Representation in Extended Cognitive Systems: Does the Scaffolding of Language Extend the Mind?

Robert D. Rupert, University of Colorado at Boulder*

To Appear in R. Menary (ed.), *The Extended Mind*

How might it be established that the human mind is, in fact, extended? The most promising route runs through the scientific investigation of humans' cognitive skills and capacities – through cognitive science. Some authors (Dennett, 1991, 1996, Clark, 1997, 1998) claim that consideration of language-use, clearly one of humans' central cognitive skills, is particularly revealing: the use of language, as a system of extraorganismic marks and sounds, creates minds that extend beyond the boundaries of the human organism. This essay examines a number of ways in which this might occur, concluding that no compelling case exists for language's mind-expanding effects. Language profoundly influences our thoughts and greatly affects the development of the human cognitive system. Nevertheless, if we understand 'cognitive system' in such a way that the location of a cognitive system bears on the mind's location, external bits of language do not become part of that system.

I do not here plumb detailed models of language acquisition and use, of the sort one finds in cognitive science's academic journals. Instead, I lay out and criticize what I take to be the most promising general lines of argument that move from empirical observations about language and cognition to the conclusion that the mind is extended – an argument-strategy henceforth referred to as the 'language-based inference'. Operating at this level of abstraction forces the discussion toward overarching questions about the explananda of cognitive science and about what counts as a cognitive system the capacities and states of which stand in need of explanation

* I would like to thank Edward Averill and Douglas Kutach for comments on an earlier draft of this paper. Thanks also to Aaron Meskin for helpful discussions of some of this material.

by cognitive science. From this discussion emerge two substantive concerns, which I now preview.

Arguments for the extended mind often rely on a version of what I will call 'dependence-reasoning': if thought (mental activity, cognition) depends on factor X in some especially strong or clear way, then X is literally part of the thinker's cognitive system. Dependence-reasoning is not, in general, a reliable form of inference. There is simply too much dependence in the world for it to ground the individuation of systems, cognitive or otherwise.[1] Thus, some special consideration must drive its use in the language-based case. I argue that, appearances to the contrary, most such considerations fall flat. Even where they seem to weigh in favor of the language-based inference, my second general concern tips the scales against them. This second concern draws our attention to the explananda of cognitive science and the standard methods used to gather data in hopes of accounting for these explananda. I argue that these explananda and methods presuppose a nonextended view of cognitive systems – at least those relevant to the mind's location[2] – and this overpowers the language-based inference even in those rare cases where it might seem otherwise well motivated. The fundamental difficulty is this: the capacities investigated by cognitive science cannot be squared with the often fleeting nature of extended systems comprised of human organisms and external linguistic resources; the latter do not have the longevity or integrity to support the capacities in question. My appeal to the explananda and methods of cognitive science is more than parochial question-begging. I have in mind very broad characterizations of the explananda and experimental methods of cognitive science, broad

[1] Cf. Adams and Aizawa's observation that "the mere causal coupling of some process with a broader environment does not, in general, thereby, extend that process into the broader environment" (2001, p. 56).

[2] This qualification serves to preempt confusion that might arise from equivocation on 'cognitive system.' One way of thinking about systems is very liberal: anything that is of any causal importance in a particular context counts as part of the system of interest in that context. This liberal view does not, however, license any inference from extended cognitive systems to an extended mind. Hereafter, I omit this qualification, except where the temptation to equivocate seems especially strong.

enough to apply to the research programs of many who conceive of themselves as radical or at least as falling outside the mainstream of cognitive science. Although some might hope to reinterpret explananda and methods of cognitive science in keeping with the extended view, I argue that such reinterpretation is unnecessarily complex and ultimately unmotivated; the standard framework accommodates sufficiently the results of interest to those outside the mainstream, and thus conservatism wins the day.

I. Cognitive systems, coupled systems, and nontrivial causal spread

What would it be for external language to become part of the cognitive system in a robust enough way to warrant the language-based inference? Begin with the idea that in at least some cases the human organism and external linguistic resources constitute a coupled system exhibiting nontrivial causal spread.[3] The concept of a coupled system is rooted in dynamical systems theory (Port and van Gelder, 1995). Two distinct systems become coupled, thus becoming a single, coupled system, when their courses of evolution are mutually interdependent: changes in some important aspect of the behavior of each of the systems is affected by the changing state of some aspect of the other system. A formal representation of this mutual interdependence is typically effected by differential (or difference) equations in which the value of a term in one equation determines the value of a term in the other, while a term in the latter also determines the value of some term in the former (van Gelder's example of the Watt governor is a nice illustration – see van Gelder, 1995). For present purposes, the relevant form of a coupled system involves the human organism (or one of its subsystems) and some system (using the term loosely) of linguistic resources external to the human organism.

[3] See Wheeler, 2004, for explicit discussion. The idea's components are found in earlier pieces by Clark and Chalmers (1998, for an emphasis on coupled systems as extended cognitive systems, and Clark, 1997, on what he there calls 'continuous reciprocal causation') and Wheeler and Clark (1999, for a discussion of nontrivial causal spread); also see van Gelder, 1995.

As a sufficient condition for the existence of an extended cognitive system, the criterion of coupling alone would be too liberal. In many cases, the human organism is coupled to some external aspect of its environment, yet that external factor plays only a trivial role in accounting for the cognitive process in question. Thus, we need a nontriviality clause (Wheeler and Clark, 1999, p. 110; Wheeler, 2004, p. 703): if an organism is coupled to an external subsystem that contributes in a nontrivial way to the production of some significant, cognitive aspect of a cognitive process, then that process is extended. There is nontrivial causal spread in such cases because the range of factors constituting or determining some distinctively cognitive aspect of the process in question include states or processes beyond the boundary of the organism.

Emphasis on coupling with nontrivial causal spread takes us only a short distance toward a convincing language-based inference. In cases involving the organism's interaction with external linguistic resources, there is insufficient coupling: too much cognition takes places in the absence of the external linguistic resources that have affected those cognitive processes. This is not to deny that linguistic input significantly influences how and what humans think. All the same, humans frequently exercise the relevant cognitive skills offline: humans reason, categorize, and remember even though no spoken or written words are actively affecting the human organism (see Wheeler's discussion of off-line language-use – 2004, pp. 707ff.; cf. Carruthers and Boucher, 1998a, p. 15). This alone does not provide a counterexample to the sufficiency claim (i.e., that coupling with nontrivial causal spread is sufficient for an extended cognitive process), but it does limit the power of a language-based inference from the sufficiency claim. In many of the cases in which language plays a role in cognition, it does so indirectly, by the traces it has left behind. The sufficiency claim at issue does not explain how the existence of such cases supports the language-based inference. The cases to which our condition does apply

involve only short-lived systems the existence of which does not support an inference to an extended *mind*.

