macrospatial and allocentric properties that vision can, there are many ways in which the visual is better situated to provide information about those very properties. For example, since touch requires extensive exploration to identify, say, the shape of a complex object, we should expect such identification to rely heavily on memory, which, we should expect, decreases its reliability—imagine what would be required of memory to use touch to identify the shape of Kawamata's chair installation. A second point concerns speed. Even if the tactile provides comparable information, it does so, again, only through extensive, time-consuming exploration. Vision, by contrast, provides this information in a way that affords smooth and rapid interaction with the environment. This is an epistemic advantage of vision over touch. Although these points could be developed and complemented further, the basic problem with the first horn should be clear: the potential to provide information about a property is not the only relevant epistemic good. Regarding the second horn: even if vision can provide information about *P* that touch can't, tactile stimuli can be relevant to detecting *P*. In the aforementioned Sathian et al. experiment, for example, tactile stimuli is required to identify the orientation of the gratings even though one can't identify the orientation through touch alone. Relatedly,

there may be a property *P* such that one can't identify *P* through touch alone even though one can identify *P* on the basis of tactile stimuli.[20] This objection, therefore, fails.

One might be tempted to object, instead, along the following lines. "Since vision can mislead us when exercising perception-perception dominance, as in the Rock and Victor experiments, the process by which vision exercises perception-perception dominance is unreliable. So, experiences produced by that dominance don't provide even prima facie justification. So, visual perception-perception dominance is not epistemically good." This objection has the surprising result that beliefs produced by visual *perception*-perception dominance of touch are unjustified, even though beliefs produced by visual *imagery*-perception dominance are, presumably, justified. That should already make one suspicious. The objection, nonetheless, merits two replies.

First, if one individuates cognitive processes rather coarsely, such that the process that produces perceptual experiences in which vision exercises perception-perception dominance over touch *just is* the process that produces perceptual experiences in general, then the objector faces a dilemma. The objector must accept either that tactile experiences that result from dominance provide prima facie justification, even when they mislead us, or that visual experiences *never* provide justification because the process that produces them is *sometimes* unreliable, as when it leads to visual perception-perception dominance of touch in Rock-Victor-style experiments. Neither horn should be embraced.

Second, and more importantly, suppose that one individuates cognitive processes more finely. Does it follow that the process by which vision dominates is unreliable? No. Ernst and Banks (2002) argue rather effectively that visual perception-perception dominance of touch is "optimal". Their conclusion is based on several findings.

- First, in ideal conditions, vision alone allows subjects to accurately judge which of two sequentially presented bars is taller: even with only a small difference in height, subjects judge the taller bar as taller. So, in effect, the range of different bars that vision judges to be the same is very small. In ideal conditions, then, vision exhibits relatively *low variance*.

[20] Sometimes, moreover, properties other than *P* may not be identifiable without identifying *P*, as one might not be able to identify the shapes of overlapping objects without identifying various facts about orientation.

- Second, in ideal conditions, subjects using touch alone to judge the relative heights of the bars are much less accurate: the difference in height must be relatively large before subjects judge accurately. So, in effect, the range of different bars that touch judges to be the same is comparatively large. In ideal conditions, then, touch exhibits relatively *high variance*.
- Third, as viewing conditions are degraded for vision, pairs of bars that would have been judged accurately in ideal conditions are judged inaccurately, suggesting that the range of different bars that vision judges to be the same has increased. So, as viewing conditions are degraded, vision's *variance increases*.
- Fourth, given the first three findings, by systematically degrading viewing conditions for subjects using vision alone, one can identify how degraded vision must be before the variance for vision becomes as great as the variance for ideal touch: before one reaches what Ernst and Banks call the *point of subjective equality*. As it turns out, vision must be quite degraded before this point is reached, before vision is as poor at discriminating heights as ideal touch always is. To emphasize the epistemic significance of this point, we call it the point of subjective *epistemic* equality.[21]
- Fifth, when subjects see and touch the bars simultaneously, vision has a greater influence on judgment than touch does (at least at the level of all-things-considered judgment, and presumably at the level of experience) right up to the point of subjective epistemic equality, the point at which (degraded) vision alone exhibits as much variance as ideal touch alone.

