ple spatial centers, each processing and extracting distinct spatial information. As a result, space is represented in the mind not once but multiple times, not in a unified manner but segmented. Each representation is a (distorted) salience map with distinctive frames of reference (see Wang et al. 2001).

While we agree that emulation is yet another type of mental representation, we maintain that it is the goal but not the means. After all, when we have an emulator of the universe, to a certain degree we have understood how it works (see also Lewontin 2001 who suggests that the essence of science is to seek metaphor/emulation). In early development children are capable of interacting with the environment, but it is implausible to assume that they possess an emulator of the environment. Learning to interact with a model/emulation of the environment is certainly desired, but children must start with a model-less interaction mode and gradually learn to acquire such a model (e.g., Sutton & Barto 1998).

We would like to clarify our argument further by raising the second issue, which is related to the author's attempt to use amodal environment emulation to explain Johnson-Laird's mental models in human reasoning. First of all, according to Johnson-Laird (e.g., Johnson-Laird 2001), each mental model represents a possibility, and people by default represent only what is true but not what is false. In this sense each mental model is a fragmented representation of what is true, and constructing mental models involves considerable information processing at the semantic level. Treating mental models as emulations raises questions about how the emulators are smart enough to perform these semantic processes. Therefore, it seems that the claim does not solve the problem but simply gives it a different name.

In addition, treating mental models as space/object emulations against sentential representations would encounter difficulty in representing certain types of ordinary inferences that involve compound statements (Yang & Bringsjord 2003). The current mental model theory is powerful in representing prepositional reasoning but it has limited power in quantified predicate reasoning. Mental models theory provides initial models for truth connectives such as disjunction, and provides initial mental models for quantified atomic sentences such as: "(For all x) Ax." In combination, it can represent a compound statement such as: "(For all x) Ax or (for all x) Bx." It is difficult, however, for the current mental models theory to have a spatial/object style presentation for a quantified compound statement such as: "(For all x)(Ax or Bx)." For statements of this type, Yang and Bringsjord (2003) argue that some special syntactic structures would have to be constructed. Though one may still find ways to represent these mental models as space/object emulations (e.g., $[x]^{Ax}_{Bx}$, as a quantified disjunction), they apparently make mental models less efficient. At this point, the emulation hypothesis would face a philosophical question: Does cognition need to be maximally efficient?

Small brains and minimalist emulation: When is an internal model no longer a model?

Barbara Webb

Institute for Perception, Action and Behaviour, School of Informatics, University of Edinburgh, Edinburgh EH1 2QL, United Kingdom.
bwebb@inf.ed.ac.uk
http://www.informatics.ed.ac.uk/people/staff/Barbara_Webb.html

Abstract: Many of Grush's arguments should apply equally to animals with small brains, for which the capacity to internally model the body and environment must be limited. The dilemma may be solved by making only very approximate predictions, or only attempting to derive a "high-level" prediction from "high-level" output. At the extreme, in either case, the "emulation" step becomes trivial.

Grush makes a good case that concepts of emulation from control theory can illuminate a range of issues when applied to the brain. But this approach should apply to a wider range of brains than the human ones that form the focus of the article. Animals with small brains, such as insects, face many of the same behavioural challenges that the emulation theory is proposed to solve. For example, flies need to distinguish the visual slip that results from self-initiated steering, from that caused by external disturbances if they are to use the latter for stabilisation. It was just this sort of problem that first led to the explicit formulation of the concept of efference copy by von Holst and Mittelstaedt (1950; simultaneously described by Sperry 1950 as "corollary discharge").

Grush notes, in the field of motor imagery, a surprising lack of appreciation of the fact that an internal copy of the motor output is not in itself sufficient to explain the prediction of sensory experience. This failure to deal with the problem of the difference between the format of the output and input signals is pervasive in the biological literature on efference copy. Von Holst and Mittelstaedt, in their early formulations, describe efference copy as being compared to the sensory reafference like a "photo-negative," without any discussion of the potential differences between an "image" of the motor output and the "image" resulting from sensory input. In principle, for the example of the fly, it would be necessary for the fly to model the exact dynamics of how muscle commands affect wing movements, how wing movements turn the body, how the turning body affects the visual input, and how the visual sensors will respond to this input, to be able to produce a prediction that is spatio-temporally accurate. Indeed, as the response of its visual system to image velocity is known to be highly dependent on the contrast and spatial frequency of the scene, the fly must in theory have advance knowledge of the scene it will encounter as it turns if it is to calculate an exact prediction. It seems unlikely that all this is actually occurring in the small brain of the fly.

