with multiple depth values at every single (x,y) location, or to represent the experience of empty space between the observer and a visible object" (sect. 3, para. 5). These same models also ignore that every voxel of "empty" space contains light of some intensity and chromaticity.

This confusion probably results from naïvely accepting the popular notion that humans care only about the location and qualities of objects, making the perception of illumination irrelevant. This assumption is so prevalent that much of color research is devoted to determining how the visual system "discounts the illuminant." However, a viable solution to the Gestalt problem of color constancy will emerge only with a more complete description and understanding of how we subjectively *experience* illumination. Ironically, Lehar's aspiration to describe the subjective experience of spatial vision in terms comparable to those of color vision reveals that current color vision research is also in peril. That is, he claims that color phenomena are reducible to hue, intensity, and saturation because that is how the brain represents them physiologically (sect. 2.3). Yet models of hue, intensity, and saturation cannot be the "primitives of raw conscious experience" (sect. 4, para. 3), in that these qualities remain invariant as illumination changes across space.

This confound is apparent when Lehar discusses his Figure 1 as containing "explicit volumes, bounded by colored surfaces, embedded in a spatial void" (sect. 5.1, para. 2), where "every point can encode either the experience of transparency or the experience of a perceived color at that location" (sect. 6). His more accurate intuition is that there are also intermediate states between transparent and opaque "to account for the perception of semitransparent surfaces" (sect. 8.1, para. 1). I suggest that Lehar consider filling these semitransparent voxels with "potential illumination" "at multiple depth values at every single (x,y) location." This would also strengthen his second and third conclusions that "volumes of empty space are perceived with the same geometrical fidelity as volumes of solid matter" and that "multiple transparent surfaces can be perceived simultaneously" (sect. 10, points 2 and 3). Having semitransparent voxels contain "potential illumination" is a more parsimonious description of the void between your eyes and this page. You can actualize the "potential illumination" of these voxels by placing your finger in front of any shadow cast on the page. More accurately, Lehar's phenomenological model allows *only* the plane of voxels directly in front of a given surface to contain cast shadows (i.e., less illumination), because the voxels that compose the surface must be the color of the opaque surface itself (sects. 5.1, 8.1).

Note that this concept is not merely peripatetic (Aristotle 1976) or an ether explanation, in that we are always subjectively aware of the illuminant. For example, by looking from your illuminated reading room into a dark hallway, your subjective experience is not only that the hallway walls are under less illumination but also that the space itself contains less light. In this way, "potential illumination" can also address why color constancy differs in two- versus three-dimensional scenes. For example, Gilchrist (1977) had observers look through a pinhole into a room containing a doorway into a second room. The near room was dimly illuminated and the far room was highly illuminated. Attached to the door frame were several papers, arranged so that a mid-gray paper appeared either adjacent to the door frame or (with its corners removed) on the far room's back wall. The lightness of the paper shifted in the direction of lightness constancy depending on whether it appeared on the door frame or on the far wall. Schirillo et al. (1990) generated equivalent stimuli in two dimensions using a stereoscopic cathode ray tube (CRT) and stereoscope, yet this replication produced only a fraction of Gilchrist's constancy. I hypothesize that this occurred because stereoscopic space does not contain the actual voxels of either high (e.g., near room) or low (e.g., far room) illumination. In essence, Schirillo and colleagues failed to preserve Lehar's necessary condition of "volumetric mapping" (target article, Fig. 1D).

The ubiquitous use of CRT images reduces scenes to Alberti's window, which retains perspective cues but eliminates Lehar's requirement that space be volumetric. This obfuscates the color constancy paradox, in that these voxels contain illumination. For example, Adelson's (1993) famous wall-of-blocks illusion contains cubes of identical luminance that appear dissimilar in lightness and concomitantly under illusory transparent stripes. Logvinenko et al. (2002) eliminated both the appearance of transparency and the lightness illusion by constructing a three-dimensional version of Adelson's two-dimensional display. I hypothesize that the visual system does not add a transparent veil to Logvinenko's display because it already ascribes illumination to every voxel in space. However, when Adelson eliminates such volumes but retains the same spatial geometry via X-junctions, the visual system reconstructs the volume to contain regions of illusory transparency (i.e., illumination). Consequently, Lehar's improved spatial model requires a phenomenal description of *empty* space that includes "potentially illuminated" voxels.

