Research Article

Developmental “Roots” in Mature Biological Knowledge

Robert F. Goldberg and Sharon L. Thompson-Schill

Center for Cognitive Neuroscience, University of Pennsylvania

ABSTRACT—Young children tend to claim that moving artifacts and nonliving natural kinds are alive, but neglect to ascribe life to plants. This research tested whether adults exhibit similar confusions when verifying life status in a speeded classification task. Experiment 1 showed that undergraduates encounter greater difficulty (reduced accuracy and increased response times) in determining life status for plants, relative to animals, and for natural and moving nonliving things, relative to artifacts and nonmoving things. Experiment 2 replicated these effects in university biology professors. The professors showed a significantly reduced effect size for living things, compared with the students, but still showed greater difficulty for plants than animals, even as no differences from the students were apparent in their responses to nonliving things. These results suggest that mature biological knowledge relies on a developmental foundation that is not radically overwritten or erased with the profound conceptual changes that accompany mastery of the domain.

I wandered lonely as a cloud
That floats on high o’er vales and hills,
When all at once I saw a crowd,
A host of golden daffodils;
Beside the lake, beneath the trees,
Fluttering and dancing in the breeze.

—Wordsworth (1869, p. 144)

Keil (1989; Keil, Smith, Simons, & Levin, 1998) has drawn parallels between developmental trends and research in mature category structures. But the question remains whether cognitive development leaves its imprint within the adult conceptual system. The strong claim is that children and adults hold qualitatively different, perhaps incommensurate, domain theories (Carey, 1985, 1988, 1999). However, if concepts are acquired and maintained by constant processes throughout development, then certain types of instances should remain ambiguous regardless of how well formed a category, or an associated theory, might be. Outcomes of developmental effects are not usually examined into adulthood, but the end state of an adult model is necessary for a more complete theory of conceptual development (Coley, 2000).

The biological domain is thought to be foundational for the developing conceptual system of young children (Wellman & Gelman, 1992). In a domain as diverse as biology (e.g., consider the range of features among cows, daisies, lichen, and wasps), children must learn the causal mechanisms that underlie life processes and conceptually bind all living things. If children initially use certain features to categorize things and assign linguistic labels (Woodward & Markman, 1998), perhaps via attentional weighting (Goldstone, 1998), then semantic categories are likely to form around superficial similarities. One should therefore expect young children to make gross overgeneralizations when drawing their early categories. Such a tendency has been noted in young children’s early biological classifications.

Piaget (1929/1960) recognized that young children ascribe life to inanimate things but neglect to see plants as living things. Because plants don’t seem to “do” much in the child’s view, they are generally denied “life.” Yet in the young child’s eyes, the sun is alive because it moves or shines. The following questioning of a young child (Zimm, age 7 years 9 months) is typical of Piaget’s early work in this area:

Is a cat alive?  Yes.
A snail? Yes.
A table? No.
Why not? It can’t move.

Address correspondence to Robert F. Goldberg, Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, 77 Massachusetts Ave., 46-5115, Cambridge, MA 02139, e-mail: grob@mit.edu.
Is a bicycle alive?
Yes.
Why?
It can go.
Is a cloud alive?
Yes.
Why?
It sometimes moves.
Is water alive?
Yes, it moves.
Is it alive when it doesn’t move?
Yes.
Is a bicycle alive when it isn’t moving?
Yes, it’s alive, even when it doesn’t move.
Is a lamp alive?
Yes, it shines.
Is the moon alive?
Yes, sometimes it hides behind the mountain. (Piaget, 1929/1960, p. 199)

Piaget used the term animism to describe this aspect of young children’s thinking. As this example readily shows, young children tend to state that moving things and nonliving natural kinds are alive.

