remembered knowledge – and we will do so *exactly* when we perform operations on inner analogues that we would normally perform in the world.

How common is this? The key question is whether imagination *in general* is an active process. Perception is an active process of saccading and foveating. If the imagination has taken its cue from perception, as the emulator theory suggests, then it would seem that we regularly saccade and foveate onto inner analogues of external objects to acquire empirical knowledge, as needs dictate. When we ask ourselves whether frogs have lips or whether the top of a collie's head is higher than the bottom of a horse's tail, we foveate onto inner images, just as we foveate onto real frogs and real horses and collies. These kinds of inner operations may be more common than we had thought.

Grush's framework shows how it is possible to have offline empirical knowledge. It also complements the extended-mind thesis. If something counts as cognitive when it is performed in the head, it should also count as cognitive when it is performed in the world (mind leaks into the world). But also, if a process gives us an empirical discovery when it is performed in the world, it will also give us an empirical discovery when it is performed in the head (the world leaks into the mind). I think that Grush's emulator framework shows us how this is possible.

Where in the brain does the forward model lurk?

Opher Donchin[a] and Amir Raz[b]

[a]*Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD 21205;* [b]*Department of Psychiatry, Division of Child and Adolescent Psychiatry, Columbia University College of Physicians and Surgeons and New York State Psychiatric Institute, New York, NY 10032.*
opher@bme.jhu.edu **ar2241@columbia.edu**
http://www.bme.jhu.edu/~opher

Abstract: The general applicability of forward models in brain function has previously been recognized. Grush's contribution centers largely on broadening the extent and scope of forward models. However, in his effort to expand and generalize, important distinctions may have been overlooked. A better grounding in the underlying physiology would have helped to illuminate such valuable differences and similarities.

Despite the length of this piece, Grush's goal is modest: He attempts to show how seemingly disparate fields can be unified under the conceptual construction of the forward model, or emulator. In his conceptual framework, Grush argues that modeling is a common theme in activities that involve fashioning our own behavior, predicting the behavior of others (i.e., theory of mind), or expecting changes in the environment. Grush implies that this general network manifests in converging neurophysiological mechanisms.

Whereas this idea is not entirely novel, it is interesting to compare Grush's presentation with like accounts that were originally raised more than a decade ago with the advent of a cerebellar role in cognitive functions (Ito 1993; Kawato 1997). Those discussions related the idea of emulation to specific anatomical and physiological details, making testable predictions that are fruitful to this day. In contrast, the target article generally avoids a discussion of the underlying mechanisms, leaving the reader unclear as to the practical significance of the emulation theory.

Grush says that, at least for motor control and motor imagery, the forward model is likely implemented by the cerebellum. The target article would have benefited from a review of evidence suggesting that other modeling functions are also cerebellum-dependent (e.g., theory of mind [ToM]). The cerebellum is one of the brain structures consistently abnormal in autism (Courchesne 1997), concomitant with impairment in ToM (Frith 2001). Moreover, the cerebellum has occasionally been implicated in functional magnetic resonance imaging (fMRI) studies pursuing the locus of ToM (e.g., Brunet et al. 2000). On the other hand, ToM is usually associated with the prefrontal cortex or, possibly, the amygdala (e.g., Siegal & Varley 2002), and most neuroimaging studies do not find cerebellar activation (e.g., Castelli et al. 2002). If mechanisms of ToM are cerebellum independent, does it not have implications for Grush's theory? We feel the author should have addressed the physiological literature much more extensively, perhaps at the expense of other points.

By way of an intellectual detour relevant to issues of the forward model and ToM, we point out the view that impairment of the forward model for motor control may be key to inappropriate behavior (e.g., in psychopathology). In the case of delusions of control (e.g., schizophrenia), abnormal behavior may arise because failure of the forward model causes a perceived difference between expected and veridical consequences of motor commands (Frith & Gallagher 2002; Frith et al. 2000). The role that the forward model of one system might play in the behavior of another system seems relevant to Grush's sweeping theory.