I return below to concerns regarding the short-lived nature of extended systems, but first consider a possible way to weaken the condition at issue, in hopes of validating a broader range of language-containing extended cognitive systems. Perhaps the coupling need only be current *or historical*. The second disjunct allows that genuine coupling at some past point in the process by which language affected thought suffices for the current existence of an extended cognitive system. This weakening does not clearly advance the case for the language-based inference. In the case of language *learning*, which one would expect to be the home turf for an historical clause, the organism is often passive (Bloom, 2000, pp. 8-9, 26ff.). The child's environment contains external linguistic resources and they affect the child, but the child and the resources do not become a coupled system *at any point in such learning processes*, for the child exerts no active causal control over the external resources. This sort of one-way dependence is a general phenomenon. In many cases, features of the environment drive the activity of the organism in such a way that the organism's behavior or cognitive processes depend in an especially sensitive way on the continuing effects of some environmental features, but not vice versa: the subject watches the airplane move across the sky but her watching does not affect the plane.[4] To be sure, I have not offered a counterexample to the proposed sufficient condition for the existence of an extended cognitive system; rather, these observations are meant to minimize the range of the condition's application and thus minimize its importance to cognitive scientific explanation:

[4] Compare Wheeler's example of a system that exhibits nontrivial causal spread, the system involving a robot that locks onto a white triangle on the wall (2004, 703-5); the robot is sensitive, in an ongoing fashion, to the triangle's reflectance of light, but the state of the triangle does not depend on the changing states of the robot (although the robot's movement does indirectly affect the values of the robot's input sensors, by changing the robot's position relative to the triangle).

coupling does not dominate cognitive development or activity in a way that might obviously compel an extended view of those phenomena.

This suggests the following even weaker formulation of a sufficient condition for extended cognitive processing, which combines the historical condition and the requirement of nontrivial causal spread while allowing only one-way dependence during the historical "coupling" period: if the states of one subsystem depend, or have at one time depended, on the changes in the states of another (e.g., states of the organism having depended on states of the world) in a way that nontrivially grounds some cognitive phenomenon (in virtue of either the former system's current or past dependence on the latter), then an extended cognitive system currently exists. In the remainder of this section, I argue that, as a general condition for the existence of extended cognitive systems, this should be rejected. In the section to follow, I consider attempts to salvage something of the historical aspect of this criterion in the language-specific case, by considering ways in which the *content* of human thoughts might depend on past interactions with linguistic resources.

One problem is that, given some plausible ancillary assumptions, the weakened condition entails unnecessary and inelegant cognitive spread. Cognitive processing frequently depends on external features in a nontrivial way, but in many of these cases, it is simply beyond credulity that we should treat the resulting systems as cognitive systems of the sort that would justify the language-based inference. Every act of visual perception in natural light involves the sun, and it does so in a nontrivial way. As the intensity of light from the sun changes, so change the states of the visual system. There is not perfect stability in the visual image, and large changes make significant differences in the cognitive results. The fact that the sun is present and in its current state explains why the resulting perceptual state has many of its cognitively relevant features;

take away the sun and the content of the perceptual state changes, as do the opportunities for further processing. Yet, anyone who claims that the proper system of study in cognitive science is an organism-star system provides a *reductio* of the extended approach to cognitive science – at least if the notion of a cognitive system is meant to admit of inference to the location or constitution of a mind (after all, no part of my mind is ninety-one million miles from earth).

Note, however, that the reasoning could run in a different direction: cognitive systems are extended, even as far as the sun; the cognitive system realizes the mind; an object individuated by its causal-functional role, as the mind is, is located wherever its realizer is located; therefore, the mind is extended, sometimes as far out as the sun! This form of argument deserves consideration, but its first premise requires support, presumably from a principle of cognitive-systems demarcation: some fact must determine what is a *genuine part* of a cognitive system. Furthermore, our principle must plausibly entail something about *minds*. What might do the trick?

In his discussion of cognitive systems and their realizations, Robert Wilson rightly exhorts his readers to defer to nature to determine what is and what is not a genuine cognitive system (2004, p. 132). Humans do not simply get to stipulate that, for example, a table seen is part of the total realization of a cognitive process and that the sun illuminating the table is mere background condition. Wilson goes on to claim that, as a matter of empirical fact, the individuals instantiating mental properties are subjects traditionally conceived of, i.e., human organisms. He rests this claim partly on the observation that "Individuals – and here, as always, our paradigms are individual people and individual organisms – are spatio-temporally bounded, relatively cohesive, unified entities that are continuous across space and time" (2004, p. 142). He also emphasizes the view's intuitive appeal (*ibid.*, pp. 142-43). Given some of Wilson's

other views, it is not clear why he settles on this position; there seems to be a mismatch between his views about the location of cognitive systems and the location of minds.[5] Nevertheless, I think his emphasis on the coherence and persistence of systems is on the right track and should be applied equally to cognitive systems (Rupert, 2004, pp. 425-28).

Generally speaking, we estimate which systems exist objectively by combining some intuitive starting point (e.g., that human organisms are the seats of cognition), methodological principles (e.g., simplicity and conservatism), and empirical results to date. Enquiry in a particular discipline must be driven initially by defensible judgments regarding the proprietary subject matter of that discipline – among these, judgments about which systems are at all likely to exhibit the properties investigated by that science. As work proceeds, these judgments are refined and sometimes altered in radical ways. At issue, then, are questions about the sorts of properties that cognitive science sets out to investigate, what systems are taken to instantiate those properties, and to what extent the inquiry in question has been successful; furthermore, the answers to these questions must be evaluated through the lens of established methodological principles such as simplicity and conservatism.

Taken together, these considerations favor a nonextended view of real-world human cognitive systems.[6] Of great importance is the way cognitive science characterizes its explananda. Humans categorize, perceive, remember, use language, reason, make sense of the actions of others – these and more are all persisting abilities of persisting systems; they do not consist merely in the activities of short-lived coupled systems (or short-lived systems that exhibit

[5] Wilson argues that human cognitive-cum-computational systems are often wide – i.e., include parts beyond the boundaries of the organism. Furthermore, he takes this to bear on the question of an extended mind: "Wide computational systems thus involve *minds* that literally extend beyond the confines of the skull into the world" (2004, p. 165, emphasis added), which would seem to entail that the individual is extended rather than organismically bound.

[6] Some support for this conclusion is offered in what follows; see Rupert, 2004, for further arguments.

one-way dependence; let this be understood for the remainder of the discussion).[7] The importance of systems that persist and cohere, even through change, is especially clear in developmental psychology: we want to know how *that* system – that single developing human – came to be the way it is and how a similar course of development happens, on average, for the relatively homogeneous multitude of such persisting human systems. We want to understand how and why the capacities and abilities of individual persisting systems change over time, eventually taking a stable form. If the systems to be investigated were fleeting, coupled systems, developmental inquiry would seem incoherent. We want to be able to explain why, for example, the child categorizes on the basis of appearance at age two but pays more attention to insides at age five. How can this question be sensibly posed – and in such a way that it might motivate a research program – if all that exists are ephemeral coupled systems, some behaving in one way and others exhibiting different behavior at a different time? If there is only a multitude of significantly differing systems, thousands of systems consisting of different *cognitive* components, there seems little reason to catalogue differences among them. Note how difficult it is to describe the behavior to be explained if one takes coupled systems to be cognitive systems whose presence constitutes the location of a mind. How might we describe, for example, appearance-based categorization within a framework of thousands of fleeting coupled systems? Do we say that there are lots of coupled systems (organisms-plus-stimulus-items), some of them having parts with, say, a certain shape (the stimuli), and that these coupled systems make judgments different from those made by other systems having older organisms put in place of the younger ones while the other parts (the stimuli) remain the same?