If vision is "veridical hallucination," what keeps it veridical?

Peter Ulric Tse

Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755. **Peter.Tse@dartmouth.edu**
http://www.dartmouth.edu/artsci/psych/faculty/tse.html

Abstract: If perception is constructed, what keeps perception from becoming mere hallucination unlinked to world events? The visual system has evolved two strategies to anchor itself and correct its errors. One involves completing missing information on the basis of knowledge about what most likely exists in the scene. For example, the visual system fills in information only in cases where it might be responsible for the data loss. The other strategy involves exploiting the physical stability of the environment as a reference frame with respect to which the eyes and body can move.

> *[S]pace and time are only forms of sensible intuition, and hence are only conditions of the existence of things as phenomena . . . we can have no cognition of an object, as a thing in itself, but only as an object of sensible intuition, that is, as phenomenon*
> —Immanuel Kant (*Critique of Pure Reason,* 1781)

Lehar develops the Kantian insight that perception is (1) entirely a mental construction; (2) lacks access to the world-in-itself to determine the accuracy of its representations; and (3) is only possible given an internal framework of space-time that permits sensory input to be interpreted as occurring in an external space-time. Here I focus on how the brain can construct true information about the world when there is no way to judge objectively whether that information is true by comparing that information to the world-in-itself.

To create veridical information, the visual system must compensate for errors, data loss, and processing bottlenecks imposed by its imperfect design. It has nothing but the ambiguous, incomplete, and noisy image to determine whether it has made an error. It must therefore know what types of image cues indicate errors and it must have strategies to correct errors. The visual system corrects itself only when it is responsible for errors or data loss. It compensates for its own likely errors using two strategies. One is to rely on world knowledge, and the other is to assume that the world is stable.

For example, when does the visual system fill in missing phenomenal features and when does it merely note that completion takes place without filling-in (see Fig. 1)? Filling-in occurs when the information that is missing from the image is missing because of the visual system's own failure to detect it. The visual system follows the principle "no news isn't necessarily bad news when there was no way to get the news in the first place." The visual system functions as if it knows that it does not always have adequate in-

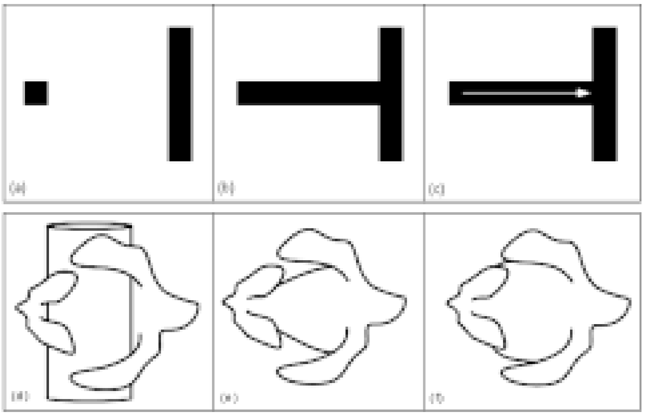

Figure 1 (Tse). When (a) is replaced all at once by (b) a smooth apparent motion (indicated by an arrow) is filled in (Tse & Logothetis 2002). No filling-in occurs in (d) cases of amodal completion (Tse 1999b). The shape behind the occluder, whether (e) or (f), is not completed.

formation in a particular domain to determine the structure of the world. Hence, when information is missing from an image, this is not necessarily regarded as contradictory information or information that the undetected thing does not exist in the world. The information might be undetected because of poor viewing conditions or because of the inherent limits on detection imposed by a noisy perceptual apparatus that has limited sensitivity. A more precise formulation is:

> In the absence of direct (but presence of appropriate indirect) image evidence for the existence of *x*, under viewing conditions where *x* would not be detectable in the image even if it were present in the world, the visual system may not only not reject *x*, it may assume *x* to be the case, and interpolate *x* so that *x* is seen as if it were visible.

