2009

Can Our Experiments Illuminate Reality?

Brian MacWhinney
Carnegie Mellon University, macw@cmu.edu

Can Our Experiments Illuminate Reality?

The papers by Ahn & Luhmann, Bowerman, Gentner, and Gershkoff-Stowe present a detailed picture of current research on the many processes involved in children's concept learning. In this commentary, I will first summarize and critique the findings of these papers. Then I will discuss possible future directions in the study of children's categorization. In particular, I will argue that we should now begin to study actual instances of real concept learning in homes and schools.

Ahn & Luhmann present a commonsense analysis of artificial concept learning that emphasizes the role of prediction from properties and goals to concept membership. The basic idea is that we can best categorize results by looking at their causes. Good categories are those that maximize the probability of a result, given a cause or p(result|cause). Bayesian theory views the ongoing updating of this cue validity as involving the modification of the subjective probability of p(result|cause) on the basis of new evidence. This same traditional cue validity model also sits at the root of the Competition Model – a model that has been used to account for the learning of grammar , morphology , and concept learning.

Ahn & Luhmann's work focuses primarily on the role of causes in the prediction of novel categories. For example, when given Jane who is depressed because she has low self-esteem, Susan who is depressed because she has been drinking, and Barbara who is defensive because she has low self-esteem, subjects tend to group together Jane and Barbara because they both have low self-esteem, rather than Jane and Susan because they are depressed. The basic idea here is that we focus more on causes than symptoms. After all, we have learned that a given underlying cause can produce a variety of effects that change over time. To form a consistent concept, it is best to focus on the causes, rather than the effects.

However, one might note that the shape of the novel concept learning task in the studies reported by Ahn & Luhmann, as well as those used in similar studies by Gelman, Keil, Rips, and others tends to emphasize the salience of causes, perhaps more than in our natural experiences with concept learning. In these studies, causes are directly presented to the subject. There is no process of causal inference or discovery. Ahn & Luhmann conclude from their studies that "theory use is as fast as, if not faster than, comparable similarity use." This may well be generally true for studies of this type. However, in the natural world, the situation seems to be reversed. In nature, we are typically confronted with results from which we have to infer causes. We may see that Jane drinks too much or that Frank yells at his children. It takes awhile before we learn that they both have been having trouble at their jobs. We may notice that both lizards and snakes like to sit on rocks in the middle of the day. It takes awhile to learn that they do this because they are both cold-blooded.

Ahn & Luhmann suggest that their causal status hypothesis can be used to provide a natural resolution of the form vs. function debate in the child categorization literature. However, the resolution they present seems to me biased in favor of the causal status account. They note that children only make clear use of function in acquiring new categories, when the causal basis of the function is plausible. But in other studies (Kemler Nelson, Frankenfield, Morris, & Blair, 2000), the experimenter makes sure that the child is fully aware of the functions of the objects. This full disclosure of cause short-circuits the normal process of discovery of causal structures that is involved in concept learning. Although I thoroughly agree with Ahn & Luhmann that categorization is designed to maximize p(result|cause), it seems to me that the normal process of concept formation does not follow the royal road of immediate causal discovery provided by this experimental literature. Instead, if we take a close look at interactions between mothers and children or teachers and their students, we will see a complicated discourse and interaction designed to explore the uses of concepts in various contexts precisely with an eye toward elucidating causes. Thus, it should not be surprising to find that the initial stages of concept learning involve considerable attention to "superficial" perceptual attributes.

However, Ahn and Luhmann also point to another view of concept learning that seems to tap more directly into early causal learning. This view points to a direct mapping between a new concept and a first-person intention. For example, Searle says that he knows immediately that he takes a drink of water because he is thirsty. If we see a horse drinking water, we can infer the cause directly by relating the horse to our own first-person perspective . I am not suggesting that all new categories are acquired through a mapping to first-person perspective. However, the idea that this type of embodied mapping stimulates concept coherence and productivity is certainly worth pursuing. Moreover it seems to me that the use of first-person perspective for causal discovery fits in well with an extended version of Ahn & Luhmann's causal status hypothesis.

