

# Knowledge Enrichment and Conceptual Change in Folkbiology: Evidence from Williams Syndrome

Susan C. Johnson

*University of Pittsburgh*

and

Susan Carey

*New York University*

Ten participants with Williams syndrome (WS) (average verbal mental age of 11;5) were compared to two groups of normally developing children (average mental ages 10;11 and 6;7 years) with respect to intuitive biological knowledge about people, animals, and plants. Participants in the older control group were individually matched to the participants with WS on verbal mental age. The probes for biological understanding were drawn from the existing literature on the development of folkbiology and were divided into two batteries based on the hypothesized distinction of (1) general knowledge consistent with the conceptual repertoire of normally developing preschool children (the T1/T2-Neutral Animal Knowledge battery) and (2) folkbiological concepts normally acquired between ages 6 and 12 which require conceptual change for their construction (*life, death, people-as-one-animal-among-many, species kind* as determined by origin of the animal; the T2-Dependent battery). The two task batteries were equated for task demands, differing only in the content of the concepts probed. It was hypothesized that if this distinction is a false one, and the construction of folkbiology is accomplished entirely by enrichment of the preschooler's knowledge, there should never be a population with differential performance on these two batteries. People with WS were nonetheless found to be differentially impaired on the T2-Dependent battery. They performed at the level of the older control group on the T1/T2-Neutral battery, but at the level of the 6-year-olds on the T2-Dependent battery. These data support the distinction between

This work was supported by PHS predoctoral training Grant 2T32 MH18823-16 to the Department of Brain and Cognitive Sciences at the Massachusetts Institute of Technology. The authors thank Ned Block, Peter Gordon, Giyoo Hatano, Karen Levine, Helen Tager-Flusberg, Debbie Zaitchik, and three anonymous reviewers for their comments; we also thank Milena Koziol, Steven Luke, Patrice Washington, the teachers, parents, and students of the Cambridge, MA public schools and MIT summer daycamp and especially the Williams Syndrome Association and families for all their assistance with data collection.

Address correspondence and reprint requests to Susan C. Johnson, 405 Langley Hall, Department of Psychology, University of Pittsburgh, Pittsburgh, PA 15260.

two types of conceptual knowledge acquisition: acquisition of new knowledge formulated over an existing conceptual base (enrichment), on the one hand, and knowledge acquisition that results in genuine conceptual change, on the other. The implication of these results for a precise characterization of how concepts of people with "cocktail party syndrome" may be "superficial" is also discussed. © 1998 Academic

Press

People have tacit intuitive theories—cognitive structures that determine a person's deepest ontological commitments (the kinds of things presupposed by the person's thinking) and that provide causal understanding of real world phenomena. Describing the path through which children come to share the intuitive theories of the adults in their culture is an important part of characterizing cognitive development. Notable progress has been made in describing development in intuitive mechanics, psychology (theory of mind), biology, cosmology, and theory of matter (Au, 1994; Carey, 1985, 1988, 1991, 1995; Gopnik & Wellman, 1994; Hatano & Inagaki, 1994; Keil, 1989, 1994; Leslie, 1994; Perner, 1991; Spelke, 1991; Vosniadu & Brewer, 1992; Wellman & Gelman, 1992).

Central to this descriptive enterprise are the controversies over whether changes in these intuitive theories are always a matter of enrichment, or whether there is ever something more—conceptual change. Take a given theory at time 1, T1, and the descendent of that theory at time 2, T2. At issue is what it might mean for T2 to have a different structure from T1, as opposed to simply being an enriched, elaborated, version of T1 with the same structure.<sup>1</sup> In this paper, we will be concerned with the descriptive comparison between two successive intuitive theories for the phenomena revolving around life and the body; the preschooler's T1 and the 10-year-old's T2. We will argue that this development involves conceptual change.

We are not making an explanatory claim about the underlying processes that effect either enrichment or conceptual change. Furthermore, there are a variety of strong interpretations of "conceptual change" we do not hold. First, we do not claim that the difference between knowledge enrichment and conceptual change is sharp. There are a variety of intermediate cases. Second, we do not endorse the views of Kuhn (1962) and Feyerabend (1962) to the effect that theories before and after conceptual change are radically incommensurable (see Carey, 1985, 1991; Kitcher, 1988; Kuhn, 1982; Hacking, 1993). Finally, we do not claim that conceptual change occurs suddenly. Rather it takes time for concepts to change, sometimes centuries in the history of science, always years in the individual scientist or student or child engaged in knowledge restructuring (Kuhn, 1977; Gruber on Darwin, 1974; Nersessian on Maxwell, 1992; Carey, 1985; Chi, 1992).

Conceptual changes take several forms. Perhaps the most common is dif-

<sup>1</sup> Summarizing the debate within cognitive science and philosophy on the topic of conceptual change is beyond the scope of this introduction. See Carey (1985, 1991), Chi (1992), Hacking (1993), Kitcher (1988), Kuhn (1977), and Thagard (1993) for relevant discussions.

ferentiation. In conceptual differentiations implicating conceptual change the undifferentiated parent concept from T1 no longer plays any role in T2. Examples include Galileo's differentiation of *average* from *instantaneous velocity* (Kuhn, 1977), Black's differentiation of *heat* from *temperature* (Wiser & Carey, 1983), and the child's differentiation of *weight* from *density* (Smith, Carey, & Wiser, 1985; Carey, 1991). Another common type is the coalescence in T2 of concepts which were considered fundamentally different kinds in T1. An example is Galileo's abandonment of Aristotle's distinction between *natural* and *artificial* motions (Kuhn, 1977). A further type of conceptual change is the reanalysis of a concept's basic structure. An example is the Newtonian reanalysis of *weight* from a property of objects to a relationship between objects. And finally, on the common analysis of concepts having a core/periphery structure, changes in the concept's core constitute examples of conceptual change (Kitcher, 1988). Carey (1988, 1991) provides examples of each of these types of conceptual change in the course of normal cognitive development.

Consider the changes within the ontologically central concepts *person* and *animal* between ages 4 and 10. Older infants and preschoolers have an elaborate concept *person* for reasoning about human behavior (for review see Spelke, Phillips, & Woodward, 1995; Wellman & Gelman, 1992). Both also have a concept *animal* (Carey, 1985; Mandler, Bauer & McDonough, 1991; Wellman & Gelman, 1992). They distinguish animals from non-animals, and use this distinction productively in similarity-based inductive reasoning. Nevertheless, there is ample evidence that the preschooler's concepts *animal* and *person* differ from the 10-year-old's, being embedded in very different framework theories (Carey, 1985, 1988, 1995).

According to Carey's analysis, the core of the preschooler's concept *animal* is that of a behaving being, in essence a simplified variant on the prototypical behaving beings—people. The young child understands and interprets the body in terms of the role body parts play in supporting behavior. That is, the preschooler's framework theory (T1) in which the concepts *person* and *animal* are embedded is an intuitive psychology rather than an intuitive biology. Others disagree, characterizing T1 as an intuitive biology, organized around central explanatory concepts such as essentialism or functional explanation (Gelman, Coley, & Gottfried, 1994; Keil, 1994).

On either characterization of T1, by age 10 (recent work has revised this estimate downward closer to 7 or 8; see Carey, 1995, for a review), the child has constructed a new intuitive theory of biology, T2, with *animal* and *plant* coalesced into the single core ontological kind *living thing* and people reanalyzed as just one animal among many (Carey, 1985). Inagaki and Hatano (1993) characterize this new biology as a vitalist biology. Crider (1981) characterizes it as the container theory of the body. We characterize it as a theory of the life cycle and the role of bodily function in maintaining life.

For it to be true that there are conceptual changes within the concepts *person* and *animal*, there must also be changes in a host of interrelated con-

cepts. And indeed, there are. These include the differentiation of *not* alive into *dead*, *inanimate*, *unreal*, and *non-existent* (Carey, 1985, 1988) and the differentiation of *family* into *biological* family and *social* family (Solomon, Johnson, Zaitchik, & Carey, 1996; but see also Springer & Keil, 1989; Springer, 1992). Others are the reanalysis of *death* from a behavioral interpretation to one based on the collapse of the bodily machine (Carey, 1985; Koocher, 1974; Nagy, 1948), and the reanalysis of *baby* from small, helpless animal to reproductive offspring (Carey, 1985, 1988; Goldman & Goldman, 1982). The core features of the concept *species kind* shift away from physical and behavioral characteristics toward origins of the animal (Keil, 1989; Johnson & Solomon, 1997).

This conceptual change view has been challenged by an “accretionist” view (cf. Atran, 1994; Coley, 1995; Keil, 1992, 1994; Springer, 1992, 1995; Springer & Keil, 1989, 1991; Wellman & Gelman, 1992).<sup>2</sup> These two views share several assumptions. Both views allow for the enrichment of existing conceptual structures via inferential processes. For example, Springer (1995) suggested that children work out a core principle of biological inheritance (that babies’ physical properties are derived from the mother) by inference over the learned premises (1) babies grow in their mother’s tummies, and (2) nothing external to the mother influences the baby while it is growing. Similarly, Carey (1985) suggested that young children use patterns of inter-correlated animal properties productively in similarity-based (inductive) inference. Furthermore, the accretionist view, as articulated by Keil (1994), Springer (1995), and Spelke (1991), like the conceptual change view, is committed to T1 being a genuine intuitive theory; determining ontological commitments, representing causal mechanisms, and constraining both the child’s interpretation of phenomena and the child’s further learning in the domain.

However, according to the accretionist view, the core concepts encompassing animals and plants in the 10-year-old’s T2 are the same as those in the preschooler’s T1. According to this view new kinds may be discovered (e.g., *aardvark*), but no new basic ontological kinds are constructed (e.g., *living thing* is already available); differentiations may occur (e.g., *dog* into *poodle*, *collie*, *terrier*, . . .) but none in which the previously undifferentiated concept no longer plays any role in T2. There are no changes in concepts’ cores, no construction of new explanatory mechanisms. There is no new T2 formulated over concepts not available in T1.<sup>3</sup> In short, there are no incommensurabilities between the two theories.

<sup>2</sup> The enrichment view of theory construction is not restricted to researchers of intuitive biology. It is also held by Spelke (1991) for the case of intuitive physics and Leslie (1993) for theory of mind.

<sup>3</sup> We do not deny that theory enrichment of the sort characterized here under the label “accretionism” occurs. Rather, the issue is whether theory changes involving conceptual change also occur during childhood, and in particular, whether the development of folk biology between the preschool years and roughly age 10 provides such a case.

*Williams Syndrome, a Test Case*

In this paper we explore a new empirical approach to bringing data to bear on this debate. According to accretionism, there is only one type of change between T1 and T2: enrichment. Theory enrichment is itself limited only by the amount of relevant information available to the child rather than the existence of incommensurabilities between the theory the child holds (T1) and the theory the child is acquiring (T2) (e.g., Springer, 1995). Therefore, on the accretionist view, at any given point in development there should be a correlation between the general biologically relevant knowledge possessed by children and the core biological concepts typical of children of the same age. The conceptual change view however makes a different prediction. If the development of folkbiology involves conceptual change as well as (and distinct from) the enrichment of knowledge, then in principle, these two aspects of development could become dissociated resulting in differential degrees of achievement within the domain. Carey's claim that conceptual change results in a new biology, T2, by the age of 10, points to that age as a relevant age for such a comparison.

In the absence of large normative samples, it is difficult to test this prediction in normally developing populations. However, it may be possible to identify special populations whose enrichment abilities are stronger than their conceptual change abilities, resulting in individuals who display relatively advanced general knowledge of animals in the absence of a biological interpretation.

As proposed in the literature, conceptual change is likely to require both analytic and metacognitive abilities (e.g., comprehension monitoring, metacognitively aware uses of analogical reasoning, limiting case analysis, same/different comparisons) (Nersessian, 1992; Carey & Spelke, 1994; Smith, Snir, & Grosslight, 1992; Wiser, 1988). These abilities are known to develop during the early school years and are impaired in mental retardation (see Campione, Brown, & Ferrara, 1982, for a review.) Therefore, if the construction of T2 requires conceptual change, it is likely that any population of mentally retarded individuals will be markedly impaired in the acquisition of the core theoretic concepts of T2: *life, death, reproduction, the bodily machine*, etc.

On the other hand, some forms of retardation such as that caused by spina bifida (Cromer, 1991) or Williams syndrome (Udwin & Yule, 1990) are known to give rise to individuals displaying "cocktail party syndrome" (the tendency to talk in great detail with only superficial understanding). This phenomena suggests the relative preservation of enrichment learning abilities (associative memory, mechanisms for adding vocabulary and new beliefs to long term memory, automatic inferential processes), in the absence of analytic and metacognitive knowledge or skills. This consideration led us to hypothesize that mentally retarded individuals with cocktail party syndrome

would demonstrate a dissociation between our two proposed classes of animal-centered knowledge, resulting in (1) the general, theoretically neutral knowledge about animals typical of 10-year-olds predicated on (2) the core theoretical concepts typical of T1.

Williams syndrome (WS) is a neurodevelopmental disorder of genetic origin (Frangiskakis et al., 1996; Morris, 1994). Among other physical, neurological, and cognitive consequences of WS,<sup>4</sup> WS results in overall mild to moderate mental retardation, in the face of characteristic strengths in vocabulary and a high incidence of cocktail party syndrome (Bellugi, Birchle, Neville, Jernigan, & Doherty, 1993; Udwin & Yule, 1990). Researchers have documented both poor analytic and metacognitive skills in individuals with WS. Bellugi and her colleagues (Bellugi et al., 1993) report that adolescents with WS have marked difficulties on same/different judgments and are unable to formulate definitions for words despite the ability to generate appropriate instances and uses. In adults with WS, Bertrand and Mervis (1994) have documented poor metacognitive skills, including poor mnemonic strategies. On the other hand, Levine (1993) has documented a relatively preserved island of knowledge achievement in this population, centered around general encyclopedic and world knowledge, the types of detailed knowledge that most likely require the ability to remember facts and make associations among them. In addition, many researchers have documented the disproportionately large receptive vocabularies of participants with WS, relative to their full scale IQs (Bellugi et al., 1993; Levine, 1993; Scott, Mervis, Bertrand, Klein, Armstrong, & Ford, 1994).