There are several ways out of this problem. The first is the recognition that the emulation may not be complete and accurate (briefly referred to by Grush in sect. 5.2), and that it does not need to be so to be useful. A simple scaling of the strength of the turn signal might be a good enough approximation of the size of the resulting optomotor signal for the fly. An issue that then arises is the following: How "schematic" (to use Grush's term) can the emulator be yet still count as an emulator? If it is defined by its role in the control cycle, as pictured in Figure 4 of the target article, then arguably even a simple inhibitory signal sent from motor areas to sensory areas should count, such as the "saccadic suppression" found in locusts (Zaretsky & Rowell 1979). On the other hand, it does not seem warranted to say that the animal is using an "internal model" in such a case. But what additional detail in the emulation would justify the ascription of a model? What if the prediction is also direction-specific, as for the *Drosophila* optomotor response (Heisenberg &Wolf 1984)? Or can be retuned online to be proportional in size (Mohl 1988)? Or is well matched in its time course, as in the inhibitory signal anticipating self-produced sound bursts in the cricket (Poulet & Hedwig 2003)? Must it estimate more than one dimension of the input signal? Or is the defining characteristic of a true internal model that it can also be run offline for planning? One might think this would eliminate examples from "small-brained" animals, although there is evidence for route-planning in jumping spiders (Tarsitano & Andrew 1999).

A second way out of the problem of having to emulate all the details of the motor system, environment, and sensors is discussed in more detail by Grush in section 5.1. The idea is to have the emulator predict higher-level representations (such as the new layout of objects in the world relative to the observer after a particular movement), rather than the raw sensations this new layout would produce. In the case of the fly, the prediction might be supposed to correspond to the output of the higher-level neurons that integrate motion across the visual field, rather than the response of each elementary motion detector. Of course, this shortcut to avoid modelling peripheral mechanisms can apply on the motor output side as well: The emulator can take as input a high-level motor command such as "rotate ten degrees left" and assume that this

https://doi.org/10.1017/S0140525X04480091 Published online by Cambridge University Press

results in a ten-degree-left rotation of the body position, rather than try to model the details of the muscle movements and resulting forces that cause the turn. However, taking this approach to its logical conclusion, the "highest level" of motor commands might be described in terms of desired goals and the "highest level" of perception might be confirmation that the goals have or have not been achieved – in which case the "emulation" required for prediction becomes trivial: It consists simply of making a copy of the goals.

Thus, in both cases – allowing approximate, or high-level, emulation – we are led to consider a continuum, from complete and detailed internal models of the plant, environment, and the measurement process, to extremely simple processes that might suffice for prediction. Much of the evidence cited to support the emulation theory does not distinguish where on this continuum the actual mechanisms might lie (Karniel 2002). Our intuitions tend to be very different when discussing human brains and insect brains, yet the problems are often common ones. The force of the theory – as a predictor of future behavioural or neurophysiological findings – lies in the assumption that the mechanisms, at least for human brains, lie towards the complex end of the continuum. The interesting question is whether there is strong evidence to support this assumption.

Two distinctions concerning emulators

Mark Wexler

LPPA/CNRS, Collège de France, 75005 Paris, France.
wexler@ccr.jussieu.fr **http://wexler.free.fr/**

Abstract: The target article distinguishes between modal and amodal emulators (the former predict future sensory states from current sensory states and motor actions, the latter operate on more abstract descriptions of the environment), and motor and environment emulators (the former predict the results of one's own actions, the latter predict all changes in the environment). I question the applicability of modal emulators, and the generalization to environment emulators.

Grush postulates two types of emulators, modal and amodal. Modal emulators operate at the sensory surface, predicting future sensory states (in some given modality) from current sensory states and motor commands; amodal emulators, on the other hand, operate on deeper descriptions of environmental parameters, predicting future parameters from current parameters and motor commands. Each of these two types of emulator runs into at least one major difficulty.

