

At Two with Nature: Agency and the Development of Self-World Dualism

James Russell

What role does agency play in mental development? Jean Piaget (e.g., 1970) famously answered that agency is the engine of mental development. However, the past 25 years of cognitive-developmental research have given rise to a widespread dissatisfaction with that answer, mainly because we have unearthed too much innate apparatus for it to be plausible. Yet I will argue that Piaget's answer is only incorrect insofar as it is too strong. There is a case to be made, in other words, for the view that without agency there can be no mental development, or at least that adequate agency is necessary for adequate mental development. I will try to give this more modest Piagetian view a run for its money.

Like many psychologists of his and the preceding generation (notably James Mark Baldwin¹), Piaget regarded mental development as the process of establishing a division between two kinds of reality: an objective reality grasped as independent of ourselves and a subjective reality constituted by our volitions and representational states. I will refer to this with the usual term 'self-world dualism', a mode of consciousness that one might regard as the symmetrical opposite of being-in-the-world. Self-world dualism is the basic human situation. For although most of us understand what it means to be at one with Nature, we only do so because our normal state is—as with the Woody Allen character who would stubbornly remain on rural excursions—"at *two* with Nature."

In presenting and defending a watered-down version of Piaget's developmental account of self-world dualism, I will proceed through the following stages. First, I will say why Piaget's theory is too strong and sketch the aims of the weaker version. Next, I will present the case for agency being necessary for the development of self-world dualism. My principal claim here will be that it is only by experiencing agency that a subject can experience the world as being resistant to her will, something which is necessary if any distinction is to be drawn between subjective and objective. I argue here that an essential feature of agency is a capacity for

willfully determining the sequence of one's perceptual inputs. The latter two sections of this paper will concern the implications that this thesis has for infants' developing conceptions of physical objects (normally called 'object permanence') and for young children's conceptions of other minds (normally called 'theory of mind'). In the first case, the argument is that object permanence must be properly assessed in terms of the actions the infant spontaneously performs upon objects. In the second case, I argue that the central role of agency in subjectivity entails that we cannot explain development of a theory of mind solely in terms of the maturation of an innately specified processing module.

1 Weakening the Piagetian Thesis

Piaget was concerned, in part, with the preverbal roots of a dualism between the subject's experience of how things appear at particular times and her conception of a world of enduring, mind-independent objects, and consequently he studied infants' understanding of object occlusion, "object permanence." The infant's senses may tell it, for example, that all that exists within its reaching space on the carpet right now is a cushion, while, in reality, there is both a rattle and a cushion, the former occluded by the latter because of the infant's relative position in space as the viewer. In essence, Piaget argued that objective experience and our conception of ourselves as experiencers located within a mind-independent spatial world emerge as our actions become progressively more self-determined, differentiated, and integrated. He opposed nativist accounts, and in his writings on infancy in particular, he strenuously rejected nativism about spatial, causal, temporal, and object concepts, a nativism that he ascribed to Kant.

There are major difficulties with this position, quite apart from the question of what is and what is not to count as an action. Take the case of spatial concepts—the fundamental concepts for Piaget and concepts that relate intimately to object permanence. His view was that through exercising progressively greater control over what it experiences, the infant becomes able to bootstrap itself towards an allocentric, rather than purely egocentric, conception of space.² That is to say, by forging links between her actions and their perceptual outcomes, it develops a conception of the spatial world that is perspective-independent, environment-centred, and maplike (Piaget 1955, 198–218).

There are good reasons for dismissing the very possibility that development could happen in this way. As James Hopkins (1987, 153–154) has argued, if the primordial state of the infant mind is entirely adualistic, then no amount of activity will make the infant's experience become that of an

objective spatial world: there is nothing on which to build. If there is no initial distinction between (say) hunger as having an inner source, and a looming object as having an outer source, then it is impossible to see how the development of agency could forge that distinction. In the second place, Piaget's (1955) picture of the infant as heroically "constructing" a maplike, perspective-neutral conception of space by acting on objects—as if this were something that the organism *wins through* to after a long apprenticeship—sits awkwardly beside evidence that an ability to code spatial relations allocentrically is both phylogenetically and ontogenetically early. Lynn Nadel (1990), for example, has demonstrated that 3-week-old rats are able to swim straight to a platform concealed in a pool of opaque liquid from a *novel* starting point, something they must achieve through coding the relations between the platform and landmarks in the environment. In fact, this finding is one of many that one could cite in support of O'Keefe and Nadel's (1978) hypothesis that animals possess abilities for spatial cognition that are innately specified (within the hippocampus). The fact is, nativism about allocentric coding is a good deal more plausible than constructivism, so there are both conceptual and empirical considerations ranged against the Piagetian view.

But the main point I need to make in this section is that a very wide scope for a broadly Piagetian account of self-world dualism still remains even after one has accepted that organisms' spatial awareness is (or must be) grounded in environment-centred coding. The reason is that, although this allocentric grounding may be necessary for self-world dualism, it is certainly not sufficient for it. And this is because self-world dualism is a more psychologically rich affair than the ability to code spatial relations as being independent of one's current location and activity. Rich in what sense? In human beings, at least, one can regard this dualism as requiring a primitive form of the distinction between appearance and reality insofar as it involves our distinguishing between our perspectives on the world, perspectives that can never tell more than partial truths, and the whole truth about the way things are. The world always seems to us to be a certain way, while we know that if we were differently located, it would seem otherwise, and we assume that what affords us these different perspectives is a mind-independent physical reality. It is fair to call this a 'theorylike' conception, in the sense that it is a matter of reflective understanding rather than of practical competence.

We are now in a position to see where a Piagetian account can enter this picture. It is possible to argue that agency is necessary for the development of self-world dualism, in the sense in which I have just described it, while being as nativist as one could wish about allocentric coding.

Henceforth I will refer to this weaker Piagetian position as 'piagetian' (with a small 'p'). And in general I shall take a piagetian position to be one in which a wide range of innately specified apparatus can be posited but in which one also argues for agency being necessary (though not sufficient) for developing a conception of objects and the self as distinct.

But as a way of placing the piagetian view more definitively, I need to say a little about the distinction between practical and theoretical abilities, again taking spatial abilities as the test case. As John Campbell (1993) has argued, the innate ability to construct perspective-neutral maps of the environment will not enable a creature to register the spatial connectedness of this environment (the fact that each place is related to every other) unless it can appreciate the causal significance of its own actions and perceptions in relation to places (e.g., by navigating)—an appreciation dubbed a "causally indexical" understanding by Campbell. This is not yet theoretical. But a "causally *non*indexical" understanding *is* theoretical in the sense that it is one through which the creature is able to reflect both on the causal relations between itself and objects (including its perceptions of them) and on the relations between physical objects in its environment.³ This, then, is how the distinction can be drawn between an organism's practical abilities for interacting with objects—an organism surely does not *require* self-world dualism to find its way around—and the kind of reflective and theorylike understanding at issue in self-world dualism. The way in which the creature's location affects its experiences and its ability to reflect upon how its actions affect these is of central importance here.