Visual dominance of touch, then, is optimal: *vision stops influencing judgment more than touch does exactly when vision ceases to be more accurate and touch tends to become more accurate*, which, as it turns out, isn't until vision is quite degraded while touch is still ideal. So, even if visual perception-perception dominance of touch occasionally misleads us, the process by which vision exercises this dominance is generally reliable.

This discussion easily extends to virtue-epistemology-inspired interpretations of reliabilism. According to virtue epistemology, agents themselves, not their doxastic

[21] Roughly, the point of subjective epistemic equality for two senses and some property P is the point at which, in ordinary circumstances, experiences in each modality provide equally reliable information about P. So, since rather seriously degraded visual experience is more reliable than touch at its best for many properties (e.g., shape), the point of subjective epistemic equality between vision and touch for those properties is not reached until visual experience is quite degraded.

commitments, are the primary targets for epistemic evaluation. Agents are virtuous, the theory goes, if and to the degree that they exhibit intellectual virtues. The potential virtue at issue here is the disposition to perceive the world in a certain way—a way that leads to visual dominance of touch. For the *virtue reliabilist*, this disposition is an intellectual virtue just in case it is truth conducive (Goldman 1993; Greco 1999, 2009, 2010; Sosa 1980, 2007). As per Ernst and Banks, this disposition *is* truth-conducive, since one is disposed to follow vision right up to the point of subjective epistemic equality, which is exactly when following vision would begin to lead one astray. More broadly, this disposition is reliable, and hence an intellectual virtue, because the visual (including imagery) is better situated than touch to teach us about macrospatial and allocentric properties, which are the very properties for which we are disposed to follow vision.

Now, consider internalism about justification, which holds that justification depends exclusively on "internal" factors.[22] While this broad claim can be specified in many ways, suppose the following rough characterization of internalism: for any proposition *p* and subject *s* , *s* is justified in believing *p* if and only if *s* can access the basis for her belief in *p*, and that basis is good. Even this rough characterization, which leaves a lot open (e.g., what counts as "good"), is controversial. Some internalists hold, for example, that a belief is justified only if its basis is actually *accessed*. We ignore such controversy here, supposing that our discussion could be adjusted to suit any plausible version of internalism.

Do beliefs that result from visual dominance over touch typically satisfy these conditions? Yes. Beliefs formed on the basis of visual dominance are a subset of beliefs formed on the basis of perception. A standard, internalist picture of perception includes that perceptual experiences provide the basis for perceptual beliefs, perceptual experiences and their contents are accessible to their subjects (at least in principle), and perceptual experiences are a good basis for belief (at least ceteris paribus). So, given internalism, beliefs formed on the basis of visual dominance are prima facie justified.

Rejecting this argument requires showing that beliefs formed on the basis of visual dominance are relevantly different from ordinary perceptual beliefs. An objector might attempt to show this in either of two ways. First, the objector can deny that beliefs

[22] Epistemic internalism can be traced back at least to Plato, and is clearly articulated by Descartes. Contemporary sources include Prichard 1950, Chisholm 1977. See Alston 1989 for general discussion of theories of epistemic justification.

formed on the basis of visual dominance are even prima facie justified. Second, the objector can insist that such beliefs are invariably defeated in a way that ordinary perceptual beliefs are not. Consider these objections in turn.

Given any internalist account of the goodness of a basis for belief, if experiences that result from visual dominance are more problematic qua basis for belief than are dominance-free experiences, then (i) the former differ subjectively from the latter, and (ii) that difference undermines the goodness of the former. Contrary to (i), we suspect that experiences that result from visual dominance are subjectively indistinguishable from relevant dominance-free experiences, at least ordinarily. So, for example, one who touches a rectangle while seeing a rectangle and consequently has a tactile experience as of a rectangle, and one who touches a square while seeing a rectangle and consequently has a tactile experience as of a rectangle (as in the Rock and Victor experiment) typically have subjectively indistinguishable tactile experiences. (Since the McGurk effect is more familiar, perhaps one should compare the auditory experience of "va" that results from visual dominance to an ordinary auditory experience of "va".)[23] Contrary to (ii), we see no reason to think that subjective differences between dominance and dominance-free perceptual experiences (if, contrary to our suspicion, there are differences) undermine the epistemic goodness of dominated experiences. After all, subjects typically use experiences that result from dominance in the same ways epistemically that they use experiences that are dominance-free (e.g., to guide reasoning and behaviour), which suggests there is no accessed epistemically relevant difference between them, which suggests that there is no accessible epistemically relevant difference. And this, finally, suggests that there is no relevant difference that internalists can acknowledge.