The difficulty of the modal emulator (of which Mel's models are examples) might be called *sensory aliasing:* Most of the time, current sensory states and motor commands can lead to many (usually infinitely many) future sensory states. Although Grush acknowledges this problem, it is probably worse than he lets on. Let us start by examining two special cases in vision where sensory aliasing *might not* be a problem.

When manually rotating a solid object, if the axis of rotation exactly corresponds to the line of sight, knowledge of the current two-dimensional image and of the motor command to the hand (assuming that this predicts the hand's angle of rotation) is enough to predict future sensory states – they are just rotations of the two-dimensional image by the given angle. This is the kind of rotation studied by Wexler et al. (1998).

In the human eye, optic and geometric centers nearly coincide. As a consequence, when the eye rotates in the head (during a saccade, for example), the two-dimensional optic array simply rotates on the retina, a rotation equal and opposite to that of the eye. Therefore, the sensory effects of eye movements can also be predicted, to some extent, by a modal emulator. (This is only true to the extent that we ignore the visual periphery, and changes in resolution and color due to a non-uniform distribution of photoreceptors.) This is the sort of prediction that Duhamel et al. (1992) found in monkey parietal cortex.

Now, these are just special cases of object manipulation and self-motion. As soon as we generalize even slightly, we come up against sensory aliasing.

When an object is rotated about any axis other than the line of sight, its image on the retina deforms because of the laws of optical projection, and undergoes other changes because previously occluded parts come into view. These deformations cannot be predicted from motor commands coupled with sensory data alone; knowledge of the three-dimensional structure of the object (the deep parameters) is required. The relationship between manual rotations in depth and corresponding mental rotations was studied by Wohlschläger and Wohlschläger (1998).

As soon as the head moves (and it does so nearly all the time, unless one's teeth are clenched around a bite bar), the eyes no longer just rotate, but undergo translations as well. This leads to motion parallax, and again the changes on the retina can be predicted only if one knows the three-dimensional layout of the environment. (Interestingly, motor action seems to play a special role in this process – see Wexler et al. 2001.) Hence, Grush's "white cube" example in section 4.4 is incomplete: Moving sideways while facing a cube leads not only to a shift of the retinal image (as would be the case for an eye rotation), but also to a deformation, a deformation that can only be predicted if one knows the three-dimensional shape of the object (a solid cube, rather than an infinite set of other objects compatible with the initial sensory data).

So situations in which modal prediction is even possible appear, at best, as special cases of much more widespread situations in which sensory aliasing precludes any possibility of predicting sensory states from other sensory states. It would be strange, therefore, if a modal prediction mechanism existed just for those special cases, only to be superseded by a much more general amodal predictor as soon as one steps out of the special case. By parsimony, we might suppose that all predictors are amodal.

Another significant distinction to make is between motor and environment emulators. Grush begins with motor emulators – ones that predict, modally or amodally, the results of one's own action – but winds up with something he calls "environment emulators," the much more general task of which is to predict changes in the environment that are due not only to the subject's own action, but to all relevant forces. Now, it might be argued that motor prediction is very special, in that one has intimate, before-the-fact knowledge of forces and goals. Indeed, the dynamic properties of the predicted trajectories of objects that one manipulates seem to be quite different from those of objects one merely observes (Wexler & Klam 2001).

The main difficulty is how the environment emulator would articulate with the motor emulators described in the first part of the target article. Indeed, it is not easy to see how the precise and well-documented mechanisms of forward models, Kalman filters, and so on, are to apply to the environment emulator. What does the environment emulator share with the motor emulator, other than the general concept of prediction? The answer to this question might prove quite interesting. For instance, it *could* turn out that to predict and to imagine changes in external objects independent of one's own action ("motion-encoded" prediction, to use Kosslyn's term), one nevertheless has to "imagine" acting on the objects on some level. In this case, the activation of the motor system would not rise to the level of consciousness, but would result in the activation of the motor predictor. However, without such a motor implication in the prediction and imagination of external events, it is hard to see how one can make the leap from motor to environment emulators.

https://doi.org/10.1017/S0140525X04480091 Published online by Cambridge University Press