[7] Clark and Chalmers's (1998) discussion of portability, as well as their criteria for extended states, seems partly motivated by a concern for persisting systemic integrity.

In response, the advocate of extended cognition will surely recommend that we reconceptualize developmental psychology. In fact, the literature on extended cognition contains much talk of reconceptualization and paradigm shifts. Reconceptualization founders on a dilemma, however: either we can pursue genuine reconceptualization of cognitive systems as fleeting coupled systems, at great cost to cognitive psychology, or we can jury-rig a method of cognitive systems individuation that preserves the successes of cognitive psychology and is consistent with a viable investigative method but is unmotivated.

Developmental psychology must group together various fleeting coupled systems, some individual groupings containing systems that exist at different times (these are what we would normally identify with individual subjects persisting over time) else much statistical analysis of the data will be lost. Thousands of experiments have yielded interesting results by assuming some privileged grouping of the various short lived coupled systems. Think of the multitude of within-subject analyses of results on short series of experiment, all data lost. This is the cost of genuine reconceptualization, for it offers no justification for our privileged groupings. The radical approach offers developmental psychologists no more reason to be interested in, for example, the series of temporal segments we normally associate with Sally from ages two-to-six rather than to be interested in, say, Sally, aged two, together with a ball she was bouncing on some particular day, Johnny, aged five, together with the book he was reading on some particular afternoon, and Terry, aged seven, plus the stimulus item he has just been shown by the experimenter. It is simply not clear how one should proceed after giving up the traditional method.

Of course, sometimes a legitimate grouping of systems is resilient to changes across subsystems of the members of a single grouping. Consider the practice of medical doctors, who

talk about the same patient over time, even though, at the biological level, that patient changes his constitution (cells die, new ones form). This, however, only places on promontory the need for a principle of organization to ground the groupings. Organismic integrity, including spatio-temporal continuity, recommends treating a person over time as a single patient, returning for visits. But insofar as an organizing principle legitimates the groupings of alleged cognitive systems into privileged sets, the principle is organism-based – which is, of course, just how traditional cognitive science identifies its systems of interest. The new developmental psychologist gives away the game by making the integrity of the traditional subject the organizing principle for the cognitive systems studied in developmental psychology. This moderate position is not, then, worth its price to the advocate of extended cognition. She buys a highly counterintuitive claim about minds (that they are extended beyond the organismic boundary) for the cost of unnecessary complications in, without any substantive departure from or improvement on, standard individuative practice.

The problems are not limited to developmental psychology. Investigations of adult capacities, for example, for memory and language use, normally presuppose that researchers investigate particular persisting systems. Some such studies are explicitly longitudinal (Bahrick, 1979, 1984), and thus much like developmental psychology in the relevant respects. Beyond these cases, however, psychologists and linguists have been interested in a great many contextual effects, some of these perceptual. It is striking that the same person behaves in one way in one context – say, when not primed – and behaves differently in a slightly different context – when, in contrast, she has been primed. There is a large body of literature filled with experiments interpreted in just this way, their explananda taken to be persisting individuals having various capacities or abilities that they exercise in different ways in different contexts.

This emphasis on a persisting individual is evident even in research on perception, where coupling, or at least one-way dependence, seems common. We would like to know why, for example, the subject perceives certain features under some conditions – say against a particular backdrop – but does not perceive those same features under other conditions (see, for example, results discussed in Treisman, 1998). The experimenter asks a single system to perform various visual tasks, and the outcome sheds light on the process by which that system sees. Perhaps, as has been recently emphasized, the subject *does* something as a way of getting information visually. Still, such results seem to reveal something about the ability of a single persisting system; it tells how *that* system gets visual information.[8] The attempt to make sense of this data, and of the explanatory project more generally, faces a dilemma analogous to the one discussed above in the case of developmental psychology.

The preceding argument against extended individuation seems to have ignored an important aspect of standard methodology: researchers frequently assign experimental subjects to different groups; in a typical experiment, these consist of a control group and an experimental group. In such experiments, researchers do not appear to be investigating the capacities of individually persisting systems as they change over time or as they exercise their capacities in varying circumstances.

This reaction misinterprets standard methodology, however. Researchers assign subjects to different groups on the assumption that the set of members of each group represents a standard distribution of skills and reactions across an homogeneous population. That is, by statistical analysis of the results, we think we discover something about the way the *standard persisting*

[8] Consider a related worry. Often in perception, and in action based on perception, humans think about or perceive the same things with which they are interacting; if the cognitive system is individuated liberally, these external objects are part of our minds (presumably because our perception depends on them). But this view should give us

human system reacts under different conditions. Of course, data is sometimes analyzed by condition or by question, but even here the role of such analyses in the larger projects in which they are typically set is unclear if the researchers are not taking each of the data points relative to that analysis to be attached to a persisting cognitive system. Analysis by condition or question is meant to reveal something about how organismically bounded cognitive systems are affected by a suspected causal factor introduced by that condition or question.

We should not pronounce in advance what a completed cognitive psychology will bring. Nevertheless, insofar as we can make out an extended-systems-based alternative to existing methodology, it introduces an unnecessarily inflated set of distinct cognitive systems, the partitioning of which into useful subgroups requires additional theoretical justification; but as we have seen this additional justification seems little more than a co-opting of the success of standard methodology. On measures of simplicity and conservatism, then, the revisionary strategy clearly loses out to the traditional taxonomy. Of course, costly revisions in theoretical frameworks can sometimes be justified, when they offer substantial gains in other respects, e.g., in explanatory power or accuracy. The shift under consideration does not, however, do so. It would seem that most of the significant results in contemporary cognitive science can be cast in terms of organismically bounded cognitive systems that interact with their environments and frequently become coupled to them. (It might be that the visual system relies more heavily on demonstratives than was previously thought – lots of deictic pointers.) Perhaps, though, there is something special about the role of language-use that paves the route from language to the extended mind, something that involves either a special form of content-dependence or processing-dependence. It is to these issues I turn for the remainder.

---

pause. It entails that humans do not, on the basis of perception, interact with these perceived objects; rather, certain parts of the human mind interact with other parts of the human mind!