Given this, piagetians claim to have a story to tell about how this theorylike conception of the organism's place in the physical world must arise, in particular about its grasp of the relations between objects and its perceptual experience of them. They argue that *our ability to change the nature of our perceptual inputs at will*—a broad definition of 'action', which includes attention shifting—is necessary for us to make the kind of appearance/reality distinction that is at the heart of self-world dualism and is at least a component of any grasp of ourselves as representers of an environment.

Just a word, before proceeding, about the intended status of what I am going to say. Piaget borrowed J. M. Baldwin's term 'genetic epistemology' to describe his enterprise, and indeed much of his work was a hybrid of developmental psychology and epistemology. Accordingly, psychologists who work within this framework tend to seek out so-called "transcendental" arguments (roughly, unless we had cognitive capacity x , we could not experience y ; we do experience y ; therefore we have x). As Patricia Kitcher (1990) has argued in her recent book on Kant's psychology, the line between philosophical and psychological approaches to questions about

the conditions for being a subject with objective knowledge (transcendental questions) is difficult to draw.⁴ As this is the kind of question I am asking here, my answers will inevitably have a rather ambiguous status.

2 What Agency Gives Us

In this section I will argue for the view that self-world dualism can only develop in agents. What are the relevant features of agency here?

In the first place, there are certain kinds of mental operation that are deeply implicated in agency, and in the second, there are modes of experience and knowledge that are only available to agents when they are behaving *as* agents. I will be concerned almost entirely with the first topic and mention one aspect of the second only in passing. Under the first heading, 'mental operations', I will discuss two integral components of agency, which I will call '*action monitoring*' and '*reversibility*'.

First, action monitoring. There are two main ways in which changes in my perceptual inputs can be brought about. There are changes originating in the world for which I am not responsible (e.g., a cat walks across my path), and there are changes for which I am responsible, insofar as they originate in my body (e.g., as my head moves to the left, a sleeping cat shifts to the right of my visual field). An organism needs to be able to register the difference between these two kinds of changes.

For illustrative purposes, I will describe the simplest and most famous case of action monitoring, that studied by von Holst and Mittelstaedt (1950, translated in Gallistel 1980). Fruit flies produce a so-called "optokinetic reaction," which means that they turn in the direction of world movements. A moment's reflection tells us that if the fly had no mechanism for distinguishing between changes in the visual flow caused by its own movements and changes caused by movements in the world, it would be paralysed every time it produced the optokinetic reaction. If, for example, the world moves to the fly's left, the fly's head moves to the left, but this leftward movement will cause the world to (apparently) move to the right, and so this in turn should cause a rightward movement, which in turn. . . . In other words, if every apparent movement of the world were taken as a real movement, the creature would be as paralysed as Buridan's ass.

What is required, therefore, is a mechanism that can record the fact that the insect has launched a movement and then use this information so as to treat the resulting visual changes as signaling an apparent change rather than a real change. The mechanism that von Holst and Mittelstaedt proposed is called *effference copying*. A copy is made of the initial turn-commanding signal. For example, if the signal was '+3' (+ meaning egocentric

right, with '3' being a unit of extent), then the efference copy would also be '+3'. The phenomenal outcome of this action will be a so-called "reafferent" signal of '-3', i.e., the world appears to move to the left. These positive and negative values cancel, and so the animal's nervous system records the fact that there has been self-movement but no world-movement. Had there been no efference copy and just the afferent signal of '-3', then it would have recorded the fact that the world had *really* moved to the left.

The mechanisms of action monitoring obviously become more complex in more complex organisms, and our understanding of them becomes predictably fuzzier. Indeed, we might want to abandon the assumption that a process as simple as efference copying can tell the whole story about how action monitoring works in higher-level cognition. But enough has been said to make two points. First, what we see here is a very primitive mechanism for distinguishing real from apparent changes in sensory input (it is an appearance/reality distinction, remember, with which the piagetian thesis is primarily concerned). Action monitoring is an integral feature of information processing in agents, and this in turn is integral to our making at least a primitive distinction between how the world appears and how it is in fact.

The second point is about what is intended by the term 'monitoring' in 'action monitoring'. The point is both a caveat and a passing reference to the kind of question, mentioned at the start of this section, about the mode of knowledge and experience uniquely available to agents when acting as agents. The caveat is that 'monitoring' does not imply the presence of some inner eye, a spectator in the 'Cartesian theater' (Dennett 1991), watching the launch and course of a movement. As I hope was clear from my description of primitive efference copying, this is a "subpersonal" mechanism with no higher-level homunculus.⁵ Indeed, the same can be said at the personal level, insofar as talk about a homunculus overseer inspires exactly the *wrong* account of agency, as we experience it. And this point brings me to the passing reference to a special mode of knowledge and experience available only to agents that I promised earlier. It is that if a subject has to observe or monitor herself to find out what she is doing, she is *not* acting intentionally: we know the nature of our intended actions nonobservationally. Personal level, inner-eye action monitoring cannot exist alongside true agency (see O'Shaughnessy 1980, 31–32).

But there is obviously more to acting intentionally than monitoring bodily movements and attention shifts. Indeed, we can imagine a creature whose action-monitoring mechanisms are in good order but that never acts willfully because all of its actions are called forth by stimuli in the external world. What needs to be added to the picture is the fact that a true

agent—as opposed to a mere reactor—determines what she does and when: an agent is in control of her behavior.⁶ This is a somewhat self-evident, if not actually tautologous, statement about agency, but my aim in making it is to draw attention to a single fact about self-determined actions: that the order in which they are produced can be *reversed*. Because actions alter perceptual inputs, this ensures that there is a class of our *experiences* that is self-determined (and reversible) rather than world-determined (and irreversible)—a distinction owing to Kant, though he did not draw it to make a point about agency.⁷ I will now explain what this means.

The sequential ordering of actions can be independent of the ordering of events in the world: I can look at *a*, *b*, then *c* and then *c*, *b*, and *a*. This means that, because actions determine perceptual inputs, the flow of the resulting percepts is independent of the flow of events in the world. For when we witness an occurrence, the flow of events is in one direction, and it happens once: *a* happens, then *b*, then *c*, and this is the end of the story. The order in which events unfold is thus world-determined and is as *irreversible* as action-generated perceptual sequences are *reversible*. For example, if I am stationary and a cat walks across my path, I have no choice but to see the animal on, say, my left before seeing it on my right. And if I want the event to occur again, I have to wait on the cat. Similarly, when somebody comes into my office, I have no choice but to see him open the door before he sits down. The order of perceptions is irreversible, and its possible repetition is something over which I have no power. However, if I am looking at a motor car or leafing through a book, I am free to experience portions of the objects in any order I wish: rear wheels then grill, index then preface, or vice versa. In this case there is relative freedom to determine what we experience at a given time and in the order we choose, while in the former case we are constrained to experience something and its features in a particular order. Moreover, agents are also free to determine *which* objects are members of the array to be ordered in experience. I can, for example, ensure that the eraser on my desk is not part of the array visible to me by moving my body to the right, occluding the eraser behind the coffee mug.