Next, an objector might grant that beliefs that result from visual dominance have prima facie justification, but insist that they will be defeated quickly because, somewhere along the processing pathway, they get marked as epistemically pernicious. The objector will struggle to accommodate the fact that naïve subjects' judgments follow their dominated experience, which they would not if such experiences were invariably marked as epistemically pernicious in a way that subjects can access. Even the weaker claim that

[23] Even if dominance and dominance-free experiences are subjectively distinguishable in principle, moreover, it is quite likely that one rarely distinguishes them, which is relevant under formulations of internalism that require actual accessing where we require accessibility.

beliefs resulting from visual dominance get marked as *unusual* is not especially plausible given how surprised subjects are when the effects of dominance are revealed.

Given internalism, then, perceptual beliefs that result from visual dominance are prima facie justified. We now turn to one final way to that visual dominance is epistemically important for an internalist. Consider an old, familiar story about vision: visual experience provides *direct awareness* of material objects in that, when having (veridical) visual experiences, one is immediately acquainted with material objects. This story is called *direct realism*. *Intentionalism* counters that visual experience requires a representational intermediary, present in both veridical and non-veridical experiences, that stands between those seeing and seen material objects. Disagreements between advocates of direct realism and intentionalism run deep.

The old familiar story can be qualified in a way that avoids commitment to either theory, but acknowledges important insights that motivate each. Specifically, rather than holding that visual experience provides direct awareness of material objects, one can hold that visual experience *seems to the perceiver* (from the first-person perspective) to provide direct awareness of material objects.[24] Since this is a *phenomenological* claim, not a claim about the structure of vision, it is neutral between direct realism and intentionalism. It captures, nonetheless, part of the motivation for direct realism—while remaining acceptable to advocates of intentionalism.[25] We can express this claim by saying that visual experience provides *phenomenologically direct awareness* of material objects, which is to say it seems to those seeing as if vision provides direct acquaintance with chairs, tables, cars, people, dogs, and so on.

This phenomenological feature of the visual, we contend, is a consequence of vision's being richly spatial. As noted, both perceptual visual experience and visual imagery provide allocentric and macro-spatial information all-at-once. It is because of this informational feature of the visual that it strikes one, phenomenologically, that visual experience provides direct awareness of material objects, as if the objects are right there before one. Since the visual is richly spatial but the tactile is not, moreover, one should

[24] Others have made this point, more or less explicitly. Two clear statements are Sturgeon 2000 and O'Callaghan 2011.
[25] Arguably, the intentionalist (or representationalist) notion of the *transparency of experience* is introduced to accommodate this very phenomenological fact (Harman 1990). So the disagreement between direct and indirect realists does not concern this *phenomenological directness*, but instead what one might call *structural directness*: whether, in perceptual (visual) experience, there is some intermediary between perceiver and world.

expect that the tactile is not phenomenologically direct. More broadly, since, as we suggest shortly, the visual is uniquely richly spatial, one should expect that *only* the visual can provide phenomenologically direct awareness of material objects, which, loosely following many advocates of direct realism, is plausible. This alone makes visual dominance epistemically important.

This kind of dominance is important for an internalist, moreover, since internalism requires that reasons for belief be cognitively accessible. Visual perception and visual imagery provide phenomenologically direct awareness of material individuals. Accordingly, one can be aware of, can access, the appearance properties of visual experience, thereby acknowledging the reason-conferring spatial content of what one sees. So, at least in the case where the visual exercises dominance over touch just at the levels of judgment, one's visual experience provides prima facie justification for beliefs formed.