Causal induction also figures prominently in Gentner's account of the acquisition of relational categories, such as *gift*, *robbery*, or *loan*. Gentner's account of the learning of relational categories focuses on the fact that category labels trigger a process of structural alignment. When a child hears two very different objects being described by the same term, the sharing of a label triggers a "cue search" process. Gentner thinks of this search as involving the alignment of two items. The technique of Namy and Gentner (2002) specifically involves the alignment of two objects, just as required by the model.. One might argue that Gentner's structural alignment model overemphasizes pairwise similarities without providing a role for larger group alignment of the type one would find in neural network models. However, in defense of Gentner's position, I would note that there is evidence from the adult literature suggesting that learners often try to focus on a small set of exemplars and hypotheses, perhaps because of working memory limitations. From this viewpoint, Gentner's emphasis on pairwise alignment seems motivated. However, one must keep in mind that here, as in the case of

Ahn & Luhmann, the particular experimental method being used is emphasizing the presentation of pairs of objects or relations to be aligned.

In accord with my earlier comments regarding Ahn & Luhmann, Gentner emphasizes the difficulties children face in discovering causal and functional properties underlying relational categories. However, instead of viewing perceptual features as a stable initial organizational system, she believes that they are only important to the extent that they induce structural alignment. Citing the results of Samuelson and Smith (Samuelson & Smith, 1998), she notes that "children's use of perceptual features as a basis for word learning may in fact be highly adaptive. When children lack knowledge about a concept, perceptual commonalities may serve as the initial 'hook' that encourages them to engage in comparison and extract deeper relational commonalities." Note that the emphasis here is on structural alignment as the sole target with perceptual features only serving to direct the learner's attention. However, Gentner would probably not deny the possibility that perceptual features serve additional roles in terms of allowing for the storing of episodic memories of objects and providing rich features for later support of structural alignment. Thus, what is at issue here is the microgenesis of the transition from initial perceptual alignment to deeper causal alignment. That is a worthy topic for future investigation.

The new work of Gentner and Klibanoff seeks to induce structural alignment in 3-6 year-olds. Inventing words such as *blick* with a meaning like "cutter", they were able to teach 4-year-olds, but not 3-year-olds, to relate cutting paper to chopping down a tree. One possible account of this finding is that the older children were quicker to map *blick* onto the concept of cutting. In this account it is the verbal and conceptual mediation and not the deep structural alignment that facilitated success in this task. Manipulations that overtly vary the easy of mapping a new action to an existent English word might control for this effect, although one would then worry about the status of actions that cannot be described by English words. Perhaps the children could be tested by a probe recognition technique to determine the degree to which they activated the word *cut* during the relevant set of *blick* stimuli.

More generally, we have only this one study in this new line of research to defend Gentner's general claim that relational categories can be learned by the general learning mechanism of progressive structural alignment, as triggered by label similarity. However, at this point, even without further work, I would wonder a bit about the extent to which this general claim is falsifiable. It would seem difficult to deny that concept learning is facilitated by the presence of verbal labels. Theories might differ in the way they see this facilitation as operating mechanistically, but it would be difficult to deny the overall effect. Similarly, it would be difficult to imagine that pointing out the similarities between two objects would fail to help concept formation. We would only expect this type of manipulation to fail when the similarities involve entities and relations that are outside of the experience of the child. So, it is not clear to me yet how the theory can be tested. This is unfortunate, since it seems largely true.

Bowerman takes a rather different approach to the word learning problem. Decrying the uniformity of the "cognitivist climate of the last forty years," she suggests that it is worth pursuing the idea that children induce categories from the linguistic input. The fact that languages differ so markedly in their structuring of verbs and classifiers opens up great opportunities for crosslinguistic research on early word learning. Surveying a continually growing literature, Bowerman shows that children are indeed sensitive to the details of the meanings of the target language in areas such as spatial relations, clothing, dressing, and a variety of causative actions.

