Chapter 3

The Mind beyond Itself

Robert A. Wilson

1. Individualism: What It Is and Why Care

Individualism is a view about how mental states are taxonomized, clas-
sified, or typed and, it has been claimed (by, e.g., Stich, 1983; Fodor,
1980), that individualism constrains the cognitive sciences. Individual-
ists draw a contrast between the psychological states of individuals and
their physical and social environments. Psychological states are not just
"in the head" but the distinction between what is inside the head and
what is outside of it is of taxonomic and explanatory significance for cog-
nitive science. Individualism is thus sometimes called "internalism,"
and its denial "externalism."

One formulation of individualism is *methodological solipsism*, most
often associated in contemporary discussions with Putnam (1975), who
attacks it, and Fodor (1980), who defends it. Methodological solipsism
is the doctrine that psychology ought to concern itself only with *narrow*
psychological states, where these are states that do not presuppose "the
existence of any individual other than the subject to whom that state is
ascribed" (Fodor, 1980, p. 244). An alternative formulation of individu-
alism offered by Stich (1978; 1983), the *principle of autonomy*, says that "the
states and processes that ought to be of concern to the psychologist are
those that supervene on the current, internal, physical state of the organ-
ism" (Stich, 1983, pp. 164–165). To put it in plain terms, organisms that
do not differ in their "current, internal, physical state" cannot differ in
the mental states they have, and so the taxonomies we formulate for psy-
chological explanation should lead us to ascribe the same states to those
individuals.

Part of the attraction of individualism for many philosophers is its
perceived connection to functionalism in the philosophy of mind. Those
committed to functionalism as a view of the mind liberal enough to
countenance the multiple realization of mental states have often seen a
functionalist commitment to individualism as what makes such liberal-
ism compatible with a thorough-going materialism. The idea is that by

31

characterizing mental states in terms of individualistically individuated perceptual inputs, other mental states, and behavioral outputs, one can view mental states as states of an organism's central nervous system, even if they are not type identical with such states. And that should be materialism enough. Even allowing for the multiple realization of mental states in different physical states, physically identical individuals must be functionally identical and, given functionalism, functionally identical individuals must have the same psychological states. Thus, physically identical individuals must be psychologically identical – hence, individualism.

Another attraction of individualism, particularly for practicing cognitive scientists (who rightly point out the unlikelihood of encountering physically identical individuals), is its perceived connection to the representational theory of mind, which holds that we interact with the world perceptually and behaviorally through internal mental representations of how the world is (as the effects of perceiving) or how the world should be (as instructions to act). Provided that the appropriate, internal, representational states of the organism remain fixed, the organism's more peripheral causal involvement with its environment is irrelevant to cognition, since the only way in which such causal involvement can matter to cognition is by altering the mental states that represent that environment.

Many of the arguments for, and many of those against, individualism in the philosophical literature have drawn on general considerations from metaphysics, epistemology, and the philosophy of mind, of language, and of science. For example, it has been claimed by Ned Block (1986) and Jerry Fodor (1987, chap. 2) that denying individualism involves making a mystery (perhaps, more of a mystery) of mental causation, since to do so would be to posit causal mechanisms governing the operation of mental states that can vary without any corresponding variation in an individual's brain states. To take another example – deriving again from Fodor but accepted by many others including Crane (1991), McGinn (1991), and Owens (1993) – individualism in psychology is simply an instance of a general constraint on scientific taxonomy : taxonomic individuation in science must be "by causal powers," where an object's causal powers are determined by its intrinsic, physical states. I have argued elsewhere (e.g., Wilson, 1995, chaps. 2, 5–6) that all such arguments are unsound, and have called into question whether there are any general, *a priori*, valid arguments for individualism.

Given the sweeping range of philosophical issues that the debate over individualism raises, those with more empirical interests in the mind might well maintain a sceptical distance from the fray here. But cognitive scientists should care about this debate over individualism because (1) as I have suggested above, many are antecedently committed

to individualism, seeing it as a trivially correct view of psychology, (2) this commitment is reflected in, and even structures, the range of research questions, approaches, and outcomes deemed relevant to the cognitive sciences, and (c) individualism is false. In short, those with a sustained, empirical interest in the mind should care about individualism for much the reason that psychologists in the 1950s should have cared about *behaviorism*: it represents a view of what is important to progress in cognitive research that is mistakenly motivated and normatively questionable.

2. Individualism in Cognitive Science

To give the flavor of the sort of individualistic tendencies that pervade much cognitive science, consider the following quotation from Ray Jackendoff:

> Whatever the nature of *real* reality, the way reality can look *to us* is determined and constrained by the nature of our internal mental representations ... Physical stimuli (photons, sound waves, pressure on the skin, chemicals in the air, etc.) act mechanically on sensory neurons. The sensory neurons, acting as transducers in Pylyshyn's (1984) sense, set up peripheral levels of representation such as retinal arrays and whatever acoustic analysis the ear derives. In turn, the peripheral representations stimulate the construction of more central levels of representation, leading eventually to the construction of representations in central formats such as the 3D level model. (Jackendoff, 1992, pp.159–161)

Jackendoff calls this view the "psychological" (versus philosophical) vision of cognition and its relation to the world and it is perhaps unsurprising that Jackendoff's scepticism about the "philosophical" vision parallels the disdain for "philosophical" approaches to language that Chomsky (e.g., 1991; 1995) expresses in his distinction between the "I-language" and the "E-language" and his insistence that only the former is suitable as an object of scientific study. To take a third example of an individualistic perspective on cognition, consider this extract from the Foreword to Simon Baron-Cohen's *Mindblindness*, written by Leda Cosmides and John Tooby.

