

On an Argument for Functional Invariance

Michael Pelczar

Received: 27 August 2007 / Accepted: 30 June 2008 / Published online: 15 July 2008
© Springer Science+Business Media B.V. 2008

Abstract The principle of functional invariance states that it is a natural law that conscious beings with the same functional organization have the same quality of conscious experience. A group of arguments in support of this principle are rejected, on the grounds that they establish at most only the weaker intra-subjective principle that any two stages in the life of a single conscious being that duplicate one another in terms of functional organization also duplicate one another in terms of quality of phenomenal experience.

Keywords Consciousness · Functionalism · Dualism · Naturalism

Naturalistic dualism is the view that although conscious experience is not an entirely physical phenomenon—or at any rate, not a physical phenomenon reducible to already recognized physical phenomena—it is nonetheless a *natural* phenomenon, in the sense that it can be brought into the scope of a systematic and broadly scientific view of the world. One can argue for dualism without arguing for naturalistic dualism, and the most influential arguments for dualism—the modal arguments and the knowledge argument—conclude only that conscious experience is not an entirely physical phenomenon, leaving open the question of exactly how, if at all, conscious experience fits into the natural scheme of things.¹

One way to get from dualism to naturalistic dualism is by showing that conscious experience relates to the physical world, or certain aspects of the physical world, in a law like way. If such law like behavior can be demonstrated, we may reckon

¹ For the modal arguments, see Campbell (1970, 100), Kirk (1974), and Chalmers (1996, 93–140); for the knowledge argument, see Broad (1925), Nagel (1970), and Jackson (1982).

conscious experience a natural phenomenon, on the grounds that whatever stands in a law like relationship to a natural phenomenon should itself count as a natural phenomenon.

One way to establish the existence of a law like connection between conscious experience and the physical world is by arguing for a *principle of functional invariance*. According to this, it is a law of nature that conscious beings that resemble one another in terms of their functional organization also resemble one another in terms of the phenomenal quality of their conscious experience. A more precise version of the principle would be more specific about the relevant level of functional organization (is it the cellular level, or the molecular, or even deeper than that?) and more explicit about how much phenomenal resemblance a given degree of functional resemblance is supposed to guarantee. These are details that need not detain us here, although for the principle to be of much interest, the degree of resemblance it entails between complete functional duplicates must be at least very high, if not total. For the purposes of what follows, we may deal with the following relatively weak version of the principle of functional invariance (henceforth “PFI”):

It is a natural law that conscious beings that duplicate one another in terms of their functional organization also duplicate one another in terms of the phenomenal qualities of their conscious experiences.

There are two ways to argue for the PFI. One is by arguing that it plays an indispensable role in our best explanation of the correlations between conscious experience and functional organization that we observe in our own respective cases.² This is not the argument I am interested in here. My interest lies with a more colorful and direct argument for the PFI. This argument runs as follows:³

- (1) If the PFI is false, then it is naturally possible for the phenomenal quality of someone’s conscious experience to change as a result of his undergoing physical changes that preserve his functional organization.
- (2) If someone undergoes physical changes that preserve his functional organization, then he undergoes changes that preserve his judgements about the phenomenal quality of his conscious experience.
- (3) So, if the PFI is false, then it is naturally possible for the phenomenal quality of someone’s conscious experience to change, without any consequent change in his judgements about the phenomenal quality of his experience. (1, 2)
- (4) But such a thing is not naturally possible.
- (5) So, the PFI is not false. (3, 4)

Call this the *aberrant qualia argument*. Its first premise states, in effect, that if it is naturally impossible for a conscious being to undergo a change in the phenomenal quality of his conscious experience without undergoing some change in his functional organization, then the PFI is true. Premise (2) states that constancy of

² See Chalmers (1996, 243–46). This way of arguing for the principle is not without dangers for a dualist, however, who must find a way for our possession of certain functional features to explain our possession of corresponding phenomenal features, without the latter simply reducing to the former. Whether this is possible is a large question beyond the scope of this paper.

³ See Chalmers (1996, 247–75).

functional organization entails constancy of judgements. Thus, in order for one's phenomenal experience to change while one's functional organization remained the same, one would have to undergo a change in the phenomenal quality of one's experience, and yet make no judgement to the effect that this was going on. For example, one's visual experience might switch back and forth from normal to photographic negative mode, without one's judging that this was happening; or one's overall sensory experience might gradually fade in intensity, without one's judging that this was so. Since, according to (4), none of these scenarios is naturally possible, we must accept the PFI.

Much of the debate over the aberrant qualia argument focuses on whether the scenarios just described really are naturally impossible. For example, one can try to argue that the described phenomenal aberrations are accompanied by corresponding shifts of conscious perspective, so that the judgements of phenomenal constancy are natural and correct (albeit brief).⁴

But there is a more basic problem with the argument. This is that premise (1) is false, or at least far from obviously true. Let us grant that it is naturally impossible for a being to undergo a change in the phenomenal quality of his experience without undergoing a change in his functional organization. The most that this establishes is an intrasubjective version of the PFI (henceforth "i-PFI"):

It is a natural law that any two stages in the life of a single conscious being that duplicate one another in terms of functional organization also duplicate one another in terms of phenomenal experience.

This principle states a law like correlation between a given conscious being's functional organization and phenomenology over time, but it says nothing about how *different* conscious beings compare to one another phenomenally. In particular, the i-PFI does not imply that if two conscious beings duplicate one another in all functional respects, they duplicate one another in terms of the phenomenal quality of their conscious experience. It implies only that *if* you have two conscious beings who, at some time, have the same functional organization and the same quality of experience, then those two beings duplicate one another phenomenally at *all* times that they duplicate one another at the level of functional organization.