A story about learning might buttress this internalist account. Early in human development, we learn that visual experience represents macrospatial and allocentric properties quite well. Presumably somewhat later, we learn that visual imagery also represents such properties quite well. We come to implicitly understand, then, that vision's being richly spatial makes relying on the visual to detect such properties a good idea. In this way, an internalist might provide an account of both knowledge and metaknowledge acquired through the visual.[26]

Both broad views on epistemic justification, then, suggest that visual dominance of touch is epistemically good. Visual dominance, then, makes vision special in a way that should interest philosophers.

7. The Picture Revisited

Given this discussion, we can specify the picture of visual dominance over touch as introduced at the close of §3. That picture appeals to the vague properties *being an appropriate target property* and *being much more degraded than*. The discussion in §5 suggests that appropriate target properties include at least the macrospatial and

[26] Of course, this understanding needn't be of any technical sophistication, and might not itself be explicit in the agent's awareness. For many or most of us, it might instead be a procedural or skillful understanding. Instead, the idea here is just the following. If one knows a bit about vision or visual imagery (namely, that it tells one about spatial features of the environment), then one can know that one knows that P, since one can access one's acquired beliefs about the nature of the visual vis-à-vis rich spatial nature, and identify the fact that one is forming one's belief that P on the basis of that visual faculty.

allocentric. The discussion in §6 qualifies the relevant notion of degradation: visual experience is "much more degraded than" tactile experience if and only if the visual becomes more degraded than it is at the point of subjective epistemic equality, which is quantitatively defined by Ernst and Banks. So, we get the following conditions for perception-perception and imagery-perception dominance, respectively:

- Given conflicting stimuli that aren't wildly discrepant, and a macrospatial or allocentric target property *P*, visual experience is likely to dominate touch with respect to *P* at each of the three levels, unless visual experience is so degraded that one passes the point of subjective epistemic equality for visual and tactile experience.
- Given that a macrospatial or allocentric target property *P* stimulates the skin but not the eyes, visual imagery is likely to dominate touch with respect to *P* at least at the level of all-things-considered judgment and experience-based judgment, unless visual imagery is so degraded that one passes the point of subjective epistemic equality for visual and tactile experience.

Each kind of dominance is significant psychologically and epistemically, and there is no reason to believe that other senses dominate the visual with respect to comparably significant properties, or a comparably wide range of properties.

8. A conclusion: Other Senses, Other Properties

Does this picture generalize to other senses? Consider the standard Aristotelian senses. Even if olfactory and taste experiences present macrospatial and allocentric properties, neither presents multiple properties all-at-once. While it is more plausible that auditory experience presents macrospatial and allocentric properties, it doesn't present them all-at-once either: try to imagine learning all-at-once from auditory experience alone, the size, shape, and orientation of Kawamata's sculpture, or how various chairs are related to one another in space. This brief consideration already shows that visual experience is *uniquely* richly spatial, which implies that the picture generalizes: the visual

dominates *all* other senses in similar ways. Further considerations support this conclusion.[27]

Does vision often dominate given target properties other than the macrospatial and allocentric? A number of results suggest that the answer is "yes": the McGurk effect doesn't involve a macrospatial or allocentric target property, and yet vision dominates; Royet et. al 1999 found that vision dominates smell when judging whether an odour comes from a comestible object, which is not obviously a judgment about a macrospatial or allocentric property either.[28] We suspect that our picture can be extended to these cases, at least in broad outline, although working out the details requires an independent project.

We close by suggesting another broad class of property for which we expect the visual to dominate other senses, although empirical evidence here is scant. Recall that Zhang et al. (2004) found that visual imagery affects judgments based on tactile-stimuli when we identify objects by high-level kind, such as egg or dog. We predict that this holds more broadly: that the visual is likely to dominate other senses when one aims to identify the source of a stimulus by such "medium-sized-dry-good" kind. We suspect that this dominance is as pervasive and significant as it is for macrospatial and allocentric properties. In order to support this speculation, we close by asking that you *imagine* ...

