



Categories, concepts, and conceptual development

Vladimir M. Sloutsky & Wei (Sophia) Deng

To cite this article: Vladimir M. Sloutsky & Wei (Sophia) Deng (2017): Categories, concepts, and conceptual development, Language, Cognition and Neuroscience, DOI: [10.1080/23273798.2017.1391398](https://doi.org/10.1080/23273798.2017.1391398)

To link to this article: <https://doi.org/10.1080/23273798.2017.1391398>



Published online: 26 Oct 2017.



[Submit your article to this journal](#)



Article views: 142



[View related articles](#)



[View Crossmark data](#)



Categories, concepts, and conceptual development

Vladimir M. Sloutsky^a and Wei (Sophia) Deng^b

^aDepartment of Psychology, The Ohio State University, Columbus, OH, USA; ^bDepartment of Psychology, University of Macau, Macau SAR, People's Republic of China

ABSTRACT

Concepts (i.e. lexicalised classes of real or fictitious entities) play a central role in many human intellectual activities, including planning, thinking, reasoning, problem solving, and decision making. How do people *acquire* concepts in the course of development and learning and *use* them in their thinking about the world? In this article, we attempt to provide an overview of conceptual development. We suggest that concepts can originate (1) in interactions with the world and get lexicalised later or (2) in the language and get grounded later. The first route is from category learning to a concept, and we discuss this route by focusing on the mechanisms of category learning and developmental changes in these mechanisms. The second route is from a word to a concept, and we discuss this route by focusing on inferring word meanings without visual referents. We then consider proposals of how concepts get organised into networks and hierarchies.

ARTICLE HISTORY

Received 2 August 2016
Accepted 28 September 2017

KEYWORDS

Categorization; conceptual development; semantic development; learning; cognitive development

Introduction

There are many abilities that reflect the remarkable intelligence of humans: people make inferences, develop and use scientific theories, make laws, preserve knowledge and pass it onto new generations, write fiction, reason about past and future, and make counterfactual arguments.

For example, upon learning that all animals are heterotrophs (i.e. they use other organisms as a source of energy) one may conclude that cats are heterotrophs as well. Furthermore, this conclusion follows with logical necessity from knowledge of taxonomy (e.g. *cat* is a proper subset of *animal*) and of the logic of classes (e.g. if class X is properly included in class Y, then whatever is attributable to Y is attributable to X). Therefore, all properties of class Y (e.g. animals) is shared by class X (e.g. cats). One can also entertain a counterfactual (which in this case follows with logical necessity): *if cats were not heterotrophs, they would not have been animals*.

All these abilities are based on *concepts*, and, in fact, it is difficult to imagine reasoning (or any intellectual activity) without concepts. Therefore, one of the most interesting challenges in the study of human cognitive development is to understand how people *acquire* concepts in the course of development and learning and *use* them in their thinking about the world.

In this article, we attempt to provide a brief overview of conceptual development. Because concepts are

classes (rather than individuals) they are more *general* or *abstract*. We use these terms interchangeably: *abstract* or *general* versus *concrete* or *specific* is a dimension reflecting how inclusive the class is. More inclusive classes are also more abstract and general. For example, cats is a more abstract class than an individual (e.g. my cat *Fluffy*), but less abstract than a more inclusive class of *mammals* or *living things*. More abstract concepts can be formed by the process of *abstraction* or *generalisation*, whereby some (presumably more important) properties are preserved, whereas other (presumably less important) properties are dropped. For example, it is easy to see that *triangle* is more abstract than *equilateral triangle*, whereas *polygon* (a shape with any number of edges and vertices) is more abstract than *triangle*.

To distinguish concepts from other classes and sets (such as categories), we define concepts as *lexicalised* classes of real or fictitious entities. We also suggest that concepts can originate (1) in interactions with the world and get lexicalised later or (2) in the language and get grounded later. For example, it is conceivable that a toddler has enough encounters with dogs to form a category before learning the word for it. In contrast, the concept *germ* cannot originate in experience and has to originate in language. Therefore, concepts can be acquired in a bottom-up manner (i.e. originating in experience) or in a top-down manner (i.e. originating in language).

Because bottom-up concepts (i.e. those acquired in a bottom-up manner) are tightly linked to experience, there are a number of constraints as to what these concepts might be. In general, many of these are “embodiment” constraints on the concepts (cf., Yu & Smith, 2012, for ideas on the role of embodiment in the early word learning). Among these constraints are (1) these concepts should primarily include perceptible (and possibly actionable) entities, (2) these concepts should be mostly based on objects, (3) these objects should predominate in a particular child’s experience, (4) these objects’ sizes should be within a certain range, and (5) there should be common and frequently used labels within the child’s native language to denote these objects.

In contrast, the top-down concepts (those acquired in a top-down manner) have a different set of constraints. Because these concepts originate in language, the primary constraints are (1) parents’ education and vocabulary, (2) topics of conversations with and around the child, (3) access to books and media and types of available books and media, and (4) access to formal education. Second language learning may offer an interesting illustration of these latter constraints, as both authors of this paper learned English as a second language. For example, because of the literature we read while learning English, we learned words like *pride*, *prejudice*, *vanity*, *expectation*, and *curiosity* long before learning words for *eggshell*, *faucet*, *finger nail*, *eyelash*, or *tire*.

