Synthese (2006) 151: 403–409
DOI 10.1007/s11229-006-9032-1

ORIGINAL PAPER

What's behind a smile? the return of mechanism: Reply to Schaffner

Luc Faucher

Published online: 17 August 2006

In his 1993 book, *Discovery and Explanation in Biology and Medicine*, Schaffner compared the merits of two models of intertheoric reduction: the Causal Mechanical (CM) model and the General Reduction Replacement (GRR) model. The GRR model, a modified version of Nagel's model of intertheoric reduction,[1] was seen as the "appropriate framework in terms of which deeper logical questions", like the establishment of synthetic identities, "could be pursued" (1993, p 498); while the CM model, according to Schaffner, could "not be the whole story" by any means because it was occluding "important deep structural issues such as specific points at which identities need to be formulated between levels" (1994, p 293). In the Paris' conference on Reduction and Emergence, Schaffner made the following comment concerning his GGR model:

> "The GRR model, itself an extension of the Nagel model, is more suited to fully axiomatized sciences, *of which there do not appear to be any*" (emphasis added).

It is not clear how to understand this statement. Should we conclude that Schaffner has significantly moved away from the GRR model, i.e. should we interpret him as now viewing it as a model of situations that never occur? In a way, "yes". But this has always been the case as Schaffner has presented GRR as forcing actual ongoing reduction into a "rational reconstructive" mode (1994, p 336):

> "The GRR is [. . .] useful in providing us with a kind of systematic summary and regulatory ideal, but *it should not in general be confused with the process of establishing reductions in the ongoing elaboration of connections that typically unite the reducing and the reduced theory*" (1993, p 497; my emphasis).

[1] The modification is intended to avoid the derivability problem faced by Nagel's theory, that is, the fact that most cases of actual reduction would imply deriving a false theory from a true one as in the case where the lower-level theory corrects the higher-level theory

L. Faucher (✉)
Départment de philosophie, Université du Québec á Montréal, Case postable 8888,
Succursale Centre-ville, Montréal, Que., Canada H3C 3P8
e-mail: faucher.luc@uqam.ca

A more plausible interpretation of Schaffner's statement is that his optimism about the prospects of systematic reduction in the biomedical sciences has dwindled over the past few years. Reduction, when it occurs at all (see for instance his doubts about the reduction of behavior to genes in his 1998), is local and patchy or as he puts it, "creeping". For this kind of reduction, the CM model is more adequate. "Sweeping" reductions are a "myth" so there is no need to lose time constructing a model of reductions for a never-to-be completed science.

According to the CM model, reduction occurs when a phenomenon at one level of aggregation is explained by reference to mechanisms described at a lower level. The CM model abandons the idea that reduction occurs between theories and involves synthetic identities between entities and properties at different levels. On the CM model, reduction consists in identifying causal relations between lower-level mechanisms and higher-level processes: mechanisms occurring at the lower-level are causally responsible for the phenomena occurring at the higher-level. If CM seems better at describing the process of establishing reductions, up to recently Schaffner has been reluctant to abandon GRR altogether, as Wimsatt (1976) once suggested him. One reason he invoked at the time to motivate his decision was that proponents of this model have used "mechanism" as an unanalysed term. Times have changed. The neo-mechanists are now providing such an analysis (Machamer, Darden, & Craver 2000; Craver 2002; Glennan 2002) and their position is gaining a certain momentum in philosophy of science as it is seen as providing a more adequate description of how science "in the wild" is practiced (I borrow the phrase from Craver 2002).

What I would like to do in the rest of this paper is to quickly summarize the neo-mechanistic position and to try to apply some of its insights to the understanding of psychopathology, a domain Schaffner has explored lately in several papers (2001, 2002; see also his 1992). In doing so, I hope to show some possible avenues to explore by a mechanistic philosophy as well as some of the limits of this approach for reductionist philosophy. My remarks won't be critical of Schaffner's effort to contribute to a truer description of scientific explanation "in the wild" for instance, through his notion of "field elements" and "preferred model system"), for I think that his account is quite adequate.