I should add in passing that my use of the term ‘reversible’ here is in fact a synecdoche for the various respects in which perceptual inputs may be self-determined. Other characteristics of self-determined perceptual sequences are that simple sequences can be combined into more complex ones and that detours and short-cuts en route to a location can be made. Nothing would be gained by listing them.⁸ Conversely, ‘irreversible’ is supposed to apply here to *any* kind of perceptual sequence over which we have no power. The term ‘reversibility’, as many readers know, also plays a

central role in Piaget's theory, where it was taken to be the central distinguishing feature of "intelligence," as opposed to "perception." But again, nothing would be gained by listing the similarities and differences between my usage and Piaget's.

The (rather paradoxical) moral I wish to draw here is this: the reversibility of self-generated perceptual sequences can reveal to us how the scope of our experience is *constrained*. It is difficult to imagine how a conception of the irreversible, world-dependent nature of some modes of experience could arise if we never experienced the reversible, action-dependent nature of others. Moreover, the freer we are to determine the nature and order of our inputs (and the less our behavior is called out by stimuli and the less our attention is captured rather than directed), the *more possibilities for resistance to our will* we encounter. The more we can do, that is, the wider the variety of ways in which our will can be resisted and its experiential outcomes limited. The freer we are to alter our perceptual inputs, the more we learn of the *refractory* nature of the world and, correlatively, the richer the conception we gain of ourselves as determiners of our immediate mental life. This refractoriness, therefore, sets limits on what our agency can achieve in determining our experiences, thereby engendering a conception in us of something as setting these limits, as causing them to be set. (The word 'refractory' is borrowed from J. M. Baldwin [1906], who used it in much the same way as I am using it, to describe how objectivity in the child's experience emerges partly as a function of the world's resisting the child's will [see Russell 1978, part 1.2, for a summary of his theory].)

Clearly, not *all* kinds of actions are reversible—many kinds of actions cannot be performed backwards—but this only serves to highlight the special status of actions that *are* reversible and points up the intimate relationship between reversible bodily movements and attention shifts. Attention shifts are always reversible for the mature thinker. For the class of reversible actions contains all those actions that change the subject's perspective on the environment, where we take these perspective changes to encompass everything from moving around in the world, to changing visual fixations, to shifting auditory attention. Recall that at the end of the first section I said that the piagetian interprets 'action' in a very broad sense to refer to the changing of perceptual inputs at will and that this includes attention shifts.

Also recall that I earlier described the achievement of self-world dualism in terms of an appearance/reality distinction, which in this context is a distinction between how the world appears to us from instant to instant and our conception of a mind-independent reality that is the condition of these changes of appearance. On the present view, it is the experience of

reversibility, the experience of changing our inputs by changing our physical or mental (attentional) relation to objects, that affords us such changes of appearance, while the manifest constraints on how appearances can be altered (e.g., by moving round an object) show us that limits are set on how one appearance can follow from another. And recall that this experience of changing our inputs is not only a matter of changing *how* things are perceived; it is also a matter of changing *what* we perceive as we change perspective. It is in this latter sense that the notion of refractoriness—resistance of the world to our will—has the clearest application, simply because the term ‘will’ has such a clear application when we are dealing with *what* is experienced. A baby desperately desiring the nipple and turning to its left when the nipple is on its right is a potent example of the experience of refractoriness.

Arguing from reversibility is not the only route to the thesis of refractoriness. So I will describe another route to it to round out my account. David Hamlyn (1990, 105–106) and Thomas Baldwin (1995), for example, have both argued that experiencing the resistance of substantial objects to the willed movements of our *bodies* (when we touch things, lie on them, push against them, and so forth) is necessary if we are to regard objects as mind-independent. Baldwin (1995) begins his argument from a position different from the one used here. Rather than arguing from a form of experience that agency affords us, he starts with the thought that if we are to view ourselves as perceivers of an objective world, then we must regard objects as the *causes* of our experiences. This in turn requires us to acknowledge the *modal* nature of causality (i.e., that causes *necessarily* precede events, while it is *possible* that events could have unfolded differently). He then argues that if we adopt a view of volition in which voluntary action is taken to be the exercise of a bodily power, we can explain how this modal conception can arise: when substantial objects impede the exercise of our will, we experience the *impossibility* of certain bodily attempts, we encounter “forces acting upon us to place limits on our bodily power” (Baldwin 1995, 116).⁹

According to my position, by contrast, felt resistance is not necessary for the modal conception to emerge. That is to say, out of *nonphysical* interaction between one’s body and the world (such as moving the eyes in relation to scenes, or moving the body in relation to objects) refractoriness will emerge—I argue—from the tension between how and what things *must* appear to us and how and what things *may* appear to us. This said, the similarities between our two views are far more notable than their differences: they are both arguments for objectivity from refractoriness via agency. Indeed, toward the end of his paper Baldwin leaves the door open

for a thesis not unlike mine in which “it is by finding that the content of visual experience, *unlike its direction*, is not subject to the will that such a subject encounters a kind of impossibility within visual experience (a visual analogue of tactile resistance), and is thus led to the thought that the content concerns objects whose existence is independent of the experience and can therefore be employed to explain the experience” (Baldwin 1995, 121; my italics).

If either of these arguments from refractoriness establishes that knowing oneself to be the subject of objective experience requires experiencing oneself as an agent, then the link between agency and self-world dualism is more or less complete. However, the implications that this thesis has for objectivity (to be considered developmentally in section 3) are much clearer than its implications for subjectivity (to be considered developmentally in section 4). Accordingly, before ending this section, I will say a few words about how refractoriness theses bear on the question of self-awareness. ‘Self-awareness’ implies a capacity richer than the capacity to regard oneself as a subject of objective experiences. It implies an ability to regard oneself as not merely existing *in relation* to objects. It implies a *reflective* orientation to one’s subjective experiences and the possibility of entertaining first-person thoughts about them, as well as about one’s intentions.

Accounts of self-awareness that place agency at the center of the picture suggest that an essential component here is a primary awareness of oneself as a being that wills (which we are to take as encompassing everything from self-determination of one’s experiences, to instinctual striving, to having rational goals). But even Schopenhauer—the philosopher, above all others, who took a volitional view of the self—did not believe that willing is *all* there is to selfhood. He accepted that the self both wills and apprehends (believes, knows, perceives, etc.), saying that the self’s willing and the self’s knowing “flow together into the consciousness of one I,” which he pronounced to be “the miracle par excellence” (Schopenhauer 1844, 243). As Janaway (1989) has recently commented, this “miracle” is at least a truth about and condition for selfhood, in the sense that without the coreference of the ‘I’ in ‘I will’ and ‘I perceive’, there is no possibility of either action or knowledge. The subject is conscious of herself as a being that strives to alter the world (and her experiences of it, one should add) and at the same time as a knower. “This,” Janaway writes, “is because the point of acting is to change something about the world that *I* perceive, or about *my* relation to it, while the capacity to perceive essentially informs, through beliefs and desires, the way I actively modify myself in response to what is perceived” (1989, 89). If willing is indeed necessarily present in self-awareness, as Janaway suggests, then an adequate conceptualization of

one's mental states will depend on adequate experience of oneself as an agent—to which I will return in section 4.