In what follows, we will overview concepts and conceptual behaviours, introduce category learning as an important step in conceptual development, discuss both ways of acquiring concepts (i.e. the bottom-up and the top-down), and consider how concepts get organised into coherent networks that promote understanding of the world as well as thinking and reasoning about it.

What are concepts?

In the simplest possible way, concepts can be defined as lexicalised *categories*, or *equivalence classes*. What is an equivalence class? In his chapter focusing on concepts (Chapter XII of the *Principles of Psychology*), William James (1983/1890) wrote: “Our principle only lays it down that the mind makes continual use of the notion of sameness, and, if deprived of it, would have a different structure from what it has.” In other words, the mind can treat different things as if they were equivalent in some way. Once equivalence is established, it could be marked by lexicalisation: by calling two different dogs d_1 and d_2 a *Dog*, we can express that as $d_1 \equiv d_2$. When such an

equivalence class (or category) is lexicalised, it becomes a *concept*. Lexicalisation allows (a) accumulation of new (often unobservable) information about categories and (b) communicating and sharing this information with others. Examples of concepts vary from *chairs* (obviously, chairs are non-identical, but merely equivalent in some way) to *odd numbers* to extremely abstract concepts, such as *cause* or *effect*. If concepts are lexicalised categories, then development may proceed either from forming a category first to lexicalisation or from acquiring a lexical item first to the formation of a category. As we discuss below, both types of progression can be observed in individual development.

We also suggest that regularities in both the world and in language are sources of conceptual development. In the world, members of many categories (e.g. cats or birds) share multiple observable features and therefore these categories can be learned without language. On the other hand, language also provides important input to conceptual development: many related categories (e.g. wolfs and dolphins) share some observable and many unobservable features and knowledge of these unobservable features comes from language. In what follows, we discuss both developments. We first review some principles of category learning – principles that may apply to learning of pre-linguistic categories. We then review how language can be a rich source of conceptual development by contributing to the formation of conceptual hierarchies.

Perceptual groupings, categories, concepts, and conceptual networks

If concepts are lexicalised categories (some of which can be learned prior to language acquisition and some cannot be learned without language), then, conceptual behaviours may vary substantially in levels of complexity ranging from simple perceptual groupings of arbitrary categories, to full blown lexicalised concepts that are linked to other concepts which thereby form conceptual networks. The study of each type of conceptual behaviour requires somewhat different research paradigms.

First, people can learn perceptual groupings or equivalence classes that are based on purely perceptual properties. Such groupings may include imposing categorical boundaries on sensory continua (known as categorical perception, e.g. Eimas, 1994), learning dot patterns coming from a single prototype and generalising learning to distortions from the studied prototype, or forming a category based on image properties (see Bhatt & Quinn, 2010, for a review). Perceptual groupings are the simplest form of categorisation because they allow extending category membership on the basis of

global familiarity. Therefore, if members of category A share some features, a novel item would be judged as a member of A to the extent that it has these features.

A more complicated variant of conceptual behaviour requires one to learn two or more mutually exclusive categories (e.g. cats vs. dogs) at the same time. The categories are mutually exclusive because there are no members common to A and B (i.e. $A \cap B = \emptyset$). This problem is more difficult than simple perceptual groupings because a decision of whether a novel item belongs to A or to B cannot be made on the basis of global familiarity (i.e. both A and B are equally familiar). The studied categories can be based on multiple correlated features (birds have wings, feathers, and beaks, whereas fish have scales, fins, and gills), few features (e.g. squirrels have a long, fluffy tail, whereas hamsters have a small tail), or relations among features (e.g. rectangles can be grouped into tall if the aspect ratio is less than 1, and wide if the aspect ratio is more than 1). The categories may be also deterministic (such that there is a subset of features that is sufficient to predict category membership with a 100% accuracy) or probabilistic (such that any feature or a combination of features predicts category membership only with a degree of probability). Therefore, to make a categorisation decision, at the very minimum, some processing of two category structures is required. This task has been used in some studies with infants and animals, and in many category-learning studies with children and adults.

An even more complicated variant of conceptual behaviour is the ability to lexicalise categories and use them in reasoning, inference, prediction, or judgment. Such lexicalised categories can be defined as concepts proper. Lexicalisation is critical as it enables acquiring knowledge that may not be directly observable in a given situation (e.g. dogs are friendly pets, they like meat, and are taken to a vet for a physical exam). In other words, having a word for a category allows accumulation of knowledge from sources that are not based on direct observation of category members. These sources include conversations with others, books and other media sources, and formal education. Such concepts proper can be studied in a variety of tasks, including grouping of items, property listing, picture naming, and category judgment among others. A grouping task may require participants to put together items of the same kind (e.g. toys versus animals), whereas an attribute listing task may require a participant to list properties of categories (e.g. of cats, birds, or animals).

