Response to Phil Gerrans

Joëlle Proust

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Consciousness and Cognition 12 (2003) 513–514
Consciousness and Cognition
www.elsevier.com/locate/concog

Response to Phil Gerrans☆

Phil Gerrans's comments raise two interesting points. One has to do with the value of sceptical arguments, the other with what a self can know of himself.

1. A proposal aiming at offering a "reduplication-proof" approach of the self is mainly inspired by sceptical worries, Gerrans writes; providing a dynamic criterion does not help however to allay these worries, because dynamic memory can also be replicated. Before coming to this point, a prior question needs to be raised: why is reduplication a problem? In my opinion, the problem is not simply of addressing a general sceptical worry—is one justified in believing that one knows what a self is, if one does not know how a self is at all identifiable? The problem is rather to take a stance on the reality of individual selves, independently of what we can know about them. This issue is addressed in the literature on the basis of a dilemma. Either selves are defined as structured bundles of representations, and can "essentially" be copied (as a matter of definition). Or selves have an additional, non-representational, property securing numerical identity on a conceptual basis. Influential naturalistic views tend to chose the first horn of the dilemma (such as Dennett, 1991), while the second horn is generally selected in non-reductionist views on the self (Campbell, 1994). I suggest that the alternative is not forced on us. One can be a reductionist about the self, and explain what a person is in more fundamental terms (taking the first horn of the dilemma), while also recognizing that the self is a numerically unique entity.

My proposal is that the additional property on which numerical identity depends consists in a disposition to reflexive affection that pertains to a class of minds. What reflexive affection achieves is revising the workings of *this* mind on the basis of prior learning. The indexical nature of the targetted system expresses the fact that revision operates on a set of token states whose properties have to be monitored and adjusted. We can compare the twin-earth fates of two replicae in the Leibnizian anti-Locke story and in the present view. Although prior learning is indeed subject to duplication, insofar as dynamic memory, functionally conceived, allows type-copying—as Gerrans correctly observes-, metacognition targets on token-mental/cerebral operations in a way that is not purely representational. Metacognition operates both on content (mental states) and on vehicle (cerebral states). If you duplicate a metacognitive organism, the replica will immediately get busy adjusting her *own* states using her own connectivity. If several replicae are created with a same learning background and a functionally similar mental organization, it will leave each replica with her own numerical identity. In other words, one might say that reduplication of such

☆ Reply to Commentary on Proust, J. (2003). Thinking of oneself as the same. *Consciousness and Cognition*, *12*, 495–509. This article is part of a special issue of this journal on Self and Action.

1053-8100/$ - see front matter
doi:10.1016/j.concog.2003.08.001

systems cannot operate in a Cartesian way. Each clone becomes a distinct candidate for self as a result of the very structure of metacognition.

2. A naturalist philosopher obviously must confront the problem of what a person can know of herself. Gerrans comments on the fact that a forward model "has no access to its own states." What does this expression mean exactly? A disturbance can only be appreciated by a forward control system if it can map its own commands to a particular environment, and retrieve information from the reafferences. Such perception monitoring cannot be achieved if the control system has no informational access to various other mental and bodily states (e.g., sense of knowing, head, trunk, and limb position). Now of course, the self is not to be confused with a controlling agent. The self does not exert control, it is constituted by a specific level of the control capacity—the social level, or the long-term evaluative level. Expressions such as "the self is aware," or "the self is in charge..." suggests that the self does something, acts in order to eliminate or coordinate. In fact, these expressions invite the homuncular fallacy.

But maybe Gerrans is rather attracting our attention to the fact that, given the architecture of a control system, the self-control loop cannot locate itself in a contrast space of other selves, because all it monitors is the contrast space of plans. How can we get over this limitation? In the present paper, it is suggested that, if access to oneself is understood through domain specific monitoring, as in any other control level, no higher level is required: accessing the self is simply getting reafferences from one's ability to select, execute, and revise plans of action. Explicit self-narratives do not need to be imagined for selves to come to existence. Nor is it necessary for accessing oneself that one models a group of selves with alternative evaluative and dynamic properties. Obviously, watching others—cooperating with them, obeying them, rebelling against them—allows social experience to build up. Metarepresentational capacities may develop, helping the individual make explicit the reflexivity that is implicit in the control structure.[1] But there is no need to have an additional control loop inside our mind to develop self-awareness. Such a loop does develop of course, but outside the individual's brain; specialized institutions (sets of representations and practices) have the function of controlling and monitoring the control systems in which selves consist—harnessing them to each other in a normative hierarchy, distributing their respective goals and values, and constraining revisions accordingly. The reafferences consist in verbal and emotional reactions and social rewards as structured by the norms in a specific culture with specific institutions. As Gerrans observes, reward is crucial at each control level.

References

Campbell, J. (1994). *Past, space and self*. Cambridge, MA: MIT Press.
Dennett, D. (1991). *Consciousness explained*. Boston: Little, Brown and Co.
Proust, J. (to appear). Does metacognition inecessarily involve metarepresentation? *Behavioral and Brain Sciences*.

Joëlle Proust
Institut Jean-Nicod, CNRS
Paris, France
E-mail address: jproust@ehess.fr

[1] On this issue, see Proust (to appear).