

tor by detecting their interaction with pure kinesthetic sensation without movement.

It is known that pure kinesthesia without movement can be elicited by vibration of the tendon with a specific frequency (83 Hz; Craske 1977; Goodwin et al. 1972a; 1972b; Naito et al. 1999). Using this fact, Naito et al. (2002) showed that motor imagery affected pure kinesthetic sensation, generated by tendon vibration without overt movement. They found that the motor imagery of palmar flexion, or dorsiflexion of the right wrist, psychophysically influenced the sensation of illusory palmar flexion elicited by tendon vibration. Motor imagery of palmar flexion psychophysically enhanced the experienced illusory angles of palmar flexion, whereas dorsiflexion imagery reduced it in the absence of overt movement. This finding indicates that the emulator, driven by the mental imagery, outputs the “mock” sensory signals in a proprioceptive format, which interferes with the real (but artificially generated) proprioceptive sensory information from the musculoskeletal system.

Another prediction of Grush’s emulation theory is that the articulated emulator is a functional organization of components (articulants), whose interaction is comparable to that within the musculoskeletal system, and hence their neural representations are expected to be common. This point was also demonstrated by the study of Naito et al. (2002). Regional cerebral blood flow was measured with O-15 labeled water ($H_2^{15}O$) and positron emission tomography in ten subjects. The right tendon of the wrist extensor was vibrated at 83 Hz (ILLUSION) or at 12.5 Hz with no illusion (VIBRATION). Subjects kinesthetically imagined doing wrist movements of alternating palmar and dorsiflexion at the same speed with the experienced illusory movements (IMAGERY). A REST condition with eyes closed was included. The researchers identified common active fields between the contrasts of IMAGERY versus REST and ILLUSION versus VIBRATION. Motor imagery and the illusory sensation commonly activated the contralateral cingulate motor areas, supplementary motor area, dorsal premotor cortex, and ipsilateral cerebellum. The researchers concluded that kinesthetic sensation associated with imagined movement was generated during motor imagery by recruiting multiple motor areas, which were also activated by the kinesthetic sensation generated by tendon vibration. These commonly activated areas may constitute the articulants of the emulator driven by the efferent copy during motor imagery.

In conclusion, generation of kinesthetic sensation during motor imagery, and its neural representation common to kinesthesia without movement, can be interpreted as “emulated kinesthetic sensation” in the framework of the emulation theory by Grush.

Modality, quo vadis?

K. Sathian

Department of Neurology, Emory University School of Medicine, WMRB 6000, Atlanta, GA 30322. ksathia@emory.edu

Abstract: Grush’s emulation theory comprises both modality-specific and amodal emulators. I suggest that the amodal variety be replaced by multisensory emulators. The key distinction is that multisensory processing retains the characteristics of individual sensory modalities, in contrast to amodal processing. The latter term is better reserved for conceptual and linguistic systems, rather than perception or emulation.

Grush develops his emulation theory as a unified account of perception, imagery, and motor control, with the prospect of extension to diverse other neural functions. This theory is an advance over previous, less systematic formulations of simulation and imagery as being important in sensorimotor function. It makes the claim that particular neural elements work together in an emulation of perceptual or motor tasks, running in a special mode in which they are disconnected from external inputs/outputs. Vari-

ous emulations differing in their characteristics can hence be run, based on which the organism can select the best one to implement in interaction with the environment. The appeal of the theory stems from its unifying potential, and hence its success will be measured to a large extent by how well its binding of seemingly disparate streams of thought bears up over time. In this commentary, I focus on the relationship between the proposed sensorimotor emulator and sensory modality.

Grush argues for modality-specific as well as amodal emulators in the nervous system. Modality-specific emulators are relatively easy to understand, in terms of the operation of modality-specific sensory or motor systems. For instance, the findings reviewed in section 4.3 of the target article are compatible with a role for a motor emulator during visual imagery. However, the concept of a strictly amodal emulator, one that is entirely independent of any sensory “tags,” is less clear.