II. Linguistic content and thought content

There is widespread, although not unanimous, agreement that the notions of representation and, correlatively, representational content continue to be of use in the study of cognition. Parties to this prevailing view include many philosophers and cognitive scientists who generally work outside the rules-and-representations tradition (Clark, 2004, p. 719, Churchland, 1998, p. 31; Elman, 1995, pp. 221-22; Wheeler, 2001, Clark and Toribio, 1994; Wheeler and Clark, 1999, Grush, 1997, 2003).[9] This suggests fleshing out the language-based inference via a thesis of content-dependence, the idea that the content of organismically internal representations depends in some specially strong way on the content of external linguistic representations.

In this section, I examine three ways to appeal to content-dependence: (1) it might be that, in an important range of cases, external representations carry the contents of our thoughts (Houghton, 1997); or (2) it could be that, in some cases, thoughts inherit the contents of external representations after which those thoughts are in some sense patterned; or weaker still, (3) it might be that the content of thought is determined by the structure of internal cognitive processes where the structure in question is shaped by the causal or temporal structure of external linguistic resources (even though content is not directly inherited).

---

[9] These authors all seem to accept that representation consists in some mind-world relation(s) grounding application of some such notion as that of reference, truth-conditions, accuracy, correctness, or fit. The discussion in the main text is appropriately ecumenical, yet there are limits to what will count as representational content. Robert Wilson's view of content as exploitative and enactive (Wilson, 2004) most likely falls outside those limits. Wilson allows external states, such as the rigidity of bodies in the environment, to play a role in computations – this is part of the reasoning that leads him to endorse extended cognition. At the same time, though, Wilson asserts that there is no computation without representations (2004, p. 177), which makes it a bit difficult to pin down his notion of representation. What does the state of a rigid body represent? What is it a representation *of*? Something more helpful might be suggested by Wilson's talk of enactive representation: "Representation is not something implanted in individuals but something that individuals do by exploiting the rich structures of their environments in cycles of perception and action" (2004, p. 178). Wilson claims his view is not behaviorist (2004, pp. 184), but he also seems to want representation to play its standard sort of role, representing "objects, properties, events, and propositions" (2004, p. 222). What is there, on this account, to representing *that P* other than behaving in certain ways in certain circumstances; i.e., what is there to make Wilson's view nonbehaviorist? Nothing, it seems to me; but if it is not behaviorist, it seems eliminativist. It is not clear, then, how much of what I say below about content applies to content as Wilson conceives of it.

Consider cases of the first sort, in which an external linguistic representation carries the content of someone's thought. Persons frequently make lists, write down their ideas in personal journals and professional papers, and so on. In such cases, one might think that the content-bearing external resources play a sufficiently integral role in the subject's cognitive life as to warrant the language-based inference: the external resources seem to be part of the realization of the subject's thought; and since a thought is wherever its realization is, and a mind is at least partly wherever its thoughts are, we might conclude that the subject's mind is extended to the external matter that carries her thought's content.

A natural question to ask here concerns the source of externally represented content. In virtue of what do the external markings (or auditory forms) have content? In many cases, the content clearly originates with the subject of the mental states in question. The examples given above generally fall into this category: the subject first thinks that he needs some milk, then writes it on his shopping list. In such cases the subject's thought content is prior to the content of the relevant external linguistic tokens; thus, these cases do not seem to implicate language, in any particularly strong way, in the determination of thought content.

Perhaps, though, once content is off-loaded – i.e., encoded in the external linguistic tokens – the content takes on a life of its own. The subject no longer tokens any persisting, internal representation carrying the content of the state, but instantiates only a memory that there is, say, a shopping list (even this might be missing in some cases; the subject might not remember the list until he sees it lying on the table).

We should not be too quick to externalize subjects' thoughts in these cases, given the systems-based arguments of the preceding section. The individual *is* the cognitive system in the relevant sense, the sense pertaining to the location of minds. If an individual has the relevant

belief (say, that he needs milk) at a given time, yet not because he tokens persisting internal representations carrying the content of that belief, it is most likely in virtue of the appearance of pointers, i.e., internal representations that take fleeting values relative to the task being performed (Ballard et al., 1997); such pointers can carry the contents represented on the list by deferring to the content of what appears on the list, which, in turn, has its content in virtue of the subject's past investment. This picture seems especially compelling when we consider the causal efficacy of the mental states in question. The external resources have the kind of causal efficacy we expect a subject's thoughts to have only insofar as the external resources have intervening effects on the subject, which effects are then causes of further actions. The role of language as a repository of knowledge and a tool for communication seems clear enough. Nevertheless, cases where external linguistic units directly cause the subject's behavior, without affecting the organism as intermediary are much more rare and difficult to produce.

Turn now to another way of understanding (1): the content of the thought has its source not in the subject but in external linguistic units. The idea here is that the linguistic content appears in the linguistic token prior to the subject's use of it or internalization of it. If it could be added that the subject's thought content is somehow constituted by the content of the linguistic tokens, this would seem to support the language-based inference.

There are two ways to interpret this suggestion, the first strictly in keeping with (1) and the second amounting to alternative (2) listed above (according to which thought-content is inherited from the external linguistic resources after which the internal mental representations are patterned). According to the first interpretation, the subject proceeds absent a thought with the content *P*. The subject comes upon external linguistic units having the content *P*; she picks up a book, for example. She then "docks up" to those external resources, coupling to them in a way

that creates an occurrent mental state in the subject, e.g., a belief that *P*, without there being an internal vehicle with the content that *P* (Hurley, 1998). These cases exhibit one-way dependence with nontrivial causal spread, but there is something more important at work as well: the *content* of the external resources seems to *constitute* the content of the subject's mental state; and this lends special credence to the language-based inference in such cases.

Two reasons recommend against this version of the language-based inference. First, we should bear in mind methodological and empirical considerations concerning systems individuation discussed above. These considerations speak in favor of a general understanding of the human cognitive system as organismically bounded. If reading constitutes an unusual case, one in which we are not sure whether to posit internal representations that carry thought contents, then it might well be considered "spoils to the victor," i.e., a borderline case to be decided in favor of whichever theory is better supported by more weighty, independent considerations. Second, return to my earlier worry about mental causation. What do the mental states of a subject who is reading cause? One possibility, commonly explored in research on reading comprehension,[10] is that these states cause the subject to give particular answers to questions asked after the reading material has been put away. This, however, presupposes that the subject, at some point prior to questioning, forms internal representations (whatever form these take) of the text or its meanings. Presumably, the internal representations are not formed after the reading material was taken away. Thus, during reading, the subject forms internal representations of either the text or its meanings. Therefore, we have independent reason to think that when a subject engages with text, she forms some sort of internal representations that

[10] See, for example, the discussion of reading comprehension studies – using measures of literal memory for details and also inferences drawn from them – in Gathercole and Baddeley (1993, p. 228). Differences in performance on such tasks is best explained by differences in various subjects' construction and maintenance of internal structure, for the reading material has been taken from the subjects at the time the capacities are tested.

carry the content of the text being read. This leaves the advocate of the current interpretation of (1) little basis for a language-based inference.