Bowerman presents brilliant detailed analyses of target linguistic structures and a view of the child as soaking up every detail of the linguistic input, unguided by universalist biases. Given the child's remarkable sensitivity to the input, the challenge is to explain the fact that children still make some errors. For example, English-speaking children may say that they are *opening* some Lego blocks, rather than *taking apart* some Lego blocks. She refers to several mechanisms that could be operative here. First, there is a cue strength tuning process that can increase attention to things such as the sharpness of the instrument used for English *cut* as opposed to Spanish *cortar*. This process may well tap into the types of comparisons that Gentner has described in the theory of structural alignment. Second, there is a process of competition (MacWhinney, 1987) that eventually leads the child to extend *take apart* into the semantic area previously occupied by *open*. The final stages of fine-tuning of this competition appear to involve what Bowerman, Brown (1958), MacWhinney, and others have called cue search. In this process, the detection of errors in the use of terms triggers a search for dimensions that can predict correct usage. For example, a child may come to realize that you can only *open* something that you can also *close*. As a result, errors involving opening things like nuts and oranges will drop out relatively early. However, some errors involving the separation of assemblies such as Legos may remain for still further cue search.

Gershkoff-Stowe applies dynamic systems theory to conceptual development. According her view, self-organizing processes centered around the child's actions shape cognitive development across multiple times scales within particular contexts. Here, Gershkoff-Stowe focuses on the ways in which children use repeated grouping and touching of objects to group them dynamically into categories. She finds that, the sooner children begin manipulating the objects, the sooner they begin to group them into categories. Of course, one could argue that this is simple a "time on task" effect . However, the point is that children are engaged here in self-defined dynamic interactions with these objects that allow them to construct new categories. It is not the mere exposure to objects across a period of time that is important. It is what children do with the objects, both physically and cognitively, that shapes the emergence of the relevant categories.

In a second study (Namy et al, 1997), children further benefited from opportunities to compare and contrast between object classes, using a shape sorter. In effect, this study can be viewed as a dynamic externalization of what Gentner calls structural alignment or what others call cue search. In a final study, Gershkoff-Stowe shows how dynamic pressures shape categorization across multiple time scales. Children begin categorization in this experiment with a long-term shape bias. When confronted with evidence of the importance of the function of the object, children then demonstrate flexibility within the shorter time frame of the experiment

Together, these four studies illustrate the extent to which we now have a clear set of cognitive mechanisms that, in principle, should be sufficient for the acquisition of a conceptual world. We have clear ideas about structural alignment, causal status prediction, cue validity, cue search, competition, flexibility, and dynamic systems. We can use these concepts to predict the outcomes of carefully controlled experiments. But can we extend this thinking to the study and exegesis of concept learning in real-life situations?

In a sense this extension has already begun. The task used by Gershkoff-Stowe is clearly a dynamic one similar to what children do in real play sessions. However, this task alone does not tell us much about the extraction of real categories. Providing another clue, Bowerman reports, at the conclusion of her chapter, an interaction between M and C (7;11). M describes a box with some broken toys and two scraps of felt as having “some other broken things” and C observes disdainfully that, “I wouldn’t call felt broken, I would call it ripped.” Presumably, M’s use of “broken” is licensed by the fact that she includes the quantifier “other” which forces a comparison of a subset of the objects in the box to some previous set of objects. Why does C decide to ignore this interpretation of “other” and correct M (who is presumably an adult) at this point? Having access to a full video record of the interaction along with preceding context would certainly help use delineate the level at which C is constructing a cue search to delineate precise lexical borders or perhaps simply joking with M.