We can drive the point home by reference to the old problem of the inverted spectrum. If the PFI is true, then natural law prevents conscious functional duplicates from having inverted phenomenal spectra: if two conscious beings duplicate one another functionally, then, according to the PFI, if one of them has bluish experience when looking at the sky, the other has the same kind of bluish experience when looking at the same sky. The i-PFI does not have this implication. For all that the i-PFI says, two conscious functional duplicates looking at the same sky may have radically different qualities of phenomenal experience—one might have (what I call) bluish experience, and the other (what I call) yellowish experience. *If* two conscious functional duplicates have the same quality of

⁴ See Greenberg (1998) and van Heuveln et al. (1998). These arguments suffer from a plausibility problem, inasmuch as it does seem possible to survive sudden and unusual changes in the quality of one's conscious experience.

experience while looking at the same sky, then they will continue to have the same quality of experience for as long as they keep looking at it (remaining both in a constant functional state). This much the i-PFI guarantees. But it does not guarantee that they have the same quality of experience to begin with, as the principle of functional invariance tells us they naturally must.

An opponent of the PFI need not hold that *every* functional architecture (of a conscious being) is capable of instantiating *every* combination of phenomenal properties. If the fading-qualia thought experiment is supposed to involve some phenomenal properties changing while others stay the same, like the pale yellow squares on washed-out table cloth getting paler and paler with every wash, until all that is left is a pure white expanse, then such “faded” experience may well be impossible in a functional duplicate of me. But this would not rule out a functionality-preserving fading that does not affect how *many* phenomenal properties are instantiated, or where in phenomenal space they are instantiated, but only *which* phenomenal properties are instantiated in the relevant pattern. There may be many different but functionally equivalent distributions of phenomenal properties compatible with a given scheme for distributing phenomenal properties across phenomenal space, like different ways of filling in a page of a coloring book.

If someone has a “faded” version of the experience I have, but is just like me functionally, he makes the same judgements about the phenomenal qualities of his experiences as I make about mine. I judge my experience to be rich in phenomenal detail. So, if I have a “faded” duplicate, he judges his experience to be as rich and varied as I judge mine to be. That, Chalmers contends, is (naturally) impossible, which he takes as a *reductio* of the idea that a phenomenally faded functional duplicate of a given conscious creature is naturally possible.⁵

Whether this argument is sound depends on how we understand “phenomenally faded” experience. If the fading is analogous to the table cloth example, we may grant that it must have functional ramifications; but that should come as no surprise. (No more would anyone suggest that increasing the volume of just one of the sounds you hear at a given time would have no functional repercussions—certainly it would, as it drowned out all the other sounds.) If the fading *preserves* the relative intensity (or saturation, or volume, or whatever) of different regions of the phenomenal field, however, there is no reason to think it cannot occur in a functional duplicate of me, who, as a functional duplicate of me, judges his experience to be quite rich—as rich as I judge my experience to be.

After all, if my faded twin’s experience is fainter than mine (in the latter, contrast-preserving sense) by a factor of two, perhaps my own experience is fainter by a factor of two than some other creature’s. This possibility does not cast doubt on any of the judgements I make about the richness of detail of my conscious experience. It may be that many of the phenomenal judgements I make differ in *content* from the corresponding judgements my faded (or super-saturated) twin makes. Some are demonstrative judgements of the form *I am having an experience with the salient phenomenal property of THIS sensation*, and what I demonstrate with my ‘THIS’ has different phenomenal properties from what my twin

⁵ Chalmers (1996, 256).

demonstrates with his. This, though, is just a difference in the basic phenomenal properties of our experiences—differences of color scheme, for example, such as the original argument for the PFI tried to capitalize on. The judgements my twin makes about the richness of his experience may differ in content from mine, but this may be only in ways that are parasitic on our more basic phenomenal differences, and that leave both our judgements true. Again, we need not deny that functional organization plays *some* constraining role on what sorts of experiences my twin and I can have. But the existence of such constraints does not mean that one's functional organization naturally necessitates every (or, nearly every) phenomenal quality of one's experience, as the PFI maintains.

The principle of functional invariance is plausible, and may be true. But it is no argument for it that any change in the quality of a single individual's phenomenal experience must be accompanied by a corresponding change in his functional organization. If I had a genuine doubt about whether other people saw the world the way I do, as opposed to in a phenomenally inverted way, my doubt would not be allayed by an argument to the effect that I am bound to see the world the same way I do now, as long as my functional organization does not undergo any relevant change. For the same reason, no one who doubts the principle of functional invariance should accept that principle on the strength of the aberrant qualia argument.

References

- Broad, C. D. (1925). *The mind and its place in nature*. London: Kegan Paul, Trench, Trubner & Co.
- Campbell, K. (1970). *Body and mind*. London: Macmillan.
- Chalmers, D. J. (1996). *The conscious mind: In search of a fundamental theory*. Oxford: Oxford University Press.
- Greenberg, W. (1998). On Chalmers' principle of organizational invariance and his "Dancing Qualia" and "Fading Qualia" thought experiments. *Journal of Consciousness Studies*, 5(1), 53–58.
- Jackson, F. (1982). Epiphenomenal qualia. *The Philosophical Quarterly*, 32(127), 23–36.
- Kirk, R. (1974). Sentience and behavior. *Mind*, 83(329), 43–60.
- Nagel, T. (1970). *The possibility of altruism*. Princeton: Princeton University Press.
- van Heuveln, B., Dietrich, E., & Oshima, M. (1998). Let's dance! The equivocation in Chalmers' dancing qualia argument. *Minds and Machines*, 8(2), 237–49.