> Although it is a modern truism to say that we live in culturally constructed worlds, the thin surface of cultural construction is dwarfed by (and made possible by) the deep underlying strata of evolved species-typical cognitive construction. We inhabit mental worlds populated by the computational outputs of battalions of evolved, specialized neural automata. They

segment words out of a continual auditory flow, they construct a world of
local objects from edges and gradients in our two-dimensional retinal ar-
rays, they infer the purpose of a hook from its shape, they recognize and
make us feel the negative response of a conversational partner from the roll
of her eyes, they identify cooperative intentions among individuals from
their joint attention and common emotional responses, and so on.
(Cosmides and Tooby, 1995, pp. xi–xii)

While Cosmides and Tooby clearly do assign the environment of the or-
ganism a role in the evolutionary history of species-typical capacities,
the cognitive capacities themselves are individualistic. In all three cases,
the idea is that we should investigate the mind and its functions by
bracketing the world to which it is connected; the taxonomies we derive
from such a methodology will be individualistic.

One way to argue that individualism is a mistaken view of taxon-
omy and explanation in cognitive science would be to reject – lock, stock,
and barrel – research programs that adopt an individualistic view. This
is not my way. Rather, my general strategy is to accept the insights that
such research programs offer but show how they can and should be di-
vorced from a commitment to individualism. In this chapter, I will de-
ploy this strategy with respect to metarepresentation by proposing that
the metarepresentational systems that we possess and use are not indi-
vidualistic but *wide*. There are two basic ideas.

The first is that metarepresentation inherits its width from the men-
tal representations that are its objects. The second is that metarepresen-
tation often involves operating on both internal *and* external symbols,
and this suggests that our cognitive systems extend beyond the heads
that house them. Before turning to this argument, I want to draw atten-
tion to two entrenched views, the first about psychological capacities
and the second about mental representation, that are incompatible with
the position I shall advocate.

3. Minimal Reductionism
and the Encoding View

The first of these views is *minimal reductionism*, the view that psycholog-
ical capacities are nothing over and above capacities of the brain and cen-
tral nervous system. We can see how minimal reductionism supports in-
dividualism by considering a standard way of interpreting the familiar
method of homuncular functional decomposition (HFD) and how it ap-
plies to metarepresentational cognitive abilities.

HFD with respect to cognition begins with some initial cognitive ca-
pacity, C, characterized at level n, which both requires intelligence and

is representational (e.g., problem-solving, recall, visual word recogni-
tion, mind-reading), decomposes C into constituent capacities $c_1 \ldots c_j$ at
level $n - 1$, and then re-applies this initial step recursively to $c_1 \ldots c_j$ until
one reaches operations that are neither intelligent nor representational.
Materialists have typically assumed that a relation of *realization* exists
between any pair of adjacent levels, such that capacities specified at
level n-1 provide a realization of those specified at level n. Realization
is *determinative* in the following sense: where some object, process,
event, or capacity A determines B, the presence of A is sufficient for the
presence of B. Given that the properties posited at the most basic ho-
muncular level, b, are individualistic, and that b provides a realization
of properties specified at level $b + 1$, such properties must also be indi-
vidualistic. Since in moving from $b + 1$ to n the same is true of each pair
of adjacent levels, the properties specified at n must also be individual-
istic. Given that *metarepresentational* states are higher-order mental states
that refer to, or control, other (in the first instance, first-order) mental
states, they, too, inherit their individualistic character from those lower-
order mental states.

To put this point another way, since the capacities at any level m are
realized by those at level $m - 1$, there is nothing *more* to instantiating the
former than instantiating the latter, since realization is a determinative
relation. This minimal reductionism allows us to understand HFD as a
materialist or physicalist view, for it is clear that deploying this strategy
of explanation does not require positing anything in addition to physical
– in this case, neural – stuff.

The second apparently harmless view that supports individualism
is a familiar view of mental representation, which I shall call the *encoding
view* of mental representation. Simply put, it is the view that to have a
mental representation, M, is to encode information about some object,
property, event, or state of affairs m. A well-known version of the encod-
ing view is the picture or copy theory of mind, where to have a mental
representation of m is to have a mental picture or image of m in your
head, where the picture is "of M" just because it looks like m. A version
of the encoding view prevalent in cognitive science is the language-of-
thought hypothesis, according to which to have a mental representation
of m is to have a token in your language of thought, M, that stands for
or refers to m. Unlike the copy theory of mental representation, this view
requires no resemblance between the representation and the repre-
sented. On either view, because mental representations encode informa-
tion about the world, cognitive scientists can (and should) explore the
properties of these representations rather than the relationships that ex-
ist between organisms and environments. This is particularly clear in the
case of metarepresentation, where the object of the cognitive process is
itself some internal, mental representation.

I shall argue that psychological capacities are sometimes something over and above capacities of the brain and central nervous system – thus, minimal reductionism is false – and that mental representation should not be thought of exclusively as the encoding of information – thus, the encoding view is mistaken. I shall show how this view applies to metarepresentation. Furthermore, I think that the threat of two types of regress is more serious for individualistic (versus non-individualistic) accounts of metarepresentation, a view I express as a mere opinion in this paper; in the next section, I briefly gesture at the sorts of regresses I have in mind.