Here we are concerned with what it might mean for people with cocktail party syndrome to have "superficial" knowledge. We hypothesize that the knowledge of adults with WS is superficial *because* it remains formulated over the core concepts of T1, despite the general level of information characteristic of older normally developing people who have constructed T2. There has yet been little work relevant to the suggestion that the lexical representations of people with WS are superficial insofar as normal processes of conceptual change have not occurred. The following anecdote is nonetheless

<sup>4</sup> Williams syndrome typically results in mental retardation as well as a variety of other physical problems including heart defects, metabolic problems of calcium and calcitonin, failure to thrive, hyperacusis, and characteristic facial and dental features (Williams, Barrett-Boyes, & Lowes, 1962; Jones & Smith, 1975; Udwin & Yule, 1991). Neuroanatomical studies of WS reveal no localized lesions in the neocortex, although there is evidence of reduced cerebral volume in general, together with unusual preservation of neocerebellum (Jernigan & Bellugi, 1990). People with Williams syndrome also have the reputation of verbal loquacity and hypersociability (Jones & Smith, 1975; Udwin & Yule, 1990) and are often characterized as displaying cocktail party syndrome; the propensity to talk prolifically about nothing. Within individual domains, researchers have documented dissociations within morphology (Bromberg, Ullman, Marcus, Kelly, & Levine, 1994) and between face recognition and other aspects of spatial and visual organization and memory (Birchle, Bellugi, Delis, & Marks, 1989).

suggestive: SK, a 21-year-old woman with WS, with a verbal IQ of 69, could read and particularly liked novels about vampires, of which she had read several. When asked what a vampire is, she replied, "Oooh, a vampire is a man who climbs into ladies' bedrooms in the middle of the night and sinks his teeth into their necks." When asked why vampires do this, she was visibly taken aback—she hesitated and said, "I've never thought about that." She then thought for a long time before finally answering, "Vampires must have an inordinate fondness for necks."<sup>5</sup>

These responses suggest that SK had indeed read several vampire books and had accumulated a store of information about vampires without ever constructing a concept of vampires as sucking blood, of being between dead and alive themselves, or of killing their victims and thereby creating new vampires. That is, this articulate 21-year-old with WS didn't really know what a vampire is supposed to be, in just those deep conceptual respects that implicate T2.

Biology is a particularly appropriate conceptual domain to test for the knowledge achievements of people with WS. Biological knowledge depends minimally on either visuospatial or numerical abilities, both of which are implicated in WS (Bellugi et al., 1993; Frangiskakis et al., 1996). Furthermore, the domain of biology lends itself to encyclopedic knowledge. Many facts, such as that animals breathe, could plausibly be learned without having to first understand that animals belong to the ontology of living things, or without understanding the role of breathing in maintaining life, i.e., without being embedded in T2.

## THE EXPERIMENT

### Constructing the Batteries

Drawing on the published literature and a reanalysis of the position of Carey (1985), we constructed two batteries of tasks based on whether or not T2-equivalent performance on each task in fact entails T2.

#### *T1/T2-Neutral Animal Knowledge Battery*

The tasks of the first battery, the T1/T2-Neutral Animal Knowledge battery, are built on the concepts *animal*, *animal properties*, and *animal parts*, concepts found in both T1 and T2. Although normal development of these concepts undoubtedly involves shifts in the core/peripheral properties of animals, contrary to the original claim of Carey (1985), adult-like performance

<sup>5</sup> SK's responses exemplify several aspects of the language of people with WS; they are syntactically well formed and complex, and display conversational hooks (Reilly, Klima, & Bellugi, 1991), such as the drawn out "oooh" that began SK's description of vampires, and her exaggerated "inordinate," which is also an example of complex vocabulary used appropriately.

is, in principle, possible in the absence of that shift having occurred. This is because the basic scope of the animal category (excepting the role of person) is already firmly in place in T1 (Carey, 1985; Keil, 1989). There is nothing proposed by either view currently discussed to prevent the acquisition or construction of new information predicated on these concepts through pure enrichment processes.

Take for example the property “breathing” as an observable bodily phenomenon of animals. The phenomenon of breathing itself can be interpreted in multiple ways. For instance, within the context of a biological theory (T2), breathing can be understood as a mechanism for maintaining life, by which a living organism exchanges oxygen and carbon dioxide with the surrounding air—the reasoning then goes, insofar as an organism has a life to maintain, it probably breathes in one way or another. On the other hand, in the context of a behavioral theory (Carey’s T1), breathing can be understood as the simple act of taking air in and out of the nose and mouth—in this case, insofar as an organism has a nose or mouth, it probably breathes. Or even more simply interpreted—breathing is something people do, therefore insofar as an object is like a person, it is likely to breathe—leaving open entirely the dimension of similarity invoked.

Three tasks assessed knowledge built on the T1/T2-Neutral concepts of animal, animal parts, and animal properties: attribution of bodily properties to animals and to non-living objects, the projection of a novel property taught on people, and size of the animal lexicon.

*Animal lexicon.* Preschool children have the concept *kind of animal* (Keil, 1989). Adding new exemplars of kinds of animals to one’s knowledge base, and learning their names, should not require conceptual change. Category naming tasks, such as this one, reflect the size, organization, and accessibility of the category in question, as well as some aspects of executive function (e.g., monitoring for repetitions, metamemorial, self-cueing strategies). Such a production task underestimates the total number of animals known, of course, since participants would surely retrieve many more animal names if cued with pictures, or verbal probes (e.g., zoo animals, sea animals, farm animals, pets, . . .).<sup>6</sup>

<sup>6</sup> Scott, Mervis, Bertrand, Klein, Armstrong, and Ford (1994) recently analyzed lists of animals produced in an untimed, category fluency task by 11 WS, 11 Down Syndrome matched on mental age and chronological age, and 22 normally developing children, half matched for mental age and half for chronological age with the two retarded samples. They found that the three groups matched on mental age (but not the chronologically matched normals) produced lists of comparable lengths, and also comparable content and organization. The likelihood of producing any given animal name as a function of judged typicality, frequency, or basic level was the same for each group. The data of Scott et al. (1994), like ours, failed to replicate those reported in Bellugi et al. (1993) in which the WS lists were longer and contained more low frequency and nonprototypical animals than matched controls.

*Attribution of bodily properties to animals.* This task assesses knowledge of properties of animals, tapping both directly retrieved knowledge and knowledge derived from similarity based inductive inference. It is modelled on those reported in Carey (1985; see also Inagaki & Hatano, 1987; Inagaki & Sugiyama, 1988). The participant is asked a series of simple yes/no questions about each of several animals and nonliving inanimates. The questions include probes of animal body properties such as "breathes" and "has a heart." Knowledge that people have these properties is most likely directly retrieved from memory, but some decisions, for example, whether birds or worms have hearts, depend on inference. Participants are unlikely to have learned such facts explicitly. Indeed, most adults do not know for sure whether worms have hearts or not. Carey (1985) argues that such inferences are similarity based, reflecting the distribution of known properties of animals (see Osherson, Smith, Wilkie, Lopez, & Shafir, 1990).

Normal performance on the attribution of bodily properties is characterized by both general and developmental effects (Carey, 1985; Inagaki & Hatano, 1987; Inagaki & Sugiyama, 1988). In general, normal participants at all ages previously studied use the distinction between animals and non-animals to constrain their inferences. Even children who hold T1 refrain from attributing animal properties to inanimate objects, including those they independently rate as highly similar to people (such as dolls or stuffed animals; Carey, 1985). This effect reflects the general importance of the concept *animal* in the conceptual structure of children as young as age 2 (Gelman & Coley, 1990).

There are, in addition, two developmental effects on the pattern of animal property attributions *within* the category of animal. First, young children attribute animal properties with decreasing frequency as the object becomes less similar to people. Preschoolers are less likely to judge that dogs breathe than to judge that people do, less likely still to judge that birds breathe, etc. Older children and adults, in contrast, attribute at least some animal properties (e.g., breathes, has babies) to all animals (Carey, 1985; Inagaki & Hatano, 1987). Second, preschoolers provide roughly the same pattern of attribution for all bodily properties. Older children and adults, on the other hand project some properties universally across the category animal, whereas others (has bones, hears, has a heart) comes to be restricted to vertebrates.

Carey (1985) interpreted these two developmental changes in patterns of attribution of bodily properties as reflecting the transition from T1 to T2. Indeed, categorical attribution of babies and breathing to all animals may reflect reanalysis of animals as living things with a universal life cycle and reanalysis of bodily parts and functions as supporting life. But it is also possible that nearly categorical attribution of such properties could be based on patterns of intercorrelated properties plus similarity-based inductive inference alone (see Osherson et al., 1990; Slomon, 1993); hence the inclusion of this task in the T1/T2-Neutral battery.

*Projection of a novel property from people.* This is an inductive inference task also modelled on Carey (1985). The participant is taught a novel property of people (has an omentum) and probed as to what other entities also have omenta. The use of a novel property tests the degree to which participants' knowledge is productive, and that the patterns generated in the attribution task are true inferences reflecting the role of the concept *animal* in a participant's conceptual system (whether T1 or T2), as opposed to the recitation of a list of memorized facts.

Carey (1985) interpreted developmental changes in inductive projection from people as reflecting the transition from T1 to T2. But as is the case of attribution of familiar animal properties to animals and inanimate objects, such changes could simply reflect a richer representation of the bodily properties of people and animals, still formulated over the concepts of T1 (people the prototypical animal; bodily properties and processes behavioral at their core.) Hence, this task is included in the T1/T2-Neutral battery.

### *T2-Dependent Battery*

Tasks in the T2-Dependent battery were selected to tap concepts thought to require conceptual change, including *life*, *death*, *the body-machine*, *person-as-one-animal-among-many*, and *species kind identity* based on origins (Carey, 1985; Keil, 1989; Johnson & Solomon, 1997). Unlike the tasks in the T1/T2-Neutral task, we propose that enrichment processes could not readily lead to enriched versions of preschoolers' concepts capable of supporting the adult-like (T2) performance. Five tasks were selected from the literature for this battery.

*Animism.* Childhood animism is the term Piaget (1929) used for a young child's propensity to claim that objects such as the sun, the moon, the wind, fire, cars, bicycles, etc., are alive. Carey (1985) interpreted this phenomenon as reflecting the child's undifferentiated concept *living/existing/animate/real*, which the child has mapped onto the term "alive." According to Carey (1985), it takes theory change within intuitive biology, not achieved during normal development until well into middle childhood, for these alternative interpretations of the word "alive" to be distinguished.

Younger children justify their judgments by appeals to usefulness, activity, existence; slightly older children by appeals to movement, older still by appeals to autonomous movement; and finally children's judgments begin to reflect the adult concept of *alive* when they invoke biological criteria such as bodily function or reproduction. Analysis based on levels of performance considers these justifications, as well as attribution of life to plants (a reflection of the conceptual change which results in the coalescence of plants and animals into the single ontological category of living things).

*Death.* The preschool child has not yet differentiated two senses of "not alive" into *dead* and *inanimate* (Carey, 1985). Preschool children interpret death in behavioral terms; a dead person has gone away, is somewhere else,

never to return. Sometimes they see the dead person as having fallen into a sleep-like state. Adult understanding entails the construction of a *body-as-biological-machine* concept, in terms of which the concept *death* is then reanalyzed. While properties such as absence remain important attributes of the dead even in the adult understanding, they are derived from a new core concept centered around the biological notion of the breakdown of the bodily machine.

*Projection of a novel property from dogs.* Carey (1985) found that when a novel bodily property is taught on a non-human animal preschool children project that property to other animals less than when it is taught on people. They are especially reluctant to project a novel property from a non-human animal to humans. For example, 4-year-olds project a novel property such as having an omentum from people to dogs around 75% of the time, but if the property were taught on dogs, project it to people less than 20% of the time (Carey, 1985). Normal 10-year-olds (and adults) project a novel bodily property with equal likelihood from dogs to people as from people to dogs. In the T2 of 10-year-olds, people and dogs apparently have equal status within the category animal, at least for the purposes of inferences about bodies. In contrast, in the T1 of preschool children people are the prototypical animal, in which *behaving being* rather than *living being* is at the core.

*Attribution of bodily properties to a tree.* The central ontological kind of T2 is living thing. The construction of this category is an example of the coalescence of two previously unrelated categories, animals and plants (Carey, 1985; but see also Backscheider, Schatz, & Gelman, 1993). The formation of the concept *living thing* includes reanalyses of bodily properties, such as breathing in terms of the role taking in air plays in supporting life, and reanalysis of the concept *baby* as reproductive offspring. These reanalyses support (though do not necessitate) attribution of breathing and babies to trees. T2 also sometimes supports attribution of hearts to trees, since participants sometimes reason that circulation of food and air to all parts of a body is required for life, and this is the function of hearts. Even attribution of hearing to trees could be licensed by T2, on some interpretation of the importance of sensitivity to sound to all living things. Note that the issue here is not the *correctness* of the attributions, but rather whether they show that children have analyzed trees as the same kinds of things as animals with respect to such properties. Preschoolers in the grips of T1 rarely attribute these properties to trees, since trees are not seen as behaving beings.

*Species transformations.* In order to assess the core of the concept of *species kind identity*, Keil (1989) studied the transformations participants judged would change the species of a given animal. Keil showed his participants a picture of a target animal (say a raccoon), and told a story of how this animal was transformed into one that looked like a member of another species (say a skunk). He found that the likelihood that the participant would judge that the raccoon had actually become a skunk depends upon the participant's age

TABLE 1  
Predictions of Both Views for Each Battery

	WS vs MC	WS vs YC	MC vs YC
<i>T1/T2-Neutral battery</i>			
Accretionist view—enrichment only	=	>	>
Conceptual change view—enrichment plus conceptual change	=	>	>
<i>T2-Dependent battery</i>			
Accretionist view—enrichment only	=	>	>
Conceptual change view—enrichment plus conceptual change	<	=	>

*Note.* =, no significant difference; >, significantly better than; <, significantly worse than.

and the nature of the transformation, which ranged from costumes, temporary surface paint, permanent plastic surgery, or an injection or pill given early in development.