As defined by Machamer et al. (2000), "mechanisms are collections of entities and activities organized in the production of regular changes from start or setup conditions to finish or termination conditions" (p 3). Mechanisms exhibit what they call "a productive continuity", that is, a sequence of steps without gap going from the setup conditions to the termination conditions. Together, the entities and activities are organized to produce something, the molar behavior of the mechanism.

A mechanical explanation works by providing a mechanical model or schema of the molar behavior of the mechanism. Such a model includes: "(1) a description of the mechanism's behavior; and (2) a description of the mechanism which accounts for that behavior" (Glennan 2002, p 347), the latter is understood as describing the "guts of the mechanism". An ideal explanation would give us a description of the mechanism without gap in productive continuity. In practice, a complete explanation is rarely available at first, what we have instead is a mechanistic sketch, a partial description of the mechanism, with gaps to be filled by further research.

Craver (ms) distinguishes between two kinds of explanation (in fact three, but for reasons of space, I won't present the third one here): constitutive explanation and etiological explanation. Constitutive explanations are explanations "that proceed by describing internal mechanisms—organized lower (−1) level activities and

entities—by virtue of which some aspect of the complex IO [input–output] function is produced" (Craver ms). Kandel's molecular cartoon explaining the working of memory (the distinction between short and long-term memory for instance) is an example of this type of explanation. Etiological explanations are "are typically offered in response to questions concerning the origins of some item, its path of development, or its historical trajectory" (Craver ms). Etiological explanations describe the working of etiological mechanisms, that is, "the organized entities and activities antecedent to and productive of the phenomenon to be explained" (Craver 2002, p 70). The explanation of the development of certain types of behavior in *Caenorhabditis elegans* as described by Schaffner (1998) is an example of etiological explanation.

Schemata often describe mechanisms organized in multi-level hierarchies, where one higher level activity is realized by a lower level mechanism whose activities are realized by still lower-level mechanisms, "with the important additional restriction that lower-level entities and activities are components of higher-level mechanisms" (Craver & Darden 2001), so that the relation is not simply a "brute" part-whole relationship (like you would have in a relation between the grains of sand and a pile of books the particular books that constitute it).

Finally, explanations are bottoming-out when a further description is seen as unnecessary, unavailable or inappropriate. The decision of bottoming-out is a pragmatic one and we can expect that different scientists belonging to different fields or traditions decide to bottom-up at different level. This is not only due to limitation of knowledge, but also to the kind of questions addressed.

This question of the level of bottoming-out is a good place to make a first remark concerning reduction. If the level at which the inquiry stops is determined by the type of questions one is asking, attempts at reduction are more motivated by preferences concerning the place where to look for an answer (preferences that may be motivated theoretically or more pragmatically) than by ontological motives of the type postulated by the logical positivists (need for reduction or unification of science). And it is not obvious at all that the answers will always require the scientist to look down. For instance, it seems that psychiatry has recently been interested in identifying the biological substrates of mental disorders like schizophrenia, autism and ADHD. But this means something different for different researchers. Some are trying first to identify at a quite abstract level the cognitive systems that are defective in patients with mental disorders. Chris Frith, for instance, explains delusion of control and voices in schizophrenia by invoking a problem with the mechanism in charge of monitoring movement (Frith 1992), and then tries to locate the defective structure underlying that function (like the cerebrellum, Wolpert, Miall, & Kawato 1999). Some are looking for defective genes and try to reconstitute the story from the genes up.[2] Others, not necessarily denying the fact that genes are involved and that cognitive systems might be defective because of those genes, are underscoring instead the role of the environment in the production of the defective phenotype. Panksepp (1998), for instance, has proposed that Attention Deficit Hyperactivity Disorder (ADHD) might be the result of the fact that certain individuals need to play more than others and since this is not always possible in our regimented society, they develop a pathology of attention (the most compelling hypothesis is that ADHD symptoms tend to emerge from some type of slow development or dysfunctions of the frontal lobe, which is in charge of

[2] This strategy has encountered many problems, some of which caused by the interpretation of the results of some methods (like twin studies, see Robert 2000), some other by the fact that the diagnostic of schizophrenia does not seem to capture a natural kind (see Poland 2006; Schaffner 2002).