4. An Aside: Metarepresentation and the Threat of Regress

There are two directions – "up" and "down" – in which accounts of the mind in general and metarepresentation in particular threaten to invoke an infinite regress. The downwards threat of an infinite regress in decomposing psychological capacities is familiar: it is the threat of never quite managing to exorcise the homunculus in the head and thus failing to provide an adequately mechanistic account of the mind. In accounts of metarepresentation, the threat is to end up with a decomposition of metarepresentational capacities that is itself metarepresentational in nature. The threat "upwards" is the threat of a never-ending story and it arises in attempting to account for mental states and operations that appear to be complexes of simpler mental states and operations. Clearly, as views of a type of higher-order mental operation, accounts of metarepresentational capacities face this threat. It arises, for example, in Gricean accounts of verbal communication, where there appears to be an intricately nested hierarchy of mental states (intentions, beliefs, desires) that are encoded by speaker and hearer. This threat also arises in attempting to account for knowledge that is "implicit" in terms of inactive encodings in the brain, which may or may not be metarepresentational. The problem here is not one of failed homuncular exorcism but of failed finitude.

Why should these threats of regress be more serious given individualism? Consider the case of metarepresentation. In the case of the threat downwards, individualism makes it more likely that we will, in effect, posit a "metarepresentation box" (whether or not it has that name), that is, an undischarged homunculus that itself is metarepresentational. In the case of the threat upwards, the tie between individualism and the encoding view of representation makes it tempting for individualists to ascribe what are essentially implicit and context-sensitive metarepresentational abilities as due solely or primarily to the exercise of internal capacities.

5. Two Ways of Being Wide-Minded:
Taxonomic and Locational Width

Let me distinguish two ways in which a mind can be wide. The first is familiar, at least to philosophers, and relatively unproblematic in its application to metarepresentation. The second is less familiar and might be thought to be precisely the sort of view that individualists have been concerned (rightly, one might add) to warn us away from. However, the gap between these two ways to be wide-minded is not as great as might be thought and those interested in metarepresentation should attend to both types of width.

Since the distinction between narrow and wide mental states is one drawn with respect to how psychological states should be individuated or taxonomized, it is natural to think of wide psychological states as those psychological states, located in the organism's head, whose taxonomy presupposes reference to features of that organism's social or physical environment. Thus, though an organism's beliefs and other mental representations are located in its head, when we individuate them *as* beliefs, we do so by reference to their *content* or *intentionality*; this typically involves reference to the world beyond the organism's head (or, more generally, body). When we say that a given vervet monkey emits a certain type of alarm call because it believes that there is a leopard about, we are not giving an individualistic characterization of that monkey's mental state; it is not individualistic in the strict sense specified above because it is possible for an individual in a radically different environment to fail to have that belief, *even if* that individual were, by hypothesis, molecularly identical to the vervet in our example. (Suppose that vervets had evolved in environments where tigers fill the ecological niche that leopards, in fact, fill.) Whether there is some notion of "narrow content" with which we can reconcile this apparent rift between folk psychology and individualism is philosophically contentious.

Call the type of wideness of mind above *taxonomic width*. Apart from being taxonomically wide, psychological states may also fail to be narrow by being *locationally wide*. That is, the cognitive states of an individual may be wide because the cognitive system to which they belong is not fully instantiated or located in that individual. I have previously defended the related idea that some of the computational states that organisms instantiate are locationally wide in just this sense: they belong to wide computational systems, computational systems that extend beyond the boundary of the individual, and inherit their width from that of the system of which they are a part (see Wilson 1994; 1995, chap. 3). A simple example is that of arithmetic done with pencil and paper, where the calculations performed involve operations on mental symbols

in the head as well as on written symbols on the page. Here the relevant representational system extends beyond the boundary of the individual performing the calculation.

Given that metarepresentational states are mental states located in an individual's head whose content is another mental state, then if those mental states are *taxonomically* wide, so too are the corresponding metarepresentational states. That is, taxonomic width (like taxonomic narrowness) is inherited up the homuncular hierarchy, in this case from the first-order states of folk psychology to the metarepresentational states that operate on them. Alternatively, if the mental states that are the object of one's metarepresentational states are located somewhere other than within the boundary of one's body, then those metarepresentational states are *locationally* wide, since the metarepresentational loop they form extends beyond oneself. In either case, minimal reductionism is false: in the former case, because the kinds of metarepresentational capacities we have are determined by factors outside of the head; in the latter case, because our metarepresentational systems themselves extend into the world.

6. Locational Width and the Locus of Control

Metarepresentations are usually thought of as mental representations of other mental representations. This might make the idea of locationally wide metarepresentations seem a contradiction in terms, for the idea that organisms capable of higher-order mental representations instantiate locationally wide cognitive systems seems to fly in the face of the internal locus of control that such organisms have. In fact, we can distinguish between organisms that are increasingly sophisticated in their cognitive abilities – reactive, enactive, and symbolic organisms – in terms of where the locus of control for their cognitive systems lies, as shown in Table 1.

As Table 1 makes clear, talk of "types of organisms" here is merely a graphic way of talking about types of cognitive systems. As we move

Locus of Control	Type of Organism or Representational System	Example in Humans
environmental	reactive	reflexes
bodily	enactive	mimetic skills
cranial	symbolic	beliefs, desires

Table 1. Locus of Control and Representational Type

from reactive through enactive to purely symbolic cognitive systems, we have a shift in the locus of control from the environment through the body to the mind. Given that this is so, how can the idea of locational width get purchase in thinking about creatures with metarepresentational (and thus symbolic) capacities?

The basic answer is that there are a variety of metarepresentational processes whose locus of control is internal but whose cognitive loop extends into the world beyond the organism to whom they belong. Symbolic representational systems with an internal locus of control can be both locationally and taxonomically wide. Consider two ways to develop this answer.