Let me make it absolutely clear that I am not arguing against amodal representations in the brain. Such representations must exist for abstract concepts that can be encoded linguistically, or “propositionally,” rather than directly in the workings of sensory systems. Indeed, as a vital part of human thought and communication, they are among the most important abilities that evolution has conferred on our species, compared to the other species that live or have lived on this planet. It is the characterization of abstract, amodal representations as imagery, and, by extension, as substrates of emulation strategies, that I am not comfortable with. Rather than “amodal” emulators, I suggest invoking “multisensory” emulators to provide the link between modality-specific systems and between these systems and abstract representations. I must emphasize that this is not a merely semantic distinction. By “multisensory,” I mean a system that receives inputs from more than one sensory modality. The existence of multisensory processes is well established, as is their neural implementation. The functions of multisensory processing include integration between the senses, cross-modal recruitment of sensory cortical regions, and coordinate transformation. Each of these functions has been studied in some detail.

Coordinate transformation in multisensory neurons of the posterior parietal cortex (PPC) has been intensively studied by Andersen and colleagues. This work indicates that multiple reference frames are represented in different regions of the PPC (Buneo et al. 2002; Cohen & Andersen 2002; Snyder et al. 1998). Reference frames may be allocentric, as in Brodmann’s area 7a; eye-centered, as in the lateral intraparietal area (LIP) and parietal reach region (PRR); body-centered, as in LIP; and both eye- and hand-centered, as in Brodmann’s area 5. Further, the eye-centered neuronal responses in LIP and PRR are gain-modulated by a variety of other factors such as eye, head, body, or hand position (Cohen & Andersen 2002). This effectively allows for a distributed representation of multiple reference frames simultaneously, and hence, for the coordinate transformations that are required for particular tasks, for example, between the retinocentric reference frame of visual stimuli or the head-centered reference frame of auditory stimuli and the body-centered reference frame of reaching arm movements, so that motor outputs may be appropriately directed. Multisensory emulators, then, could be engaged for specific coordinate transformations to allow planning of motor behavior as dictated by the organism’s current goals.

Another function of multisensory neurons is to integrate perceptual processes across the different senses. Such multisensory integration has been studied at the level of single neurons in the superior colliculus (Stein & Meredith 1993) and more recently in human cerebral cortex using functional neuroimaging. A case in point is the integration of auditory and visual information during perception of speech, which appears to depend importantly on cortex in the superior temporal sulcus (Calvert 2001). Moreover, Freides (1974) suggested three decades ago that, regardless of the modality of sensory input, the task at hand, especially if it is complicated, will recruit the sensory system that is most adept at the kind of processing required.

One means of such cross-modal recruitment is imagery. For instance, visual imagery may accompany tactile perception and could play a role in the engagement of visual cortical areas during tactile perception. Such recruitment of visual cortex has now been demonstrated in a variety of tactile tasks involving perception of patterns, forms, and motion, and appears to be quite task-specific, with areas that are specialized for particular visual tasks being recruited by their tactile counterparts (Sathian et al. 2004). An alternative interpretation of this type of cross-modal sensory cortical activation is that the regions involved are truly multisensory rather than unimodal. There is, in fact, increasing evidence that cortical regions traditionally considered to be unimodal are actually multisensory, receiving projections from other sensory systems in addition to their “classic” sources (e.g., Falchier et al. 2002; Schroeder & Foxe 2002). Multisensory emulators could clearly be employed to facilitate such cross-modal recruitment and synthesis.

My point is that, in all these examples of multisensory and cross-modal processing, specific modality tags appear to accompany the relevant sensory representations, which are associated with corresponding coordinate systems. This differs from Grush’s account, in which there is an amodal system, devoid of specific modality tags, that is used for perception and for internal emulation. I suggest that such amodal, propositional systems are conceptual and linguistic rather than being perceptual or the substrate for either imagery or sensorimotor emulation. It will be important for future empirical and theoretical research to attempt to distinguish clearly between multisensory and amodal neural systems.

Brains have emulators with brains: Emulation economized

Ricarda I. Schubotz and D. Yves von Cramon

Department of Neurology, Max Planck Institute for Human Cognitive and Brain Sciences, 04103 Leipzig, Germany. schubotz@cbs.mpg.de
cramon@cbs.mpg.de

Abstract: This commentary addresses the neural implementation of emulation, mostly using findings from functional Magnetic Resonance Imaging (fMRI). Furthermore, both empirical and theoretical suggestions are discussed that render two aspects of emulation theory redundant: independent modal emulators and extra measurement of amodal emulation. This modified emulation theory can conceptually integrate simulation theory and also get rid of some problematic philosophical implications.