Researchers have suggested that the animate/inanimate distinction (for a recent review, see Gelman & Opfer, 2002) is the starting point for the domain of biological kinds (Wellman & Gelman, 1992). The key for developing mature biological knowledge involves determining the causes of an object’s movement and recognizing different kinds of movement, such as those exhibited by nonliving natural kinds and animals. This knowledge assists in the development of life concepts that unite plants with animals through inferred, underlying biological mechanisms (Carey, 1985; Keil, 1989). Yet even though young children recognize that plants and animals share important characteristics, such as the capacities to grow, reproduce, and die (Hatano & Inagaki, 1994; Nguyen & Gelman, 2002; Springer & Keil, 1991), they encounter difficulty forming the common category of living things. Because plants are seen as stationary objects that are acted upon, they are denied life (Richards & Siegler, 1986). If young children learn to reinterpret objects’ behavior in terms of similarities in underlying causes (e.g., animals moving toward water, plants growing toward water; Opfer, 2002), they readily reshape their life concepts to include plants (Opfer & Siegler, 2004).

Given the vast differences between the child and the adult frameworks for biology, Carey (1985, 1988, 1999) proposed that the two are incommensurate, noting that the “child’s conceptual system includes undifferentiated concepts that no longer play any role in the adult conceptual system” (Carey, 1988, p. 180). The difficulty with this view is that the exact age at which this shift to the adult system takes place cannot be firmly established. Some data suggest an early transition (Carey, 1985), whereas other research suggests that the shift occurs in much older children and may be impaired by a neurodevelopmental disorder (Johnson & Carey, 1998). In the strongest case, if adults and children differ qualitatively in both their differentiated concepts (e.g., alive is differentiated from movement and natural) and their coalesced concepts (e.g., animal is coalesced with plant; Carey, 1988, 1999), there should be no connection between the adult category of living things and children’s misconceptions about biology.

By examining the verification of life status for a diverse set of word stimuli, we tested whether the adult framework for biological knowledge relies, to some extent, on conceptual structures laid down during development. We used a classic category-verification task (e.g., Rips, Shoben, & Smith, 1973) to test adults under speeded conditions. With this paradigm, it is possible to compare conceptual structures from the developmental trajectory with mature conceptual structures by using the difficulties of young children to predict the response characteristics of adults. We hypothesized that an underlying developmental foundation biases the mature representation of life status, so adults should encounter difficulty in verifying the life status of instances similar to the ones that confuse young children, including plants, moving things, and nonliving natural kinds. We tested movement and naturalness as separate features because pilot studies, and the early Piagetian work, suggested that they are dissociable dimensions among nonliving things.

**EXPERIMENT 1**

The first experiment was designed to test whether adults encounter difficulty in verifying the life status of items that typically cause confusions for young children. We expected adults to demonstrate greater difficulty in classifying plants as living things than in classifying animals as living things. We also expected that performance in classifying nonliving things would be degraded if items had the features of movement, naturalness, or both.

**Method**

**Participants**
Thirty participants (15 females, 15 males; ages 18–22 years old, \( M = 18.6 \)) were recruited from the undergraduate subject pool of the Department of Psychology at the University of Pittsburgh. Data from 1 participant were excluded from the analyses reported here because of an average accuracy level below 75% across all trials.

---

1 We use italics to indicate when we are referring to concepts, rather than individual tokens.
Materials
Word items were chosen on the basis of young children’s difficulties in learning mature biological knowledge. The set of living things comprised 30 animals and 30 plants. The set of nonliving things included 20 items from each of the following categories: (a) nonmoving artifacts (e.g., broom, towel), (b) nonmoving natural kinds (e.g., stone, crater), (c) moving artifacts (e.g., truck, ferry), and (d) moving natural kinds (e.g., comet, river). Ratings of movement and naturalness obtained from a separate group of participants were used in selecting the non-living items.