This leads us to view (2) from our earlier enumeration of theses. According to (2), there are internal, mental representations active in the relevant cases, but those representations inherit their content from the content of external linguistic units. A significant amount of thought is, on this view, the use of internalized language. This picture obviates concern about systems-identification by claiming a privileged role for linguistic content *vis-à-vis* thought content without requiring that external linguistic resources be present at the time the subject employs the relevant internal resources.

Presumably, neural or syntactic facts determine whether the same mental representation is active on two different occasions (modulo, perhaps, some architectural facts that help to determine which units are active); and the origin of at least some of these units lies with the external linguistic units that originally caused their formation. The proponent of content-dependence might claim, then, that a significant portion of the internal units of cognition are, even when processing takes place in the absence of external linguistic resources, the same internal structures that are activated when a subject is affected by external linguistic resources. Furthermore, it might be claimed that the content of those internal units is copied, in a straightforward way, from the content of the linguistic units that caused the formation of the internal ones. If correct, this picture appears to support the language-based inference because it identifies a robust sense in which thought is no more than an aping use of external linguistic resources.

As the argument stands, it seems wanting, for a number of reasons. To begin with, the basic logic of the argument does not lead very neatly to an extended mind, at least not without

reliance on a dependency premise of the sort criticized above. We simply should not be moved by arguments that proceed from a premise of the form "Condition *c* is (given certain background assumptions) necessary for individual *i* to be in state *b*" to a conclusion of the form "*c* is a proper part of *i* (or of *b*)." Being surrounded by air is a necessary condition for my continued survival; we should not conclude on this basis that the surrounding air is a proper part of me or a proper part of my continued existence.[11] Exposure to language might be a necessary condition for having certain thoughts (Carruthers, 2002, p. 659; Carruthers and Boucher, 1998a, pp. 2, 10), but we should not conclude on this basis that external language is literally part of the resulting system. As Carruthers notes, the relation here is merely diachronic (*op. cit.*, p. 660). These concerns would hold even if a fairly strong version of the Sapir-Whorf hypothesis were correct:[12] even if the range of thoughts available to a subject are limited to those expressible in her natural language and even if she cannot think those thoughts until she has had sufficient interaction with that language, thought proceeds absent actual tokens of that language (visual images and subjectively heard sounds are *not* part of any external natural language).

Furthermore, (2) does not jibe very well with our best accounts of language learning (see Bloom, 2000, for extensive discussion of these issues). If one takes the internal representations to have the same content as the external units, and thinks this is so *because* the internal units were patterned after or caused by the external units, one has excluded from the picture the mental content necessary for the child to learn language. For example, it appears that the child uses pragmatic hypotheses, including suppositions about what other people are likely to do, in

[11] Elsewhere I have criticized an epistemic version of this principle (Rupert, 2004, pp. 395-96).
[12] There is good reason to reject the Sapir-Whorf hypothesis in its stronger forms (Bloom, 2000, Goldin-Meadow and Zheng, 1998, Carruthers, 2002), but many important questions about the relation between thought and language remain open; see the essays in Carruthers and Boucher, 1998, and Gentner and Goldin-Meadow, 2003.

order to learn the names of things. The child can frame such hypotheses only if she has at her disposal a significant conceptual repertoire – with content fixed, not inherited from language.

Lastly, even if we allow that the internal units operate by some kind of principle of deference, whereby they inherit their content from independently content-laden linguistic units, reasoning from that assumption to an extended mind seems perilous. Such reasoning would seem to proceed by a special form of dependence premise: the content of the subject's mental representation depends on the content of the linguistic unit. But this dependence-based reasoning leads equally to the inclusion of *many other minds* in the subject's mind. Insofar as the external units possess independent content, it is in virtue of the content of the mental states of ancestors, elders, and current-day speakers other than the subject. Language does not suddenly appear in the world, its content in place. If our subject's mind extends into the world because the content of her thoughts derives from the content of the external units that caused the development of some of her internal resources, then our subject's mind should also extend to encompass the minds of those ancestors and the like, the mental states in whom are responsible for the current external units' having the content they have. I admit that I do not fully understand the principles guiding inferences from dependence to constitution, but I see no reason to deny that such principles warrant transitive inferences, so long as the same form of dependence holds throughout the inference. Thus, however the dependence-reasoning is supposed to proceed here, it seems unprincipled to include external linguistic resources as part of the extended mind, while excluding the minds that give rise to the content of those external linguistic resource; in both cases – moving from current subject to linguistic units and from linguistic units to other subjects

– content-dependence is the issue.[13] But to include all of the other minds in the extended mind is a *reductio* of the view.

Alternative (3) offers a more roundabout route to content-dependence and from there to the extended mind. On this view, internal units are structured or processed in certain ways, and they, or their governing processes, take on that structure as an effect of interacting with external linguistic resources; and it is on account of this structure that the internal units possess the content they do.

This possibility can be fleshed out in a variety of ways. Think first in terms of a Fodorian language of thought (LOT) at least some of the terms in which are subject to an externalist semantics (for the various aspects of this picture and arguments for them, see Fodor, 1975, 1987, 1990, 1994). Fodor's claim that the system of mental representations is innate (Fodor 1975, 1981) strikes many authors as implausible, largely because these authors resist the idea that our concepts could, in any sense, be present at birth. So far as I can tell, however, this results largely from misunderstanding (Rupert, 2001, 1996, chapter 4). Fodor's conception is very weak, requiring only that there is no cognitive psychological explanation of the acquisition of our many innate concepts (where what counts as such an explanation is defined restrictively, in terms of hypothesis-testing); this does not preclude there being innate concepts the acquisition of which occurs, for humans, under only very constrained conditions. This view allows that many concepts – even innate concepts – are *difficult* to acquire; the triggering conditions for acquisition might elude most of us for some time, maybe for life (Fodor, 1998, pp. 156-161).

The difficulty of acquiring certain concepts can be made clearer by attending to the conditions of concept acquisition: part of what is required to acquire concept *C* is that the subject

[13] A counterfactual test does not help to isolate language from users other than the subject; there are no nearby worlds where external words have their content but that content does not derive from minds.

come to have a mental representation with the content *c*. Now assume that coming to have a mental representation with such content is largely a matter of coming into the right causal relation to the property or individual *c*. A mental representation's coming to stand in such a relation may require, at least for the typical human, much mediation (Fodor, 1987, pp. 121-22). It might be that, in some cases, the structure of external linguistic resources provides a model of some sort – a compositional or inferential model – after which internal processing can be patterned, eventuating in the acquisition of the concept in question.