First, let us return to the idea that mental states are realized as physical states of the brain. Consider humans, who possess reactive, enactive, and symbolic cognitive systems. What is it, then, that constitutes a realization[1] of their various cognitive capacities? In the case of "purely mental" capacities, it is the brain and, more particularly, the cognitive arrangement of the brain. In the case of what I shall call *enactive bodily skills*, such as rehearsing a tennis stroke or flossing one's teeth, it is the parts of the brain responsible for planning and executing motor control and the relevant parts of one's body. In the case of what I shall call *wide symbol systems*, involving conventionally established codes and symbols, it is the brain plus the physical tokens of the corresponding external symbols. In short, creatures like us, who possess cognitive systems with an internal locus of control, can also instantiate cognitive capacities with core realizations that are either internal, bodily, or world-involving (see Table 2).

Table 2 should help to clarify a point that I intimated at the end of section 5: the locational width of world-involving cognitive capacities is incompatible with minimal reductionism about cognition in general. It should also indicate what is problematic about the encoding view of mental representation, for neither in the case of enactive, bodily skills nor in that of world-involving capacities do parts of the brain encode for the other constituents of the realization of that capacity. Rather, in both cases what is inside the head and what is outside of it are related as parts of an integrated whole, with information flowing between those parts.

Cognitive Capacities in Symbol-Using Creatures	Realization of the Capacity
purely internal	internal cognitive arrangement of the brain
enactive bodily	cerebral + bodily configuration
world-involving	cerebral arrangement + external symbol tokens

Table 2. Cognitive Capacities and Their Realizations

		Locus of Control	
		Internal	**External**
Wide	Locational	arithmetical abilities animal navigation (*Gallistel* 1990)	cyborg cognition, e.g., human navigation (Hutchins 1995)
	Taxonomic	folk psychology	sustained manipulation of others
Narrow		most subpersonal psychology and cognitive science	stimulus-bound perceptual performance

Table 3. Some Examples of Narrow and Wide Psychology

Second, suppose we simply distinguish between the case where the locus of control is internal to the organism and that in which the locus of control is external. That, together with the distinction between taxonomic and locational width, gives us four possible types of wide cognitive systems. I think that there are examples of each of these types of cognitive system (see Table 3), despite the fact that the bulk of cognitive science has been conceived of as involving only narrow cognitive systems.

In the remainder of this chapter, I shall indicate how the distinctions summarized in Tables 2 and 3 illuminate contemporary work with a metarepresentational edge. Central to these illustrations is the idea that we should locate some of the psychological capacities typically identified as "purely internal" (in Table 2) and as "narrow" (in Table 3) elsewhere in the respective tables.

7. Memory

A focus of current debates over memory has been the relationship that research into memory bears to everyday memory; my illustrations of the ways in which memory is neither "purely internal" nor "narrow" are made with this focus in mind. In particular, debates over the classic "storehouse" model of memory and alternatives to it suggest ways of extending our traditional conception to show how memory itself is wide.

First, Koriat and Goldstein (1996) propose that two metaphors have structured research into memory over the last 100 years: the familiar storehouse metaphor, and what they call the "correspondence metaphor," according to which it is of the essence of memory to correspond

to some past state of affairs, rather than simply to act as a storehouse for readily identifiable chunks of data. This characterization of the alternative view to the dominant conception of memory invokes a taxonomically wide conception of memory. That is, what individuates memory in general from other types of cognitive processes (e.g., imagination, fantasy, wishes) is the relationship memory bears to past, experienced states of affairs in the world. Moreover, on this view, what individuates a particular memory from others is, at least in part, what it is a memory of or about – that is, its intentionality. This represents a second way in which use of a correspondence metaphor relies on a taxonomically wide conception of memory.

Second, and by contrast, consider Neisser's (1996) claim that the underlying metaphor that structures research into "real-life" memory is one of "remembering as doing." On this view, such research may even adopt a *locationally* wide conception of memory since it views memory as performative or enactive, where what is enacted does not simply stop at the skin but involves engaging with the world through cognitively significant action. To adapt Neisser's own examples, to tell a joke or recite an epic tale is not simply to make certain mouth and body movements but to "impress and entertain" an audience. One way to state this point is to say that, if memory is a bodily, enactive skill and those skills are individuated widely, then memory is taxonomically wide. But, we might also conceive of such memory itself as extending into the world through the way in which it engages with and appropriates external systems, treating them in just the way that internal symbols are treated.

The idea that procedural memory may involve doing things with one's body, while itself old hat, does suggest an idea that seems more novel: that one may remember *by doing things with one's environment*. Perhaps even this idea is old hat; after all, we all know that we can use environmentally cued mnemonics, such as tying a piece of string around one's finger or leaving a note on the refrigerator. My suggestion is that, apart from these promptings to remember, there are ways of remembering that involve a sustained, reliable, causal interaction between an organism and its symbolic environment. The magnitude of our symbol-laden environments should be taken seriously and to do so is to see the mind as extending beyond itself, that is, as being constituted by such symbols and thus as locationally wide.

This wide perspective is most compelling in cases in which systems of external symbols come to change in significant ways the cognitive capacities of individuals who interact with them. Most striking here are cases in which an organism, such as a bonobo, develops in a symbol-enriched environment and subsequently displays massively increased symbolic capacities (Savage-Rumbaugh & Lewin, 1994). For example, given Kanzi's actual developmental environment, Kanzi plus a 256-

symbol keyboard forms a cognitive system with memory and other cog-
nitive capacities that far exceed those of just Kanzi. (Much the same
holds true of Alex, Irene Pepperberg's African grey parrot.) My point
here is not the trivial one that enriched environments can produce
smarter "critters"; rather, it is that what makes at least some "critters"
smart is their being part of wide cognitive systems. Again, we have the
mind beyond itself.