Within the set of living things, the plant and animal items were matched on number of letters, number of syllables, and average word frequency. To measure frequency, we used log-transformed Hyperspace Analogue to Language (HAL) frequency norms (Lund & Burgess, 1996), which we downloaded from the English Lexicon Project Web site (Balota et al., 2007). These norms were calculated from 131 million words gathered across 3,000 Usenet newsgroups during February 1995 and are preferred over other frequency norms given the much larger number of items sampled (Balota, Cortese, Serpent-Marshall, Spieler, & Yap, 2004). The four subsets of nonliving things were matched on the same three variables as the plants and animals. Table S1, in the supporting information available on-line (see p. 487), lists the items used.

Procedure
Participants were instructed to indicate, as quickly and as accurately as possible, whether each presented word referred to a living or nonliving thing. In this category-verification task, each trial began with the presentation of a fixation cross, which appeared for a random interval that ranged from 500 to 1,500 ms. Following fixation, the stimulus word, in black on a white background, was presented until the subject responded or for a maximum of 1,000 ms. The presentation window was adjusted through pilot testing and was designed to elicit accuracy differences, in addition to response time variations, within the stimulus set. If the participant did not respond in time, “no response detected” was briefly flashed on the screen in red. The trial ended with a visual mask presented for 500 ms. The task was implemented using the E-Prime software package (Psychology Software Tools, Inc., Pittsburgh, PA) and was administered via standard computers. On each trial, participants responded by pressing one of two keys on the keyboard.

Statistical Analyses
We conducted a priori statistical comparisons between the matched word lists. Accuracy comparisons included all trials, whereas only trials on which participants responded correctly were included in analyses of response latencies. The animal and plant lists were compared using paired-sample t tests. The lists of nonliving things were compared using two-factor (movement, naturalness) analyses of variance (ANOVAs). In addition, to confirm the generalizability of effects across items, we conducted item analyses (averaging across subjects, treating each item as a random variable in the analysis) for each of the main effects. Table S2, in the supporting information available on-line, presents item-wise data.

Results
As Figure 1 shows, accuracy was 16% lower for plants than for animals, t(28) = 8.13, p < .001, and participants were 70 ms slower to respond correctly to plants than to animals, t(28) = 9.48, p < .001. Comparisons of the subsets of nonliving things demonstrated main effects of movement and naturalness, but no interaction. Accuracy was 5% lower for moving nonliving things than for nonmoving nonliving things, F(1, 28) = 16.12, p < .001, and was about 6% lower for natural kinds than for artifacts, F(1, 28) = 16.73, p < .001 (see Fig. 2a). Correct responses to moving nonliving things were significantly slower (by 25 ms)
than correct responses to nonmoving things, $F(1, 28) = 6.00, p < .05$, and correct responses to natural kinds were significantly slower (by 35 ms) than correct responses to artifacts, $F(1, 28) = 12.11, p < .001$ (see Fig. 2b).

Item analyses confirmed the three main effects. Responses to plants, as compared with responses to animals, were less accurate, $t(58) = 5.18, p < .001$, and slower, $t(58) = 5.57, p < .001$. Responses to moving things, as compared with responses to nonmoving things, were less accurate, $t(78) = 2.47, p = .016$, and slower, $t(78) = 2.47, p = .016$. And responses to nonliving natural kinds, as compared with responses to artifacts, were less accurate, $t(78) = 2.68, p < .01$, and slower, $t(78) = 3.67, p < .001$.

Discussion

We asked undergraduate students to verify the life status of items that do and do not tend to cause confusion for young children and found that the students encountered much more difficulty, as indexed by decreased accuracy and increased response times, in responding to the items that tend to cause such confusions. These results suggest that features used early in development to help constrain classification of biological kinds also influence the conceptual processing of adults. However, it is possible that undergraduate students, despite possessing mature knowledge of biology, have not acquired conceptual resources sufficient to foster radical conceptual change in this domain. To test this hypothesis, we presented the same verification task to university biology professors. If such experts demonstrate difficulties in verifying the life status of items that tend to cause confusions early in childhood, then a residual developmental foundation would seem to underlie further knowledge acquisition within the biological domain.

