Color, Consciousness, MUI

Don Hoffman, Cognitive Science, UC Irvine 92697

Locke's Question

Is it possible that "the idea that a violet produced in one man's mind by his eyes were the same that a marigold produced in another man's, and vice versa"?

Essay Concerning Human Understanding, 1690

Abstract

Locke's question is still debated because of its implications for functionalist theories of conscious experience. Here we prove that the answer to Locke's question is yes: Jack's color experiences could differ from Jill's, yet no experiment would reveal this difference. This entails that conscious experiences are not identical with functional relations. We propose a new theory.

Scrambling Theorem

Scrambling Theorem. Denote by X the color experiences of one person and by Y the color experiences of a second. Let the following describe the functional relations among color experiences of the first: (1) $f\colon X_1 \times \ldots \times X_n \to W$, where $X_1, \ldots, X_n$ are copies of X, n is a positive integer and W is any set, (2) $X' \subseteq 2^X$, and (3) $f'\colon X'_1 \times \ldots \times X'_N \to Z$, where $X'_1, \ldots, X'_n$ are copies of X', m is a positive integer and Z is any set. Let $b\colon X \to Y$ by any bijection, i.e., any scrambling, between the two sets of color experiences. Then there exist (1) $g\colon Y_1 \times \ldots \times Y_n \to W$, where $Y_1, \ldots, Y_n$ are copies of Y, (2) $Y' \subseteq 2^Y$, and (3) $g'\colon Y'_1 \times \ldots \times Y'_m \to Z$, where $Y'_1, \ldots, Y'_m$ are copies of Y', such that (1) $f = gb$, (2) $X' = b^{-1}(Y')$ and (3) $f' = g'b$.

Proof. Let $g = fb^{-1}$; $Y' = b(X')$; $g' = f'b^{-1}$.■

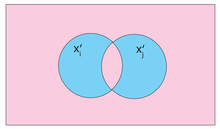
Meaning of Theorem

Suppose we show Jack a display, and he has color experiences $x_1, \ldots, x_n$. If we show Jill the same display, then she will have color experiences $b(x_1), \ldots, b(x_n)$, since b describes how Jill's color experiences are scrambled relative to Jack's. Now Jack's performance in color tasks is determined by the functional relations f, X', f', and Jill's by g, Y', g'. The Scrambling Theorem says it is always possible to choose g, Y', g' so that Jill responds identically to Jack in all color tasks, despite their different color experiences. Therefore their conscious experiences cannot be identical to the functional relations.

Example of Theorem

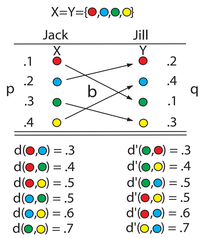
Let f' be a metric that gives distances, d, between color experiences. Suppose that d is induced by symmetric difference from a probability measure, p, on (X, X'), where X' is a σ-algebra on X:

$$d(x'_i, x'_j) = p(x'_i \triangle x'_j) = p(x'_i \cup x'_j - x'_i \cap x'_j), \quad x'_i, x'_j \in X'.$$

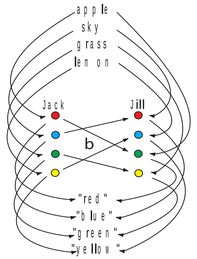


For any scrambling $b\colon X \to Y$, let the probability, q, on Y be the distribution of b, i.e., for any measurable set A of Y, $q(A) = p(b^{-1}(A))$. Then taking g' to be the metric d' on Y induced by q satisfies the Theorem.

Scrambling Colors, But Not Functions



Even though Jill sees red what Jack sees blue and Jack sees red what Jill sees green, Jack and Jill still both call grass green and the sky blue.



Reductive Functionalism

Reductive functionalism claims that the type identity conditions for mental states refer only to relations between inputs, outputs and each other (e.g., Block & Fodor, 1972). That is, it claims that each mental state is numerically identical to some functional relation. The Scrambling Theorem proves that reductive functionalism if false: Color experiences, a type of mental state, are not identical to any functional relations.