3 Object Occlusion and Object Permanence

In this section I will discuss what piagetian views about the role of agency in cognition imply about how object permanence should be assessed and about how it develops. The second topic cannot be addressed without regard to the first.

What counts as evidence for prelinguistic children knowing that objects continue to exist when they are completely occluded from view? A piagetian must take a very strong and distinctive line on how this question should be answered, and if this line cannot be sustained, the theory loses much of its force. That is to say, if you think that an infant's awareness of the division between itself and the physical world is determined in part by its agency, you will tend also to think that its knowledge of the mind-independent existence of objects should be assessed in terms of what it *does*.

Piaget's (1955) own research showed—and it has been massively confirmed since—that there is a long period in development during which infants will not retrieve completely occluded objects *although they have the motor capacity to do so*. For example, a baby of 6 months can lift an inverted polystyrene cup, but if it sees a trinket that it strongly desires being hidden under such a cup, it does nothing. There is no sign of frustration, and the infant acts as if the trinket has ceased to exist. Why? Piaget's answer was that this is a manifestation of *egocentrism*, which broadly means a fusing of one's current experience with reality, a failure to grasp how the appearance of the world is relative to one's perspective on it, an incomplete self-world dualism, in fact. A perspective-relative conception is supposed to be achieved through the experience of willfully changing one's perspectives.

I will mention only in passing Piaget's actual account of how egocentrism was supposed to be overcome. The only explanatory tools that he allowed himself were the "circular reactions" of J. M. Baldwin (1906), movements that can be repeated at will, unlike reflexive movements evoked by stimuli. As the infant grows older, circular reactions become progressively directed toward having effects in the physical environment (unlike, say, thumb sucking) and become more integrated. For instance, means-end behavior is described as an example of the 'coordination of secondary circular reactions', with 'secondary' roughly meaning having an environment-centered effect rather than a bodily-centered effect. Furthermore, the significance for development of what I am calling 'the experience of refractoriness' is represented within Piaget's theory of learning, which proposes

an interplay between the child's attempts to *assimilate* new data to preexisting action schemes and the *accommodations* she has to make to adjust the mental contours of the old scheme, where the greater the novelty, the greater the degree of accommodation required. As James McClelland has indicated, there is a strong parallel between Piaget's principle of accommodation and the connectionist back-propagation algorithm, insofar as both are founded on the principle "Adjust the parameters of the mind in proportion to the extent to which their adjustment can produce a reduction in the discrepancy between expected and observed events" (McClelland 1989, 20). If there were no experiences of refractoriness, there would be no need for accommodation. (Some of the earliest connectionist simulations were of Piagetian tasks [e.g., Papert 1963].)

However, piagetians need not be committed to Piaget's theory of learning, and as we saw in the first section, they can avail themselves of nativist accounts of spatial coding. That said, my view about why younger infants fail to search for occluded objects is entirely at one with Piaget's. The assumption is that the baby fails to search because, given its inadequate grasp of how its visual experience depends on its actions, the question of whether something that has ceased to be perceptible can be rendered perceptible again through action *cannot arise*. It is *not* that the baby "believes"—insofar as very young infants believe anything—that the object has ceased to exist; rather, any conception of an unexperienced but existing object is beyond its grasp.

There is, however, a natural and plausible objection to this inference from lack of searching to lack of knowledge. It is entirely possible that very young infants are able to mentally represent the continuing existence of presently unperceived objects while being unable to organize, for whatever reason, a successful search. Given this, many contemporary developmental psychologists would argue that the information available to infants about object permanence can best be detected by experimental techniques that do not require action. That is, we need to find out how infants *react* to different kinds of physical events. The reaction typically studied is that of surprise (strictly, recovery of interest), the rationale being that if an infant witnesses an anomalous event, an event that *violates the principle of object permanence*, and shows surprise at this, then we can infer that it was coding the event *as* anomalous and therefore had expectations about permanence.

The most discussed study in this area, and the one that speaks to the existence of such early knowledge most strongly, was carried out by Renée Baillargeon (1987).¹⁰ Baillargeon showed that infants between 3.5 and 4.5 months of age are surprised (the index being *dishabituation*—recovery of looking toward the display) when a screen, swinging like a

drawbridge toward and away from the infant *fails to be impeded* by a wooden block that has been temporarily rendered invisible by the screen on its backward journey. This age is about 4 months before infants search for completely occluded objects and about 14 months before they attain Piaget's criterion for possession of the concept of an object, namely, searching after invisible displacements of an object.

Still, piagetians can place this finding within their theoretical framework by drawing the following distinction. On the one hand, there is the maintenance of a representation of an object that happens to be currently occluded, a representation that drives reactive behavior (such as dishabituation) and, on the other, there is the conception of a physical datum as distinct from and external to the self. I refer to the first as 'representation permanence' and to the second as 'externality' (see Russell, *in press*, for a defence of this distinction).

Let us consider first what representation permanence amounts to. It would surely be a badly designed nervous system that routinely extinguished representations of objects at the instant that they ceased to be visible (and it is difficult to imagine how such extinguishings would be achieved within the constant flux of the perceptual input). Baillargeon's experiment is an elegant demonstration that such representations are indeed maintained, but the demonstration does not inform us about how the very young infant conceives of the relation between itself and the wooden block. Studying the situations in which such infants retrieve objects does, however, tell us something about this, because a subject who searches for an occluded object is manifesting a degree of knowledge about where she is located in relation to the occluder and the object. Surprise at nonresistance can occur, however, without any knowledge of how one is spatially related to the occluder and the object.

The piagetian position, then, is that search is an appropriate diagnostic criterion for self-world dualism because there is a conceptual linkage between what one knows about one's spatial relation to an object and what one is inclined to do, given certain desires. Consider an infant of, say, 6 months of age who wants its occluded toy back. If it knows the relation between itself and the occluder, the relation between the toy and the occluder, and the relation between the toy and itself, and if it is able motor-wise to remove the occluder, then why does it not search? There would seem to be a contradiction between saying that the infant knows where a still-existing object is and saying that the infant does not know how to act on that knowledge.

Note that the contradiction just described is present only if we take 'knows where the still-existing object is' to mean 'knows where the still-existing object is *in relation to me*'. My point is that if it can truly be said of

somebody that she knows where something is and that she knows where that location is in relation to her present location, then it must also be true that she knows how to act to gain sight of that thing. And note also that I am not discussing here the development of knowledge that an object may continue to exist when its location is unknown (e.g., a departed parent or a lost toy). I think that it is fair to regard this as a more sophisticated and later-developing form of externality. Perhaps it requires what Piaget (1955) called 'mental representation'—the ability to evoke conscious representations of absent objects—whose arrival he timed at about 18 months, *after* the development of searching for currently occluded objects was complete.