In T1, physical characteristics are central to an animal's kind. Indeed, the youngest children (3-year-olds) judge that any transformation, including a costume change, yields an animal of the new species (DeVries, 1969).

Accepting that surgery can change species kind, whereas costume changes cannot, is an intermediary stage in normal development between T1 and T2. This pattern of responses indicates that the core of the child's concept of species kind now includes genuine *bodily* properties (not just surface appearances due to something the animal is wearing), but does not yet include consideration of the origin of those bodily properties. Keil (1989) found that kindergarteners judge that plastic surgery but not costume change yields a new species.

From the point of view of the 10-year-old's T2, origins of the animal and its properties are central to kind identity while its actual physical properties are peripheral. Keil (1989) found that by the fourth grade children reject transformations resulting from plastic surgery or costumes (though they still accept more internally mediated transformations caused by slow-acting injections.)

#### *Information Processing Demands of the Tasks of the Two Batteries*

The WS participants are predicted to perform at the same level as a group of matched controls on the T1/T2-Neutral Animal Knowledge battery, but much worse than their matched controls on the T2-Dependent battery (see Table 1). Compared to a group of younger controls, the WS participants are predicted to perform much better on the T1/T2-Neutral battery, but at the same level or worse on the T2-Dependent battery. To ensure that this pattern of results, if obtained, reflects the predicted differences in concepts and conceptual structure between the two populations, it is important that the task

demands of the two batteries are equated. This has been accomplished in several ways.

First, identical tasks are included in both batteries, differing only in the knowledge structures they call upon. For example, the inductive projection of a novel property of dogs to people is part of the T2-Dependent battery but the nearly identical task, inductive projection of a novel property from people is part of the T1/T2-Neutral battery. Similarly, the attribution of properties to animals is a T1/T2-Neutral task, but the same task construed over plants is a T2-Dependent task. Finally, the costume transformation task places identical information processing demands on the participant as does the surgery transformation task, yet young children always perform better on the costumes task than the operations task because the latter requires the T2 concept of species kind.

Secondly, wherever possible, the T2-Dependent tasks are scored on the bases of patterns of yes/no or forced choice judgments (animism, projection from dog, attribution to plants, species transformations). Finally, when justifications and explanations are analyzed, the analysis does not depend upon the ability to provide explanations per se, but on the content of what is said (animism, death). On these tasks, young preschool children provide scorable data, so it is unlikely that the task demands will defeat the WS adolescents and adults.

#### *An Important Caveat about the Two Batteries*

None of the tasks employed here provide operational definitions of the concepts they are being used to diagnose. It is possible that some of the more "advanced" response patterns of the T2-Dependent battery could, however unlikely, be achieved without true understanding of the underlying concepts. For instance a T2 judgment pattern on the animism task, in which only animals and plants are judged to be alive, could be achieved if a participant happens to have learned, as a piece of isolated knowledge, that both plants and animals, and only plants and animals, are said to be alive, even though they have not yet constructed the biological interpretation of life that unites these two into a coherent superordinate category. Similarly, participants could judge that plants have babies because they know that there are little ones, even though they have not yet constructed the concept of baby that is part of T2, involving a biological theory of reproduction.

One way to tell whether apparent T2 performance actually reflects T2 is to analyze consistency across tasks. For example, one WS participant, who provided a T2 pattern of judgments on the Animism task had almost certainly simply learned that only plants and animals are called "alive." This is shown by the fact that 62% of her justifications were appeals to existence, utility, activities, or irrelevant facts about the items (typical of preschoolers' justifications on this interview, see Carey, 1985) in contrast to 9% of the justifications of the matched controls who also achieved a T2 performance. Similarly,

if this participant had constructed the T2 concept of life, she should understand death as the cessation of life, due to breakdown of the bodily machine. However, her performance on the death interview contained no references to the bodily machine, indicating a concept of death closer to that of preschoolers.

Therefore, in addition to measure-by-measure comparisons of the performance of the three groups on the tasks, analyses of consistency of performance across tasks will also be presented.

Our main hypothesis could best be tested by matching WS participants to normally developing participants on the basis of performance on one of the two batteries of tasks, then comparing their performances on the other battery to test for correspondence or lack thereof. We adopted an indirect route to this end. We reasoned that a test of lexical knowledge, such as the Peabody Picture Vocabulary Test-Revised (PPVT-R; Dunn & Dunn, 1981), would be a good predictor of performance on the T1/T2-Neutral Animal Knowledge battery, since lexical growth of the type tapped by the PPVT-R generally requires no conceptual change. That is, the harder items on the PPVT-R are harder just because they are rare; hard items are from the same ontological categories (tools, animals, vehicles, clothes, human physical or social interactions, human occupation, etc.) as are easier ones. All these categories are available to preschool children. In addition, the PPVT-R provides standardized scoring which allows fine-grained matches between WS participants and normally developing participants predicted to have comparable general knowledge to the WS participants. Furthermore, in order to maximize the chance that the WS participants would possess T2, and that they would possess the general knowledge of typical 9–10-year-olds, we selected only those who performed at least as well as normally developing 9-year-olds on the PPVT-R (i.e., had PPVT-R mental ages of at least 9).

## METHODS

### Participants

#### *Williams Syndrome Participants*

Ten individuals with Williams syndrome were recruited through the National Williams Syndrome Association and Boston's Children's Hospital. They ranged in chronological age (CA) from 10 years 7 months to 32 years 1 month ( $M = 24$  years 3 months.) They ranged in PPVT-R mental age (MA) from 9 years 1 month to 15 years 11 months ( $M = 11$  years 5 months.) Standard scores (SS) ranged from 59 to 102 ( $M = 71$ ). The majority of participants were tested at their homes and were paid for their participation. Three people volunteered their participation at the Association's national conference in San Diego.

#### *Matched Control Participants (MC)*

Each WS participant was matched to one of ten normally developing children ( $M_{CA} = 9$  years 10 months, range 8 years 5 months to 12 years 7 months;  $M_{MA} = 10$  years 11 months,

range 8 years 7 months to 14 years 7 months;  $M_{SS} = 109$ , range 105 to 114). MC participants were selected for mental ages within one standard deviation of both (1) the average for their own chronological age and (2) that of their WS matches. The average mental age difference within each matched WS-MC pair was 3 standard points. MC participants were from the Cambridge public schools and a summer day camp for the children of MIT employees. All testing was done in private sessions at the school or camp.

### *Younger Control Participants (YC)*

Nine normally developing children were selected to form a younger control group. In order to get as homogeneous a group as possible, selection criteria were strict. Only children (1) between 5 years 9 months and 7 years of age and (2) with standard scores within 10 points of average (90 to 110 inclusive) were included. The nine selected participants had a mean chronological age of 6 years 5 months (range 5 years 10 months to 6 years 11 months), a mean mental age of 6 years 7 months (range 6 years 1 month to 7 years 2 months) and a mean standard score of 101 (range 90 to 108.) Recruitment and testing procedures were the same used with the matched control group.

## Materials and Procedure

### *T1/T2-Neutral Animal Knowledge Battery*

*Animal lexicon.* Participants were simply asked to list all the animal kinds they could think of. We placed no time limit on participants. Participants were allowed to continue until they stated they had finished. If a participant did not spontaneously state this, the experimenter waited for a pause of about 15 seconds and then asked if they were done.

*Attribution of bodily properties to animals.* The participant was asked a series of simple yes/no questions about each of seven pictured objects from the categories of animals and non-living inanimates (a series of questions about trees was also included among these items but is discussed below as a part of the T2-Dependent battery). The questions included four probes of animal body properties. The objects were people, dogs, birds, worms, computers, the sun, and ragdolls. The bodily properties of animals were: breathes, has a heart, hears, and has babies. The questions about bodily properties were intermixed with filler questions designed to minimize response biases and to ensure that the task demands were within the capabilities of the participants. Fillers were questions that young 3-year-olds can answer easily, such as "Is the sun hot?" and whenever possible find amusing as well, such as, "Do dogs live in refrigerators?" or "Do worms wear clothes?"

All participants were presented with the same fixed order of items. Objects were presented in groups of four. The first group of objects appeared early in the first session and consisted of person, tree, computer, and bird. The second group appeared late in the second session (or first session if a participant finished in only one session) and consisted of the sun, worm, ragdoll, and dog. Order of property items was different for each object, but constant across participants. Questions for each object always began with an easy filler for which the answer was "yes" followed by an easy "no" question.

*Projection of a novel property taught on people.* This is a similarity based inductive inference task from Carey (1985). Participants were asked if they had ever heard the word "omentum," told they were being taught a new word, and asked to say it out loud. They were then shown a schematic sketch of a red, round thing, and told, "See this, this is an omentum. Lots of things have omentums in them. One of the things in the world that has an omentum in it is people. People have omentums right about here inside (experimenter points to the mid section of a picture of a person.))" The participant was then asked which of the objects used in the attribution task, with the added item, cow, have omentums. Objects were always probed in the following order: trees, cows, dogs, computers, birds, worms, the sun, ragdolls, and

people. As a check for memory of the information provided during the teaching, the taught-on animal, in this case people, was probed last.

### *T2-Dependent Battery*

*Animism.* The animism task we administered was adapted from the standardized version of Laurendeau and Pinard (1962). Participants were asked if they knew what it means for something to be alive and asked to provide examples of things which are and are not alive. Then, 20 pictures were presented and named, the participant judged whether each is alive, and was asked to justify each judgment. The 20 objects were drawn from the categories of animals, plants, non-biological natural kinds, and artifacts and included cat, bird, fly, snake, tree, flower, mountain, sun, fire, wind, cloud, rain, table, car, pencil, bicycle, watch, airplane, bell, and lamp.

*Death.* This interview is adapted from those of Koocher (1974) and Nagy (1948). The interview included the following questions in order: "Do you know what it means to die? Can you name some things that do die? What happens to a person when they die?" This question was followed up with more specific probes about the fate of the body or the spirit if that information had not already been volunteered, for instance: "What happens to a person's body when they die? What happens after they're buried?" The interview ended with the questions: "Do you know what might cause a person to die? Does every person die? Once something dies, is there anything anyone can do to make it live again? Can a doctor make a dead person live again?"

*Projection of a novel property from dogs.* This was a second inductive inference task exactly like the omentum task of the T1/T2-Neutral battery, with one difference. Participants were told that they were being taught a new word, "golgi," and that golgies are long green things, that many things in the world have golgies inside them, and that one of the things in the world with a golgi inside it is a dog (pointing roughly to the midsection of a pictured dog as the location of the dog's golgi.) The participant was then asked whether each of the same eight objects probed for omenta have a golgi inside it. The order of objects was the same as in the omentum task except that the positions of dog and people were reversed so that the taught-on-animal, dog, was probed last.

*Attribution of properties to a tree.* Embedded in the larger attribution task of the T1/T2-Neutral battery was a set of questions about the properties of trees, including the four biological properties: breathes, has a heart, hears, and has babies.

*Species transformations.* Our tasks, taken directly from Keil (1989) included two stories of costume transformations (a goat costumed to look like a sheep and a zebra costumed to look like a horse) and two stories of surgery transformations (a raccoon turned into a skunk-look-alike and a tiger turned into a lion-look-alike). On each trial the participant was shown a picture of the original animal (e.g., a raccoon). The story of the transformation was then told and the participant was shown a picture of what the animal looked like afterwards (e.g., a skunk), and asked "Now, when the doctor is done and the animal looks just like this (touching the picture of the transformed animal, the skunk) what kind of animal is it? Is it a skunk or a raccoon?" The participant's response was then challenged. If the participant judged that the animal was unchanged (i.e., still a raccoon), the challenge was: "Even though it looks just like a skunk, you think it is what?" If the participant judged that the animal was changed (i.e., now a skunk), the challenge was: "Even though it was a raccoon to start with, and its parents were raccoons, you think it is what?"

## RESULTS

The data were analyzed with respect to the predictions from the conceptual change hypothesis. Each task in each battery was chosen in part because it

reveals developmental change over this age range. As can be seen from Table 1, both hypotheses predict that the older group of control participants (MC) will perform better than the younger group of control participants (YC) on both batteries. In addition, both hypotheses predict that the WS participants will perform at the level of the matched controls on the T1/T2-Neutral battery. The two hypotheses differ only in their predictions for the T2-Dependent battery. The accretionist position predicts the same pattern of results for the T2-Dependent battery as for the T1/T2-Neutral battery since conceptual change is not considered a prerequisite for the construction of the 10-year-old's biology. The conceptual change hypothesis, in contrast, predicts that the WS participants will perform worse than their matched controls on this battery, at the level of the younger controls.

The results will be discussed in three stages. First we examine the predictions of the conceptual change position outlined in Table 1. Paired *t* tests are used to examine the relative performances of the WS and MC participants within the individually matched pairs. The corresponding WS/YC and MC/YC comparisons are based on unpaired *t* tests. Although multiple *t* tests are performed, this method of analysis does not raise the problem of inflated *p*-values, since the predictions are *a priori*, precise, and expected for every comparison. In each case for which the conceptual change position predicts a difference, one-tailed tests were used. In each case where no difference was predicted, two-tailed tests were used.

Secondly, we examine the specific progress of the participants on the path between T1 and T2. On the T2-Dependent tasks, MC subjects were predicted to be closer to T2 than T1 performance. Both of the other groups were predicted to be closer to T1 performance than T2.

Finally, aspects of individual performances across tasks and batteries were examined. Each WS participant was predicted to perform relatively better on the T1/T2-Neutral tasks than on the T2-Dependent tasks compared to the normally developing children. For the normally developing children in both groups we expected that performance on the T1/T2-Neutral tasks would predict performance on the T2-Dependent tasks.

## Paired Comparisons between the WS and MC Participants; 6-Year-Olds

### *T1/T2-Neutral Animal Knowledge Battery*

Summary results from the T1/T2-Neutral Animal Knowledge battery are presented in Table 2.