This offers us a way to make sense of Carruthers's and Boucher's remark that language is obviously a prerequisite for the acquisition of such concepts as ELECTRON (1998a, p. 2). Furthermore, the view makes sense even on the assumption that the mental representation ELECTRON is atomic. For even if it is atomic, it is very likely that, closely connected to it, are complex mental representations of sentences containing the word 'electron' (in the case of English speakers). To acquire a mental representation that has the content *electron*, the mental representation ELECTRON must come into the right causal (or other content-making) relation to the property of being an electron; but until the subject reads about electrons or has a model of the atom explained, it is very unlikely that she has any mental representation bearing the content-making relation to electrons. On this view, it is the very arrangement of the external units that catalyzes the acquisition of such concepts, by creating causal dependences within the subject as well as between the subject and the extra-organismic world.

Understood in the preceding fashion, (3) takes us no closer to the extended mind. The content of the resulting concepts is not so intimately linked to the content of the external resources so as even to suggest that the external resources are part of the subject's mind. As presented, (3) merely asserts causal dependence. Perhaps, though, we might try invoking a more

robust aspect of linguistic structure as the ground of (3). For example, it is plausible that mathematical language provides processing guidelines that indirectly fix the content of those mental representations disciplined to follow linguistically given guidelines (Bloom, 2000, chapter 9). In many cases, a disposition to perform certain kinds of inferences is the best way a human can acquire a particular concept, even if the inferences are not constitutive of having the concept; and the persistence of these inferential abilities can be an important way for a mental representation to retain its content. Part of what one learns when learning mathematics is how to *proceed*, in counting, adding, and solving various other sorts of problems. If the patterns of relations of language cause certain internal causal processes to be far more likely to occur, processes that help to determine the content of the mental representations involved, we will once again have found an important role for language to play in determining mental content.

This interpretation of (3) resolves into (a) a causal dependence claim and (b) a claim about the contribution of other minds to the content of the subject's mental representations. The former asserts that linguistic structure has caused the subject's brain to process internal units in a way that confers on those units particular contents; this is a straightforward claim of causal dependence and thereby advances the language-based inference not a bit. Insofar as there is *content*-dependence – the content of the subject's mental representations depending on the content of their causes, by mirroring the structure of use – this dependence reaches equally to the minds of those who first began using mathematical language in a way that gives it its content; this subjects (3), as currently interpreted, to the first and third criticisms made of (2).

III. Structural effects

In this section, I shift focus to a different part of the explanatory enterprise in cognitive science, the attempt to understand cognitive processing. This work investigates the processes leading

from one content-laden state to another, from one set of capacities to a new set, and from a task-situation to the responses given in that situation. Our best explanations of how the relevant internal cognitive processes come to take a certain form might invoke the subject's interaction with external linguistic resources; and this interaction might be so intimate that the external linguistic resources in question constitute parts of human minds.

Clark, following the lead of Dennett and others, has done the most to develop this line of thinking (Clark 1997, 1998, 2004). In what follows, I consider three kinds of argument Clark has used: one emphasizing the internalization of dynamical structure, one focusing on the role of external language as an active control structure, and the last built on the observation that language facilitates second-order thought.

Consider first the way in which a subject's experience of external, linguistic structures can have a lasting effect on the way the subject solves problems. For instance, when introducing rules of inference, some logic instructors use as examples particular natural-language applications of the rules, to which these instructors then refer repeatedly in the weeks that follow. These natural language instances might – if they do their job – stick in a student's head to the extent that, when the student has to work a proof on her own, she replays internal copies of the sound-forms or quasi-visual images of those sentences. In some cases, such mental processing might follow an internalized causal dynamics; in the cases at hand, the process is more likely to be merely temporal (it is not as if bits of external language *cause* other bits). The student replays the sound of the instructor's voice, order intact, and her reasoning follows the structure provided by the mental representation of the instructor's voice.

Although the preceding story, and many more of its ilk, seems psychologically realistic, it lends no support to the language-based inference. Dependence and systems-based considerations

must again be given their due. It might be that learning to recognize the sex of chickens requires interaction with chickens, but that does not make the chickens a part of the chicken-sexer's mind, even if some mental images of chickens persist in the mind of the chicken-sexer and guide her future determinations. Furthermore, even if there is some useful sense in which the external language is, during coupling or periods of one-way dependence (during logic lecture, say), part of a short-lived cognitive system, there is no reason to include it as part of the cognitive system in the sense in which that system is a realization of the persisting mind. Complex, ordered mental representations of words may become part of the persisting system, but this gives us no reason to think the causal "instigators" become part of the system. External language is not, after all, ever literally *in* the organism.

There is a further reason for skepticism here. In many cases of language-dependent skill acquisition, the subject learns a generalization, something that applies beyond the training cases. The best logic students understand the idea behind, for example, hypothetical syllogism (these are the students who do not have to ask whether it matters in which order two premises of a hypothetical syllogism appear in a natural deduction proof). The nature of generalization is, of course, poorly understood. Nevertheless, flexible intelligence does not exist without the capacity to generalize. In the case at hand, it involves at the very least knowledge of what in general to look for when comparing a proposed solution to a new problem with the auditory representation of the instructor's examples. Bear in mind that in the typical case, the observed problem has little in common – linguistically speaking – with the instance originally given by the lecturer. The stored mental representation of the instructor's examples would thus seem to act more as trigger, allowing the student access to the mental representation of a general rule, than as a template that guides processing.

Sometimes Clark emphasizes the role that external linguistic resources play as an active control structure (1997, pp. 195-96, 1998, pp. 173, 181). Such a structure might be produced by the subject herself (talking herself through a problem, for example), or it might be something more like a list of written instructions given to the subject. At times when this occurs, there exists a coupled system (or a system that exhibits one-way dependence); thus, the natural question to ask is whether this case should be treated any differently from the way the content-oriented case was treated above. The current case differs in that the external linguistic units play a central role in computation, the guiding role played by important aspects of the *program* in a standard computational system.

We now face the same sort of choice we faced above in the discussion of content-dependence, version (1). Surely for the external code to do its business, it must have effects on the organism. Even if very little of the code is explicitly represented at any given time, the small portion that is explicitly represented by the organism plays a causal role, working together with whatever traces or other effects remain from earlier bits of explicit representation in the organism. Thus, in competition with the model presented by advocates of an extended mind, we have available a more conservative model, one that involves the organism's ongoing use of external instructions. Our earlier considerations speak in favor of this latter model.