EXPERIMENT 2

Method

Biology professors were recruited from the Departments of Biology at Yale University and Johns Hopkins University. Twenty professors (17 males, 3 females; 40–79 years old, $M = 57$ years old) agreed to participate; they had been employed as biology faculty for an average of 24 years. Data from 1 participant were removed from further analyses because of an average accuracy level below 75% across all trials. The materials, procedure, and statistical analyses were the same as for Experiment 1.

To examine interaction effects involving subject group (undergraduates vs. professors), we conducted mixed ANOVAs on both accuracy and response times. Effects specific to the living things were analyzed by examining the interaction of category (animals vs. plants) with group. For the nonliving things, the interactions of movement (moving vs. nonmoving) with group and naturalness (natural kind vs. artifact) with group were examined, as was the three-way interaction.

Results

Figure 3 shows that biology professors were significantly less accurate, by 6%, for plants than for animals, $t(18) = 4.37, p < .001$, and responded correctly more slowly, by 25 ms, for plants than for animals, $t(18) = 4.58, p < .001$. For the nonliving things, biology professors also showed main effects of movement and naturalness, but no interaction. They were about 4% less accurate for nonliving things that move than for those that do not move, $F(1, 18) = 4.04, p < .05$, and were about 6% less accurate for nonliving natural kinds than for artifacts, $F(1, 18) = 7.52, p < .01$ (see Fig. 4a). The biology professors showed a trend for longer response times, by 21 ms, to moving nonliving things than to nonmoving things, $F(1, 18) = 2.91, p = .09$, and were significantly slower, by 30 ms, in responding to natural kinds than to artifacts, $F(1, 18) = 5.96, p = .017$ (see Fig. 4b).
Analyses based on the professors’ responses to each item confirmed the three main effects. Responses to plants, as compared with responses to animals, were less accurate, $t(58) = 2.55, p = .013$, and slower, $t(58) = 4.14, p < .001$. Responses to moving things, as compared with responses to nonmoving things, trended toward less accuracy, $t(78) = 1.82, p = .072$, and were significantly slower, $t(78) = 2.63, p = .010$. And responses to nonliving natural kinds, as compared with responses to artifacts, were less accurate, $t(78) = 2.53, p = .013$, and slower, $t(78) = 3.86, p < .001$.

The analyses directly comparing the undergraduates’ and biology professors’ performance among the living things yielded significant interactions of category (animals vs. plants) with group, for both accuracy, $F(1, 46) = 15.07, p < .001$, and response time, $F(1, 46) = 18.30, p < .001$. For both dependent variables, the difference between plants and animals was reduced by nearly two thirds in the professors, as compared with the students. By contrast, for the nonliving things, the interactions of movement with group and of naturalness with group were not significant, and neither was the three-way interaction. Figure 5 shows the effect sizes for all comparisons across the subject groups. Table S2, in the supporting information available on-line, shows the item-level effects across the two groups.

**Discussion**

This experiment examined whether biology professors would demonstrate the same difficulties in verifying life status as undergraduate students. The results suggest that extensive training within the domain does not lead to radical changes in the underlying conceptual web of biological knowledge. The same items and features that cause confusions in young children also appear to cause underlying classification difficulties in university biology professors. Compared with the undergraduate students, the biology professors showed a reliable decrease in the magnitude of the performance difference between animals and plants. This result suggests a subtle restructuring of the conceptual framework that does not appear to filter through to
nonliving things. Extensive mastery of the biology domain may bring conceptual overwriting in a limited and piecemeal fashion, without wholesale replacement of early biases.

