Available online at www.sciencedirect.com
SCIENCE DIRECT®
Consciousness and Cognition 15 (2006) 51–53
Consciousness and Cognition
www.elsevier.com/locate/concog

Reply

The Scrambling Theorem unscrambled: A response to commentaries ☆

Donald D. Hoffman *

Department of Cognitive Science, University of California, Irvine, CA 92697, USA

If we assume that conscious experiences exist and can be represented by sets, then conscious experiences are not identical to functional relations, no matter what mathematical structure one chooses to model such relations. This is the content of the Scrambling Theorem. The proof is trivial: any mathematical structure on one set of conscious experiences can be transported, intact, to any other scrambled set of the same conscious experiences. We thus conclude that reductive functionalism is provably false.

The imaginative commentary by Hardin and Hardin distinguishes two versions of reductive functionalism. The "black box" version restricts functional descriptions to behaviors of the whole organism, ignoring inner states. The second version, call it "glass box," extends functional descriptions to include inner states. The question raised by Hardin and Hardin is whether the Scrambling Theorem disproves black box, but not glass box, functionalism.

The short answer is that the Scrambling Theorem disproves both. Glass box functionalism, when finally made precise, must posit mathematical structures that model the function of inner states. All such structures are subject to the Scrambling Theorem.

This short answer holds no matter how finely one individuates inner states. They might be neural systems, individual neurons or, at the finest level, every subatomic particle in the brain. One can debate whether the subatomic level is too fine; some argue that the functional descriptions relevant to consciousness are at a coarser level, say the level of individual neurons or neural systems. The outcome of this argument is, for present purposes, irrelevant. The point here is that even if the relevant functional descriptions extend to the quantum state of every subatomic particle of the brain, the Scrambling Theorem, and its disproof of reductive functionalism, still holds. Even if Jack and Jill are functionally equivalent down to their subatomic particles, the conscious experiences of Jack could, in principle, be completely scrambled relative to those of Jill.

The only version of reductive functionalism that can escape the purview of the Scrambling Theorem is one that refuses to specify the relevant functional relations with mathematical precision. But such refusal precludes doing science.

Mausfeld and Andres nicely summarize the meaning of the Scrambling Theorem in their quote and discussion of Weyl: due to the "insurmountable boundary" of isomorphism, science cannot capture the essence of its objects. This holds for electrons no less than for conscious experiences. The wave equation of the free electron is not an electron and does not create an electron; no physicist thinks otherwise. Functional descriptions of the

☆ Reply to commentaries by Hardin, C. L., Hardin, W. J. (2006). A tale of Hoffman. *Consciousness and Cognition, 15*, 46–47 and Byrne, A., Hilbert, D., Hoffman's "proof" of the possibility of spectrum inversion. *Consciousness and Cognition, 15*, 48–50.

* Fax: +1 949 824 2307.
E-mail address: ddhoff@uci.edu.

1053-8100/$ - see front matter
doi:10.1016/j.concog.2006.02.001

brain are not conscious experiences and do not create conscious experiences; but many cognitive scientists think otherwise. The insurmountable boundary of isomorphism is understood in physics and must be incorporated into the philosophy of mind.

Mausfeld and Andres then raise an important question: do the Scrambling Theorem and reductive functionalism use the same notion of functional relations? It might appear that they do not. In the Scrambling Theorem the domain of functional relations is a set of conscious experiences, whereas in reductive functionalism the domain is a set of inputs, outputs, and mental states. These are prima facie different sets, so perhaps they entail different notions of functional relations.

The difference is harmless. Conscious experiences are mental states. So the domain of functional relations as used by the Scrambling Theorem is a proper subset of the domain of functional relations as used by reductive functionalism. This is illustrated in Fig. 3 of the original paper. Part (A) shows the functional relations (and scrambling) of the mental states alone. Part (B) shows the embedding of these functional relations within the wider context of inputs and outputs; this is done, and can always be done, without in any way altering the functional relations, or restricting the possible scramblings, of the mental states. The Scrambling Theorem, by focusing on mental states, incurs no loss of generality.