executive functions related to attention, planning and social sensitivities). If Panksepp is correct and if he can show that play therapy works by increasing attentional behavior (see his 2003) and if we had the means and will to implement that therapy, I bet that environment-oriented research would gain momentum in that domain. Reflecting on a similar case, Schaffner (1994) concludes that the reference to the environment is not an argument against in-principle reductibility, with which I agree. To use a term he introduces in his article in this volume, it is at most an argument for pragmatic emergence, for it is thinkable, and surely desirable, to discover how the interlevel causal interaction—the playing of the infant contributes to a functional frontal lobe—is implemented at the biomolecular level, but I see no reason why this level should be given any ontological precedence. What reduction achieves in this case is render more intelligible the high-level phenomena, not getting rid of them. We still might want to understand all the factors involved in the production of the disorder (why are certain persons more likely to suffer from ADHD). In a case like Panksepp's, we have to take an "*augmentationist*" stance instead of a "reductionist" one: our explanation does not consist in going down one level, but going up one level to include the environment (cultural or social, depending on the case) in which the organism lives, then it can move down again. A very similar case could be made for memory (the conclusion of this case seems more radical though). You might want to go down to the molecular level of implementation to explain the difference between short-term and long-term memory and go to the system level to explain why H.M. has troubles with his anteretrograde memory. But there would be no need to go down to the molecular level to explain why we remember better informations that are presented in a story-like fashion than randomly or why old couples remember more of their common life when asked together and can help each other out than when asked individually.

A second remark can be made using psychopathologies as an example. Progress can take many forms in a mechanistic model. It can happen by (1) replacing or reshaping the high-level phenomena to be explained; by (2) finding that a phenomenon X is explainable by a type Y of mechanisms (without knowing precisely which one); by (3) discovering part of the mechanism Y that explains the phenomena X; by (4) going from a mechanistic sketch to a complete schema; by (5) discovering a still lower mechanism to explain the activities of a lower-mechanism. I want to propose another way in which progress can come about. We talked earlier about two types of mechanistic explanations: constitutive explanation and etiological explanation. It seems that progress in some domains can be seen as a switch from one type of explanation to the other. For instance, Karmiloff-Smith (1998) has recently criticized neuropsychology for having adopted a constitutive explanation approach in the case of developmental disorder. Take for instance the way autism is traditionally conceived. Developmentalists (Baron-Cohen 1995; Leslie & Scholl 2001) and evolutionary psychologists (Cosmides & Tooby 1995) are seeing autism as the breakdown of an innate theory of mind mechanism (ToMM).[3] They thus see autism on the model of adult neuropsychopathology, that is as the breakdown of one module (ToM) with other modules unaffected (Shallice 1988). A similar story would apply for William syndrome where children have defective spatio-constructive abilities, as well as in problem resolution, planification, and numerical cognition but would have otherwise

[3] For instance Carruthers (1996, p 258) who wrote: "Autism represents a developmental failure of innate, isolable component of the mind which embodies a theory of the nature and operations of mind".

intact face recognition and language abilities, or for schizophrenia where the system of action monitoring is thought to be impaired (Frith 1992).

Karmiloff-Smith's project is to oppose the picture of intact and damaged modules (what she calls the "localization" picture) by re-injecting development in neuropsychology. Karmiloff-Smith bases her proposal on data coming from developmental cognitive neurosciences (she develops a point of view quite compatible with Schaffner's work on development 1998). Data from developmental cognitive neurosciences, she writes, suggest that "genes that are expressed in neocortex tend to do so throughout most regions, resulting in a similar six-layer structure and a similar overall pattern of intrinsic connectivity. Combinations of neuroanatomical features, cortical layers and brain cytoarchitectural regions are found to be remarkably similar in all regions of the brain from birth to 72 months" (Karmiloff-Smith 1998, p 397). To make a long story short (for a more developed story see Johnson 1999; Quartz 1999), the cortex, where many of the modules postulated by neuroscientists are supposed to be, goes through a process of progressive specialization (and focalization of its activity) or "modularization". For instance, perception of faces goes from very crude representation of faces (blob faces) to more sophisticated (6-month olds can recognize particular human faces as well as particular monkey faces and show similar psychophysical effects, like the inversion effects, see Johnson 1999) to still more sophisticated (older kids can't recognize monkey faces as well). At the same time, the part of the brain in charge of recognizing faces moves from the superior colliculus to the cortex where at first many areas are activated at the view of a face, until this capacity is taken in charge uniquely by the fusiform gyrus (this is what Johnson calls an "increasing restriction of fate" (1999, p 420)). Indeed what developmental cognitive neuroscientists are showing is that the environment (the presence of human faces in the infants' environment) and the activity of the organism play a very important role in shaping his brain (again here, the complete explanation of how we end up with a face recognition system would require to take an 'augmentationist' stance).

