Commentary/Hurley: The shared circuits model

Although AVHs are classified as a first-rank symptom of schizophrenia, they may not necessarily signify pathology and may best be understood within the wider context of the development of inner speech (Jones & Fernyhough 2007a; 2007b). In Vygotsky’s theory of the social origins of higher mental processes, inner speech represents the end point of a developmental process in which external conversation gradually becomes internalized to form verbal thought (Vygotsky 1934/1987). Like its semi-covert developmental precursor, private speech, inner speech retains the dialogic nature of the external discourse from which it derives. Fernyhough’s four-stage model of the development of inner speech as conceptualized by Vygotsky suggests two distinct forms of dialogic inner speech (Fernyhough 2004): expanded inner speech, where the give-and-take quality of external conversation permeates the verbal median; and a condensed variety of inner speech, where inner speech becomes “thinking in pure meanings” (Vygotsky 1934/1987), having lost most of the acoustics and structure of external dialogue. According to Fernyhough’s (2004) theory of AVHs, which draws on Vygotskian ideas about the developmental significance of inner speech, AVHs result from the temporary re-expansion of condensed inner speech, particularly under conditions of stress and cognitive challenge. The acoustic properties of the voices in inner dialogue are thus not attenuated but are experienced fully. The question then is how it is possible that cognition (inner voice) produced by self may be experienced as produced by other.

The cognitive dysfunction that results in the failure to differentiate self from other in inner speech may be explained by a forward model similar to the one underpinning Hurley’s layer 2. SCM relies on the forward model of motor control as proposed by Miall (2003) to postulate the subpersonal process that predicts the consequences of motor commands and compares them with the desired state. In her article on delusions of alien control, Blakemore (2003) uses this model to explain how an internal predictor uses information about intentions to enable the distinction between self-generated and externally-activated sensory events. The forward model is dysfunctional when it cannot accurately predict the sensory consequences of a movement based on the efference copy of the motor command. This results in sensory discrepancy and a failure to cancel the reafference or actual feedback, so that the self-produced movement feels externally caused (Blakemore 2003; see also Frith et al. 2000b). Although developed to explain abnormalities involving overt actions, this forward model has recently been applied to inner speech (Jones & Fernyhough 2007b). Jones and Fernyhough’s application proposes a direct causal mechanism leading from a malfunction of the predicted state to the experience of inner speech as being of alien origin. When the brain either produces a degraded predicted state or fails to produce a predicted state at all from the initial inner speech motor command, the consequence is that an emotion of self-authorship is not felt and instead the inner speech is experienced as authored by an other. For any model of the mind or cognitive functioning to be complete, it must relate to the brain. Thus, we need to understand the neural underpinnings of the predicted-state mechanism proposed by the forward model. This may require investigating networks, such as the interactions between perceptual and motor areas (Jones & Fernyhough 2007b). For example, Leube et al. (2003) have suggested that neurological activity associated with a deficit in the efference copy mechanism may involve the cortical network that de Vignemont and Fournieret (2004) found implicated in action attribution, including the prefrontal and the parietal cortex, the supplementary motor area, and the cerebellum. In terms of AVHs, Shergill et al. (2000) examined functional magnetic resonance imaging (fMRI) scans of patients with schizophrenia made while the patients were experiencing AVHs. They noted that the pattern of activation observed during AVHs was remarkably similar to that seen when healthy volunteers engaged in auditory verbal imagery (AVI), which is produced when one imagines being spoken to by another person. Specifically, Shergill et al. (2000) observed common activation of bilateral frontal and temporal gyri, along with right-sided precentral and inferior parietal gyri. Increased supplementary motor area activation was associated with healthy participants generating auditory verbal images; however, the supplementary motor area (SMA) was only weakly activated during AVHs. Other studies have suggested a role for the right anterior cingulate gyrus (see Jones & Fernyhough 2007a and studies cited therein). Given that the parietal and cingulate cortices subserve attention to internal and external bodily space and the attribution of significance to sensory information, they provide a plausible neural substrate for the misattribution of self-generated inner speech to other (see Spence et al. 1997).

Shared circuits in language and communication
do: 10.1017/S0140525X07003172

Simon Garroda and Martin J. Pickering

Department of Psychology, University of Glasgow, Glasgow G12 8QF, United Kingdom; Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, United Kingdom. simon@psy.gla.ac.uk http://www.psy.gla.ac.uk/~simon/martin.pickering@ed.ac.uk http://www.psy.ed.ac.uk/people/martin/index.html

Abstract: The target article says surprisingly little about the possible role of shared circuits in language and communication. This commentary considers how they might contribute to linguistic communication, particularly during dialogue. We argue that shared circuits are used to promote alignment between linguistic representations at many levels and to support production-based emulation of linguistic input during comprehension.