You're lying in your warm bed, early on a cold morning. Your partner prepares for work, somewhat noisily. For some sounds, the source is easy to identify: he's running a hair dryer, spraying hair spray, opening the fridge, running the clothes dryer. For other sounds, the source is mysterious: something in the dyer makes sustained metallic, clanging noises, but you can't tell what; some of the items being moved in the kitchen are identifiable, others aren't.

Suppose you want to identify the sources of the more elusive sounds? What would you do? Pause, and consider this question before reading on.

---

[27] Two notes. First, if there are multi-modal experiences, we conjecture that these can be richly spatial only if and to the extent that they have visual experiences as constituents. Second, a generalization would be: given conflicting stimuli to the eyes and the organ for any non-visual sense *S*, and a macrospatial or allocentric target property, vision is likely to exercise perception-perception dominance over *S* at each level, unless visual experiences are degraded *beyond the point of subjective epistemic equality* with experiences in *S*; likewise, mutatis mutandis for imagery-perception dominance.

[28] They exposed subjects to a wide variety of odorants. In the control condition, subjects were asked to identify each odor as familiar. In the experimental condition, subjects were asked to report whether the odor was one from a comestible object. In the experimental condition only, PET scans found increased cerebral blood flow to visual cortex, suggesting that subjects invoke visual imagery when trying to identify important facts about an object. For more research on vision and olfaction, see Djordjevic et al 2004 and Sakai et al 2005. For research on visual imagery and audition, see Intons-Peterson 1992; De Volder et al 2001; von Kriegstein et al 2005; Amedi et al 2005; Hubbard 2010. For research on vision, visual imagery, and taste/flavor, see Clydesdale 1984 and Kobayashi et al 2004.

We suspect that, if it really mattered, if something important depended on getting the answer right, you would go look, however comfortable bed may be. But suppose that you can't or won't go look; you are too tired, lazy, injured, or whatever. What would you do then? Again, pause and consider this question before reading on.

We suspect that you would begin to image, v*isually image*, candidate sources of the elusive sounds. You might run through candidate sources for the clanging in the dryer, for example, by visually imaging, in order,

(i) trousers and tops (concluding that their buttons wouldn't clang so loudly),
(ii) coins (concluding that their clanging wouldn't be so sustained), and
(iii) a pair of jeans with a belt attached (concluding that these would make just the noise you hear, and would be in the dryer, given your partner's early morning absent-mindedness).

Here, visual imagery helps you generate and assess hypotheses: running through images of candidate sound-sources generates hypotheses; "cross-referencing" images against incoming auditory information tests them. Without visual imagery, you might never identify the source of the initially elusive sound—or, worse, you might need to get out of bed. At the very least, imagery is central to what you actually do.

Perhaps you did not spontaneously think to identify the source of the clanging through visual imagery. Even if you did, moreover, others won't. The use of visual imagery surely differs from person to person, and case to case: one who is bad at visually imaging might not be disposed to use visual imagery; one who has heard the clanging every morning for years won't need to. Our suggestion that you would invoke visual imagery, nonetheless, probably did not sound bizarre. But suppose we had suggested that you would invoke smell, taste, or touch imagery, perhaps *tactilely* imaging shirts, coins, and belts. This would have seemed immediately implausible. It is not too surprising, then, that many people spontaneously say they would appeal to visual imagery when we present the belt-in-dryer case, but none say they would appeal to other kinds of sense imagery. Similarly, it would have sounded bizarre if we had described a scenario in

which you see something clearly, and then suggested that you might invoke auditory imagery to identify its source.[29]

We suspect that many properties, including "medium-sized-dry-good-kind", could be added to the list of target properties with respect to which the visual dominates other senses. We also suspect, that there are few if any properties with respect to which other senses exercise similar dominance over the visual; at the very least, there is no evidence that other senses dominate vision with respect to comparably significant properties, or a comparably wide range of properties.[30] And, as we have argued, the dominance of the visual is psychologically and epistemically significant. This dominance, then, goes a long way toward explaining the sense in which we prefer vision, in which we are visual creatures, and in which vision is special.