But the fundamental distinction here is that between representation permanence and externality, since maintaining a representation of a currently invisible object beyond the instant at which it was rendered invisible implies no grasp of the relation between itself and the object on the infant's part. This is not to deny that the 'representation' in 'representation permanence' has content, because it is possible for there to be a representation with a particular content and with certain causal liaisons to output (hence the surprise reaction) without this representation being that of something external to the self, of something that is the cause of one's actual or possible perceptual experiences.¹¹ Such a representation may even fall short of being 'causally indexical' (p. 130 above).

Recall that what we are concerned with here is the development of a subject's reflective, theorylike grasp of its place in the physical world. Studying the conditions under which a subject will search for an occluded object will inform us about this development because an integral feature of this theorylike understanding is a reflective awareness of the implications that arrays of objects have for what can be experienced, avoided, and reached for (see Campbell 1993, 88). Infants of around 8 months who search for completely occluded objects for the first time may not yet be manifesting this understanding in a reflective form; indeed, Piaget insisted that they were not.¹² But they are taking strides along the royal road to this reflective understanding if there is any truth in what I argued in the second section.

Finally, I want to consider in this section an objection to my distinction between representation permanence and externality. Despite the broad sympathy, noted above, that workers in connectionism have with Piaget's account of sensorimotor development, an objection arises out of current attempts to model object permanence in neural networks.¹³ The objection is, in effect, that there is no deep mystery about why infants are surprised by anomalous occlusion events at 4 months of age but do not search until they are twice as old. The situation for the 4-month-old is that its representation

of the occluded object is strong enough to drive a surprise reaction but too weak to drive a search. The research program for connectionist modeling of object permanence then becomes that of modelling the kind of information that strengthens the representation sufficiently to drive a search.

The reply is that it only makes sense to regard the developmental progression in this way if we treat search as a *reaction to events*. But search is not a reaction to events (and, so far as we know, there is no period of "reactive search"): it is action taken on the basis of knowledge of where something is in relation to us. It is not even strictly correct to treat it as means-end behavior, because a means may be *arbitrarily* related to an end. To thus talk of representations "driving" or "causing" search is to use a misleading metaphor that assimilates acting to reacting.¹⁴ Search is only intelligible when it is regarded as an action carried out from knowledge of how the world is configured in relation to us: representations do not cause it.

4 Self-World Dualism and a Conception of Other Minds

Having looked at the world half of self-world dualism, I now turn to the self. I will argue that because experiencing oneself as an agent is necessary for being a subject of objective experience, a conception of mentality (a "theory of mind" as it is now called in developmental psychology) will depend upon the experience of agency. That is to say, for the child to develop a conception of mind sufficient for her to explain and predict the behavior of other people, she must have experienced first-person agency. (Note that I am assuming the identity of the self in 'self-world' and 'self-other'.) As we will see, this way of arguing means that we cannot dismiss the view that the rudiments of a representational theory of mind are innate, though piagetians do not hold that such a theory *has* to be innate.¹⁵ But even were we to succeed in establishing that an innate representational "theory" exists, we would not thereby have succeeded in explaining how we come to have our conception of mentality, given that it depends on the experience of agency.

Whereas nativism about spatial and object concepts is based on foundational arguments from philosophers such as Kant and on a wealth of empirical evidence, nativism about mental knowledge is based on Platonic arguments about concept acquisition of the kind fielded by J. A. Fodor (1975, 1987), as well as on rather controversial evidence from developmental psychology and from the study of abnormal populations (Russell 1992).¹⁶ Fodor's argument is essentially that no organism can learn mental-state predicates such as 'think' and 'intend' without having the prior capacity to represent the extension of these predicates in an innate "language

of thought." Within developmental psychology, Leslie (1987), among others (e.g., Premack, 1990), has proposed that a theory of mind in which an innate, modular device matures in the second year of life, thereby enabling the child to compute relations between a subject, her propositional attitudes, and her propositional contents. According to Leslie, this can be first seen in the child's comprehension of pretence in other people, where another's deviant treatment of an object is coded by the child as her purposefully representing that object as being other than it is (e.g., a mother's pretending that a banana is a telephone receiver, in Leslie's example). *PRETEND* is taken to be a symbol in the language of thought (Leslie 1988).

Let us accept, for the sake of argument, that this view is correct, that human beings do indeed possess an innate apparatus for computing propositional contents and—in Fodor's (1987) metaphor—for placing these contents into different propositional-attitude "boxes"—belief boxes, pretend boxes, and the like. Now imagine a system that is able to perform all these operations. It observes somebody picking something up and codes this as an intentional act; observes somebody being startled and codes this as nonintentional; observes somebody putting a banana to her head and talking into it and records this as a case of pretence; observes somebody leaving his house, turning on his heels to go back indoors, and reemerging with an umbrella and computes the propositional contents of his belief and desire attitudes. In short, this mentalizing system is a prodigious parser of mental categories. However, it has no capacity for monitoring and reversing its actions, and so it is not an agent in the present sense.

Predictably, the piagetian will say that this system has no conception of other minds, despite its skills at parsing mental categories. The first reason for saying so emerges naturally from what has gone before. The refractoriness thesis (whether expressed in the form of reversibility or felt resistance) seeks to establish that experiencing oneself as an agent is necessary for knowing oneself to be the subject of objective experience—is necessary, in other words, for subjectivity. As the system in our thought experiment is not an agent, it cannot, on this view, possess subjectivity, and if it does not possess subjectivity, then it makes little sense to say that it is a mind. So we have the paradoxical state of affairs that a nonmental entity is parsing mental episodes. What's wrong with this? Cannot minds be known purely "from the outside," as it were? A negative answer to this question emerges with two further considerations.

The piagetian will say (a) that a conception of other minds depends on a conception of others as agents, and (b) that conceiving of others as agents is possible only if one can experience oneself as an agent. A proper defence of both of these claims would take a long time, but they can be

defended briefly. With regard to (a), to ascribe mentality to others means conceiving of them as *rational*, and this requires us in turn to see them as “capable of assessing positions, of following out reasons, and possibly of being critical. All these things presuppose our not being merely passive in relation to putative objects of knowledge” (Hamlyn 1990, 148). With regard to (b), agency is grounded in first-person experience unmediated by observation of oneself: we cannot gain a conception of what an agent is by watching how we and others behave—a view that bears a more than accidental resemblance to the denial that our knowledge of what we are doing could be gained by self-observation (see section 2). Something of this is captured by Thomas Nagel when he speaks of “a clash between the view of action from the inside and *any* view of it from the outside. Any external view of an act as something that happens . . . seems to omit the doing of it” (Nagel 1979, 198–199).

The basic idea here is that agency is something that can only be known “from the inside,” as it were. So the system in our thought experiment must be achieving its feat (*if* such a feat can be achieved at all) by coding (what J. B. Watson called) “colorless movements,” not by coding the actions of agents. It will have no conception of agency, on this view, because this conception depends on *the experience* of trying to achieve goals and of being in control of one’s body and thus—a point I have labored—of one’s immediate mental life. It is an experience no less than, say, pain is an experience. (See O’Shaughnessy 1980, chap. 11, on the parallels between our knowledge of our own trying and of sensory experiences like pain.) A creature that never felt pain could accurately code pain behavior, but its conception of pain would be empty. The parsings of agency by our imagined system would be similarly empty.