Neither the case of taxonomic width nor that of locational width in-
volves viewing memory as having an external locus of control. Indeed,
maintenance of an internal locus of control would seem crucial in such
cases in order to explain the sense in which a person's memories are *that*
person's memories rather than a part of some collective memory or
(worse) free-floating cognitive flotsam.[2] We do, of course, have forms of
external memory storage, such as diaries, which, while deriving from,
and often recording, in part, an individual's mental life, can exist beyond
the life of their author and come to be accessed as a form of memory by
others. Further, each time any one of us speaks or writes in a communi-
cative context, we create tokens of external symbols that constitute a
common symbol pool from which each of us draws. To be sure, acts of
communication always involve the internal mental representations each
of us houses but my point is that they also involve *in just the same sense*
public and shared representations that are not the province of any one
mind. In many such cases, the locus of control may be internal to the
speaker (in uttering) and to the hearer (in listening) but I see no incoher-
ence in the idea that the locus of control may be outside of the head.
Cases of fluent reading (e.g., speed reading) and manipulation (see §9,
below) may best be viewed as examples in which this actually occurs.

Column 3 of Table 4 presents an extension of Table 2 that summa-
rizes these ideas. As that table indicates, enactive, procedural memory
that is locationally wide is an extension of traditionally conceived pro-
cedural memory, just as external memory is an extension of tradition-
ally-conceived long-term memory. In both cases, we have a conception
of memory that takes the mind beyond itself.[3]

If the view of memory as locationally wide provides a fruitful ex-
tension of traditional conceptions of memory, then it is obvious how
metarepresentation itself may be either taxonomically or locationally
wide: it is wide in one of these senses just when some of the representa-
tions on which it operates are locationally wide in the corresponding
sense. In such cases, metarepresentational capacities are, indeed, some-
thing over and above capacities of the brain and central nervous system
(minimal reductionism is false); and metarepresentation is not simply
the encoding of information about other internal mental states but the
formation of metarepresentational loops beyond the head in which seg-
ments of those loops are realized.

Cognitive Capacities in Symbol-Users	Realization of the Capacity	Adding Memory	Adding Theory of Mind	Adding Culture
purely internal	architectural + non-architectural features of the brain	traditional forms of memory, e.g., declarative/ procedural, LTM/STM	bare-bones FP	mental image of God? moral intuitions?
bodily, enactive	cerebral + bodily configuration	extension of procedural memory	full-blown FP	dance, played melody
world-involving	internal + external symbol tokens	extensions of other forms of memory to include external symbol systems	narrative engagement, sustained deceit and manipulation	street signs, maps, instruments and tools

Table 4: Adding Memory, Theory of Mind, and Culture

8. Theory of Mind

We are mindreaders. The explosion of work over the last 15 years in both cognitive development and primatology exploring the developmental and evolutionary origins of this ability has largely construed the capacity itself as a theory of mind, a theory that attributes folk-psychological states to agents and that allows one to predict and explain an agent's behavior in terms of the relationships between those states, perception, and behavior. I want to focus my discussion of the theory of mind on the end-state of these ontogenetic and phylogenetic processes, the folk psychology that we end up sharing and relying on in everyday life, beginning with what we can think of as our bare-bones folk psychology, the psychology of belief and desire.

The capacity that normal human adults have to ascribe belief and desire to one another is both locationally narrow and taxonomically wide. It is locationally narrow because the realization of the capacity is purely internal to the individual who has the capacity. On the other hand, it is taxonomically wide because beliefs and desires are individuated, in part, by their intentional content – that is, what they are about – and such content is wide. This is so whether one thinks that this ascriptive ability operates via a theory or via acts of imaginative simulation. Matters are less straightforward, however, when one considers both the full-blown capacities that we have for engaging in folk-psychological explanation and some of our more advanced deployments of folk psychology.

First, take full-blown folk psychology, which posits not only beliefs and desires but a whole range of psychological states, such as emotions (anger, elation, fear), moods (restless, aroused, inattentive), and sensations (of pain, of experiencing red, of tickling). Although these additional states are by no means homogenous, it is much less plausible to think that the realization of the capacity to ascribe them is purely internal than in the case of belief and desire. That is because these states have a felt component, whether it be experiential or bodily (or both), and it is difficult to see how one could accurately and reliably ascribe such states to others without knowing what they were like in one's own case. Further, such knowledge itself is procedural and has a bodily realization in that it involves not simply having one's brain in some internal state but, at least, having one's brain and body in a certain state.

Prima facie, this would seem to make a simulation view more plausible than a theory view of full-blown folk psychology, since it is easier to see how such experience can be integrated into simulation-based ascription than into theory-based ascription. The most obvious ploys for proponents of the theory view of folk psychology are to argue that (1) full-blown folk psychology can be reduced to bare-bones psychology, or (2)

however important experiential and bodily aspects are to the acquisition of folk psychology, they do not form part of its realization, which is purely internal. While argument (1) appeals to minimal reductionism, argument (2) employs the encoding thesis, and so the plausibility of each of these responses will depend, in part, on the plausibility of these two views.

My claim, then, is that the move from bare-bones to full-blown folk psychology involves a shift from a purely internal mental capacity to a bodily enactive skill. I also want to suggest that some of our most sophisticated deployments of folk psychology – such as understanding a complicated narrative about the mental lives of others and manipulating another's full-blown folk psychology – involve a symbolic capacity that is world-involving. In such cases, folk psychology starts to look not just taxonomically but also locationally wide.