Filling-in occurs because the visual system in effect blames itself. The sensitivity of the visual system under given viewing conditions can be too poor to permit detection of an entity that indirect image evidence implies exists. Under such conditions the visual system creates what it "knows" must be missing. In amodal completion there is no filling-in because the visual system does not blame itself. The shape or features of the occluded portions of an object are not filled in, because under no possible viewing conditions would shape or surface features be visible through an opaque occluder. No matter how insensitive the visual system might be, it cannot blame itself for not detecting entities that are in principle undetectable under any viewing conditions.

A second strategy to overcome potential errors is to analyze image data under the assumption that the world is stable. First, the visual system does not need to store detailed information about the world because it can always sample the world for more information (O'Regan 1992). Second, the visual system can stabilize perceptual space by relying on the presumed stability of the world. For example, the retinal image usually changes en masse only when the body or eyes move. The system can exploit this stability in order to maintain the eyes and body in a constant position with respect to the world. A classic demonstration of this is the "moving room" experiment (Lee & Aronson 1974), in which a person stands in a room that is set on rollers. When the walls move, rather than assume that the world has moved, the visual system assumes that the body has moved and corrects for this by changing the body's position. Sometimes subjects even fall over. It is as if the visual system blames itself for the discrepancy caused by the moving room and compensates by relying on a world that it wrongly assumed was stable.

Another example can be found in the recalibration of perceptual space that takes place after a saccadic eye movement. Deubel and colleagues (Deubel et al. 1998) have argued that the system seeks its saccade target immediately after a saccade. If this target is found within a certain spatial and temporal window, the visual system assumes the target object to have remained stable and uses it as a reference object to determine the positions of other objects. This is true even when the target object in fact moves during the saccade. Even more surprisingly, Deubel and colleagues find that if the target has moved to the right, and a neighboring distractor has not moved at all, the visual system creates a percept of a target that has remained stationary and a distractor that has jumped to the left. Because the visual system's initial saccade lands accurately on the position where the target was at the beginning of the saccade, the visual system should know that the target has changed position. But this is true only if it assumes its saccade was infallibly correct. Instead, a corrective saccade is automatically made to the new position of the target, and the object is assumed to have remained stable. The distractor's illusory jump to the left is filled in because it is the motion that must have occurred, assuming the stability of the target and the world. Again, it is as if the visual system blames itself for the discrepancy and relies on the stability of the world to correct its presumed error.

Because the visual system has no direct access to the world, it must rely solely on the image to judge whether it has made errors in specifying the image-to-world correspondence. Error correction is only possible based on assumptions about world structure and statistics. Completion may be phenomenal or not, depending on whether the visual system "blames itself" for the data loss. In addition, the visual system takes a world that it assumes to be stable as its frame of reference. These two strategies allow the visual system to overcome the handicap that the truth of perceptual information cannot be judged by comparing that information with the world-in-itself.

Is the world in the brain, or the brain in the world?

Max Velmans

Department of Psychology, Goldsmiths College, University of London, New Cross, London SE14 6NW, United Kingdom. **m.velmans@gold.ac.uk**
http://www.goldsmiths.ac.uk/departments/psychology/staff/velmans.html

Abstract: Lehar provides useful insights into spatially extended phenomenology that may have major consequences for neuroscience. However, Lehar's biological naturalism leads to counterintuitive conclusions, and he does not give an accurate account of preceding and competing work. This commentary compares Lehar's analysis with that of Velmans, which addresses similar issues but draws opposite conclusions. Lehar argues that the phenomenal world is in the brain and concludes that the physical skull is beyond the phenomenal world. Velmans argues that the brain is in the phenomenal world and concludes that the physical skull is where it seems to be.

Is the phenomenal world in the brain, or is the brain in the phenomenal world? As William James (1904) noted, "the whole philosophy of perception from Democritus's time downwards has been just one long wrangle over the paradox that what is evidently one reality should be in two places at once, both in outer space and in a person's mind." James defended the former view, and consequently developed a form of neutral monism in which the phenomenal world can be regarded as being either "mental" or "physical" depending on one's *interest* in it. If one is interested in how the appearance of the perceived world depends on perceptual processing, one can think of it as mental (as a psychological effect of that processing). If one is interested in how some aspect of the perceived world relates to other aspects of that world (e.g., via causal laws), one can think of it as physical. Lehar, by contrast, defends "biological naturalism" (a form of "physicalism") – the view that the experienced world is literally in the brain.