Moreover, it is questionable whether this imagined system could ever entertain first-person thoughts. Because it is not an agent, it possesses only an apprehending self and not a willing self. It is the point, as it were, at which informational input terminates. But if the position sketched at the end of section 2 is correct—if, as Janaway (1989) argues, selfhood requires the ‘I’ in ‘I will’ and ‘I perceive/believe/know/etc.’ to corefer—then the system will be incapable of entertaining first-person thoughts even if it, per impossibile, developed subjectivity.

There is a natural objection to this claim that knowledge of others’ agency is grounded in first-person experience. Some might say that this view is Cartesian, insisting that there are tried and trusted arguments against the Cartesian view that knowledge of our own mental states is immediate, incorrigible knowledge that can be projected to others by analogy. Wittgenstein (1953, §243 onwards) has shown us, the objector might say,

that because the meaning of a term is grounded in public corrigibility, if one is the sole arbiter of whether a mental-state predicate has been correctly ascribed to oneself, then no predicate has been meaningfully ascribed in fact. The following argument from Strawson might also, perhaps, be recruited: If one can succeed in picking oneself out as the argument of a mental-state predicate, then one must already be in a position to pick others out as an argument for that predicate. "One can ascribe states of consciousness to oneself only if one can ascribe them to others" (Strawson 1959, 100).

But an anti-Cartesian objection is beside the point here. The child will obviously come to apply agency predicates, among other kinds of predicates, when she acquires language, but seeing other as agents, *appreciating that some observations of others are observations of their agency*,¹⁷ is not a matter of ascribing to others a predicate one first ascribes to oneself. Indeed, calling agency an 'experience' is tantamount to denying that it consists in the ascription of predicates to oneself. When a baby of 4 months extricates its rattle from the rungs of its cot, it is experiencing agency, not ascribing predicates.

To bring the discussion round to empirical issues in development, we can see that the piagetian view is well placed to describe the role of social interaction in the development of selfhood. For, in clear contrast to the nativist, the piagetian is in a position to acknowledge that self-awareness will emerge, in part, through social interaction, though in equally clear contrast to writers like Mead (1934) and Vygotsky (1962), the piagetian would *deny* that this process is one of social transmission or internalization.

Consider the executive demands that social interaction makes on the subject. Compared to inanimate objects, other people "go off like guns on the stage of [the child's] panorama of experience," to use J. M. Baldwin's evocative phrase (1906, 49). Although social psychologists and students of early communication (e.g., Bruner 1975) have tended to stress the rule-boundedness of early social interaction, here it is its *unpredictability* that must be noted. In order for them to deal successfully with other people, young children must be prodigious decision makers: they must constantly be on social guard to select new but appropriate behaviors in the light of what the other has done. As was the case for interactions with physical objects, there must be instigation, inhibition, planning, and monitoring of behavior, but the most crucial in social life is the ability to improvise. To illustrate, in their model of the "executive systems" Norman and Shallice (1986) distinguished between routine-action "scripts" (with triggering and inhibiting relations between their subcomponents) and the "Supervisory Attentional System," the executive overseer that takes control in novel

situations.¹⁸ In social interaction there is constant novelty, in the sense that new things—if not new *kinds* of things—are constantly being done, and so the supervisory system is, loosely speaking, in constant use.

What does this imply about the developing conception of other minds? From the first-person perspective, the experience of piloting itself through the relatively uncertain waters of social interaction will provide the infant with the experience of being in charge of itself. The more novelty *required*, the more one's experience becomes that of the supervisor of one's behavior and immediate mental life. From a first-person perspective, in terms of the executive demands of social interaction, the other is refractory, is uncontrollable and unpredictable—to return to this (J. M.) Baldwinian theme. As I have argued, the first-person perspective is primary, but it also makes sense to talk of self- and other-awareness as being constructed in social interaction. The more autonomous—the more *self-determined*—the subject, the more successful the process of ego development.

This account makes a very clear prediction about what the consequences would be of early impairments in agency. We can assume that one form that these impairments might take is that the “executive system”—meaning roughly the system responsible for inhibiting, monitoring, and regulating behavior and for the instigation and transformation of strategies—will be dysfunctional, and that early executive dysfunctions will lead to impaired self-awareness and thus to an impaired conception of other minds. In the syndrome of autism there are indeed such co-occurrences of executive dysfunctions (e.g., Hughes and Russell 1993; Hughes, Russell, and Robbins 1994; Ozonoff, in press; Ozonoff, Pennington, and Rogers 1991; Russell, Mauthner, Sharpe, and Tidswell 1991) and very well-documented difficulties in predicting and explaining behavior mentalistically (see the papers in Baron-Cohen, Tager-Flusberg, and Cohen 1993). This is highly consistent with the piagetian view.

By contrast, nativists about mental knowledge have tried to explain autism in terms of the delay or deviance in the maturation of the “theory of mind module” or TOMM (Leslie and Thaiss 1992), but in doing so, they have difficulty in explaining why “mentalizing” impairments should coexist with executive impairments. There have, however, been attempts. Christopher Frith (1992), for one, argues that performing executive tasks relies heavily on action monitoring, and this requires the adequate functioning of a TOMM-like system whose core is a mechanism for representing one's own mental states. Consequently, he refers to the neurological impairment that he takes to underlie the mentalizing difficulties within autism and schizophrenia as ones of “metarepresentation,” after Leslie (1987).

At this point the debate between the nativist and the piagetian must obviously be decided empirically. The piagetian would predict, for example, that administering executive tasks such as attention shifting (Atkinson, Hood, Wattam-Bell, and Braddick 1992) in order to screen infants at risk of acquiring autism (with relatives suffering from autism, autismlike disorders or schizophrenia) should successfully identify which of these at-risk infants will acquire the disorder. So, despite its derivation from an essentially philosophical position, the piagetian thesis has a clear empirical cash value.

Acknowledgments

In writing this chapter I received invaluable advice from the editors and also benefitted from the comments of the following people, to whom thanks are due: Chris Frith, Susan Goodrich, Suzanne Hala, Paul Harris, Jane Heal, Jennifer Hornsby, and Paul Noordhoff. For an extended treatment of the themes covered here, see my *Agency: Its Role in Mental Development* (in press).

Notes

1. James Mark Baldwin was an American philosopher-psychologist who produced his influential work around the turn of the century, most notably his three-volume *Thought and Things*. See Russell 1978 for an extensive discussion of Baldwin and a comparison of his theory with Piaget's.

2. To code spatial locations and relations allocentrically is to code them without regard to one's point of view ("egocentric" coding). Some relations are allocentric per se, such as place *b* lying *between* place *a* and place *c*. An "allocentric representation" is a rather theoretically loaded term (Brewer and Pears 1993). However, in experiments with animals and human infants, the term 'allocentric coding' simply means coding in terms of the relation between two points without regard to where the subject is positioned.