Consider narrative engagement that involves understanding the full-blown folk psychology of characters in a literary, dramatic, or cinematic genre. To understand, say, a novel such as *Pride and Prejudice* or *Bleak House* one must not only ascribe full-blown folk-psychological states to the characters in the novel but also understand those characters' (partial) views of the world, a world that naturally includes other people. As you read deeper into the novel, you must, of course, modify your representations of the folk-psychological representations that each character has. But since the metarepresentational load here increases dramatically with the complexity of the portrayal of the characters and their relationships to one another, it is no surprise that even partial expertise typically involves knowing how to find one's way about in the novel; it involves knowing how to locate and identify the folk-psychological representations that respective characters have and the signs of these in the novel itself. Here the representations that are the object of your own representations are located somewhere other than in your own head. In short, this understanding involves constructing a metarepresentational loop that extends beyond the head and into the minds of the fictional characters with whom you are engaged.

Much the same is true of appreciating the full-blown folk psychology of real people, especially those to whom you are close. Our (meta)representations of the mental lives of companions and friends are more sophisticated not simply because of the added internal complexity such representations have in our own heads but also because they index richer mental representations in the minds of one's companions than those in the minds of strangers. Rather than simply encoding information about these mental representations, we engage and interact with them and, in so doing, extend the system of mental representations to which we have access beyond the boundary of our own skins. As with our reliance on cognitive artifacts to bear some of the representational load carried during a complicated cognitive task, here we exploit rather

than replicate the representational complexity of our environments. But unlike at least some such tasks (e.g., navigating a vessel or an aircraft, to take an example from Hutchins, 1995), we individuals remain the locus of representational control, with our interactions with external representations augmenting our internal representational systems.

Both the case of narrative engagement and that of locationally wide, full-blown folk psychology involve metarepresentational capacities whose locus of control is still, by and large, internal. Cases in which there is sustained deception that involves folk-psychological manipulation of an individual or those involving an individual's blind trust in the views of others are plausibly viewed as operating with a locus of control that is external to that individual. As with these other forms of locationally wide folk-psychological systems, the cognitive capacity here involves the world, with the relevant folk-psychological representations being located both inside and outside of a given individual's head. Considering just the perspective of the deceiver, manipulator, or person trusted, the locus of control here remains internal. But, from the perspective of the deceived, the manipulated, or the person trusting, their metarepresentational folk-psychological states are controlled by folk-psychological states beyond their own mind.

The fourth column of Table 4 summarizes how to fit theory of mind into the wide framework; looking ahead, the fifth column of that table does the same for culture.

9. Culture, Cognition, and Evolution

The ideas about memory and theory of mind that I have been floating question the peripheral or secondary role that individualists give to an organism's environment. In particular, I have argued that some of the metarepresentational capacities we have are world-involving rather than purely internal in that the realization base for the capacity extends beyond the boundary of the individual. The idea that external symbols come to form part of an individual's cognitive system is central to this argument. Culture represents a source for such symbols.

One general concessionary response to the above suggestions regarding memory and the theory of mind is to allow the world beyond the organism a role in a broader cultural understanding of metarepresentation but no such role in metarepresentational *cognition*. There are various ways to develop this idea, all of which ultimately turn on something like the encoding view of mental representation. For example, one might allow that certain cultural conditions are necessary (e.g., numerical systems) for specific cognitive capacities to emerge (arithmetical abilities) but deny that the cognitive capacities themselves are either tax-

onomically or locationally wide. Alternatively, one might see cultural in-
novations either as causes or effects of cognitive adaptations but never
as constitutive of them. In either case, culture "beyond the head" is rel-
evant to cognition only insofar as it becomes encoded by individual
minds. These sorts of concessionary responses presuppose problematic
and impoverished conceptions of culture and of the relation between
culture, cognition, and evolution. I shall argue this point by exploring
some of the broad options open to one in adopting a naturalistic ap-
proach to culture and its relationship to cognition and evolution.

Classic sociobiological theory (e.g., Wilson, 1975; 1978) attempted to
account for "social behaviors" in a variety of animal species, including
our own, through what I will call a *direct extension* (DE) model of evolu-
tionary explanation.[4] That is, classic sociobiology took individual social
behaviors as phenotypes and then applied the theory of natural selection
to them, in effect explaining the persistence of those social behaviors as
adaptations selected for their contribution to an organism's reproductive
fitness. DE models construe heritability *genetically* and fitness *reproduc-
tively.* In addition, the theory of natural selection was typically under-
stood by sociobiologists in terms of genic selection. This meant that genes
for those social behaviors had been selected for their contributions to fit-
ness, making sociobiological accounts individualistic about the causes of
social behavior in something close to the sense in which most approaches
to the mind within cognitive science have been individualistic.

Implicit in the sociobiological approach was an *aggregative* view of
culture that corresponds to minimal reductionism about the mind: that
cultural phenomena could be conceived of as the aggregation of indi-
vidual social behaviors, and so are nothing over and above those indi-
vidual acts. Thus, to take a classic example, the existence of "homosexual
cultures" within certain societies was conceived of as an aggregation of
individual agents who engaged in homosexual behavior. To explain cul-
tural phenomena in evolutionary terms was simply to explain the be-
haviors of individual agents. Despite the recognition of various inade-
quacies in the sociobiological approach to "social behavior" (the
anthropomorphism of its taxonomies; its omission of "the missing link"
of psychology; an uncritical adaptationism; its insensitivity to philo-
sophical questions about explanation), the aggregative conception of
culture on which it relied is still predominant, particularly in other DE
models of cultural evolution. And, like the genic version of sociobiology,
alternative approaches to cultural evolution have remained individual-
istic in their orientation.