Objections and Replies

1. Conscious experiences cannot be represented mathematically.
   - If so, then there is no science of consciousness. But it is premature to claim this. Give science a chance.
2. X could represent unconscious states of robots, so the Theorem is not about *conscious* experiences.
   - Using integers to count apples does not preclude using integers to count oranges.
3. The Theorem is not adequately constrained by empirical facts about color perception.
   - The Theorem establishes the *logical* possibility of inverted spectra, not their *empirical* existence.
4. Since I can't imagine a red that looks more similar to a green than another red, it must not be possible.
   - Proof establishes logical possibility and, ipso facto, imaginability. Proof trumps weak imagination.
5. The Scrambling Theorem is mathematically trivial.
   - Yes. The real work is the proper math formulation. It is a bonus that then the answer is obvious.

Wittgenstein's Question

"Consider this case: someone says "I can't understand it, I see everything red blue today and vice versa." We answer "it must look queer!" He says it does and, e.g., goes on to say how cold the glowing coal looks and how warm the clear (blue) sky. I think we should under these or similar circumstances be inclined to say that he saw red what we saw blue. And again we should say that we know that he means by the words 'blue' and 'red' what we do as he has always used them as we do."

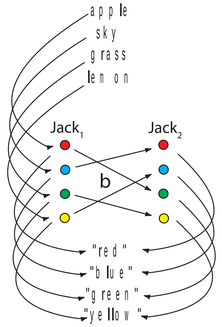
Philosophical Review, 3, 284, 1968.

Nonreductive Functionalism

Nonreductive functionalism claims that conscious experiences are determined by, but not identical to, the functional organization of a physical system. Such functionalists argue that Wittgenstein's color inversion could not happen without a change in functional organization (e.g., Chalmers 1995; see also Tye, 1995). For if Jack's color experiences changed but his functional organization did not, then his beliefs about his color experiences would not change. He would believe, e.g., that he experienced red when he in fact experienced blue. This, they conclude, is a reductio ad absurdum. But the reductio fails, for if one grants for sake of argument that experience can change without changes in functional organization, then surely one must grant that beliefs can so change as well.

Noncommuting Diagram

If Jack's color experiences suddenly change, he will report the changes. He will say, e.g., that the sky is red and the grass yellow. This contradicts the prediction of nonreductive functionalism.



The MUI Theory of Consciousness

The MUI theory of consciousness states that the conscious experiences of an agent are a multi-modal user interface (MUI) between that agent and an objective world, objective in the sense that its existence does not depend on that agent. MUI theory makes no claim about the nature of that objective world. In particular, it does not claim that the format or contents of the MUI in any way resemble or approximate that objective world. Conscious experiences of space and time, rocks and trees, colors and smells, might not resemble or approximate anything in that objective world.

A Virtual-Reality Illustration of MUI

Imagine that you and some friends play virtual volleyball at a VR arcade. You don a helmet and body suit, and find yourself at the beach on a volleyball court. You serve the volleyball and a friend across the net hits it back. You consciously experience a volleyball, the net, the sand, your body, your friends, and palm trees in the background. These are all contents of your MUI. The objective world, in this example, is a computer with megabytes of software running a VR program. The volleyball of your MUI does not resemble or approximate the diodes, resistors, or software of the computer. Nor do the net, sand, and your body. But these elements of your MUI inform you how to act so as to alter registers in the computer to win the game (Hoffman, 1998).

MUI, Bayes and Evolution

Many researchers model perceptual experience as the result of Bayesian estimation. They assume the objective world contains physical objects with determinate properties. Perceptual experiences approximate, or reconstruct, these properties. This physicalist model of the objective world is one of countless models allowed by MUI theory. On evolutionary grounds, however, the physicalist model is unlikely. Evolutionary pressures dictate that our MUI interface, this world of our daily experience, should itself be a radical simplification, selected not for the exhaustive depiction of the objective world but for the mutable pragmatics of survival. Humans, like roaches, are a species in a niche. We do not expect the perceptions of roaches to be insights into the true nature of the objective world; it would be a miracle if the perceptions of humans were such insights.