In the light of these data, Karmiloff-Smith proposes that we should see developmental disorders as "different developmental trajectories caused by initial differences at a neurocomputational level (caused by gene mutation or delation[4]). Thus, there might be differences in the microcircuitry of the brain or the firing properties of neurons, as opposed to discrete lesions to particular large-scale brain structures or pathways" (Karmiloff-Smith & Thomas 2002, p 293). In this approach, delation or mutation of genes will "have stronger effects on some outcomes and weaker effects on others. A totally specific disorder will, *ex hypothesi*, be extremely unlikely, thereby changing the focus of research in pathology" (Karmiloff-Smith 1998, p 390). This is indeed what she found when she started looking more closely at William syndrome children. She found that what looked like intact capacities where indeed results of compensation from a plastic brain: for instance, William syndrome kids are not using the same holistic strategies to recognize faces (they are better at recognizing faces identified by single features than faces requiring to compute configurations of features) or they show atypicalities in language development (they learn the lexicon in

[4] "Genes that affect brain development relatively early are likely to have wide cascading effects, often resulting in severe mental retardation [. . .] genetic defects that have their influence late in brain development are likely to affect those part of the brain that show the most prolonged growth, such as the cerebral cortex. [. . .] [they] are likely to specifically affect [. . .] dendritic and synaptic development [that] take place after cell migration in all brain structures [. . .]" (Johnson, Halit, Grice, & Karmiloff-Smith 2002, p 524).

a way different from normals and show difficulties in processing embedded relative clauses for instance).[5] The disorder has much more blurry borders than was thought by traditional neuropsychologists. Progress here came from abandoning one type of explanation for another. This forces to rethink the distinction between the "field elements" and the "preferred model system". What seems to happen in this case is a proposal by Karmiloff-Smith to reject the "*Field* Preferred Model System" for one more consistent with brain and cognitive development data.

There would be more to say about how this neuropsychological model of development challenges some forms of reductionism (the idea that there is a gene for grammar or that ToMM's mechanism and representational content is coded in the genes; Leslie and Scholl 2001), but I'll leave that for another time. All I wanted to do today was to propose some reflections on the neo-mechanistic model now adopted by Schaffner.