This seems clearest in the case of evolutionary psychology, which,
when most akin to cognitive science (e.g., Tooby & Cosmides, 1992), at-
tempts to identify "Darwinian algorithms" that underlie species-specific
adaptations in an organism's cognitive architecture. Evolutionary

psychology encapsulates a DE model of the evolutionary explanation of culture insofar as it construes the specific structures of the cognitive architecture as phenotypes that are selected for their contribution to reproductive fitness. While the process that maintains such phenotypes in the species – natural selection – is itself taxonomically wide and involves an *external locus of control, the relevant phenotypes themselves are con-*strued individualistically in that they are instantiated in brain structures taxonomized *computationally.*[5] Evolutionary psychology is aggregative about cultural phenomena just as classic sociobiology was, except that what is aggregated are not the individual social behaviors but the individual Darwinian algorithms underlying these.

Models of cultural evolution that rightly see culture related to (social) behavior through psychology need not and, in some cases, should not also adopt the aggregative and individualistic view of culture that prevails in DE models. This is both because DE models can themselves be wide and because there are alternative models of cultural evolution that make aggregativity and the accompanying individualism more difficult to sustain.

To take the former of these options first (and more briefly): DE models can be taxonomically or even locationally wide because the psychological capacities they purport to explain may be taxonomically or locationally wide. Many of an organism's adapted phenotypes are taxonomically wide (e.g., running faster than predators, being camouflaged, detecting prey effectively) and locationally wide (e.g., beavers + beaver dams, hermit crabs + the shells they inhabit, birds + the nests they build). Given that physical and behavioral phenotypes can be extended (*sensu* Dawkins, 1982), it is not clear why one should make an exception of psychological phenotypes; the idea that they are special in stopping at the skin relies on a dubious Cartesian separation between the physical and behavioral, on the one hand, and the psychological, on the other.

One reason for spending little time in pursuing wide options for DE models of cultural evolution is that DE models themselves are limited. The transmission of many aspects of culture (e.g., eating with a knife and fork, smoking, hand-shaking) has, at best, a tangential connection to reproductive fitness, and those whose promulgation *is* tied to reproductive fitness are often so tied only given the existence of wide cognitive capacities that are dependent upon culture. For example, while our capacities for sharing of food, selection of mates, and social exchange *do* bear a closer connection to reproductive success, each of these is a wide capacity (or, if you like, a narrow capacity whose adaptive significance lies in its being deployed in certain cultural environments rather than others). Likewise, I shall ignore what I call *heritability generalization* (HG) models, which add culturally mediated mechanisms of heritability (e.g., imitation, learning) to genetic mechanisms but maintain the reproduc-

tive fitness of individuals as the relevant notion of fitness. Instead, I want to comment on two other alternatives, each of which departs further from DE models than do HG models.

First are *fitness generalization* (FG) models, which as well as extending the mechanism for transmitting phenotypes also generalizes the notion of fitness. In these models there is a focus not simply on the reproductive fitness of individuals who carry certain memes but on the memes themselves and their "fitness," where this may conflict directly or indirectly with the reproductive fitness of individuals. The fitness of a meme is its capacity for transmission, where this clearly depends not only on its intrinsic properties and the cognitive architecture of individuals who transmit it, but also on features of the cultures in which the meme exists and the contexts in which it is expressed. To put it colloquially (but accurately), how catchy a meme is depends on who, how, where, and when it is expressed.

Second are what have been called *contagion* models of cultural evolution of the sort that Dan Sperber has defended over the last 10 to 15 years. Sperber's emphasis is on the ways in which the process linking cultural units is not one of replication but instead is *transformational* in its nature; this makes him sceptical of the notion of a meme, deriving as it does from that of a gene. As Sperber (1996) makes clear, there is dynamic interplay between mental and public representations in processes of cultural transmission and it is the inattention to the processes whereby mental representations are actually transmitted that gives FG models their superficial plausibility.

Although both FG and contagion models of cultural evolution give a central role to mental representations and the mental structures that underlie their processing, it is easy to see both as offering locationally wide accounts of the basic cultural units that are transmitted over time within a culture.[6] There are two ways in which these models posit locationally wide cognitive components of cultural transmission. The first is parallel to the way in which enactive procedural memory and external memory are locationally wide. These each involve extending the cognitive capacities of the mind of an individual. Likewise, we can see aspects of a culture as offering such an extension of an individual's cultural-processing hardware, the most obvious of which is technology in general and cognitive artifacts (maps, other people, computers) in particular. The second involves shifting the locus of control from the individual to some larger unit of which that individual is a part, as Hutchins (1995) has suggested is the case with human navigation. The final column of Table 4 makes these options more graphic.

Details about the cognitive architecture underlying the cognitive capacities that allow us to transmit cultural units will certainly take us part of the way, but both FG and contagion models identify more than the

cognitive architecture as the "missing link" between behavior and fit-
ness; included also are features of the resulting mental and external rep-
resentations themselves (e.g., their rhythm), acquired and non-universal
features of the mind (e.g., learning Chinese rather than English numer-
ical systems), the availability of specific technological resources (e.g., the
existence of particular writing systems), and non-representational facts
about the structure of the cultures in which representations are trans-
mitted (e.g., class structure, sexual division of labor). This makes the re-
lationships between culture and cognition more complex than implied
by the idea that the latter encodes the former and suggests a bidirectional
causal relationship between the two that is incompatible with the aggre-
gative view that individualists adopt.