[29] This illustration depends itself on imagining the very scenario in which you purportedly invoke imagery, and this kind of imagining may invite you to use visual imagery when you otherwise wouldn't. So, the illustration is less than ideal. For audio recordings that may, depending upon the listener, trigger the phenomenon we have in mind, please see here: [**UPLOAD SOUND SAMPLES TO DS WEBSITE; ADD LINK HERE].**

[30] Although we only consider the dominance of the visual, one non-visual sense sometimes dominates another; e.g., audition sometimes dominates touch (cf. De Gelder and Bertelson 2003).

Figures:

FIGURE 1: From Sathian (permissions needed):

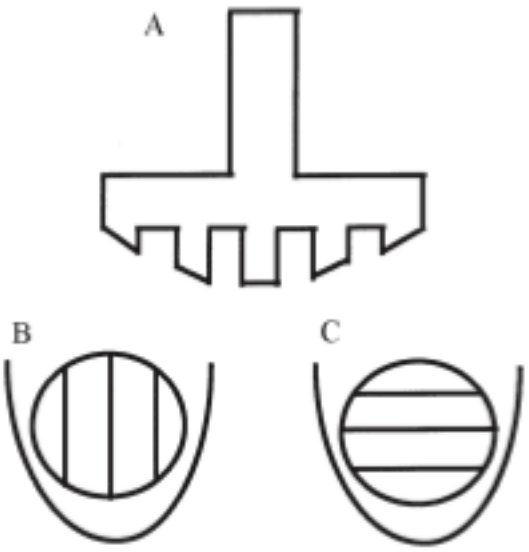

FIG. 1. Cross-sectional view of grating (**A**), which was applied to the fingerpad oriented either along (**B**) or across (**C**) the long axis of the finger.

FIGURE 2: Tadashi Kawamata: Chair installation (permissions needed)

References

Amir Amedi, Katharina von Kriegstein, N.M. van Atteveldt, M.S. Beauchamp, & M.J. Naumer. (2005). Functional Imaging of Human Crossmodal Identification and Object Recognition. *Experimental Brain Research* 166: 559-571.

Clare Batty (2011). Smelling Lessons. *Philosophical Studies* 153 (Mar.):161-174.

Craig M. Bennett, Abigail A. Baird, Michael B. Miller, & George L. Wolford (2010). Nueral Correlates of Interspecies Perspective Taking in the Post-Mortem Atlantic Salmon: An Argument for Proper Multiple Comparisons Correction. *Journal of Serendipitous and Unexpected Results* 1(1): 1-5.

Berkeley, G. (1713/1975). Three dialogues between Hylas and Philonous. In: M. R.

Ayers (Ed.), *Philosophical works, including the works on vision*. Dent, London.

David J. Chalmers (1998). On the Search for the Neural Correlate of Consciousness. In Stuart R. Hameroff, Alfred W. Kaszniak & A.C. Scott (eds.), *Toward a Science of Consciousness II*. MIT Press.

Clydesdale, F. M. (1984). The influence of colour on sensory perception and food choices. In J. Walford (Ed.), *Developments in food colours* (pp. 75–112). London: Elsevier Applied Science.

Beatrice De Gelder & Paul Bertelson (2003). Multisensory Integration, Perception, and Ecological Validity. *Trends Cognitive Science* 7(10): 460-467.

Anne De Volder, H. Toyama, Y. Kimura, Kiyosawa, H. Nakano, A. Vanlierde, M.C. Wanet-Defalque, M. Mishina, K. Oda, K. Ishiwata, M. Senda. (2001). Auditory Triggered Mental Imagery of Shape Involves Visual Association Areas in Early Blind Humans. *Neuroimage* 14(1): 129-39.

J. Djordjevic, R.J. Zatorre, M. Petrides, and M. Jones-Gotman (2004). The Mind's Nose Effects of Odor and Visual Imagery on Odor Detection. *Psychological Science* 15(3): 143-8.

Marc O. Ernst & Martin S. Banks (2002). Humans Integrate Visual and Haptic Information in a Statistically Optimal Fashion. *Nature* 415:429-433.