Fortunately, we now have a unique new opportunity to explore these issues. Over the last two years, researchers have begun to contribute large amounts of digitized video records of conversational interactions in the home and in the classroom. These sources are now directly accessible over the web through the servers of the Child Language Data Exchange System (CHILDES) at http://childes.psy.cmu.edu/data/ and TalkBank at http://talkbank.org/data/. Interactions of children in the home are at the CHILDES site and interactions in classrooms are at the TalkBank site in the “classroom” data folders. In the following comments, I will be specifically referring to the “browsable” audio and video files at those sites. Browsable corpora include either audio or video recordings that are directly linked to the transcripts. By using QuickTime SMIL technology through your web browser, you can watch the video or listen to the audio and read the transcript concurrently. When you hit upon a segment that reveals important aspects of concept learning, you can pause and repeat the playback. There are also facilities for commenting on segments of the interactions and sharing these comments with colleagues across the web.

Even at this early stage, the quantity of browsable data is so great that researchers interested in tracking out the naturalistic contexts of concept development would benefit from a set of guideposts and suggestions. Having now spent perhaps 100 hours reviewing these new resources, let me present three illustrations of how these data can teach us about the real world dynamic process of concept and word learning.

1. Consider first the case of the learning of the numerical interpretation of the word *dependable* in a 7th Grade classroom in Nashville. The relevant video was included in a *Journal of the Learning Sciences* special issue that is duplicated in the "JLS" folders on TalkBank. The class is discussing the performance of two types of batteries. Type A has a large number of long-lasting batteries, but also many failures. Type B has almost no failures, although fewer batteries were tested. In the lively discussion, several students focus only on the fact that type A has more successes, judging that therefore it is more "dependable." A dispute arises when one of the students points out that dependability is not just about successes, but also about lack of failures. This interaction not only illustrates the lateness of a full acquisition of the notion of dependability, but also some remarkably aspects of the social dynamics involved in the shifting of definitions in public spaces.
2. As a second example, consider the acquisition of the word *alert* by Mark MacWhinney at age 5;4 in the boys85.cha file on the CHILDES website. When the taping begins, the Father (Brian MacWhinney) is questioning the form *alerd bean* used by Mark. His older brother Ross, then corrects Brian, saying that Mark was saying *alert*, not *alerd*. Eventually, it requires the collaborative participation of the whole family to figure out what Mark meant to say and where he learned it. For the details, please just go to the website and replay this interaction.
3. As a final example, consider the treatment of the phrase "when Ella was very small" in Michael Forrester's recordings from his daughter Ella at 28 months. Failing to use the conjunction "when" as a temporal delimiter, Ella then protests against her father's description of her as a "tiny, tiny baby."

What do we learn from examination of materials such as this? Perhaps the most immediate and obvious lesson is that current research on concept development fails to prepare us for the obvious fact that most conceptual learning is embedded in rich social contexts. These contexts are important not only for the way they present information, but also in the ways in which they provide for the dynamic sharpening of category boundaries. Through conversation, the borders of categories are explored, discussed, and refined. In effect, one of the major goals of conversation is the sharpening of shared categories.

This is not to say that concepts such as structural alignment, competition, dynamic exploration, causal prediction, and cue search are irrelevant to this social process. On the contrary, we can study real interactions to see exactly how children instantiate these processes. We use conversation to learn about causes, to resolve competitions, to align comparisons, and to search for cues. By studying these processes in their real, natural context we can make our understanding of the actual growth of the conceptual world far more accurate and precise.

References

Brown, R. (1958). How shall a thing be called? *Psychological Review, 65*, 14-21.

Kemler Nelson, D., Frankenfield, A., Morris, C., & Blair, C. (2000). Young children's use of functional information to categorize artifacts: three factors that matter. *Cognition, 77*, 133-168.

MacWhinney, B. (1987). The Competition Model. In B. MacWhinney (Ed.), *Mechanisms of language acquisition* (pp. 249-308). Hillsdale, NJ: Lawrence Erlbaum.

Namy, L., & Gentner, D. (2002). Making a silk purse out of two sow's ears: Young children's use of comparison in category learning. *Journal of Experimental Psychology: General, 131*, 5-15.

Samuelson, L. K., & Smith, L. B. (1998). Memory and attention make smart word learning: An alternative account of Akhtar, Carpenter, and Tomasello. *Child Development, 69*, 94-104.