Epiphysicalism and the Brain

On your computer screen is an icon for a text file. The icon is, say, red and rectangular. The shape, color, and position of the icon do not, of course, resemble anything about the text file, but give you access to that file. You drag the text-file icon to the recycle icon, and your file is deleted. Did the motion of the icon on your screen *cause* the file to be deleted? No. Instead the icon informed your actions, guiding you to move and click a mouse, and your actions triggered the causal sequence that deleted the file. Likewise, the tables, chairs, rocks and other objects of your everyday MUI appear to have causal powers, but they do not. In particular, the brain has no causal powers. It is another icon of your MUI. Indeed, just as there is no cube in the figure above until someone looks and constructs a cube, so also you have no brain until someone observes it and, thereby, constructs it. Of course, if someone destroys your primary visual cortex, you go blind, just as when you drag the file icon to the recycle bin you delete the file. And if someone stimulates your somatosensory cortex, you feel a sensation in your body, just as when you click on the file icon you access the file. Nevertheless, the brain does not cause these sensations, any more than the text-file icon causes the text file. Consciousness does not supervene, logically or naturally, on the brain.

Conscious Realism

Conscious realism states that the objective world, the world behind the MUI, consists only of conscious agents. Conscious realism can be made mathematically precise, and used to derive quantum field theory (Bennett et al., 1989). It thus differs from Kant's idealism, which claims that things-in-themselves cannot be mathematically modeled. It also differs from panpsychism.

The Mind-Body Problem

Physicalism has failed to produce a scientific theory describing how brain activity could cause, or be, conscious experience. It faces a mysterious gap between (1) the experience of red and (2) ions flowing through holes in neural membranes. MUI Theory and Conscious Realism together transform the mind-body mystery into a tractable scientific problem. Consciousness is primary, and can be mathematically modeled (Bennett et al., 1989). Tables, chairs and other "physical" objects are elements of a MUI, constructed by a conscious mind according to rules now being discovered in the cognitive sciences (e.g., Hoffman, 1998). Electrons and atoms are also elements of a MUI, constructed by a conscious agent—a claim compatible with the standard Copenhagen interpretation of quantum theory.

More Objections and Replies

1. If that speeding train is just an icon of your MUI, why don't you step in front of it?
   - For the same reason I don't carelessly drag a file icon to the recycle bin. Not taking icons literally differs from not taking them seriously. Icons are there to inform our behavior.
2. How can this table be just a construct in my MUI, since you see the same table?
   - We don't see the same table, any more than we see the same volleyball in the VR volleyball game. You construct your ball and your table, I construct mine. We assume, though we cannot prove, that our constructs are relevantly similar.
3. If this table is just a construct in my MUI, why can't I put my fist through it? Shouldn't my constructs do what I want?
   - The constructs in your MUI are constrained by your rules of construction and your interactions with the objective world that triggers your construction. They are not subject to your every whim.
4. Physicalists don't say that the objective world resembles our experiences. They claim instead, e.g., that this table is mostly empty space, with protons, neutrons and electrons whizzing around. That is the true objective reality.
   - This claim makes the same mistake as someone who pulls out a magnifying glass, looks at a computer screen, and concludes that, although the file icon is just an icon, the true objective reality is the set of pixels now visible. Those pixels are not the computer to which the icon is an interface. The pixels are part of the windows interface; similarly, protons, neutrons and electrons are part of your MUI.

References and Note

1. Bennett, B.M., Hoffman, D.D., Prakash, C. (1989). *Observer Mechanics: A Formal Theory of Perception.* New York: Academic Press.
2. Chalmers, D.J. (1995) *The Conscious Mind.* Oxford Press.
3. Hoffman, D.D. (1998). *Visual Intelligence: How We Create What We See.* New York: W.W. Norton.
4. Locke, J. (1690). *An Essay Concerning Human Understanding.*
5. Tye, M. (1995). *Ten Problems of Consciousness.* Cambridge, MA: MIT Press.
6. Wittgenstein, L. (1968). Notes for Lectures on 'Private Experience' and 'Sense Data'. (Rush Rhees, ed.), *Philosophical Review, 3,* (July LXXVII): 284.

Poster presented at the 17th Annual Convention of the American Psychological Society, Los Angeles, California, May 26–29, 2005, and at the 9th Annual Meeting of the Association for the Scientific Study of Consciousnes, Caltech, Pasadena, California, June 24–27, 2005.