*Animal lexicon.* Two of the WS participants were not given this task, leaving eight matched pairs for direct comparison. Each participant was assigned a score based on the total number of animal names produced, minus any repetitions and intrusions. As predicted, there was no difference between the WS participants (average 14.4 animals listed) and their matched controls (average 14.8 animals listed),  $t(7) = -0.16$ , n.s., 2-tailed (WS = MC). Of

TABLE 2  
Summary Measures from All Tasks

	WS score (S.E.)	MC score (S.E.)	YC score (S.E.)	Value of statistical comparisons		
				WS vs MC	WS vs YC	MC vs YC
T1/T2-Neutral battery						
Animal lexicon—net number of exemplars produced	14.4 (2.1)	14.8 (1.7)	8.0 (1.4)	=	>	>
Attribution of properties to animals— general percentages	95.0 (2.2)	97.0 (2.1)	80.0 (8.3)	=	>	>
Universal/restricted distinction— difference score of percentages	50.0 (11.2)	15.0 (15.0)	8.0 (10.7)	=	>	=
Projection of novel property from people— difference score of percentages	77.0 (8.0)	77.0 (8.0)	82.0 (8.0)	=	=	=
T2-Dependent battery						
Animism—% correct on inanimates	53.0 (12.9)	86.0 (7.2)	83.0 (11.0)	<	=	=
Animism—level analysis	2.8 (0.6)	5.0 (0.4)	4.2 (0.8)	<	=	=
Death—level analysis	-0.7 (0.2)	1.1 (0.2)	-0.3 (0.3)	<	=	>
Attribution of properties to trees— difference score of percentages	20.0 (4.4)	32.5 (4.3)	8.3 (3.5)	<	=	>
Projection of novel property from dog— yes/no projection to people	20.0 (13.0)	70.0 (15.0)	11.0 (11.0)	<	=	>
Species costume transformation— degree of resistance to change	1.8 (0.2)	2.9 (0.1)	2.2 (0.2)	<	=	>
Species surgery transformation— degree of resistance to change	1.4 (0.2)	2.2 (0.2)	1.2 (0.2)	<	=	>

the eight pairs, three reflected higher scores among the controls and five reflected higher scores among the WS participants, Wilcoxon  $z = -0.07$ , n.s., 2-tailed.

The 6-year-olds listed an average of 8.0 animals in this production task. This was significantly fewer than both the WS participants,  $t(14) = 2.52$ ,  $p < .02$ , 1-tailed, and the MC participants,  $t(14) = 3.04$ ,  $p < .005$ , 1-tailed (WS > YC; MC > YC).

*Attribution of bodily properties to animals.* With the exception of one WS participant who denied that people have babies, all participants attributed all four bodily properties to people. Both WS participants and their matched controls answered an average of 99% of the filler questions correctly; YC participants answered 97% of the filler questions correctly. Both also over-attributed bodily properties to inanimates only 5% of the time, and the YC participants never did so, reflecting the normal constraint of the concept *ani-*

*mal* on performance for all groups. Thus, all groups of participants easily performed this task.

The first measure of theoretically neutral knowledge of animals derived from this task taps how widely participants attribute the bodily properties to non-human animals.<sup>7</sup> For this analysis, attribution of hearing and hearts to worms was removed, since adults do not always consider worms to have hearts or to hear (as reflected in the second measure below). The percentage of attributions of all four bodily properties to dogs and birds and of having babies and breathing to worms was calculated. The WS participants and their matched controls do not differ on this measure (95% to 97%,  $t(9) = -0.69$ , n.s., 2-tailed.) Two of the WS participants performed higher than their matched controls, three of the matched controls performed higher than their WS matches, and five pairs performed equally, Wilcoxon  $z = -0.71$ , n.s., 2-tailed.

The 6-year-olds attributed bodily properties only 80% of the time to dogs, birds, and worms (excluding heart and hear). This score reflects their tendency to rely less categorically on the concept of animal than either the WS group ( $t(17) = 1.82$ ,  $p < .05$ , 1-tailed) or the MC group ( $t(17) = 2.07$ ,  $p < .03$ , 1-tailed). Again this measure shows a normal developmental pattern (MC > YC) where the WS participants perform at the older range of the pattern rather than the younger end (WS = MC; WS > YC).

The second measure of theoretically neutral knowledge was derived from an analysis of the differentiation of the universal bodily properties (breathes, has babies) from the restricted bodily properties (hears, has a heart). Adults typically map this distinction onto the distinction between vertebrate and invertebrates (Carey, 1985). Therefore the relative attribution of the two sets of properties to worms were compared to that for dogs and birds. A difference of difference score was calculated,  $DD = [\% \text{attribution worms (breathes/babies-heart/hears)} - \% \text{attribution dog/bird (breathes/babies-heart/hears)}]$ . A higher score reflects a sharper (more adult-like) distinction. WS participants actually outperformed their matched controls (50% to 15%) on this task, but the difference did not reach significance,  $t(9) = 1.87$ ,  $p < .10$ , 2-tailed. Six of the WS participants performed better than their matched controls, two performed worse, and two pairs performed equally, Wilcoxon  $z = 1.70$ ,  $p < .10$ , 2-tailed. Indeed the WS score (50%) did not differ from that of 18 MIT undergraduates given this task (41%).<sup>8</sup> This apparent difference between the WS participants and their matched controls is due to the fact that the matched controls attributed hearing and having hearts to worms (85%) whereas the WS participants did so less (35%).

<sup>7</sup> Interested readers should contact the first author for more detailed reports of these and other data not fully described here.

<sup>8</sup> The normal adult data are not reported here; for most measures adults performed significantly better than both the WS subjects and the matched controls. That is, the equal performance of WS and MC subjects is not a ceiling effect.

The 6-year-olds showed little evidence of having distinguished between universal and restricted properties in their attribution across the universal/restricted distinction. Their average difference score of 8.0 was significantly lower than that of the WS participants ( $t(17) = 2.26, p < .02$ ). They did not differ from the MC participants on this task ( $t(17) = 0.37, n.s.$ ). Thus, on this measure, the pattern of performance was  $WS = MC$ ;  $WS > YC$ ;  $MC = YC$ .

*Projection of a novel property from people.* Each participant's performance on this task is captured by a difference score between the percentage of animals attributed omentums minus the percentage of inanimate objects attributed omentums. Both the WS and MC groups achieved a difference score of 77% ( $t(9) = 0.011, n.s., 2$ -tailed). Five WS participants performed higher than their matched controls, three control participants performed higher than their matched WS participants, and two pairs performed equally, Wilcoxon  $z = 0, n.s., 2$ -tailed. The younger controls averaged a difference score of 82% which was comparable to that of the WS group  $t(17) = -.042, n.s.$ , and the MC group  $t(17) = -0.43, n.s.$ ;  $WS = MC = YC$ .

The failure to find a developmental difference between the two normal groups suggests that there is a ceiling effect for all three groups on this task. Nonetheless this task reconfirms both the availability of the concept animal to WS participants and their ability to use it productively for normal inference.

### *Summary of Comparisons on the T1/T2-Neutral Animal Knowledge Battery*

The WS participants did not differ from their matched controls on any of the measures of theoretically neutral knowledge of animals and bodily properties, except for one measure in which they marginally outperformed the controls. These data support the prediction that PPVT-R predicts general knowledge in the domain of animals. They also ensure that any differences between the two groups on the T2-dependent measures are not due to vastly different amounts of available information about animals and their properties.

Our normally developing group of 6-year-olds performed worse than both the WS and MC groups on measures of animal naming and animal property attribution. In addition, like MC participants, they performed worse than the WS participants with respect to their sensitivity to the distinction between universal bodily properties (e.g., has babies) and restricted bodily properties (e.g., has a heart). These three results confirm that these tasks are sensitive to normal developmental patterns of knowledge enrichment. The final task in this battery—the projection from people—failed to show any developmental pattern, probably due to a general ceiling effect. It nonetheless confirms that the knowledge that WS participants have is not simply rote knowledge, but is embedded in productive conceptual structures which support novel inferences.

Overall these results confirm the first half of our hypothesis: that participants with WS are unimpaired on T1/T2-Neutral knowledge of animals relative to their PPVT-R mental ages.

### *T2-Dependent Battery*

Each task in the T2-Dependent battery diagnoses a particular concept within T2 which it is hypothesized cannot be represented in T1. As indicated on Table 1, the conceptual change hypothesis demands that the WS participants be worse than the MC participants on each task in this battery. Results for all tasks are shown in Table 2.

*Animism.* The performance of each participant was scored in two different manners. First, judgments alone were analyzed, each participant receiving a score that reflects the percentage of inanimate objects correctly judged not alive. WS participants denied life to only 53% of the inanimates while their matched controls did so 86% of the time,  $t(9) = 2.45$ ,  $p < .02$ , one-tailed. In three of the pairs, the percentage of attribution to inanimate objects was the same (0% in two cases), in one of the pairs, the MC participants had a higher overall level of attribution to inanimate objects, and in six of the pairs, the WS participants attributed life to more inanimate objects, Wilcoxon  $z = -2.03$ ,  $p < .03$ , one-tailed (WS < MC). In a sample of 500 children between ages 4 and 12 by Laurendeau and Pinard (1962), 4-year-olds correctly denied life, on average, to only 57% of the inanimate objects, a rate comparable to that of the WS participants here (53%). Laurendeau and Pinard's 10-year-olds' rate (82%) is comparable to the present matched control participants' rate (86%).

The YC participants were able to deny life to non-living objects 83% of the time. This was comparable to the MC performance,  $t(17) = -0.29$ , n.s., one-tailed, and better (fewer overattributions) than the WS participants, but not significantly so,  $t(17) = 1.69$ , n.s., two-tailed (WS = YC; MC = YC).

Performance on the animism task is better interpreted, however, when justifications are taken into account. For this purpose, children were assigned to performance levels based on pattern of judgments and the nature of explanations given (for similar coding schemes see Carey, 1985; Laurendeau & Pinard, 1962). Coding criteria for each performance level are described in Table 3. One experimenter coded all 29 participants. A secondary coder completed a subset of two-thirds of the participants and agreed with 90% of the assignments.

WS participants achieved an average level of 2.8 (range 1–6), whereas MC participants had an average level of 5.0 (range 3–6;  $t(9) = -3.60$ ,  $p < .01$ , one-tailed). Only one of the MC participants scored below a level 4; six of the WS participants scored below level 4. In one of the matched pairs the scores were tied; in eight of the matched pairs the controls outperformed the WS participant and in one pair the WS participant outperformed his or her matched control, Wilcoxon  $z = -2.44$ ,  $p < .01$ , one-tailed (WS < MC).

TABLE 3  
Coding Criteria for Each Performance Level on the Animism Task

Level 0	Response biases, all yes or all no; random judgments, some animals judged not alive, some inanimates judged alive.
Level 1	Yes to all animals plus at least one animistic judgment that an inanimate object is alive. Explanations of activity, utility, or existence.
Level 2	Yes to all animals plus at least one animistic yes. Explanations include at least one explicit reference to movement (object specific movements (e.g., air-planes fly and birds fly) were considered activities and assigned to Level 1).
Level 3	Yes to all animals plus at least one animistic yes. Explanations include at least one explicit reference to autonomous movement.
Level 4	Yes to all animals plus at least one animistic yes. Explanations include at least one reference to non-object-specific biological process such as breathing, eating, growing, etc.
Level 5	Yes to all animals, no animistic yeses, but failure to attribute life to one or both plants in the list.
Level 6	Yes to all animals and plants, no animistic yeses.

On the levels analysis, the 6-year-olds scored an average of 4.2 (range 0–6) which although in the intermediary range between the WS and MC, was significantly different from neither,  $t(17) = -1.51$ , n.s., two-tailed, and  $t(17) = .95$ , n.s., one-tailed, respectively (WS = YC; MC = YC).

*Death.* Each participant was assigned a level of understanding based on their answers to the questions: “What does it mean to die?” “What happens to a person when they die?” and its follow-ups. Four levels were identified. Level -1: Death was described as the absence, departure, or altered sleep-like state of the dead person. Level 0: No evidence was offered for any interpretation of bodily death. Level 1: Death was described as a cessation of behavioral processes such as talking, moving, or thinking. Level 2: Death was described as the cessation of bodily processes such as the heart beating, breathing, or the blood circulating. In cases where participants offered responses falling into more than one level they were credited with the higher level. All responses reported are based on the coding of a single primary coder. A secondary coder achieved 93% agreement with these assignments.

One WS participant failed to complete this task, leaving nine pairs for comparisons. As predicted the MC participants outperformed the WS participants. Scores of the MC participants ranged from 0 to 2, with an average of 1.11, whereas those for the WS participants ranged from -1 to 0, with an average of -0.67,  $t(8) = -5.49$ ,  $p < .001$ , one-tailed. In eight of the nine pairs, the MC participant outperformed the WS participant; in one pair, participants were tied, Wilcoxon  $z = -2.56$ ,  $p < .01$ , one-tailed (WS < MC).

The 6-year-olds achieved an average Level score of -0.33 on this task. This was comparable to the WS group  $t(16) = -1.16$ , n.s., two-tailed, and

significantly lower than the MC group,  $t(16) = 3.71, p < .001$ , one-tailed (WS = YC; MC > YC).

Participants' answers were also coded for a variety of response types relevant to the change in the concept of death between ages 6 and 10 that have been documented in the literature (Carey, 1985; Koocher, 1974; Nagy, 1948). These included whether or not participants explicitly defined death in terms of life, and various misconceptions including that death is not universal to people and animals, that it is reversible, and that some inanimates can die. Only 30% of the WS participants defined death as the complement of life, whereas 70% of the MC participants did this. Four of the WS participants held at least one misconception about death, such as that not all people die. In contrast only two MC participants held misconceptions and in both cases they involved uncertainty about whether certain medical techniques might be sufficient for bringing a dead person back to life.

Participants' answers were also coded for a variety of factual elaborations of the process of death. Fifty percent of the WS participants and 40% of the MC participants elaborated on the fact that the body decays after death; 70% of the WS participants also mentioned something about heaven during their interviews, as did 40% of the MC participants. Thus, the WS participants had equally or more elaborated factual knowledge (heaven, bodily decay) than did the MC participants, in the face of a preconceptual change concept of death (levels analysis, misconceptions, failure to relate death to the concept of life).