Hurley’s shared circuits model (SCM) provides a framework for investigating the role of emulation and imitation in social cognition. The SCM builds on two recent developments in cognitive neuroscience: Grush’s (2004) notion of an emulator (originating from motor control theory) and the discovery of mirror and canonical neurons in monkeys. The target article specifically concentrates on the role of shared circuits in imitation, deliberation, and mindreading. However, it says little about their role in language and communication, which presumably underpin many of the cognitive abilities that Hurley focuses on.

Section 2.3.1 of the target article discusses various hypotheses about how imitation might support language. For example, Hurley argues that the “flexible articulated relations between means and ends in imitative learning could be an evolutionary precursor of arbitrary relations between symbols and referents” (para. 2) and that “mirror systems provide a common code for actions of self and other, and thus for language production and perception” (para. 3). Finally, she suggests that the “flexible recombinant structure of ends and means in imitation may be a precursor of recombinant grammatical structure in language” (para. 4).

However, section 3 contains surprisingly little about the relationship between the SCM (and its various layers) and language processing. In fact, it is only when discussing layer 5 (the full-blown model) that language is considered at all. This is in relation to how imitative learning together with learned manipulation of external symbols could support the rich structure of language. Hurley also speculates that language could assist layer 3 circuits in taking input off-line, thereby allowing for more advanced mindreading (e.g., in multi-person strategic deliberation).

By contrast, we suggest that lower layers of the SCM may play a crucial role in language processing, in particular during
Hurley’s high level of generality suggests that a control-theoretic framework underpins all of the phenomena in question, but it is not without problems. In contrast to the action-perception domain, where the control-theoretic framework certainly applies, there is no evidence that this framework equally applies to feelings and emotions, such as pain, touch, and disgust, where mirroring and simulation are also found.

Hurley’s target article is pitched at a high level of generality. It speaks broadly of shared circuits, control, mirroring, simulation, mindreading, and so forth, giving the impression that its major theses apply equally across all applicable types of cognition. But there is good reason to doubt that this is accurate, and it is not entirely clear whether Hurley really intends it. Important sub-themes of the target article seem principally aimed at the relation between action and perception – for example, the falsity of the “classical sandwich architecture” (sect. 3, para. 1). Is everything she says about action, perception, and feedback supposed to apply equally to other domains in which shared circuits, mirroring, and mindreading are found? The article’s level of abstraction leaves the distinct impression that the theses advanced at the various layers of analysis cut across all the domains, but that is dubious.

Be that as it may, we would worry centers on the relation between shared circuits (or mirroring) and control theory. Hurley is not alone in emphasizing such a connection (Gallese 2003; Wolpert et al. 2003). However, the case for tying the control-theoretic perspective to shared circuits, mirroring, and simulation is based mainly on the action-perception domain, where there is specific physiological, theoretical, and experimental evidence for efferent copy and reafferent input. Nothing of this sort exists, however, for a number of other domains where shared circuits and simulation are found.

To be specific, mirroring phenomena exist in several areas of cognition in addition to the motoric: in sensation, including pain (Jackson et al. 2004; Singer et al. 2004) and touch (Keysers et al. 2004), and in emotion (most clearly, disgust; see Wicker et al. 2005). But in these domains, there are no established feedback or control-theoretic phenomena of comparable importance – or any sort at all. Here is a brief review of the shared circuits (or mirroring) findings across multiple domains. The shared areas or circuits for action are the prefrontal cortex and inferior parietal lobule interconnected with the superior temporal sulcus (STS)/middle temporal gyrus (MTG); for disgust, the insula; for fear, (possibly) the amygdala; for pain, the anterior cingulate cortex (ACC) and anterior insula; and for touch, the somatosensory cortices. In all cases, observing what other people do or feel is transformed into an inner representation of what we would do or feel in a similar, endogenously produced, situation. In many of these cases, moreover, evidence drawn from lesion studies and imaging studies indicates that mirroring produces mindreading of others’ mental states (Goldman 2006; in press; Goldman & Sripada 2005). However, only in the case of action is there clear evidence of feedback loops that fit the control-theoretic framework. So the notion that systematic relationships between shared circuits, simulation, and mindreading crucially depend on control-theoretic mechanisms is unsupported. Yet that is what Hurley suggests, since her architecture of social cognition is erected on a control-theoretic foundation.

Hurley writes that “the shared circuits model (SCM) shows how subpersonal resources for control, mirroring, and simulation can enable the distinctively human sociocognitive skills of imitation, deliberation, and mindreading” (sect. 3, para. 1). Her two bottom layers of analysis highlight adaptive feedback control and prediction of effects for improved control, and the three higher layers are explained in terms of these lower-level mechanisms. She makes no attempt, however, to explain how feedback and control account for simulational, empathic, or mindreading properties related to sensation and emotion. Indeed, the latter are barely mentioned. The explananda listed at her top level, the personal-animal level, all involve action and behavior; yet the