Two further points might be added, though, in support of the conservative taxonomy of cognitive systems, considerations that apply specially to the proposal at hand. First, in some of the parade cases of external linguistic control – children talking themselves through complex tasks, for instance – there is an independent argument for the internal representation of the instructions: the child must have these internally encoded else she would not be able to produce the external linguistic instructions, her own spoken words. This makes it all the more plausible

that vocalization merely strengthens the role of an internally represented code, giving *it*, the internal code, control over the computational process. At the very least, this observation counteracts a tendency to think that the external linguistic units exhaust the representational resources at work. Second, note that many advocates of the extended mind are sympathetic to connectionist models (and not merely as models of the implementation of classical theories of cognition – see Fodor and Pylyshyn, 1988). If, however, this is the correct account of the human organism's cognitive architecture, the language-based inference faces further difficulty. Connectionist views claim that the cognitive system consists of simple interconnected units, their connection strengths, and various rules for the activation of those units and for the alteration of the connection strengths holding between them (there are many variations, but this is the general picture – see Rumelhart, Hinton, and McClelland, 1986). Distinctive of such models is the inseparability of data and process (Clark, 1989, p. 135). There are no encoded instructions stored at memory addresses, waiting to be called up for execution. To the extent that they are present in a connectionist system, such things as concepts, programs, and data structures are *implicit* in the system's processes – i.e., built into patterns of connectivity, connection strengths, and activation profiles. In contrast, external linguistic units are discrete, repeatable and have local causal efficacy (Clark, 2004, p. 723); these external units would seem to function, *qua* control structures, like commands in a traditional program. Thus, advocates of connectionist theories of the human organism's role in cognition face another principled reason to draw a theoretically important distinction between the organismic cognitive system and external linguistic units.

The third proposed processing-based route through the language-based inference appeals to the way in which external language facilitates higher-order thought. This approach itself

comes in two flavors, one internalist, the other externalist. The internalist approach claims that external language provides the subject with augmented computing power and new cognitive strategies by offering to the subject fixed mental units that serve as stand-ins for her own thoughts, units which can then be the object of further reflection and manipulation. External language contributes discrete orthographic and auditory units, after which internal representations are patterned by straightforward causation. Because those bits of external language express prior thoughts of the subject, and because the internal copies inherit that content, it becomes manageable for the subject to think *about* the thoughts expressed by the sentences internally copied.

This internalist way through the language-based inference is subject to the criticisms made above of dependence-reasoning and of some of the content- and control-based approaches discussed in section II. Humans value greatly the abilities acquired in the manner described in the preceding paragraph, and it is difficult – perhaps even nomologically impossible – for humans to acquire these skills absent the past effects of language. Such observations do not, however, constitute responses to our concerns about systems-individuation, causal explanation, and dependence reasoning.

The externalist option returns our attention to coupled systems and systems that exhibit one-way dependence. Clark frequently points out the extent to which our cognitive achievements are rooted in iterated interaction with various bits of external media, including language (Clark, 1997, pp. 206-7). This is partly a matter of external storage,[14] but in the present context, it is more a matter of the extent to which we actively engage with the external linguistic units that codify our previous thought processes. Here I think Clark blends the pursuit of multiple goals. He wants partly to dispel the hubris and accompanying skin-bag prejudice

(Clark's colorful term, used throughout *Natural Born Cyborgs*) of those who think they do it all themselves. Dispelling this hubris and prejudice hardly entails that the mind is extended, however. It is quite consistent to say both "I could not have written this paper without my notes" and "my notes are not part of my mind." It would be interesting to know what explicit dependence premise gets us from the first claim to a denial of the second.

In addition, Clark is moved by the *extent* of dependence, the thought being that our cognitive lives and achievements depend on the scaffolding of language in such a deep way that the connection becomes essential or constitutive. The sheer number of times one interacts with bits of language when, say, writing a substantive paper boggles the mind. Don't the external bits of language thereby become part of the cognitive system writing the paper? As argued above, this general form of reasoning is not reliable, but we might still wonder whether there is something about the cumulative and complex nature of the case that outweighs the weakness of dependence-reasoning. It seems to me that there is not. Take one instance of a paper writer's use of her notes. The previous thoughts she recorded in those notes surely were her thoughts just prior to the time she wrote them; they were, apparently, not thoughts of an extended system. Reviewing her notes now reminds her of those prior thoughts and helps her to hold in mind a complicated structure of mental representations, even if only a complicated structure of pointers that make more readily accessible material held in long-term memory (Ericsson and Kintsch, 1995). Why should this make the external reminder part of her mind, though, especially when the explanation of the relevant phenomenon – the formulation of a new or additional thought – factors into the subject's prior contribution of nonextended content and the current causal contribution of the external symbols? Furthermore, why should the nature of the explanation change simply because the process recurs, say, two-hundred times, rather than occurring only

---

[14] For concerns about extended memory, see Rupert, 2004.

once? Think of the way the process began. The subject contributed the content to the notes she made initially, when she set out to write her paper; apparently she can have *those* thoughts without notes. If she can make further notes of the thoughts catalyzed by the first-stage notes, then clearly she can have those second-stage thoughts independently of the second-stage notes she uses to write down those thoughts. In fact, she can have those second-stage thoughts independently of the first-stage notes; it is not as if she must stay in constant contact with the first set of notes in order to have the ideas that are then expressed in the second set. Reiteration introduces no extended aspect into the explanation, nothing beyond content-dependence and causal interaction.[15]

Clark also emphasizes the ubiquity of language in the human environment. The sort of ubiquity required does not seem to hold, however. It is one thing to say that some bits or other of language are frequently in the subject's environment; it is another to say that there is some particular subset of external linguistic resources that is constant in the subject's environment – enough so as to become part of her mind. Language is frequently in the air, as it were, but it is not that any particular bit of language – say, a particular set of notes – is ubiquitous in the typical subject's environment. What is ubiquitous is the subject's ability to engage with language,

[15] Driving the language-based inference in such cases might be an inference from "the final paper consists of the (nonextended) author's thoughts" to "there was a single time when the author had, clearly in mind, all of the thoughts expressed in the paper," taken together with the view that the author simply could not have had the entire structure of the paper in mind at once. But this would be a mistake. First off, the inference is bad. The attribution of written work to a nonextended author does not entail that the author ever had the entire work in mind at a single time. (Jack built a house, but he did not build it all at once.) Second, much of the work reviewed in Ericsson and Kintsch (1995) suggests that human memory is really quite impressive. This is especially clear in cases of expertise, which is the sort of case we are addressing when we talk about professionals writing papers and books. Note, too, that impressive memory capacity is not merely a freak show trick mastered only by a few (contrary to what some of Clark's remarks suggest – 2003, p. 74); the experimental work reveals such skill among everyday people operating in their own domains of expertise, for example, servers in restaurants.

whatever bits of language happen to turn up in her environment (with some limitations, of course; she might know only one language).[16]

In this essay I have returned repeatedly to two points. The first concerns scientific methodology. We must make judgments about the properties of interest to a given discipline, and about the systems that instantiate those properties – often these judgments are built into the explananda of the discipline. They are to some extent negotiable, but in the case at hand, negotiations favor conservatism in our identification of the cognitive systems relevant to the mind's location. Second, dependence-reasoning should, in general, be rejected. When properly understood and applied, these two points speak strongly against any inference from the admittedly enormous importance of language in our lives to the conclusion that the human mind extends beyond the boundary of the organism.