*Projection of a novel property from dogs.* For the present purpose only projections to a person are analyzed. Only two of the WS participants projected golgies from dogs to people, compared to seven of the MC participants,  $t(9) = -2.24, p < .03$ , one-tailed. In six pairs of participants, the MC made the projection but the WS participant did not. In one pair the WS participant projected to people but the MC participant did not and in three pairs either both participants did or both did not, Wilcoxon  $z = -1.89, p < .03$ , one-tailed (WS < MC). Overall, the WS participants were less likely to have assimilated people into their category of animals as equivalent biological beings than were the matched controls.

The 6-year-olds performed very much like the WS participants projecting golgies to people only 11% of the time,  $t(17) = .051$ , n.s., two-tailed. Like the WS participants this was significantly less than the MC participants,  $t(17) = 3.06, p < .005$ , one-tailed (WS = YC; MC > YC).

*Attribution of properties to a tree.* Performance on this task was measured by the percentage of bodily properties a participant attributed to trees minus the percentage of bodily properties attributed to inanimates. Scores could range from -100 to 100. WS participants attributed the biological properties (breathes, has babies, has a heart, hears) to trees an average of 1.0 times, whereas their matched controls did so an average of 1.5 times. Both groups also attributed some bodily properties to inanimates. The resulting difference

scores of the two groups, 20.0 and 32.5, respectively, were significantly different,  $t(9) = -1.96$ ,  $p < .05$ , one-tailed. Difference scores were higher by the MC participant in seven of the pairs and higher by the WS participant in the remaining three pairs, Wilcoxon  $z = -1.75$ ,  $p < .05$ , one-tailed (WS < MC).

The YC participants achieved an average difference score of 8.3, differing significantly from the matched control participants,  $t(17) = -3.30$ ,  $p < .005$ , one-tailed, but not from the WS participants,  $t(17) = 1.57$ , n.s., two-tailed (WS = YC; MC > YC).

*Species transformations.* Each story was scored by the method described in Keil (1989). Participants received a score of 1 if they judged that the animal's kind was changed by the transformation and stuck to that response throughout the challenges. Participants received a score of 2 if they showed any vacillation or change of mind about the answer. Participants received a score of 3 if they judged that the animal's species identity was not changed during the transformation and resisted the challenges. Each participant received a composite score for each pair of transformation stories.

Two control participants and one of the corresponding WS participants were not administered the species transformation tasks, leaving eight pairs for this analysis. On the costume stories WS participants were unsure if the animal's kind identity would change with scores that ranged from 1.0 to 3.0 and an average score of 1.8. Matched control participants, in contrast, denied that a costume change changed species kind, with scores ranging from 2.5 to 3.0 and an average of 2.9,  $t(7) = -4.82$ ,  $p < .001$ , one-tailed. In seven of the eight pairs, control participants outperformed WS participants; in the eighth pair, the two participants were tied, Wilcoxon  $z = -2.39$ ,  $p < .01$ , one-tailed (WS < MC).

The YC participants were also more likely than the MC participants to consider the animal's kind changed with an average score of 2.2 (range 1.5 to 3),  $t(14) = 3.79$ ,  $p = .001$ , one-tailed. They did not differ from the WS participants in this respect,  $t(14) = -1.33$ , n.s. (WS = YC; MC > YC).

Although young preschoolers accept that a *costume* change can change species kind, by age 5 children have moved away from T1 in this respect (Keil, 1989), as have both our WS and YC participants. That is, few of the participants firmly accepted that costumes can change species kind. The *surgery* task, however, requires T2. On the surgery task WS scores ranged from 1.0 to 2.0, with an average of 1.4 and MC scores ranged from 1.0 to 3.0 with an average of 2.2,  $t(7) = -3.87$ ,  $p < .005$ , one-tailed. Two of the eight pairs had tied scores; in the remaining six pairs, the MC participants outperformed the WS participants, Wilcoxon  $z = -2.23$ ,  $p < .02$ , one-tailed (WS < MC). Ultimately most of the control participants *denied* that surgery could change species kind (the T2 judgment), whereas most of the WS participants *affirmed* that it could do so.

The 6-year-olds, like the WS participants, did not make the judgments

that are diagnostic of T2. Their average score was 1.1 (range 1–1.5), differing from the MC participants,  $t(14) = 4.71$ ,  $p < .001$ , one-tailed, but not the WS participants,  $t(14) = 1.25$ , n.s., two-tailed (WS = YC; MC > YC).

The parallel design of the two species tasks also helps us to address the question of whether task demands, thought of in information processing terms, rather than conceptual content, determine the developmental difference between T1 and T2, or the difference between the WS and the MC participants. The two versions of the task, costume and plastic surgery, differ only in content, yet studies with full complements of normally developing children show clear performance differences between versions (Keil, 1989). So too in our data; as can be seen from Table 2, all three groups of subjects performed about one-half of a level higher or more on the costume transformation task than on the surgery transformation task.

### *Summary of Comparisons on the T2-Dependent Battery*

The participants with Williams syndrome had failed to acquire just those concepts not available in T1, just that knowledge that requires conceptual change. The WS participants performed worse than their matched controls on every task in the T2-Dependent battery. They were largely animistic, judging that inanimate things are alive. The majority gave departure or altered state interpretations of death, saying that death means to go away or to sleep, rather than the breakdown and cessation of the bodily machine. In their inductive inferences, they did not treat people as just one animal among many, as they denied that people might share a novel internal property with animals if the property was first encountered in the context of animals. They did not treat trees like animals in the attribution of biological properties. The majority firmly believed that an animal's species identity could be changed with surgery and were unsure whether it could be changed with a costume. These beliefs mirrored those of the normally developing 6-year-olds, who also performed worse than the matched control group on every T2-Dependent measure, with the exception of those from the animism interview. Like the tasks in the T1/T2-Neutral battery, the T2-Dependent tasks showed substantial sensitivity to normal development.

In summary, the overall results of the pairwise comparisons supported the predictions of the conceptual change hypothesis. Thirty-three planned  $t$  tests were performed with precise one- and two-tailed predictions. Of those, 28 yielded the results predicted by the conceptual change hypothesis. The remaining 5 that failed to do so were in comparisons that did not differentiate between the predictions of the two hypotheses.<sup>9</sup>

<sup>9</sup> Of the 17 one-tailed predictions confirmed by the data, only 4 actually depended on the one-tailed prediction. If 2-tailed, two of the four would have resulted in  $p$ 's  $< .06$  and two in  $p$ 's  $< .09$ . (The results of the 7 one-tailed Wilcoxon's performed between the WS and MC mirrored those of the  $t$ -tests.)

Table 4  
Theoretical Level Analyses

	% Participants with T2 concepts			% Participants with T1 concepts		
	WS	MC	YC	WS	MC	YC
Animism—% correct on inanimates	20	60	67	60	30	22
Animism—level analysis	10	50	44	60	0	33
Death—level analysis	0	50	0	67	0	44
Projection of novel property from dog	20	70	11	80	30	89
Attribution of properties to trees	30	60	0	40	10	67
Species costume transformations	11	75	12	22	12	0
Species surgery transformations	11	12	0	56	12	75
Overall	15	55	23	55	12	48

### Analyses of Theoretical Level

The WS participants performed worse than their matched controls on every task in the T2-Dependent battery. There is a stronger prediction from the conceptual change hypothesis, however. If WS prevents conceptual change, allowing only knowledge acquisition by accretion, then as a group the WS participants' core concepts should still be those of T1, albeit enriched beyond those of preschoolers. Similarly, on the analysis of Carey (1985), the YC participants are at the beginning of the process of constructing T2, so they too should largely reveal T1 concepts of animals and plants. The MC participants, on the other hand, are at the age at which a life-based, vitalist, intuitive biology (T2) has been constructed and their performances as a group should reflect that.

To test these hypotheses each participant was scored with respect to whether they were at the beginning level (T1) or highest level (T2) on the T2-Dependent battery as described under Constructing the Batteries (see the Appendix for explicit coding criteria on each task). Results are shown in Table 4. Note that participants' performances can, and often do, reflect neither extreme on a given task—neither T1 nor T2 performance, but rather an intermediate level. Intermediate scores are not entered on Table 4.

Based on the literature from which these tasks were taken, we expected the matched controls to have achieved those aspects of T2 probed here. However, as shown on Table 4 they have not done so. Over all the tasks in the T2-Dependent battery, only 55% of the matched control participants performed at the T2 level. This result seems low until we recall that the estimates of ages at which developmental milestones take place in the literature on cognitive development are not placed in a psychometric context. In order to find the 10 participants for our matched controls and the 9 participants in the younger control group, we screened 72 participants from the populations we generally use for our studies on cognitive development (a university day

camp and both public and private schools in the university vicinity). Of those 72 normally developing children, only 26 had PPVT-R standard scores within one standard deviation of average, 36 had scores more than one standard deviation higher than average, and 10 had scores lower. Thus, the mental ages of children in our previous studies of folkbiology, like those of other researchers, are probably substantially higher than their chronological ages. This suggests that the true *average* age at which T2 is constructed is slightly later than estimated by the recent literature.

Nonetheless, as Table 4 shows, the MC participants are well on their way toward constructing the folkbiological concepts tapped by this battery, with half or more achieving the T2 level on every task but the surgery transformation. The WS participants, in contrast, rarely performed at the T2 level (15% over all the tasks in the T2-Dependent battery).

On the other hand, WS participants performed at the T1 level 55% of the time (Table 4). On all but two measures (the costume transformation measure (on which the normal developing 5-year-olds of Keil, 1989, perform at the T2 level) and the attribution of bodily properties to tree) over half of the WS participants were still at the T1 levels of performance. The matched control participants, in contrast, very rarely performed at the T1 level (12% overall).

Our hypothesis does not require that *all* of the WS participants perform at the T1 level on *all* the tasks. Just as T2 performance does not provide operational measures of conceptual change, T1 performance cannot provide operational measures of its absence. T1/T2-Neutral knowledge, gained by enrichment, may begin to move participants away from the T1 level. Given this, we see these data as a striking confirmation of our hypothesis. Our WS participants ranged in age from 10 to 30, and in mental age from 9 to 16, yet on over half of the tasks in the T2-Dependent battery, they performed as young preschoolers who are consistently in the grips of T1.

### Analyses of Individual Performances

So far we have addressed several questions with respect to the group data. First we showed that pair by pair the WS participants performed equally to or better than their matched controls on each of the T1/T2-Neutral tasks, but significantly worse on each of the T2-Dependent tasks. Second, we showed that for the T2-Dependent tasks as a whole, the WS participants were performing closer to the T1 level and, conversely, the MC participants were performing closer to the T2 level. For the most part the WS participants looked like the 6-year-olds who were also still largely at the T1 level. But three important questions have not yet been addressed.

First, does the claim that WS participants do not construct T2 hold at the individual level as well as the group level? That is, 16% of the WS performances on the T2-Dependent tasks were at the T2 level. Was this the result

of one or two individuals who achieved T2 levels on multiple tasks, like a normally developing matched control participant? Similarly, given that the matched controls were not themselves at ceiling on the T2-Dependent tasks, did any of them look like the WS participants?

Second, does the differential impairment across batteries seen in the WS group data accurately reflect individual performance as well? That is, does *each* WS participant perform better on the T1/T2-Neutral tasks than on the T2-Dependent tasks relative to the normally developing children?

Finally, in normally developing children, does the amount of T1/T2-Neutral animal knowledge possessed actually predict understanding of the life cycle and body function, i.e., T2-Dependent knowledge?

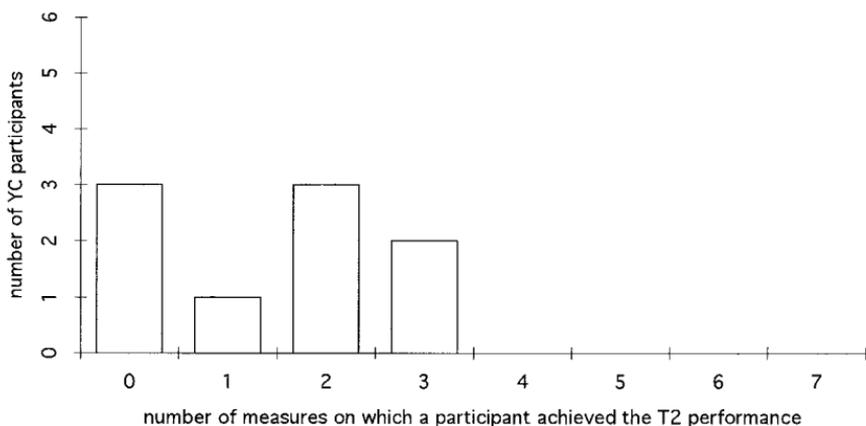
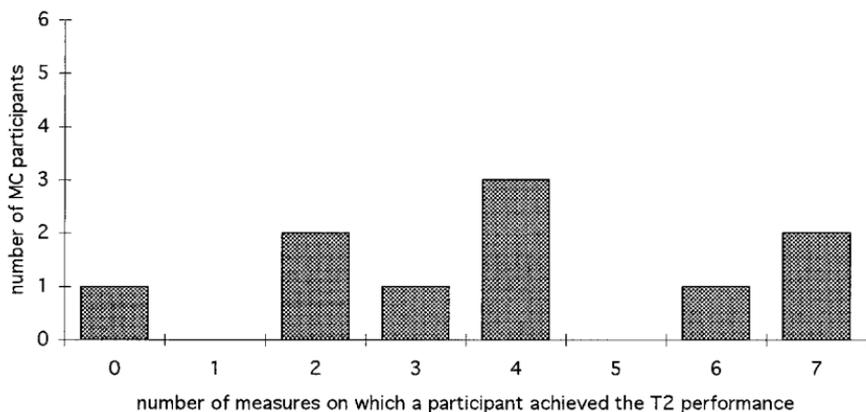
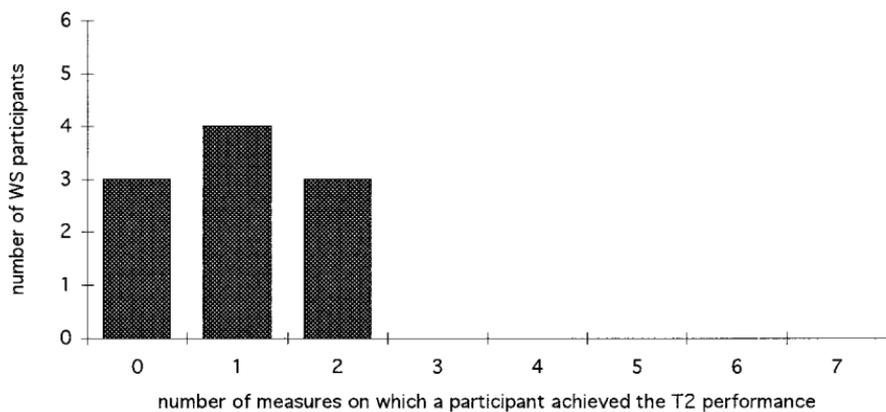
### *Do Any WS Participants Consistently Attain T2?*

To address this question, the number of T2-Dependent tasks on which each participant performed at the T2 level was tabulated. The histograms in Fig. 1 show the number of children in each group who attained T2 levels on 0, 1, 2, 3, 4, 5, 6, or 7 of the T2-Dependent measures. No single WS participant achieved the T2 level performance on more than two of the seven measures. From this we conclude that none of the WS participants had undergone the general shift from T1 to T2.