3. Causally indexical representations are those made purely in terms of the implications for the subject's own movements, "within reach," for example. A causally nonindexical representation is reflective, in contrast to the indexical variety, in the sense that its character is not entirely determined by the implications that it has for what the organism can do.

4. In her book *Kant's Transcendental Psychology*, Patricia Kitcher attempts to reinstate Kant's psychological doctrines, taking the term 'transcendental psychology' to mean "the psychology of the thinking, or better, the knowing self" (1990, 22). She treats this as being roughly equivalent to the analysis of cognitive tasks. This is not empirical psychology. But, as Kitcher points out, much of what we call psychology involves the conceptual analysis of tasks rather than experimentation, Newell and Simon's (1972) work on problem solving for example.

5. The term 'subpersonal' is a term coined by Daniel Dennett, originally in chapter 4 of his *Content and Consciousness*. The personal level is the level on which we talk about acts, intentions, beliefs, thoughts, reasons, feelings, and the like (some call this the 'folk-

psychological level'), while the subpersonal level, as Dennett puts it elsewhere, is the "behind-the-scenes *machinery* that governs speech dispositions, motor subroutines, information storage and retrieval, and the like" (Dennett 1978, 216; my italics).

6. The obvious distinction to be drawn at this point might appear to be one between reactions (or responses) and actions (see Dickinson and Balleine 1993), where we take actions to be goal-directed behaviors, movements made *for the sake of* a goal. However, I do not wish to make a strong commitment to the goal-directedness of self-determined actions. What I mean by being 'self-determined' involves more than goal-directness. Moreover, self-determined actions may be aimless.

There is a distinction to be made between what Brewer (1993) calls the "mere capacity" to get things done and the genuine ability for basic bodily action that involves control over one's body, the "ability to direct the will's extension into the body" (1993, 308). Goal-directed actions, may be the former without being the latter.

7. Kant drew this distinction in an (apparently unsuccessful) attempt to establish that world-generated causal sequences possess a form of "necessity." But this does not make it irrelevant to our concern with agency.

8. Piaget (1955, chap. 2) attempted to explain the relation between an objective conception of space and the forms of activity necessary for this conception in terms of mathematical group theory. Reversibility (or "inversion") is one of the four properties of the relation between operations and elements in a mathematical group.

9. Hume argued that such a position is untenable because in acting voluntarily, we are only experiencing the conjunction of occurrences of willing with bodily movements. Following O'Shaughnessy, Baldwin adopts a non-Humean account of action couched in terms of "the having of power or control over the limb": action is not a causal conjunction between a mental and a physical event but an act of will "coming to fruition." See pages 112–116 of Baldwin's chapter in this volume.

10. The work of Elizabeth Spelke (e.g., 1991) and her coworkers should be mentioned in this context. They have performed a number of dishabituation studies showing that infants have expectations about some principles of naive physics (e.g., inertia).

11. Chrisley (1993) has argued that infants' representations of objects prior to attaining full object permanence (on Piaget's criteria) have "nonconceptual content" (Crane 1992). He has also equated the attainment of conceptual content with meeting Evans's (1982) generality constraint, as have I (Russell 1988). However, there is a good deal of debate over how the line between nonconceptual and conceptual content should be drawn, on which question, see the chapter by José Bermúdez in this volume.

12. In fact, in light of subsequent failures to search (after visible and invisible displacements of the object), Piaget regarded search at 8 months as being little more than an instrumental procedure. Equivalent to a reflective understanding of object permanence was what Piaget called 'mental representation', which was supposed to be attained at around 18 months of age.

13. This view was put to me by Kim Plunkett, and a version of it can be found in Munakata, McClelland, Johnson, and Siegler 1994.

14. I do not think that David Hamlyn is putting it too strongly when he writes, in a commentary on the view that representations cause behavior, "it could be said that an appeal to representations at this point is something of a fraud: the appeal is made simply

to provide an otherwise missing connection between the central processes and behavior." (1990, 128–129).

15. Broadly empiricist theories of the development of a theory of mind are those in which the developing child is regarded as being in a position similar to that of a scientist testing hypotheses against the behavioral data (e.g., Gopnik and Wellman 1992). I have argued against such views in Russell 1993.

16. Some theorists (see papers in Baron-Cohen et al. 1993) argue that the fact that people with autism are specifically handicapped on tests supposed to measure "mentalizing" ability suggests that they lack the "theory-of-mind module."

17. I am grateful to Jennifer Hornsby for suggesting this way of putting it.

18. 'Executive systems' is a rather loose term used to refer to the mental functions carried out by the prefrontal cortices. These center on the control of behavior via inhibition, monitoring, planning, and selection of strategies. Some theorists, such as Shallice, stress the importance of control in nonroutine circumstances while others (e.g., Goldman-Rakic 1987) stress the control of behavior by models held in working memory.

References

- Atkinson, J., B. Hood, J. Wattam-Bell, and O. J. Braddick. 1992. "Changes in Infants' Ability to Switch Attention in the First Three Months of Life." *Perception* 21:643–653.
- Baillargeon, R. 1987. "Object Permanence in 3.5- and 4.5-Month-Olds." *Developmental Psychology* 23:655–664.
- Baldwin, J. M. 1906. *Thought and Things*, 3 vols. London: Swann and Sonnenschein.
- Baron-Cohen, S., H. Tager-Flusberg, and D. Cohen, eds. 1993. *Understanding Other Minds*. Oxford: Oxford University Press.
- Brewer, B. 1993. "The Integration of Spatial Vision and Action." In *Spatial Representation: Problems in Philosophy and Psychology*, ed. N. Eilan, R. McCarthy, and B. Brewer. Oxford: Blackwell.
- Brewer, B., and J. Pears. 1993. "Introduction: Frames of Reference." In *Spatial Representation: Problems in Philosophy and Psychology*, ed. N. Eilan, R. McCarthy, and B. Brewer. Oxford: Blackwell.
- Bruner, J. S. 1975. "The Ontogenesis of Speech Acts." *Journal of Child Language* 2:1–19.
- Campbell, J. 1993. "The Role of Physical Objects in Spatial Thinking." In *Spatial Representation: Problems in Philosophy and Psychology*, ed. N. Eilan, R. McCarthy, and B. Brewer. Oxford: Blackwell.
- Chrisley, R. L. 1993. "Connectionism, Cognitive Maps, and the Development of Objectivity." *Artificial Intelligence Review* 7:329–354.
- Crane, T. 1992. "The Non-conceptual Content of Experience." In *The Contents of Experience*, ed. T. Crane. Cambridge: Cambridge University Press.
- Dennett, D. C. 1978. "Why You Can't Make a Computer That Feels Pain." In his *Brainstorms*. Hassocks, U.K.: Harvester Press.
- Dennett, D. C. 1991. *Consciousness Explained*. Harmondsworth, U.K.: Penguin Books.