While these issues go unaddressed, Grush devotes considerable attention to his emulation theory of motor imagery (previously suggested by Nair et al. 2003 and Berthoz 1996), contrasting it with the seemingly similar simulation theory. His argument for the emulation theory depends on a critical assumption that motor planning is in either kinematic or dynamic coordinates rather than in sensory coordinates. However, Grush does not convincingly support this assumption, and there is some reason to challenge its validity. For example, recent evidence on the effect of eye position on the behavior and physiology of reaching (Batista et al. 1999; Henriques et al. 1998) has been used to argue that reaching is planned in visual coordinates (Batista et al. 1999; Donchin et al. 2003). Moreover, even if we accept Grush's assumption, he does not explore the inevitable subsequent physiological implications. Presumably, motor planning takes place in either primary motor (MI) or premotor areas, and the forward model is to be implemented by the cerebellum. Towards that end, the actual sensory experience should be in either the primary or the secondary somatosensory cortex (SI or SII). However, fMRI studies of motor imagery find activation of MI, premotor areas, and the parietal reach regions (all regions associated with motor planning), but neither SI nor SII display such compelling activations (e.g., Hanakawa et al. 2003; Johnson et al. 2002; Servos et al. 2002;).

Grush also invests in a detailed development of the Kalman filter. The Kalman filter is an important idea in motor control, where a proper mixture of estimation and feedback are necessary for performance, but it is not appropriate in the other systems. In extending the model from the world of motor control, Grush obscures the fundamental idea behind the Kalman filter: The quality of the signals is used to determine the balance between its inputs. A gating, rather than filtering, mechanism would have been more fitting for all of his other examples, and the implementation of gating mechanisms is a different problem from that of filtering.

The difference between a gated and a filtered system affects the characteristics of the required forward model. The Kalman filter theory of motor control would be effectively served by an unarticulated forward model that calculated a rough linear approximation. This forward model needs to be fast, but it does not need to be accurate (Ariff et al. 2002). In contrast, the forward model implied by the emulation/simulation theory of motor imagery is the opposite: It does not need to be any faster than the actual motor-sensory loop of the body (and evidence indicates that it indeed is not faster; Reed 2002a), but it should provide an accurate notion of the sensations that would accompany action (Decety & Jeannerod 1995). We feel that physiological accounts could speak to such differences, and a more rigorous exploration might have made them more obvious to both Grush and his readers.

In sum, like Grush we agree that modeling is an important brain function. However, we believe that Grush's generalized approach may at times blur important distinctions rather than unravel previously unseen commonalities. We feel that had Grush more closely tied his account to the physiological literature, this shortcoming might have been evaded.

Emulators as sources of hidden cognitive variables

Peter Gärdenfors

Department of Cognitive Science, Lund University, Kungshuset, S-222 22 Lund, Sweden. **Peter.Gardenfors@lucs.lu.se**
http://www.lucs.lu.se/People/Peter.Gardenfors/

Abstract: I focus on the distinction between sensation and perception. Perceptions contain additional information that is useful for interpreting sensations. Following Grush, I propose that emulators can be seen as containing (or creating) hidden variables that generate perceptions from sensations. Such hidden variables could be used to explain further cognitive phenomena, for example, causal reasoning.

I have a great deal of sympathy for Grush's emulator model. Albeit still rather programmatic, it promises a powerful methodology that can generate a multitude of applications in the cognitive sciences.

Grush presents some evidence concerning the neural substrates of the emulators. However, this evidence is based on different kinds of neuroimaging. In my opinion, one should rather be looking for functional units in the brain, described in neurocomputational terms that can be interpreted as some kind of Kalman filter. At a low level, the example from Duhamel et al. (1992) concerning saccade anticipation seems to be such a system. However, the functional units should be searched for at higher levels of cognition as well. What ought to be developed is a way of combining the modeling techniques of artificial neuron nets with the control theoretical principles of Kalman filters (see the volume by Haykin [2001] for some first steps). What is needed, in particular, is an account of how a Kalman filter can *adapt* to the successes or failures of the controlled process.