Emulators with brains. The emulation account provides a formal way to apply the idea that the brain’s default mode is not passive waiting but active prediction, not only in motor control and imagery, but also in perception and perceptual imagery – an extension which fits perfectly with a long series of fMRI studies we performed on voluntary anticipatory processes. These studies made use of the serial prediction task, which requires participants to predict perceptual events on the basis of stimulus sequences. The lateral premotor cortex (PM), pre-supplementary motor area (pre-SMA), and corresponding parietal/temporal areas are engaged in active anticipation of sensory events. Note that this network is activated in absence of motor behavior, and that perceptual input is controlled by contrast computation.

Several functional characteristics of the considered areas render them candidate components of an emulator network. First, in the aforementioned studies each PM field’s response is restricted to specific stimulus features: PM fields for vocal movements are engaged in rhythm and pitch prediction, those for manual movements, in object prediction, and those for reaching and pursuit, in spatial prediction. A simplified synopsis of the results indicates that the anticipation of sensory events activates the PM fields of those effectors that habitually cause these sensory events (Schubotz & von Cramon 2001; 2002; Schubotz et al. 2003). This

“habitual pragmatic body map” (Schubotz & von Cramon 2003) in PM may precisely reflect Grush’s description of an “articulated” body/environment emulator. Second, our findings would also be in line with an emulation network that entails both amodal and modal representations.

Grush proposes motor regions to reflect the controller, and ventral and dorsal processing streams to be the core environmental emulator. We would rather suggest that multiple PM-parietal loops (including the ventral/dorsal stream) function as emulators, with each loop linking both heteromodal and unimodal representations (following the terminology in Benson 1994). One may even hold articulated emulation to be the default mode of PM-parietal loops which are exploited for perception, action, and imageries (see Fig. 1). Visual, auditory, or somatosensory imagery might be generated by efferent signals to and feedbacks from the corresponding unimodal association cortices.

We argue that such a modal emulation cannot be considered to be independent from amodal emulation. Rather, the same signal is concurrently sent to both unimodal and heteromodal association areas, even though current internal and external requirements may then determine which feedback becomes causally effective. Visual, auditory, hand, and foot imagery may introspectively feel different possibly because the controller exploits different premotor-parietal-subcortical loops. But all these networks, first, are made of both unimodal and heteromodal cortices which, second, communicate with ease. Possibly this in turn renders an extra measurement process redundant, as we also argue. On the other hand, “controller” functions (or perhaps better, competitive filter functions) may be realized more restrictedly within pre-SMA, in turn under the influence of anterior median frontal cortices, lateral prefrontal cortex, and extensive feedback projections.

Don’t introduce independent modal emulators – even if imagery sometimes feels purely visual . . . An introspectively compelling reason for suggesting independent modal emulation is that some kinds of modal imagery (e.g., a vase) feel purely visual and not at all motor. However, our fMRI findings reveal introspective reports to be unreliable (because introspection does not tell us that motor areas are engaged in non-motor anticipation). Likewise, we are introspectively blind to the empirical fact that perceiving an object includes perceiving what is potentially done with that object (see Gibson [1979/1986] for the notion of an object’s affordance, and, e.g., Fadiga et al. [2000] for premotor responses to mere object perception in the monkey). Conversely, it is conceptually inconsistent to assume amodal emulators to be independent of modal emulators, because in the emulation account, perception is sensation, given an interpretation in terms of amodal environment emulators, whereas sensation in turn is the on-line running of modal emulators. It therefore appears that amodal and modal emulation have to be conceptualized as reciprocally dependent¹.

. . . And don’t measure the emulators – even if imagery sometimes feels proprioceptive. An introspectively compelling reason for suggesting extra measurement is that motor imagery feels proprioceptive and not at all dynamic/kinematic. This also builds the core premise for splitting emulation from simulation: A motor plan is a dynamic/kinematic plan, whereas full-blown motor imagery is (mock) proprioceptive by nature and therefore must be previously transformed from the former by intermediate emulation and measurement.² However, exactly this premise would be rejected by accounts based on the ideomotor principle (e.g., theory of event coding; Hommel et al. 2001). These take motor acts to be planned in terms of desired action effects, that is, expected sensory events, and therefore plans and effects most likely share a common neural code. Comfortingly, emulation theory is not committed to the view that efferent signals are motor by nature. To be an efferent signal is nothing more than to be a delivered signal, no matter whether motor, sensory, sensorimotor, or amodal. Let us assume that the controller speaks “Brainish,” the lingua franca spoken by every subsystem in the brain, and that “measurement” is nothing but (and therefore should be termed) feedback from