- Dickinson, A. J., and B. Balleine. 1993. "Actions and Responses." In *Spatial Representation: Problems in Philosophy and Psychology*, ed. N. Eilan, R. McCarthy, and B. Brewer. Oxford: Blackwell.
- Evans, G. 1982. *The Varieties of Reference*. Oxford: Oxford University Press.
- Fodor, J. A. 1975. *The Language of Thought*. New York: Crowell.
- Fodor, J. A. 1987. *Psychosemantics*. Cambridge: MIT Press.
- Frith, C. D. 1992. *The Cognitive Neuropsychology of Schizophrenia*. Hove, U.K.: Lawrence Erlbaum.
- Gallistel, C. R. 1980. *The Organization of Action: A New Synthesis*. Hillsdale, N.J.: Lawrence Erlbaum.
- Goldman-Rakic, P. S. 1987. "Circuitry of Primate Prefrontal Cortex and Regulation of Behavior by Representational Memory." In *Handbook of Physiology: The Nervous System*, ed. F. Plum. New York: Oxford University Press.
- Gopnik, A., and H. M. Wellman. 1992. "Why the Child's Theory of Mind Really *Is* a Theory." *Mind and Language* 7:145–171.
- Hamlyn, D. W. 1990. *In and Out of the Black Box*. Oxford: Basil Blackwell.
- Hopkins, J. 1987. "Synthesis in the Imagination." In *Philosophical Perspectives on Developmental Psychology*, ed. J. Russell. Oxford: Basil Blackwell.
- Hughes, C., and J. Russell. 1993. "Autistic Children's Difficulty with Mental Disengagement from an Object: Its Implications for Theories of Autism." *Developmental Psychology* 29:498–510.
- Hughes, C., J. Russell, and T. W. Robbins. 1994. "Evidence for Executive Dysfunction in Autism." *Neuropsychologia* 21:643–653.
- Janaway, C. 1989. *Self and World in Schopenhauer's Philosophy*. Oxford: Clarendon Press.
- Kitcher, P. 1990. *Kant's Transcendental Psychology*. Oxford: Oxford University Press.
- Leslie, A. M. 1987. "Pretence and Representation: The Orgins of 'Theory of Mind'." *Psychological Review* 94:412–426.
- Leslie, A. M. 1988. "Some Implications of Pretense for Mechanisms Underlying the Child's Theory of Mind." In *Developing Theories of Mind*, ed. J. Astington, P. L. Harris, and D. R. Olson. Cambridge: Cambridge University Press.
- Leslie, A. M., and L. Thaiss. 1992. "Domain Specificity in Conceptual Development: Neuropsychological Evidence from Autism." *Cognition*, 43:225–251.
- McClelland, J. 1989. "Parallel Distributed Processing: Implications for Cognition and Development." In *Parallel Distributed Processing*, ed. R. Morris. Oxford: Clarendon Press.
- Mead, G. H. 1934. *Mind, Self, and Society*. Chicago: University of Chicago Press.
- Munakata, Y., J. L. McClelland, M. H. Johnson, and R. S. Siegler. 1994. "Now You See It, Now You Don't: A Gradualist Framework for Understanding Infant's Successes and Failures in Object Permanence Tasks." Department of Psychology, Carnegie Mellon University, technical report PDP.CNS.94.2, May 1994.

- Nadel, L. 1990 "Varieties of Spatial Cognition." In *The Development and Neural Basis of Higher Cognitive Functions*, ed. A. Daimond. New York: New York Academy of Sciences.
- Nagel, T. 1979. "Subjective and Objective." In his *Mortal Questions*. Cambridge: Cambridge University Press.
- Newell, A., and H. Simon. 1972. *Human Problem Solving*. Englewood Cliffs, N.J.: Prentice-Hall.
- Norman, D.A., and T. Shallice. 1986. "Attention to Action: Willed and Automatic Control of Behavior." In *Consciousness and Self-Regulation*, vol. 4, ed. R. J. Davidson, G. E. Schwartz, and D. Shapiro. New York: Plenum.
- O'Keefe, J., and L. Nadel. 1978. *The Hippocampus as a Cognitive Map*. Oxford: Clarendon Press.
- O'Shaughnessy, B. 1980. *The Will*, vol. 1. Cambridge: Cambridge University Press.
- Ozonoff, S. In press. "Executive Functions in Autism." In *Learning and Cognition in Autism*. ed. E. Schopler and G. B. Mesibov. New York: Plenum Press.
- Ozonoff, S., B. Pennington, and S. Rogers. 1991. "Executive Function Deficits in High-Functioning Autistic Children: Relationship to Theory of Mind." *Journal of Child Psychology and Psychiatry* 32:1081-1105.
- Papert, S. 1963. "Intelligence chez l'enfant et chez le robot." In *La filiation des structures*, ed. L. Apostel, J. Grize, S. Papert, and J. Piaget, Etudes d'epistemologie genetique, no. 22.
- Piaget, J. 1955. *The Child's Construction of Reality*. London: Routledge and Kegan Paul.
- Piaget, J. 1970. "Piaget's Theory." In *Carmichael's Manual of Child Psychology*, vol. 1, ed. P. Mussen. New York: John Wiley.
- Premack, D. 1990. "The Infant's Theory of Self-Propelled Objects." *Cognition* 36:1-16.
- Russell, J. 1978. *The Acquisition of Knowledge*. London: Macmillan Press.
- Russell, J. 1988. "Cognisance and Cognitive Science. Part 1: The Generality Constraint." *Philosophical Psychology* 1:235-258.
- Russell, J. 1992. "The Theory-Theory: So Good They Named It Twice?" *Cognitive Development* 7:485-519.
- Russell, J. 1993. "On Leaving Your Children Wrapped in Thought." *Behavioral and Brain Sciences* 16:76-77.
- Russell, J. In press. *Agency: Its Role in Mental Development*. Hove, U.K.: Lawrence Erlbaum.
- Russell, J., N. Mauthner, S. Sharpe, and T. Tidswell. 1991. "The 'Windows Task' as a Measure of Strategic Deception in Preschoolers and Autistic Subjects." *British Journal of Developmental Psychology* 9:331-349.
- Spelke, E. 1991. "Physical Knowledge in Infancy: Reflections on Piaget's Theory." In *The Epigenesis of Mind: Essays on Biology and Cognition*, ed. S. Carey and R. Gelman. Hillsdale, N.J.: Lawrence Erlbaum.
- Schopenhauer, A. 1844. *Die Welt als Wille und Vorstellung*, vol. 1. Page references to *The World as Will and Representation*, vol. 1, trans. E. Payne. New York: Dover Press, 1966.

Strawson, P. F. 1959. *Individuals: An Essay in Descriptive Metaphysics*. Oxford: Oxford University Press.

von Holst, E., and H. Mittelstaedt. "Das Reafferenzprinzip: Wechselwirkung zwischen Zentralnervensystem und Peripherie." Translation by R. D. Martin published in *The Behavioral Physiology of Animals and Man: Selected Papers of E. von Holst*, vol. 1. Coral Gables: University of Miami Press, 1973.

Vygotsky, L. S. 1962. *Thought and Language*. Cambridge: MIT Press.

Wittgenstein, L. 1953. *Philosophical Investigations*. Oxford: Basil Blackwell.