Acknowledgments

Notes

1 I should perhaps note that here, relying on Shoemaker's (1981) distinction be-
   tween core and total realizations of mental states, I am asking a question
   about the core realization of various cognitive capacities. I have generalized
   this distinction elsewhere to argue that many "higher-order" states, includ-
   ing mental states, have a wide realization and that this has implications for
   both reductionist and non-reductionist views in the philosophy of mind. See
   Wilson (submitted).
2 This is not to suggest that being realized inside the head is sufficient for having
   an internal locus of control. Interesting, in this respect, is the discussion of
   George Graham and G. Lynn Stephens (1994) of psychopathologies that in-
   volve a feeling of a loss of control of the ideas that occur in one's own head.
3 I am indebted here to Merlin Donald (1991, esp. chap. 8) on the idea of external
   symbol storage and the conception of it as an extension to traditional concep-
   tions of memory.
4 Although the name "direct extension" and the others that I use below are my
   own, the ideas are drawn and generalized from Elliott Sober (1991).
5 Of course, if the argument of my "Wide Computationalism" is correct
   (Wilson, 1994), then this inference from computationalism to individualism
   is invalid.

6 I doubt that this is Sperber's own view of either contagion or FG models of
cultural evolution. In a number of places (e.g., Sperber, 1996, pp. 62, 96–97),
Sperber himself indicates that intraindividual representations have some sort
of priority over public representations in thinking about culture, the former
being "psychological" and the latter "ecological."

References

Block, Ned (1986). Advertisement for a semantics for psychology. In P. French,
T. Vehling, & H. Wettstein (Eds), *Midwest studies in philosophy: Vol.10. Phi-
losophy of mind*. Minneapolis, MN: University of Minnesota Press.
Chomsky, Noam (1991). Linguistics and adjacent fields: A personal view. In A.
Kasher (Ed.), *The Chomskyan Turn*. Cambridge, MA: Basil Blackwell.
Chomsky, Noam (1995). Language and nature. *Mind 104*, 1–61.
Cosmides, Leda,& Tooby, J. (1995). Foreword to S. Baron-Cohen, *Mindblindness*.
Cambridge, MA: MIT Press.
Crane, Tim (1991). All the difference in the world. *Philosophical Quarterly 41*, 1–25.
Dawkins, R. (1982). *The extended phenotype*. Oxford: Oxford University Press.
Donald, M. (1991). *Origins of the modern mind*. Cambridge, MA: Harvard Univer-
sity Press.
Fodor, J. A. (1980). Methodological solipsism considered as a research strategy
in cognitive psychology. *Behavioral and Brain Sciences 3*, 63–73. Reprinted in
his *Representations*. Sussex: Harvester Press, 1981.
Fodor, J. A. (1987). *Psychosemantics*. Cambridge, MA: MIT Press.
Gallistel, C. R. (1990). *The organization of learning*. Cambridge, MA: MIT Press.
Graham, G., & Stephens, G. L. (1994). Mind and mine. In their *Philosophical Psy-
chopathology*. Cambridge, MA: MIT Press.
Hutchins, E. (1995). *Cognition in the wild*. Cambridge, MA: MIT Press.
Jackendoff, R. (1992). The problem of reality. In his *Languages of the mind*. Cam-
bridge, MA: MIT Press.
Koriat, A., & Goldstein, M. (1996). Memory metaphors and the real-life/labora-
tory controversy: Correspondence versus storehouse conceptions of mem-
ory. *Behavioral and Brain Sciences 19*, 167–188.
McGinn, C. (1991). Conceptual causation: Some elementary reflections. *Mind,
100*, 573–586.
Neisser, Ulric (1996). Remembering as doing. *Behavioral and Brain Sciences 19*,
203–204.
Owens, J. (1993). Content, causation, and psychophysical supervenience. *Philos-
ophy of Science 60*, 242–261.
Putnam, H. (1975). The meaning of "meaning." In K.Gunderson (Ed.),*Language,
Mind, and Knowledge*. Minneapolis, MN: University of Minnesota Press.
Reprinted in his *Mind, language, and reality*. New York: Cambridge Univer-
sity Press.
Savage-Rumbaugh, S., & Lewin, R. (1994). *Kanzi: The ape at the brink of the human
mind*. New York: Wiley and Sons.
Shoemaker, S. (1981). Some varieties of functionalism. *Philosophical Topics 12* (1),

83–118. Reprinted in his *Identity, cause, and mind*. New York: Cambridge University Press, 1984.

Sober, E. (1991). Models of cultural evolution. In P. Griffiths (Ed.),*Trees of life*. Dordrecht, NL: Kluwer Academic Publishers. Reprinted in E. Sober (Ed.),*Conceptual issues in evolutionary biology* 2nd ed. Cambridge, MA: MIT Press, 1994.

Sperber, D. (1996). *Explaining culture*. Cambridge, MA: Basil Blackwell.

Stich, S. (1978). Autonomous psychology and the belief-desire thesis.*Monist 61*, 573–591.

Stich, S. (1983). *From folk psychology to cognitive science*. Cambridge, MA: MIT Press.

Tooby, J., & Cosmides, L. (1992). *The psychological foundations of culture*. In J. Barkow, L. Cosmides, & J. Tooby (Eds.),*The adapted mind*. New York: Oxford University Press.

Wilson, E. O. (1975). *Sociobiology*. Cambridge, MA: Harvard University Press.

Wilson, E. O. (1978). *On human nature*. Cambridge, MA: Harvard University Press.

Wilson, R.A. (1994). Wide computationalism. *Mind 103*, 351–372.

Wilson, R.A. (1995). *Cartesian psychology and physical minds: Individualism and the sciences of the mind*. New York: Cambridge University Press.

Wilson, R.A. (submitted). Two views of realization.