Gareth Evans (1980). *Varieties of Reference*. Oxford: Oxford University Press.

S.M. Fishkin, V. Pishkin, & M.L. Stahl (1975). Factors involved in visual capture. *Perceptual Motor Skills* 40 (2): 427-34.

Mathew Fulkerson (this volume). What Counts as Touch? In D. Stokes, S. Biggs, and M. Matthen (eds.), *Perception and Its Modalities*: Oxford University Press.

Alvin I. Goldman (1967). A Causal Theory of Knowing. *Journal of Philosophy* 64 (12):357-372.

---Alvin I. Goldman (1993). Epistemic Folkways and Scientific Epistemology. *Philosophical Issues* 3:271-285.

John Greco (1999). Agent Reliabilism. In J. Tomberlin (ed.), *Philosophical Perspectives 13: Epistemology*. Atascadero: Ridgeview.

---(2009). *Epistemic Value.* Oxford: Oxford University Press.

---(2010). *Achieving Knowledge*. Cambridge: Cambridge University Press.

Gilbert Harman (1990). The Intrinsic Quality of Experience. *Philosophical Perspectives* 4:31-52.

Morton Heller (1982). Visual and Tactual Perception: Intersensory Cooperation. *Perception and Psychophysics* 31: 339-344.

Morton Heller (1983). Haptic Dominance in form Perception with Blurred Vision. *Perception* 12: 607-613.

Timothy Hubbard (2010). Auditory Imagery: Empirical Findings. *Psychological Bulletin* 136(2): 302-29.

David Hume (1739/2000). *A Treatise of Human Nature.* Oxford University Press.

Margaret Jena Intons-Peterson, Wendi Russell, & Sandra Dressel (1992). The Role of Pitch in Auditory Imagery. *Journal of Experimental Psychology. Human Perception & Performance* 18(1): 233-240.

R.L. Klatzky, S. Lederman S and C.J. Reed (1987) There's more to touch than meets the eye: The salience of object attributes for haptics with and without vision. *Journal of Experimental Psychology: General* 116: 356-369.

M. Kobayashi, M. Takeda, N. Hattori, M. Fukunaga, T. Sasabe, N. Inoue, Y. Nagai, T. Sawada, N. Sadato, Y. Watanabe (2004) Functional imaging of gustatory perception and imagery: "top-down" processing of gustatory signals. *Neuroimage*: 23(4):1271-82.

Stephen Kosslyn, W.L. Thompson, I.J. Kim and N.M. Alpert (1995) Topographical representations of mental images in primary visual cortex. *Nature* 378: 496–498.

Zoe Kourtzi & Nancy Kanwisher (2001). Representation of Perceived Object Shape by the Human Lateral Occipital Complex. *Science* 293: 1506-1509.

Katharina von Kriegstein, A. Kleinschmidt, P. Sterzer, & A.L. Giraud. (2005). Interaction of Face and Voice Areas During Speaker Recognition. *Journal of Cognitive Neuroscience* 17: 367-376.

S.J. Lederman & S.G. Abbott S. G. (1981). Texture Perception: Studies of Intersensory Organization Using a Discrepancy Paradigm, and Visual Versus Tactual Psychophysics. *Journal of Experimental Psychology: Human Perception and Performance* 7: 902–915.

William Lycan (2000). The slighting of smell. In N. Bhushan & S. Rosenfeld (Eds.), *Of minds and molecules: New philosophical perspectives on chemistry*. Oxford: Oxford University Press.

D. L. C. Maclachlan (1989). *Philosophy of Perception.* Cliffs Prentice-Hall.

Michael Martin (1992). Sight and Touch. In Tim Crane (ed.), *The Contents of Experience*. New York: Cambridge University Press.

Colin McGinn (2004). Mindsight. Cambridge, MA: Harvard University Press.

Harry McGurk & John MacDonald (1976). Hearing Lips and Seeing Voices. *Nature*, 264: 746–748.