The WS distribution is markedly different from that of the MC participants, where 70% achieved T2 levels on three or more of the T2-Dependent measures ( $p < .005$ , Fisher exact test, one-tailed). One MC participant did in fact fail to reach the T2 level on a single task. This participant was the youngest control participant, with a mental age of only 8 years 9 months, which we now have reason to believe is still too young to have constructed T2. The YC group results are again almost indistinguishable from the WS group.

Was there any factor within the WS group which predicted even their minimal successes? Three simple regressions were computed between number of tasks on which T2 performance was achieved in the T2-Dependent battery and three possible factors: mental age (MA) as measured by the PPVT-R, PPVT-R standard score, and chronological age (CA), resulting in correlation coefficients of .16, .49, and .70, respectively. Only the final correlation, between performance and chronological age, reached significance at  $p < .03$ .<sup>10</sup> Given that all of our participants had a minimum MA of 9 years it is interesting that it is the oldest of the WS participants who succeed on

<sup>10</sup> The same regressions computed for the matched control group yielded the predictable results of effects of MA ( $r = .68$ ,  $p < .04$ ) and CA ( $r = .68$ ,  $p < .04$ ) but not PPVT-R standard score, which was constant in our sample,  $r = .16$ , n.s.. Again these results confirm the sensitivity of our tasks to the normal developmental factors of MA and CA. The same analyses for the younger control group yielded no effects of MA ( $r = .45$ , n.s.), CA ( $r = .27$ , n.s.), or PPVT-R standard ( $r = .50$ , n.s.), thus confirming that, as intended, we did in fact select an homogeneous group with respect to these factors.



**FIG. 1.** Individual pattern analyses on tasks in the T2-Dependent battery for WS, MC, and YC participants.

the occasional task or two. It seems to be the case that the extra years (as many as 15–20 years in some cases) that the older participants have over the younger participants gives them a slight advantage. Given the local nature of their successes, it is likely that they reflect the fact that the enrichment process itself can eventually yield some knowledge consistent with, if not actually reflecting, T2 understanding.

### *Does Each WS Participant Show the Differential Impairment?*

To address the second and third questions about individual performance the number of advanced performances on the T1/T2-Neutral battery was tabulated for each participant. Technically speaking these tasks did not yield theory-dependent performances, which could be scored as T1 or T2. They did however produce age-dependent performances for the normally developing children. Participants were therefore credited with a level 2 (L2) performance for each task if they achieved within one standard error of the MC average on that task or better. These criteria translated into the following scores for each T1/T2-Neutral task: (1) thirteen or more exemplars in the animal lexicon task; (2) 90% or greater general attributions in the bodily properties measure; and (3) a difference score of 4 or greater on the universal/restricted properties attribution measure. (The novel property projection task was not used here because of the uninformative nature of its ceiling effect.)

A difference score was computed (%L2 (Neutral battery)–%T2 (Dependent battery)) for each participant to determine whether s/he performed comparably on the two batteries. The participants' difference scores were then rank ordered. The distribution of participants within the ranking was not random with respect to population groups. Nine of the 10 individuals with the highest difference scores had WS and 18 of the 19 remaining individuals were children developing normally, Mann–Whitney  $U = 20$ ,  $p < .01$ . The single WS participant achieving a difference score in the normal range did so due to failure on both batteries rather than success on both. Thus we see that the differential impairment on the T2-Dependent tasks is restricted to the participants with WS.

### *Does T1/T2-Neutral Knowledge Really Predict T2-Dependent Knowledge in Normally Developing Children?*

In order to confirm the assumption that in normal development, knowledge acquisition through enrichment and conceptual change proceed hand in hand, the relationship between performance on the two batteries among the normally developing children was examined. A partial correlation, controlling for age, between the percentage of tasks in the T1/T2-Neutral battery on which a participant achieved L2 and the percentage of T2 performances in the T2-Dependent battery revealed a significant relationship,  $r = .49$ ,  $p < .05$ . Simply put, the more knowledge of animals a normally developing child

possessed, the better that child also understood the life cycle and bodily functions.

## GENERAL DISCUSSION

### Summary of Results

Not astonishingly, these data confirm acquisition of knowledge with increasing age. The younger controls revealed less knowledge on both batteries than did the older controls. Furthermore, within normally developing children, the degree of T2 performance on the T2-Dependent battery was positively correlated with the degree of level 2 performance on the T1/T2-Neutral battery, even when age was partialled out of the correlation. These findings in normally developing children are consistent with either of two interpretations: (1) the accretionist interpretation: there is no real distinction between the types of knowledge development tapped by our two batteries; the acquisition of biological knowledge is of a piece, accomplished as the child grapples with the biological world and his/her own culture's beliefs (Atran, 1994; Coley, 1995; Keil, 1992, 1994; Springer, 1992, 1995; Springer & Keil, 1989, 1991; Wellman & Gelman, 1992); or (2) the conceptual change view: there is a real distinction, but because enriched knowledge is an input to the processes that produce conceptual change, the two appear seamless in normal development.

The data from the WS participants in this study provide empirical support for the second possibility. In spite of having the general information typical of normally developing 10-year-olds, the WS participants revealed a conceptual understanding of central biological concepts (*life, death, the determinants of species identity, and people-as-one-animal-among-many*) typical of preschool children. This pattern of results held for every task, and the differential performance on the T2-Dependent battery was found in virtually all of the WS participants and none of the normally developing children.

The WS participants' general knowledge of animals showed several qualities consistent with a mental age of approximately 10. Their knowledge of bodily properties was sensibly organized and mental-age-appropriate with respect to the categories of animals and inanimate and animate objects. It also reflected use of a subtle distinction between universal and restricted bodily properties commonly used by lay adults. They were able to produce as many exemplars of the category animal as their mental age matched controls. Importantly, their categorical knowledge of animals was productive, supporting novel inductive inferences. However, having elaborate knowledge of animals is not synonymous with having a theory of living kinds. The further performance of the WS participants makes this point clear.

When probed for deeper explanatory knowledge about living things, WS participants demonstrated relatively immature conceptual understanding,

more like that of normally developing preschoolers (T1), than that which characterizes the folkbiology of 10-year-olds (T2). WS participants claimed that cars are alive because they move, are useful, or can be seen. They conceived of death in terms of the departure of the dead person rather than in terms of the breakdown of a bodily machine. They held no superordinate category of living thing which includes both animals and plants. They refused to project a novel bodily property of dogs to people, thus revealing that they do not conceive of people as one-animal-among-many with respect to bodily structure. And they claimed that a raccoon, whether surgically altered to look like a skunk or simply dressed up in a skunk costume, was in fact a skunk, revealing no understanding of the role of reproduction and origins in determining biological kind identity. All of these conceptions are characteristic of the intuitive framework for understanding animals (T1) typical of preschoolers.

This pattern of results was predicted on the basis of our analysis of the relations between beginning and later knowledge in the two batteries. We argued that in the course of cognitive development, there are two different types of relations between successive conceptual states (CS1 and CS2). Sometimes CS2 is an enriched version of CS1, as is largely the case for the WS knowledge tapped by the T1/T2-Neutral battery; and sometimes CS2 is a new framework theory, the construction of which has required conceptual change away from CS1, as is largely the case for the knowledge tapped in the T2-Dependent battery. We predicted that WS participants, on the basis of aspects of their mental retardation, would lack the ability to undergo conceptual change in the face of relatively preserved capacities for acquiring new enriched knowledge formulated over their current concepts.

### *Alternative Interpretations*

*General task demands.* For these data to support the distinction between knowledge enrichment and conceptual change it is important that there exist no difference between the batteries, other than the targeted conceptual content, that can explain why WS participants performed differentially on them. Differential task demands was one possibility we attempted to rule out with the design of the two batteries such that they often employed the same experimental methods, varying only in the content of the knowledge probed. For example, attribution of the properties "has babies," "breathes," "hears," and "has a heart" to animals and inanimate objects was part of the T1/T2-Neutral battery, but attribution of the same properties to plants was part of the T2-Dependent battery. Projection of a novel property of people was part of the T1/T2-Neutral battery; projection of a novel property of dogs was part of the T2-Dependent battery. Forced choice or yes/no answers were required in both batteries. Finally, those tasks that did require formulation of explanations (the animism and death interviews) do not exceed the information processing capacities of normally developing 4-year-olds. WS partic-

ipants, like normally developing children, provided consistent, interpretable responses. We conclude, then, that the dissociation obtained in WS between the knowledge tapped by the two batteries cannot be explained in terms of differences in the on-line information processing demands.

*Abstractness.* Of course, it remains possible that some other contrast accounts for the selective difficulties the WS participants had with the T2-Dependent battery relative to the T1/T2-Neutral battery. Perhaps the concepts in the T2-Dependent battery are more abstract than those in the T1/T2-Neutral battery and because retardation impairs abstraction, it naturally impairs the acquisition of those concepts. This alternative however suffers from the vagueness of the notion of abstractness. If the interpretation is that superordinate concepts (e.g., *living thing*) are more abstract than subordinate concepts (e.g., *collie*), there is no evidence that abstract (superordinate) concepts are harder to acquire than concrete (subordinate) ones; that is, both superordinate and subordinate concepts pose problems for young children (Markman, 1989). If, on the other hand, inferential distance from sense data is a measure of abstraction, there is also no evidence that conceptual development moves towards greater abstractness (Simons & Keil, 1995). For example, the theory of mind constructed by preschoolers is abstract—want/belief explanation is far removed from sense data. Importantly, WS individuals far younger than those interviewed for this study have shown quite robust knowledge of “abstract” mental states like beliefs and desires (Tager-Flusberg, Sullivan, & Zaitchik, 1994.) Finally, later theories are often considered more abstract than earlier ones in the sense of offering unified explanations for what were first seen as distinct phenomena. On this final metric of increasing abstraction, the alternative difference between the two batteries collapses onto our own.

*Complexity.* The alternative notion of complexity may also characterize the difference between the batteries. But again, for this account to have force, we must have a metric of complexity. Probably all knowledge acquisition, resulting either from enrichment or from conceptual change, increases complexity (more concepts with more beliefs formulated over them). With respect to folkbiology, we certainly agree that T2 is more complex than T1 (more phenomena represented and explained), and it is also true that only the T2-Dependent battery uniquely probes for the concepts and understanding of T2. So, again, on one interpretation of “more complex” this alternative account collapses onto our own. The conceptual change view has the advantage of greater precision, however, in that it proposes exactly *which* aspects of increasing complexity pose problems for people with Williams syndrome—namely those involving incommensurabilities and conceptual change.

*Age.* Although the current data may suggest that people with Williams syndrome are ultimately incapable of conceptual change, they fall short of fully establishing such a strong claim. Recall that the one predictor of progress on the T2-Dependent battery among WS participants was chronological

age. Perhaps still older WS individuals would have achieved T2. Although possible, we find it unlikely, given the local nature of these older participants' successes. Examination of individual performances both within and across the T2-Dependent tasks showed that the rare adult-like scores achieved by participants with WS reflected performances quite unlike the performances of normal adults. Nonetheless, the degree to which it is impossible for WS individuals to achieve these conceptual changes is still an open question, one that further studies, possibly training studies, may help resolve.

*Relevant biological input and the T1/T2-Neutral battery.* Clearly, the WS participants have not constructed the biological knowledge tapped by the T2-Dependent battery. However, our interpretation that they are unable to do so depends upon the assumption that they have the factual knowledge that is the input to conceptual change for normally developing people. This assumption is supported by the results from the T1/T2-Neutral battery, designed to diagnose knowledge of animals that might be formulated in the concepts of either T1 or T2. On this battery, WS participants performed as well or better than their matched controls.

However, the T1/T2-Neutral battery explored a very limited subset of biological knowledge, and may have failed to tap the actual knowledge that leads to conceptual change. For instance, it did not probe for knowledge of plants formulatable in terms of T1, such as the simple knowledge that plants grow, or that plants die, or that plants need water (see Backscheider et al., 1993; Inagaki & Hatano, 1996). If just such knowledge is crucial to the coalescence of animals and plants into the single category of living things, we are not in a position to tell if the WS participants failed to construct T2 because they lacked the relevant information rather than the ability to accomplish conceptual change.

Similar relationships have been shown between individual facts and conceptual change in children's understanding of other biological phenomena. In training studies, Springer (1995) reported a relationship between children's knowledge that babies grow in their mothers' tummies and their understanding of kinship. Solomon and Johnson (submitted) showed a similar effect of teaching children information about "genes" and their later ability to differentiate biological and social aspects of family. Slaughter, Jaakkola, and Carey (in press) recently showed that knowledge of specific functions of bodily organs (again, acquirable in terms of the concepts of T1) was correlated, within 4- to 6-year-olds, with the beginnings of the T2 differentiated concept of death and the beginnings of a vitalist biology. It is possible that the WS participants in the current study lacked some or all of such relevant T1/T2-Neutral knowledge.