**GENERAL DISCUSSION**

The results of these two experiments suggest that rapid categorical processing in the domain of biology relies on a developmental foundation. Feature biases within biological knowledge appear to reflect developmental “roots” that cannot be completely overwritten, or replaced, with the acquisition of more advanced knowledge. The most obvious alternative explanation is that the task we used required participants to access a prototype representation (Rosch & Mervis, 1975) of living things biased in favor of movement and naturalness. However, this explanation leaves unanswered why these particular features should have affected performance when many other features are much more relevant for distinguishing between living and nonliving things. For instance, moving natural kinds (e.g., “cloud” and “river”) seem to have little in common (i.e., very few overlapping features) with most living things. It is therefore difficult to place the performance decrements associated with these items within traditional prototype accounts based on the central tendency of features within categories (E.E. Smith, Shoben, & Rips, 1974; Rosch & Mervis, 1975). Similarly, to verify life status quickly, participants may have used an identification heuristic (Armstrong, Gleitman, & Gleitman, 1983) in which prototypical features served as a reliable guide to category membership. Yet, as Armstrong et al. noted, it is not clear what types of features should be useful when simply identifying an item, as opposed to making theory-based distinctions. That is, even if an identification heuristic is used, it is still not clear why that feature-based processing should reflect childlike confusions, especially in the case of domain experts with decades of accumulated “core” knowledge.

Our results suggest that initial category processing relies on superficial features whose importance is established early in the developmental trajectory, and more mature knowledge appears to act at a secondary stage to correct or inhibit feature-based confusions. This characterization is similar to seminal models of category processing (McClosky & Glucksberg, 1979; E.E. Smith et al., 1974), though even in these models the developmental influences on mature structures remain underspecified. The current results suggest that conceptual relatedness involves perceptually grounded similarity (Goldstone, 1994) based on coarsely coded interfeature associations and weights (Hampton, 1997) set early in development. Similarity-based mechanisms may be used to readily interpret, or identify, instances on the basis of perceptual features (Goldstone, 1998), whereas theory-based knowledge may operate to constrain inferences and explanations (Gopnik & Wellman, 1994; Murphy & Medin, 1985; E.E. Smith & Sloman, 1994). Hybrid models examining both similarity- and theory-based mechanisms are not often the focus of empirical work (Hahn & Chater, 1998), but such work is necessary for a more complete understanding of conceptual processing.

The problem in drawing a clear distinction between living things and nonliving natural kinds (Kelemen, Widdowson, Posner, Brown, & Casler, 2003) highlights the difficulty in explaining how language is used to partition the world, especially across cultures (Berlin, 1992). For instance, the polysemy of “alive” and related biological concepts introduces challenges to English-speaking children that are less problematic in other languages (Anggoro, Waxman, & Medin, 2008). Cross-cultural studies using the current paradigm could indicate how developmental traces arise from, and persist in, the context of linguistic and perhaps even religious (Inagaki & Hatano, 2004) pressures during early childhood. If a developmental foundation grounds concept acquisition throughout the life span, then cross-cultural differences observed among children should also be apparent in adults performing speeded tasks.
Continuity between the initial and end states may reflect more general mechanisms of cognitive development. However, it is unclear whether the constraints of development within mature semantic processing are homologous, and not simply analogous (J.D. Smith & Kemler Nelson, 1984), to the ways that children derive meaning from the world. Similar residual effects of development in adults have been found with false-belief tasks, which indicate persistence of early biases in theory of mind (Keysar, Lin, & Bar, 2003), and in object-tracking tasks, which show fixed limits on the number of objects that can be tracked (Carey & Xu, 2001). Further work is needed to determine the extent to which our conceptual representations rely on developmental roots well past the time when we should “know” better.

Acknowledgments—The authors would like to thank the biology departments at the University of Pennsylvania, Yale University, and Johns Hopkins University for their support. M. Chi, J. Fiez, C. Johnson, J. McClelland, C. Olson, J. Opfer, C. Perfetti, and R. Siegler offered especially valuable advice and discussions at different stages of this work. Support from the National Institutes of Health (R01 MH070850) is gratefully acknowledged.

REFERENCES


(SUPPORTING INFORMATION)

Additional Supporting Information may be found in the on-line version of this article:

**Table S1**

**Table S2**

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

(Received 3/8/08; Revision accepted 9/16/08)