Emmanuel Mellet, Nathalie Tzourio, Fabrice Crivello, Marc Joliot, Michel Denis, & Bernard Mazoyer (1996). Functional Anatomy of Spatial Mental Imagery Generated from Verbal Instructions. *The Journal of Neuroscience* 16(20):6504–6512.

E.A. Miller (1972). Interaction of Vision and Touch in Conflict and Non-Conflict from Perception Tasks. *Journal of Experimental Psychology* 96:114-123.

Casey O'Callaghan (2011). Lessons From Beyond Vision (Sounds and Audition). *Philosophical Studies* 153 (1):143-160.

Brian O'Shaughnessy (2009). The location of a perceived sound. In M. Nudds & C. O'Callaghan (Eds.), *Sounds and perception: New philosophical essays*. Oxford: Oxford University Press.

Mark R. Stoesz, Minming Zhang, Valerie D. Weisser, S.C. Prather, Hui Mao, & Krish Sathian (2003). Neural Networks Active During Tactile Form Perception: Common and Differential Activity During Macrospatial and Microspatial Tasks. *International Journal of Psychophysiology* 50: 41–49.

Ray Over (1966). An Experimentally Indiuced Conflct between Vision and Proprioception. *British Journal of Psychology* 57: 335-341.

Roderick Power (1980). The Dominance of Touch by Vision: Sometimes Incomplete. *Perception* 9: 457-466.

Roderick Power & Anne Graham (1976). Dominance of Touch by Vision: Generalization of the Hypothesis to a Tactually Experienced Population. *Perception* 5: 161-166.

Thomas Reid (1764/1997). *An Inquiry into the Human Mind on the Principles of Common Sense*. Derek R. Brookes (ed.), Pennsylvania State University Press.

P.E. Roland, B. Gulyas, R.J. Seitz, C. Bohm, and S. Stone-Elander (1990). Functional anatomy of storage, recall, and recognition of a visual pattern in man. *NeuroReport* 1, 53–56.

Irvin Rock & Jack Victor (1964). An Experimentally Created Conflict between the Two Senses. *Science* 143: 594-596.

Lawrence Rosenblum, Mark Schmuckler, & Jennifer Johnson (1997). The McGurk Effect in Infants. *Perception and Psychophysics* 59 (3): 347-357.

Nobuyuki Sakai, Sumio Imada, Sachiko Saito, Tatsu Kobayakawa and Yuichi Deguchi (2005). The Effect of Visual Images on Perception of Odors. *Chemical Senses* 30: i244-i245.

Krish Sathian, Andro Zangaladze, John Hoffman, & Scott Grafton (1997). Feeling with the Mind's Eye. *Neuroreport* 8(18):3877-81.

Ernest Sosa (1980). The Raft and the Pyramid: Coherence versus Foundations in the Theory of Knowledge. *Midwest Studies in Philosophy* 5: 3–25.

---(2007). *Apt Belief and Reflective Knowledge, Volume 1: A Virtue Epistemology*. Oxford: Oxford University Press.

Peter F. Strawson (1959). *Individuals*: *An Essay in Descriptive Metaphysics*. Methuen and Co Ltd.

Nigel Thomas. (2010). Mental Imagery. In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy.* Stanford, CA: CSLI.

David Warren (1979). Spatial Localization under Conflict Conditions: Is There a Single Explanation? *Perception* 8 (3): 323-37.

Robert Welch & David Warren (1980). Immediate Perceptual Response to Intersensory Discrepancy. *Psychological Bulletin* 88: 638-667.

Ludwig Wittgenstein (1967). *Zettel*. Ed. G.E.M. Anscombe & G.H. von Wright; Trans. G.E.M. Anscombe.). Oxford: Blackwell.

Andro Zangaladze, Charles Epstein, Scott Grafton, & Krish Sathian (1999). Involvement of Visual Cortex in Tactile Discrimination of Orientation. *Nature* 401: 587–590.

Minming Zhang, V.D. Weisser, R. Stilla S.C. Prather, & Krish Sathian (2004). Multisensory Cortical Processing of Object Shape and its Relation to Mental Imagery. *Cognitive Affective Behavioral Neuroscience* (4): 251–259.