This is a serious possibility and we plan to explore it in further work on WS. Though we have not yet shown it, we predict that people with WS do indeed know that babies grow in their mummies' tummies, that lungs are for breathing, that plants grow, and so on. However, we also predict that

the close interconnections among different aspects of biological knowledge demonstrated in the above training studies and studies of coherence will not be found among WS adolescents and adults. Here we merely note that the failure to include probes of T1/T2 neutral knowledge of plants is not decisive, for the conceptual differences between T2 and T1 do not center solely on the inclusion of plants in the domain of intuitive biology. The vitalist understanding of bodily function (Inagaki & Hatano, 1993), the understanding of people as one animal among many with respect to bodily structure (Carey, 1985), the understanding of animal species kind in terms of origins (Keil, 1989; Johnson & Solomon, 1997), and the understanding of death as the breakdown of the bodily machine (Carey, 1985; Slaughter et al., in press) all concern developing conceptions of animals alone. Although the T1/T2-Neutral battery provided only the merest snapshot of relevant knowledge of animals that might be input to the formation of T2, we believe it is representative.

*Did we underestimate T2 knowledge in WS?* The argument in this paper does not depend on the general characterization of T1 as a nonbiological theory of animals; we argue merely that it differs from T2 in terms of the differentiations, coalescences, and reanalyses of concepts we have specified that implicate conceptual change. Not all researchers working on children's intuitive biology, however, agree on which tasks in the literature in fact tap T1 versus T2 concepts. For instance, not all agree that children have not differentiated the social from the biological family until age 7 or older (see the debate between Springer (1992, 1995) and Hirschfeld (1995) on the one hand and Solomon et al. (1996) and Johnson and Solomon (1997) on the other). If further work on WS corroborates the findings presented here, studies of individuals with WS could bear on this issue. We would predict that the findings reflecting biologically relevant knowledge in preschool children (e.g., Backscheider et al., 1993; Hirschfeld, 1995; Rosengren et al., 1991; Springer, 1992, 1995) would be manifest in WS adolescents and adults, in the face of continued failure at tasks that tap conceptual change. Such a pattern of results would support the view that these phenomena are not markers of T2, although they may well be important aspects of the knowledge that is the input to the formation of T2.

*General theorizing abilities.* A final alternative interpretation of the present results has often been suggested to us. Rather than being differentially impaired in cases of theory development requiring conceptual change, perhaps people with Williams syndrome are impaired in all theory acquisition. Many domains of conceptual knowledge, not just biology, are organized as intuitive theories. It could be that even if T2 related to T1 only via enrichment relations, a lack of "scientific" curiosity or causal appreciation might still prevent T2 construction.

Data on this issue in people with WS are scarce. Based on the work of Tager-Flusberg et al. (1994), we know that very young children with WS

are able to reason about and explain human behavior in terms of belief/desire causality. However, under some construals, theory of mind is not considered a true intuitive theory in the same sense in which folkbiology is (see Carey & Spelke, 1994; Fodor, 1992; Leslie, 1994, 1995). If it is not, the successes of WS participants in that domain become irrelevant to the current debate.

It is certainly possible that an impaired theory-building ability is the source of the current findings. It is quite likely that causal generalizations are particularly important inputs both to theory-building in general and conceptual change in particular. If individuals with WS are generally impaired in forming causal generalizations, and in seeking explanatory coherence, they should certainly find conceptual change impossible, and they should also be relatively impaired in other aspects of general theory development.

This is an empirical question that should be addressed in further study. However, to do so, one must separate aspects of knowledge acquisition that are "theory development" from aspects of knowledge acquisition that are not. On the "theory-theory" point of view, in which concepts are all ultimately grounded in intuitive theories (cf. the argument of Keil (1989) that there is no "original sim"; Murphy & Medin, 1985; Wellman & Gelman, 1992), all conceptual knowledge acquisition is theory development. A particularly promising distinction to make, we believe, is between the acquisition of causal/explanatory knowledge and other types of knowledge. It may well be that WS participants are impaired in building causal/explanatory knowledge in general, and thus in theory development that does not require conceptual change as well as theory development that does.

### *What Have We Learned about Williams Syndrome?*

We chose Williams syndrome individuals as a population to compare with normally developing children on the basis of (1) prior considerations of the analytic and metacognitive skills supporting conceptual change (Carey & Spelke, 1994; Nersessian, 1992; Smith et al., 1992; and Wiser, 1988); (2) evidence that retardation in general, and WS in particular, affects those analytic and metaconceptual abilities (Bellugi et al., 1993; Bertrand & Mervis, 1994; Campione et al., 1982); and (3) evidence that people with WS are relatively spared, compared to overall IQ, on tests of vocabulary (Bellugi et al., 1993) and general information such as the Faces and Places subtest and the Riddles subtest of the Kaufman battery (Levine, 1993). However, this study did not directly address any of the underlying mechanisms responsible for these findings. As such, the underlying reasons that WS participants were markedly more impaired on the T2-Dependent than the T1/T2-Neutral battery are as yet unknown. Because of the above mentioned literature we assume that different learning mechanisms underlie conceptual change and enrichment. But of course, it is also possible that a single learning mechanism underlies both types of knowledge acquisition, with conceptual change requiring a quantitatively more powerful version than is available to individu-

als with retardation. Perhaps similar studies with gifted children (predicting the opposite pattern of results) as well as training studies with Williams syndrome participants might shed light on this question. In either case, these data demonstrate the consequences of such an impairment for theory development within everyday conceptual structures like folkbiology.

Although our WS participants performed at the level of preschool children on the tasks that diagnosed T2 knowledge, their overall responses would never actually be confused with those of preschool children. They displayed a great deal of general knowledge concerning people and animals that preschoolers do not yet have (witness their performance on the T1/T2-Neutral battery). Furthermore, they express their knowledge in language quite different from that of a preschooler. Take for instance the response of SS, a 32-year-old man with WS and a PPVT-R mental age of 15-11, when asked what happens to people and their bodies when they die (punctuation added), "They go to heaven, depending upon their quest in life, how they made their decisions and how they felt about other people and how much they have given . . . Their body remains for it is a living, and it is the shell that stays with us. But the soul goes to heaven. To me when a person dies it's not the same as when they're alive. It's a different part of life. When you see them in that casket, they're there but their spirit is always with you. The body is either buried or cremated depending on the faith or religion or depending on the family or loved ones that are nearest them make arrangements and it's by the request of the loved one that had died made the decisions before dying." SS has elaborated knowledge of religious practices, and of the relevance of intentional decisions for the fate of a body after death that preschool children do not have. Nonetheless, neither here nor in other parts of the interview does he refer to bodily mechanisms that cause death or result from it; he still conceptualizes death as a different state of living "it's a different part of life."

SS's description of death illustrates a characteristic often ascribed to people with WS: "cocktail party syndrome." Although this ascription captures salient social/emotional features of WS (they are often extremely outgoing and use many pragmatic linguistic devices to keep conversational partners engaged, see Reilly, Klima, & Bellugi, 1991), "cocktail party syndrome" has a cognitive connotation as well, referring to a tendency to talk at length with little depth. The current study offers an interpretation of this cognitive connotation. The concepts of individuals with WS, at least in the domain of folkbiology, are superficial in a sense we can make precise—they are still embedded in T1, the intuitive theory of animals that is held by preschool children. We predict that this will be equally true of concepts in other domains that undergo conceptual change in normal development, in both the elementary school years as well as later in development.

Bellugi et al.(1993) report other data consistent with this prediction. In

that work WS adolescents failed Piagetian tasks including conservation of number, weight, substance, and quantity. Although performance on conservation tasks is often taken as a reflection of reasoning ability (e.g., understanding that a greater extent in one dimension can be compensated for by a lesser extent in another), conservation of weight and matter also reflect important conceptual changes within the child's intuitive theory of matter (Carey, 1991; Piaget & Inhelder, 1941; Smith et al., 1985). Again, the concepts of the WS adolescents appear to remain embedded in the T1 typical of preschool children.

The characterization of this sense in which the concepts of people with WS may be superficial does not yet bear on what makes WS unique. We predict that other retarded populations, if impaired in metacognitive and analytic abilities, will also fail to restructure theories requiring conceptual change. Whether or not the dissociation between enrichment-based and conceptual change-based knowledge acquisition documented here is unique to WS remains to be seen. Based on the reasoning outlined above, we would predict to find it only in other retarded populations displaying cocktail party syndrome (e.g., some people with retardation due to spina bifida, see Cromer, 1991). These are empirical questions, of course, which have not yet been addressed in the literature on mental retardation.

#### *Why Does PPVT-R Predict Performance on the T1/T2-Neutral Animal Knowledge Battery?*

Of course, WS participants are not *unimpaired* on the T1/T2-Neutral battery. With one exception (the differentiation between universal and restricted bodily properties), they never perform at the level typical of their chronological age. Rather, the T1/T2-Neutral battery patterns are consistent with mental age as measured by the PPVT-R, on which WS participants are also impaired relative to their chronological age. This result was predicted, and vindicates the choice of the PPVT-R as a matching instrument. However, it was by no means a foregone conclusion. The T1/T2-Neutral battery—production of animals' names, inductive projection of a newly taught bodily property of people, and attribution of bodily properties to animals and inanimate objects—contains tasks very different from a receptive vocabulary task. Nonetheless, we expected that the PPVT-R would predict performance on this battery, for vocabulary acquisition of words such as those on the PPVT-R, like vocabulary acquisition of names of animal kinds, does not generally require conceptual change. Words on the PPVT-R name examples of kinds of which some exemplars are known to preschool children. What predicts difficulty on the T1/T2-Neutral battery, like the PPVT-R, is rarity of the knowledge, likelihood of having encountered it. We assume that retarded people are less likely than normally developing people to encounter any given piece of information and are generally slower to form associations.

*Further Implications for Normal Development and Folkbiology*

The differential impairment on the two batteries in the participants with WS justifies our inclusion of the attribution and projection from people tasks in the T1/T2-Neutral battery. Clearly, it is possible to present adult-like patterns of attribution in the absence of T2, for this is what the WS participants did. Although categorical attribution of universal bodily properties to all animals and the differentiated attribution patterns of universal and restricted bodily properties might well reflect T2 in normally developing individuals (Carey, 1985), such patterns of attribution need not do so. Thus, the (1985) interpretation of Carey (1985) that the developmental changes in these attribution patterns are diagnostic of the shift from T1 to T2 was not justified.

The data from the normally developing control participants also have implications in their own right for researchers studying children's intuitive biology. This study is the first in which a large collection of intuitive biological concepts has been examined both within individuals and in a psychometric context. To our surprise, the older controls had not completed the construction of T2, despite an average chronological age of nearly 10 and a mental age of nearly 11. Only two of the older participants (MA = 11;0 and 14;7) displayed T2 understanding across the board. This result is consistent with the claim of Carey (1985) that the construction of the framework theory of living kinds is a long and slow process and is not completed until the end of the first decade of life. It is inconsistent with much of the recent work pushing the age of acquisition of the living kind ontology younger, even into the preschool years. We suggest that much of the recent work, ours included, has ignored psychometric issues in participant selection. If as we suggest, mental retardation results in the delay or prevention of conceptual change, then the reverse is also conceivable. That is, high IQ children, who seem to be overrepresented in university participant pools, may be relatively advanced in their progress compared to psychometrically average children. The only other study from this battery of tasks that draws from children in the normal distribution of IQ is the animism study of Laurendeau and Pinard (1962). The data reported here replicate their findings almost exactly.

## CONCLUSIONS

Adolescents and adults with WS were found to be differentially more impaired in the acquisition of concepts posited to require conceptual change than those which do not. When probed for concepts of life, death, living thing, species kind, and people-as-one-animal-among-many, highly articulate, seemingly well-informed individuals with WS were found to possess the conceptual understanding of normally developing 6-year-olds. We have argued that this differential impairment in the WS participants favors the conceptual change position over the accretionist position in accounting for

conceptual development within intuitive biology, while simultaneously demonstrating the dramatic consequences such an impairment can have on the developmental outcome of everyday conceptual structures.

## APPENDIX

### Criteria for Analyses of Theoretical Levels

#### *Animism*

*T2 performance.* For the measure based on the percent of correct denials, T2 performance required no attributions of life to inanimate objects.

On the level of performance analysis, T2 performance was Level 6.

*T1 performance.* For the percent denials measure, judgments that 15% or more of the inanimate objects were alive were considered consistent with T1 performance.

On the level of performance analysis, Level 0, 1, or 2 was considered T1.

#### *Death*

*T2 performance.* Level 2 was the T2 performance on this task. Level 2 reflects analysis of death in terms of the breakdown of the bodily machine and the cessation of bodily functions like breathing and circulation.

*T1 performance.* The T1 level was Level -1, reflecting interpretations based on departure or altered state.

#### *Projection of a Novel Property of Dogs*

There were only two patterns on this task: projection from dog to people (T2 performance) or no projection (T1 performance). Thus, on this task alone, all participants necessarily fell into the T2 performance or the T1 level pattern.

#### *Attribution to a Tree*

*T2 performance.* Based on adult performance (unpublished data), a difference score of 33 or more was considered the T2 performance.

*T1 performance.* A difference score of 10 or less was considered the T1 performance.

#### *Species Transformation*

*T2 performance.* As noted above, T2 is not required for success on the costume task, but a participant with T2 certainly would succeed on it. T2 performance on each of these tasks, then, is to deny that the transformation (costume or surgery) changed the species of the animal, and not to waver in this judgment.

*T1 performance.* The T1 level on these tasks is to affirm, unwaveringly, that the animal species is changed by the transformation.