[16] Dennett makes the further suggestion that experience with language transforms the architecture of the cognitive system, by causing the formation of virtual machines in the connectionist wetware of the human brain; the structure of external language itself causes some of the thought processes of language-users to take the form of serial operations on discrete units (Dennett, 1991, pp. 224-25). If Dennett is right about the effects of language on language-users, this a testament to the transformative power of language use, not an argument for including language in the physical system that realizes the mind. Many factors have profound and lasting effects on children's development – parents' political attitudes, for example. But if one takes seriously the idea that the mind *has* a location – because, say, one wishes to adhere to naturalistic and materialist scruples – this transformative effect carries us no distance toward extended minds. The physical part of the parent's brain that carried her political attitudes has not literally become part of the physical system that is the child's mind.

Works Cited

Adams, F., and K. Aizawa. 2001. “The Bounds of Cognition,” *Philosophical Psychology* 14: 43-64.

Bahrick, H. P. 1979. “Maintenance of Knowledge: Questions about Memory We Forgot to Ask,” *Journal of Experimental Psychology: General* 108: 296-308.

Bahrick, H. P. 1984. “Semantic Memory Content in Permastore: Fifty Years of Memory for Spanish Learned in School,” *Journal of Experimental Psychology: General* 113: 1-29.

Ballard, D. H., M. M. Hayhoe, P. K. Pook, R. P. N. Rao. 1997. “Deictic Codes for the Embodiment of Cognition,” *Behavioral and Brain Sciences* 20: 723-42.

Bloom, P. 2000. *How Children Learn the Meanings of Words*. Cambridge: MIT Press.

Carruthers, P. 2002. “The Cognitive Functions of Language,” *Behavioral and Brain Sciences* 25: 657-74.

Carruthers, P., and J. Boucher. 1998a. “Introduction: Opening up Options,” in Carruthers and Boucher, 1998b, pp. 1-18.

Carruthers, P., and J. Boucher (eds.). 1998b. *Language and Thought: Interdisciplinary Themes*. Cambridge: Cambridge University Press.

Churchland, P. M. 1998. “Conceptual Similarity across Sensory and Neural Diversity: The Fodor/LePore Challenge Answered,” *Journal of Philosophy* 95: 5-32.

Clark, A. 1989. *Microcognition: Philosophy, Cognitive Science, and Parallel Distributed Processing*. Cambridge: MIT Press.

Clark, A. 1997. *Being There: Putting Brain, Body, and World Together Again*. Cambridge: MIT Press.

Clark, A. 1998. “Magic Words: How Language Augments Human Computation,” in Carruthers and Boucher, 1998, pp. 162-83.

Clark, A. 2003. *Natural-Born Cyborgs*. Oxford: Oxford University Press.

Clark, A. 2004. “Is Language Special? Some Remarks on Control, Coding, and Co-ordination,” *Language Sciences* 26: 717-26.

Clark, A., and D. Chalmers. 1998. “The Extended Mind,” *Analysis*

Clark, A., and J. Toribio. 1994. “Doing without Representing?” *Synthese* 101: 401-31.

Dennett, D. 1991. *Consciousness Explained*. Boston: Little, Brown and Company.

Dennett, D. 1996. *Kinds of Minds: Toward an Understanding of Consciousness*. New York: Basic Books.

Elman, J. 1995. “Language as a Dynamical System,” in Port and van Gelder, 1995, pp. 195-225.

Ericsson, K. A., and W. Kintsch. 1995. “Long-Term Working Memory,” *Psychological Review* 102: 211-45.

Fodor, J. 1975. *The Language of Thought*. Cambridge: Harvard University Press.

Fodor, J. 1981. “The Present Status of the Innateness Controversy,” in J. Fodor, *Representations*. Cambridge: MIT Press, pp. 257-316.

Fodor, J. 1987. *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge: MIT Press.

Fodor, J. 1990. *A Theory of Content and Other Essays*. Cambridge: MIT Press.

Fodor, J. 1994. *The Elm and the Expert*. Cambridge: MIT Press.

Fodor, J. 1998. *Concepts: Where Cognitive Science Went Wrong*. Oxford: Oxford University Press.

Fodor, J., and Z. Pylyshyn. 1988. “Connectionism and Cognitive Architecture: A Critical Analysis,” *Cognition* 28: 3-71.

Gathercole, S. E., and A. Baddeley. 1993. *Working Memory and Language*. Hillsdale: Lawrence Erlbaum Associates.

Gentner, D., and S. Goldin-Meadow (eds.). 2003. *Language in Mind: Advances in the Study of Language and Thought*. Cambridge: MIT Press.

Goldin-Meadow, S., and M. Zheng. 1998. “Thought before Language: The Expression of Motion Events Prior to the Impact of a Conventional Language Model,” in Carruthers and Boucher, 1998b, pp. 26-54.

Grush, R. 1997. “The Architecture of Representation,” *Philosophical Psychology* 10: 5-23.

Grush, R. 2003. “In Defense of Some ‘Cartesian’ Assumptions Concerning the Brain and Its Operation,” *Biology and Philosophy* 18: 53-93.

Houghton, D. 1997. “Mental Content and External Representations,” *Philosophical Quarterly* 47: 159-77.

Hurley, S. 1998. “Vehicles, Contents, Conceptual Structure, and Externalism,” *Analysis* 58: 1-6.

Port, R. F., and T. van Gelder (eds.). 1995. *Mind as Motion: Exploration in the Dynamics of Cognition*. Cambridge: MIT Press.

Rumelhart, D., G. Hinton, and J. McClelland. 1986. “A General Framework for Parallel Distributed Processing,” in D. Rumelhart, J. McClelland, and the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*. Vol. 1, *Foundations* (Cambridge: MIT Press, 1986), pp. 45-76.

Rupert, R. 1996. “The Best Test Theory of Extension,” Ph.D. dissertation, University of Illinois at Chicago.

Rupert, R. 2001. “Coining Terms in the Language of Thought: Innateness, Emergence, and the Lot of Cummins’s Argument against the Causal Theory of Mental Content,” *Journal of Philosophy* 98: 499-530.

Rupert, R. 2004. “Challenges to the Hypothesis of Extended Cognition,” *Journal of Philosophy* 101: 389-428.

Thelen, E., and L. B. Smith. 1994. *A Dynamic Systems Approach to the Development of Cognition and Action*. Cambridge: MIT Press.

Treisman, A. 1998. “The Perception of Features and Objects,” in R. D. Wright (ed.), *Visual Attention*. Oxford: Oxford University Press, pp. 26-54.

van Gelder, T. 1995. “What Might Cognition Be, If Not Computation?” *Journal of Philosophy* 92: 345-81.

Wheeler, M. 2001. “Two Threats to Representation,” *Synthese*, 129: 211-31.

Wheeler, M. 2004. “Is Language the Ultimate Artefact?” *Language Sciences* 26: 693-715.

Wheeler, M., and A. Clark. 1999. “Genic Representation: Reconciling Content and Causal Complexity,” *British Journal for the Philosophy of Science* 50: 103-35.

Wilson, R. 2004. *Boundaries of the Mind: The Individual in the Fragile Sciences*. Cambridge: Cambridge University Press.