As used in traditional control theory, Kalman filters operate with a limited number of control variables. In his presentation in section 2.3, Grush presumes that the emulator has the same set of variables as the process to be controlled. Although he notes that this is a special case and mentions that the variables of the emulator may be different from those of the process itself, he never presents alternative versions of the filters.

Now, from the perspective of the evolution of cognition, the distinction between sensation and perception that Grush makes in section 5.1 is of fundamental importance (Gärdenfors 2003; Humphrey 1993). Organisms that have perceptions are, in general, better prepared for what is going to happen in their environment. My proposal is that perceptions are generated by emulators and they function as forward models.

One important property of an emulator is that it does not need to rely exclusively on the signals coming from sense organs; it can also *add on* new types of information that can be useful in emulating. As a matter of fact, Grush (1998) has written about this possibility himself:

> The emulator is free to "posit" new variables and supply their values as part of the output. A good adaptive system would posit those variables which helped the controller [. . .] They are variables which are not part of the input the emulator gets from the target system. They may be the actual parameters of the target system, they may not. But what is important is that *the emulator's output may be much richer than the sensory input it receives from the target system.* (emphasis in original)

It does not matter much if the added information has no direct counterpart in the surrounding world as long as the emulation produces the right result, that is, leads to appropriate control signals.

The information provided by these variables is what generates the difference between sensations and perceptions. For example, when the system observes a moving object, its sensations consist only of the positions of the object, whereas the *forces* that influence the movement of the object are not sensed. However, if the system has been able to extract "force" as a hidden variable and relates this to the sensations via something like Newton's Second Law, then the system would be able to make more efficient and general, if not more accurate, predictions.

In section 2.2, Grush makes the point that emulators must have a certain degree of plasticity. This is not sufficient: A general theory must also account for how an emulator can *learn* to control a system. Supposedly, it slowly adjusts its filter settings (and set of variables) on the basis of some form of reward or punishment feedback from the process to be controlled. This would be analogous to how artificial neuron networks learn. Such a form of learning may pick up higher-order correlations between input and output. These correlations may be expressed by the hidden variables of the emulator.

The hidden variables of the multimodal emulators that Grush discusses in section 6.1, may provide the system (the brain) with cognitive abilities such as object permanence. More generally, one would expect the multimodal emulator to represent the world in an *object-centered* framework, rather than in a viewer-centered one (Marr 1982). As Grush (1998) writes: "[S]pace is a theoretical posit of the nervous system, made in order to render intelligible the multitude of interdependencies between the many motor pathways going out, and the many forms of sensory information coming in. Space is not spoon-fed to the cognizer, but is an achievement." Another speculation is that phenomena related to *categorical perception* are created by the hidden variables of the emulator.

More generally, different kinds of emulators may produce the variables that are used in *causal reasoning.* An interesting finding is that there is a substantial difference between humans and other animal species. As has been shown by Povinelli (2000) and others, monkeys and apes are surprisingly bad at reasoning about physical causes of phenomena. Tomasello (1999, p. 19) gives the following explanation of why monkeys and apes cannot understand causal mechanisms and intentionality in others: "It is just that they do not view the world in terms of intermediate and often hidden 'forces,' the underlying causes and intentional/mental states, that are so important in human thinking."

On the other hand, even very small human children show strong signs of interpreting the world with the aid of hidden forces and other causal variables. Gopnik (1998, p. 104) claims that "other animals primarily understand causality in terms of the effects of their own actions on the world. In contrast, human beings combine that understanding with a view that equates the causal power of their own actions and those of objects independent of them." Apparently, humans have more advanced causal emulators than other animals.

Finally, as Grush mentions in section 6.3.2, another relevant area is our "theory of mind," that is, the ability of humans to emulate (yes, not simulate) the intentions and beliefs of other individuals. An important question for future research then becomes: Why do humans have all these, apparently very successful, emulators for causes and a theory of mind, and why do other species not have them? A research methodology based on emulators and Kalman filters may provide the right basis for tackling these questions.