## REFERENCES

- Atran, S. 1994. Core domains versus scientific theories: Evidence from systematics and Itz'at Maya folkbiology. In L.A. Hirschfeld & S.A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture*. New York: Cambridge Univ. Press.
- Au, T.K. 1994. Developing an intuitive understanding of substance kinds. *Cognitive Psychology*, **27**(1), 71–111.
- Backscheider, A.G., Schatz, M., & Gelman, S.A. 1993. Preschoolers' ability to distinguish living kinds as a function of regrowth. *Child Development*, **64**, 1242–1257.
- Bellugi, U., Bihle, A., Neville, H., Jernigan, T., & Doherty, S. 1993. Language, cognition, and brain organization in a neurodevelopmental disorder. In M. Gunnar & C. Nelson (Eds.), *Developmental behavioral neuroscience (Minnesota Symposium)*. Hillsdale, NJ: Erlbaum.
- Bertrand, J., & Mervis, C. 1994. *Metamemory in people with Williams syndrome*. Poster presented at the 1994 National Williams Syndrome Professional Conference, La Jolla, CA.
- Bihle, A., Bellugi, U., Delis, D., & Marks, S. 1989. Seeing either the forest or the trees: Dissociation in visuospatial processing. *Brain and Cognition*, **11**, 37–49.
- Bromberg, H., Ullman, M., Marcus, G., Kelly, K., & Levine, K. 1994. *A dissociation of memory and grammar and grammar: Evidence from Williams syndrome*. Paper presented at the annual meeting of the Linguistic Society of America, Boston, MA.
- Campione, J., Brown, A., & Ferrara, R. 1982. Mental retardation and intelligence. In R.J.-Sternberg (Ed.), *Handbook of human intelligence*. Cambridge: Cambridge Univ. Press.
- Carey, S. 1985. *Conceptual change in childhood*. Cambridge, MA: MIT Press.
- Carey, S. 1988. Conceptual differences between children and adults. *Mind and Language*, **3**, 167–181.
- Carey, S. 1991. Knowledge acquisition: Enrichment or conceptual change? In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition*. Hillsdale, NJ: Erlbaum.
- Carey, S. 1995. On the origins of causal understanding. In D. Sperber, J. Premack, & A. Premack (Eds.), *Causal cognition: A multidisciplinary debate*. Oxford: Clarendon.
- Carey, S. & Gelman, R. 1991. *The epigenesis of mind: Essays on biology and cognition*. Hillsdale, NJ: Erlbaum.
- Carey, S., & Spelke, E. 1994. Domain specific knowledge and conceptual change. In L.A. Hirschfeld & S.A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture*. New York: Cambridge Univ. Press.
- Chi, M.T.H. 1992. Conceptual change within and across ontological categories: Examples from learning and discovery in science. In R. Giere (Ed.), *Cognitive models of science: Minnesota studies in the philosophy of science*. Minneapolis, MN: Univ. of Minnesota Press.
- Coley, J. D. 1995. Emerging differentiation of folkbiology and folkpsychology: Attributions of biological and psychological properties to living things. *Child Development*, **66**, 1856–1874.
- Crider, C. 1981. Children's conceptions of the body interior. In R. Bibace & M.E. Walsh (Eds.), *Children's conceptions of health, illness, and bodily functions*. San Francisco: Jossey-Bass.

- Cromer, R. 1991. *Language and thought in normal and handicapped children*. Oxford: Blackwell Sci.
- DeVries, R. 1969. Constancy of genetic identity in the years three to six. *Monographs of the Society for Research in Child Development*, **34**.
- Dunn, L.M., & Dunn, L.M. 1981. *Peabody picture vocabulary test—Revised*. Circle Pines, MN: American Guidance Service.
- Feyerabend. 1962. Explanation, reduction, and empiricism. *Minnesota Studies in the Philosophy of Science*, **3**, 28–97.
- Fodor, J. 1992. A theory of the child's theory of mind. *Cognition*, **44**, 283–296.
- Frangiskakis, J.M., Ewart, A.K., Morris, C.A., Mervis, C.B., Bertrand, J., Robinson, B.F., Klein, B.P., Ensing, G.J., Everett, L.A., Green, E.D., Proschel, C., Gutowski, N.J., Noble, M., Atkinson, D.L., Odelberg, S.J., & Keating, M.T. 1996. LIM-kinase1 hemizyosity implicated in impaired visuospatial constructive cognition. *Cell*, **86**, 59–69.
- Gelman, S.A., & Coley, J.D. 1990. The importance of knowing a dodo is a bird: Categories and inferences in 2-year-old children. *Developmental Psychology*, **26**, 796–804.
- Gelman, S.A., Coley, J.D. & Gottfried, G.M. 1994. Essentialist beliefs in children: The acquisition of concepts and theories. In L.A. Hirschfeld & S.A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture*. New York: Cambridge Univ. Press.
- Goldman, R., & Goldman, J. 1982. How children perceive the origin of babies and the roles of mothers and fathers in procreation: A cross-national study. *Child Development*, **53**, 491–504.
- Gopnik, A., & Wellman, H.M. 1994. The theory theory. In L.A. Hirschfeld & S.A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture*. New York: Cambridge Univ. Press.
- Gruber, H. 1974. *Darwin on man: A psychological study of scientific creativity*. London: Wildwood House.
- Hacking, I. 1993. Working in a new world: The taxonomic solution. In P. Horwich (Ed.), *World changes: Thomas Kuhn and the nature of science*. Cambridge, MA: MIT Press.
- Hatano, G., & Inagaki, K. 1994. Young children's naive theory of biology. *Cognition*, **50**, 171–188.
- Hirschfeld, L. 1995. Do children have a theory of race? *Cognition*, **54**, 209–252.
- Hirschfeld, L., & Gelman, S. 1994. *Mapping the mind: Domain specificity in cognition and culture*. New York: Cambridge Univ. Press.
- Inagaki, K., & Hatano, G. 1987. Young children's spontaneous personification as analogy. *Child Development*, **58**, 1013–1020.
- Inagaki, K., & Hatano, G. 1993. Children's understanding of the mind-body distinction. *Child Development*, **64**(5), 1534–1549.
- Inagaki, K., & Hatano, G. 1996. Young children's recognition of commonalities between animals and plants. *Child Development*, **67**, 2823–2840.
- Inagaki, K., & Sugiyama, K. 1988. Attributing human characteristics: Developmental changes in over- and underattribution. *Cognitive Development*, **3**, 55–70.
- Jernigan, T., & Bellugi, U. 1990. Anomalous brain morphology on magnetic resonance images in Williams syndrome and Down syndrome, *Archives Neurology*, **47**.
- Johnson, S.C., & Solomon, G.E.A. 1997. Why dogs have puppies and cats have kittens: Young children's understanding of biological origins. *Child Development*, **68**(3), 404–419.
- Jones, K., & Smith, D. 1975. The Williams elfin facies syndrome: A new perspective. *Journal of Pediatrics*, **86**, 718–723.

- Kaufman & Kaufman. 1983. *Kaufman assessment battery for children (K-ABC)*. Circle Pines, MN: American Guidance Service.
- Keil, F.C. 1989. *Concepts, kinds, and cognitive development*. Cambridge, MA: Bradford Books.
- Keil, F.C. 1992. The origins of an autonomous biology. In M.R. Gunnar & M. Maratsos (Eds.), *Modularity and constraints in language and cognition: The Minnesota symposia on child psychology* (Vol. 25). Hillsdale, NJ: Erlbaum.
- Keil, F.C. 1994. The birth and nurturance of concepts by domains: The origins of concepts of living things. In L.A. Hirschfeld & S.A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture*. New York: Cambridge Univ. Press.
- Kitcher, P. 1988. The child as parent of the scientist. *Mind and Language*, **3**, 217–227.
- Koocher, G.P. 1974. Talking with children about death. *American Journal of Orthopsychiatry*, **44**, 404–410.
- Kuhn, T.S. 1962. *The structure of scientific revolution*. Chicago: Univ. of Chicago Press.
- Kuhn, T.S. 1977. *The essential tension: Selected studies in scientific tradition and change*. Chicago: Univ. of Chicago Press.
- Kuhn, T.S. 1982. Commensurability, comparability, communicability. *PSA*, **2**.
- Laurendeau, M., & Pinard, A. 1962. *Causal thinking in the child: A genetic and experimental approach*. New York: International Universities Press.
- Leslie, A.M. 1994. ToMM, ToBy, and Agency: Core architecture and domain specificity. In L.A. Hirschfeld & S.A. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture*. New York: Cambridge Univ. Press.
- Leslie, A.M. 1995. A theory of Agency. In D. Sperber, D. Premack, & A.J. Premack (Eds.), *Causal cognition: A multidisciplinary debate*. Oxford: Clarendon.
- Levine, K. 1993. *Cognitive profiles in Williams syndrome*. Unpublished manuscript, Boston's Childrens' Hospital.
- Mandler, J., Bauer, P., & McDonough, L. 1991. Separating the sheep from the goats: Differentiating global categories. *Cognitive Psychology*, **23**(2), 263–298.
- Markman, E. 1989. *Categorization and naming in children*. Cambridge, MA: MIT Press.
- Morris, C. 1994. *The search for the Williams syndrome gene*. Paper presented at the 1994 National Williams Syndrome Professional Conference, La Jolla, CA.
- Murphy, G.L., & Medin, D.L. 1985. The role of theories in conceptual coherence. *Psychological Review*, **92**(3), 289–316.
- Nagy, M.H. 1948. The child's theories concerning death. *Journal of Genetic Psychology*, **73**, 3–27.
- Nersessian, N.J. 1992. How do scientists think? Capturing the dynamics of conceptual change in science. In R.N. Giere (Ed.), *Cognitive models of science: Minnesota studies in the philosophy of science* (Vol. 15, p. 344). Minneapolis: Univ. of Minnesota Press.
- Osherson, D., Smith, E., Wilkie, O., Lopez, A., & Shafir, E. 1990. Category-based induction. *Psychological Review*, **97**(2), 1985–2000.
- Perner, J. 1991. *Understanding the representational mind*. Cambridge, MA: MIT Press.
- Piaget, J. 1929. *The child's conception of the world*. London: Routledge & Kegan Paul.
- Piaget, J., & Inhelder, B. 1941. *Le développement des quantités chez l'enfant*. Neufchatel: Delchaux et Niestle.
- Reilly, J., Klima, E., & Bellugi, U. 1991. Once more with feeling: Affect and language in atypical populations. *Development and Psychopathology*, **2**, 367–391.
- Rosengren, K.S., Gelman, S.A., Kalish, C.W., & McCormick, M. 1991. As time goes by: Children's early understanding of growth in animals. *Child Development*, **62**, 1302–1320.

- Scott, P., Mervis, C., Bertrand, J., Klein, S., Armstrong, S., & Ford, M. 1994. *Lexical organization and categorization in people with Williams syndrome*. Poster presented at the 1994 National Williams Syndrome Conference, La Jolla, CA.
- Simons, D.J., & Keil, F.C. 1995. An abstract to concrete shift in the development of biological thought: The insides story. *Cognition*, **56**(2), 129–163.
- Slaughter, V., Jaakkola, R., & Carey, S. in press. Constructing a coherent theory: Children's biological understanding of life and death. In M. Siegal & C. Peterson (Eds.), *Children's understanding of biology and health*. Cambridge: Cambridge Univ. Press.
- Sloman, S.A. 1993. Feature-based induction. *Cognitive Psychology*, **25**, 231–280.
- Smith, C., Carey, S., & Wiser, M. 1985. On differentiation: A case study of the development of the concepts of size, weight, and density. *Cognition*, **21**, 177–237.
- Smith, C., Snir, J., & Grosslight, L. 1992. Using conceptual models to facilitate conceptual change: The case of weight-density differentiation. *Cognition and Instruction*, **9**(3), 221–283.
- Solomon, G.E.A., & Johnson, S.C. Submitted. Conceptual change in the classroom: Teaching young children to understand biological inheritance. *British Journal of Developmental Psychology*.
- Solomon, G.E.A., Johnson, S.C., Zaitchik, D., & Carey, S. 1996. Like father, like son: Young children's understanding of how and why offspring resemble their parents. *Child Development*, **67**, 151–171.
- Spelke, E. 1991. Physical knowledge in infancy: Reflections on Piaget's theory. In S. Carey & R. Gelman (Eds.), *Epigenesis of mind: Essays on biology and cognition*. Hillsdale, NJ: Erlbaum.
- Spelke, E., Phillips, A., & Woodward, A. 1995. Infants' knowledge of object motion and human action. In D. Sperber, J. Premack, & A. Premack (Eds.), *Causal cognition: A multidisciplinary debate*. Oxford: Clarendon.
- Springer, K. 1992. Children's beliefs about the biological implications of kinship. *Child Development*, **63**, 151–171.
- Springer, K. 1995. How a naive theory is acquired through inference. *Child Development*, **66**, 547–558.
- Springer, K., & Keil, F. 1989. On the development of biologically specific beliefs: The case of inheritance. *Child Development*, **60**, 637–648.
- Springer, K., & Keil, F. 1991. Early differentiation of causal mechanisms appropriate to biological and nonbiological kinds. *Child Development*, **62**, 767–781.
- Tager-Flusberg, H., Sullivan, K., & Zaitchik, D. 1994. *Social cognitive abilities in young children with Williams syndrome*. Paper presented at the 6th International Conference of the Williams Syndrome Association, San Diego.
- Thagard, P. 1993. *Computational philosophy of science*. Cambridge, MA: MIT Press.
- Udwin, O., & Yule, W. 1990. Expressive language of children with Williams syndrome. *American Journal of Medical Genetics Supplement*, **6**, 108–114.
- Udwin, O., & Yule, W. 1991. A cognitive and behavioral phenotype in Williams syndrome. *Journal of Clinical and Experimental Neuropsychology*, **13**(2), 232–244.
- Vosniadu, S., & Brewer, W.F. 1992. Mental models of the earth: A study of conceptual change in childhood. *Cognitive Psychology*, **24**, 535–585.
- Wellman, H.M., & Gelman, S.A. 1992. Cognitive development: Foundational theories of core domains. *Annual Review of Psychology*, **43**, 337–375.
- Williams, J.C.P., Barrett-Boyes, B.G., & Lowe, J.B. 1962. Supravalvular aortic stenosis. *Circulation*, **24**, 1311–1318.

- Wiser, M. 1988. Can models foster conceptual change? The case of heat and temperature (Educational Technology Center Technical Report). Cambridge, MA: Harvard University.
- Wiser, M., & Carey, S. 1983. When heat and temperature were one. In D. Gentner & A. Stevens (Eds.), *Mental models* (pp. 267–297). Hillsdale, NH: Erlbaum.
- Accepted August 27, 1998