Why is this the case? The functional relations used by the Scrambling Theorem in Fig. 3 are modeled by a distance metric on a set of possible color experiences. This metric specifies the network of relationships of relative similarity and discriminability among the color experiences. This network of relationships determines, for each pair of colors presented to an observer, the possible judgments of similarity and discriminability that can be made by that observer. That is, the proper distance metric on color experiences correctly determines all possible outputs to all possible inputs. Thus it determines the functional relationships over the entire domain considered by reductive functionalists.

The reductive functionalist claims that the network of relationships just described is all there is to conscious experiences. As Clark (1983) puts it, "To specify the content of the sensory state is just to specify its place in a network of relationships of relative similarity and discriminability. To query its qualitative content in any more absolute way is as meaningless as querying location after coordinates have been given." Clark goes on to assert, "Once two points are shown to have the same respective location in the color solids of two individuals, there is no clear sense to the further suggestion that their associated sensations are somehow qualitatively dissimilar ... There may be such a sense, but it has never clearly been articulated." There is a clear sense, and the Scrambling Theorem articulates it.

Byrne and Hilbert have posted a commentary entitled 'Hoffman's "Proof" of the possibility of spectrum inversion,' at this URL: http://web.mit.edu/abyrne/www/hoffmansproof.pdf. In their commentary, they try to construct a simple counterexample to the Scrambling Theorem. Recall that the Scrambling Theorem models functional relations between color experiences X by a function, $f: X \times \cdots \times X \rightarrow W$. For instance, f might be a distance metric and W, the nonnegative reals. Then $f(x_1, x_2)$ is the distance between colors x_1 and x_2, where smaller distances indicate greater similarity. In trying to construct a counterexample, Byrne and Hilbert take W, instead of f, to model functional relations. Doing so, they discover, leads to bizarre conclusions. I reply as follows:

Addition and 2 are distinct. The number 2 is in the range of addition, as in $1 + 1 = 2$. But elements of a function's range are not the function: 2 is not addition. A function $f: X \times \cdots \times X \rightarrow W$ is not its range W.

The Scrambling Theorem models functional relations by a function $f: X \times \cdots \times X \rightarrow W$. The proof of the Scrambling Theorem then follows trivially. Byrne and Hilbert conflate a function and its range—by using W, not f, to model functional relations. They discover that bizarre consequences follow. Indeed, if one conflates 2 and addition, strange consequences follow. If one conflates a distance metric and a real number, strange consequences follow. This says nothing, of course, about the Scrambling Theorem, only about the dangers of conflation. Conflation of function and range misleads Byrne and Hilbert to conclude that the Scrambling Theorem conflates epistemic and metaphysical possibility because, in part, it "makes no use of the special features of *functional* states." Well, here is one special feature of functional states that the Scrambling Theorem uses: they have the complex structure of functions; they are not unstructured points in the range of a function. If one ignores this crucial feature of functional states, as Byrne and Hilbert do, then of course problems arise.

On a separate issue, Byrne and Hilbert question the notation $b(X')$. The question is surprising. This is standard notation, e.g., for defining a continuous function: if (X, X') and (Y, Y') are topological spaces, then a function $b : X \rightarrow Y$ is continuous if, for all open sets $A \varepsilon Y'$, $b^{-1}(A) \varepsilon X'$. Thus, in this standard notation, $b(X') = \{b(C) \mid C \varepsilon X'\}$, where $b(C) = \{b(x) \mid x \varepsilon C\}$. Byrne and Hilbert propose that the proper definition of $b(X')$ is $\{b(x) \mid x \varepsilon X'\}$. This is an elementary mistake: x does not index X'.

Byrne and Hilbert close by quoting Kripke (1976, p. 419) saying, ''There is no mathematical substitute for philosophy.'' Yes. But one virtue of mathematics in philosophy is precision. A branch of philosophy made sufficiently precise might, we can hope, beget a science.

References

Clarke, A. (1983). Spectrum inversion and the color solid. *The Southern Journal of Philosophy, 22*, 431–443.

Kripke, S. (1976). ''Is there a problem about substitutional quantification?'. In G. Evans & J. McDowell (Eds.), *Truth and meaning*. Oxford: Oxford University Press.