References

Baron-Cohen, S. (1995). *Mind-Blindness*. New York: Oxford University Press.
Carruthers, P. (1996). Autism as mind-blindness. In P. Carruthers & P.K. Smith (Eds.), *Theories of Theories of Mind*, (pp. 257–273) Cambridge: Cambridge University Press.
Cosmides Tooby. (1995). *Preface to Baron-Cohen mindblindness. Mindblindness*. New York: Oxford University Press.
Craver, C. F. (2002). Structures of scientific theories. In P. Machamer & M. Silberstein (Eds.), *The Blackwell Guide to the Philosophy of Science* (pp. 55–79). London: Blackwell.
Craver, C. F., (ms). Functions and mechanisms in contemporary neuroscience. To appear in P. Poirier & L. Faucher, *Des Neurones à la Philosophie: Neurophilosophie et philosophie des neurosciences*.
Craver, C. F., & Darden, L. (2001). Discovering mechanisms in neurobiology: the case of spatial memory. In P. K. Machamer, R. Grush, & P. McLaughlin (Eds.), *Theory and method in neurosciences* (pp. 112–137). Pittsburgh: Pittsburgh University Press.
Frith, C. (1992). *The neuropsychology of Schizophrenia*. East Sussex: Lawrence Erlbaum associates.
Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of Science, 69*, 118–137.
Green, M.F. (2001). Schizophrenia revealed: From neurons to social interactions. New York: W.W. Norton.
Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in Cognitive Sciences, 2*(10), 389–398.
Karmiloff-Smith, A. & Thomas (2002). Developmental disorders. In A. Arbib (Ed.), *The handbook of brain theory and neural networks*, 2nd edn. (pp. 292–294). Cambridge: MIT Press.
Johnson, M.H. (1999). Cortical plasticity in normal and abnormal cognitive development: evidence and working hypotheses. *Development and Psychopathology, 11*, 419–437.
Johnson, M. H., Halit, H., Grice, S. J., & Karmiloff-Smith, A. (2002). Neuroimaging of typical and atypical development: a perspective form multiple levels of analysis. *Development and Psychopathology, 14*, 521–536.
Leslie & Scholl. (2001). Minds, modules and meta-analysis. *Child Development, 72*, 696–701.
Machamer, P., Darden, L., & Craver, C. F. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1–25.
Panksepp, J. (1998). Attention deficit hyperactivity disorders, psychostimulants, and intolerance of childhood playfulness: a tragedy in making? *Current Directions In Psychological Science, 7*(3), 91–98.
Panksepp, J., Burgdorf, J., Turner, C., & Gordon, N. (2003). Modeling ADHD-type arousal with unilateral frontal cortex damage in rats and beneficial effects of play therapy. *Brain and Cognition, 52*, 97–105.
Poland, J. (2006). Steps toward and schizophrenia-free world. In L. Faucher (Ed.), *Philosophiques: numéro special sur la philosophie des psychopathologies*, *31*(1), 99–129.
Quartz, S. (1999). The constructivist brain. *Trends in Cognitive Sciences, 3*(2), 48–57.
Robert, J. S. (2000). Schizophrenia epigenesis? *Theoretical Medicine and Bioethics, 21*, 191–215.

[5] A similar story can be told about schizophrenia (Green 2001) as well as for autism (see Shultz et al. 2000).

Schaffner, K. (1992). Psychiatry and molecular biology: reductionistic approaches to schizophrenia. In J. Z. Sadler, O. P. Wiggins, & M. A. Schwartz (Eds.), *Philosophical perspectives on psychiatric diagnostic classification* (pp. 279–294). Washington: John Hopkins University Press.

Schaffner, K. (1993). *Discovery and explanation in biology and medicine*. Chicago: University of Chicago Press.

Schaffner, K. (1994). Philosophy of medicine. In M. H. Salmon et al. (Eds.), *Introduction to the philosophy of science* (pp. 310–345). Engelwood Cliffs: Prentice Hall.

Schaffner, K. (1998). Genes, behavior and developmental emergentism: one process, Indivisible? *Philosophy of Science. 65*, 209–252.

Schaffner, K. (2001). Nature and nurture. *Current Opinion in Psychiatry, 14*, 485–490.

Schaffner, K. (2002). Clinical and etiological psychiatric diagnoses: do causes count? In J. Sadler (Ed.), *Descriptions and prescriptions: values, mental disorders, and the DSMs* (pp. 271–290). John Hopkins Universtity Press.

Schultz, R. T., Gauthier, I., Klin, A., Fulbright, R. K., Anderson, A. W., Volkmar, F., Skudlarski, P., Lacadie, C., Cohen, D. J., & Gore, J. C. (2000). Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and asperger syndrome. *Archives General of Psychiatry, 57*, 331–340.

Shallice, T. (1988). *From neuropsychology to mental structure*. Cambridge: Cambridge University Press.

Wimsatt, W. (1976). Reductive explanation. In R. S. Cohen et al. (Ed.), *Proceedings of the 1974 philosophy of science association* (pp. 671–710). Dordrecht: Reidel.

Wolpert, D. M., Miall, C., Kawato, M. (1998). Internal models in the cerebellum. *Trends in Cognitive Sciences, 2*(9), 338–347.