Finally, a conceptual network involves not only knowledge of concepts, but also of relations among these concepts. Take, for example, Newton's second law ($F = ma$) that acceleration of a body is directly proportional to

the net force acting on the body and inversely proportional to the mass of the body. Here, the concepts of *mass*, *force*, and *acceleration* are linked and are part of a broader conceptual network. For example, *force* is linked to *work* and *power*, whereas acceleration is linked to *time* and *space*. In other words, all these concepts are linked and each can be derived from others.

Such networks can be organised in a variety of ways; for example, networks of naturally occurring categories often have hierarchical, or taxonomical, organisation (e.g. grey hound \rightarrow dog \rightarrow mammal \rightarrow animal \rightarrow living thing). One way of detecting such hierarchies is a classification task in which a diverse set of items is partitioned into N mutually exclusive and exhaustive sub-sets. These subsets can then be further partitioned into smaller groups or combined into larger groups. Although it has been argued that classification tasks may underestimate children's concepts (the fact that a child may put together a dog and a bone does not mean that the child considers the two to be the same thing, see Fodor, 1972), classification tasks are useful in that they may reveal a limit on the kinds of concepts children may form.

In short, humans exhibit multiplicity of conceptual behaviours, some are universal and shared with other animals, whereas others are uniquely human. Overall, human conceptual repertoire ranges from perceptual groupings (something that can be also achieved by certain non-mammalian species) to conceptual networks that are likely to be unique to humans.

Early conceptual development: from categories to early concepts

Conceptual behaviours come in various forms: they range from more simple, universal, and early emerging forms (i.e. establishing equivalence between non-identical percepts) to rather complex, uniquely human, and late emerging forms (i.e. forming a conceptual network in a knowledge domain).

Category structure and category learning

Are all categories the same? Perhaps not: although there is little doubt that categories differ in content, the most interesting distinctions pertain to category structure. Structural differences identified by researchers include syntactic differences (nouns versus verbs; e.g. Gentner, 1981), ontological differences (natural kinds versus artifacts; e.g. Barton & Komatsu, 1989), taxonomic differences (i.e. basic-level versus superordinate-level; e.g. Rosch & Mervis, 1975), differences in organisational principle (entity categories versus relational categories; e.g.

Gentner & Kurtz, 2005), differences in concreteness (concrete versus abstract categories; e.g. Barsalou, 1999), differences in category coherence and confusability (e.g. Homa, Rhoades, & Chambliss, 1979; Rouder & Ratcliff, 2004; Smith & Minda, 2000), and some other distinctions (for a review, see Medin, Lynch, & Solomon, 2000).

Kloos and Sloutsky (2008) proposed another structural distinction, one that could form the basis for many of the above distinctions. They proposed the idea of statistical density, that is a measure of category structure that (a) can (in principle) be measured independently rather than be inferred from participants' patterns of response and (b) provides a continuous measure rather than a dichotomous one (which makes it well suited for capturing the graded nature of differences between categories). Conceptually, statistical density is a ratio of variance relevant for category membership to the total variance across members and non-members of the category. Intuitively, statistical density is a measure of how members of a category are separated from non-members (see Kloos & Sloutsky, 2008, for a detailed discussion). For example, a category of small racing cars is dense (even when contrasted with other categories of vehicles) because there are multiple correlated features that distinguish this category. In contrast, a category of red things is sparse as there is a single feature common to the category members and distinguishing this category from any contrasting category.

The idea of statistical density has important implications for the development of category learning. One possibility is that category learning progresses from spontaneous learning of highly dense categories (i.e. when multiple dimensions are correlated within a category) to less spontaneous (and more guided or supervised) learning of sparser categories (i.e. only few dimensions are relevant; for example, members of a category are all red, but vary on multiple dimensions, such as shape, colour, texture, and size).

Category learning: what is the mechanism and what develops?

Category learning is the process by which one or more equivalence classes of discriminable entities are formed. How do people form these classes? One of the first ideas was that category learning is a variant of stimulus generalisation: if a new item is sufficiently similar (i.e. exceeds some threshold value) to a member or members of an identified category, it would be included in this category. However, this simple and compelling idea may have difficulty explaining learning of categories based on a single dimension. For example, Shepard, Hovland, and Jenkins (1961) demonstrated that people easily

learn single-dimension categories (e.g. black shapes vs. white shapes), even though within-category similarity (measured as stimulus confusability) may be small. Shepard et al. (1961) concluded that categories can be learned by selectively attending to a relevant dimension. These ideas have been captured in several influential models of categorisation (e.g. Kruschke, 1992; Love, Medin, & Gureckis, 2004; Nosofsky, 1986).

Similar to Shepard et al. (1961), these models suggested that categories of the same structure can be learned either by allocating attention to few relevant dimensions or by distributing attention across many dimensions. The way attention is allocated is consequential for how quickly the category will be learned and how the dimensions will be represented in memory. In particular, the former way of category learning is fast and efficient, but it may result in inattention to (and consequently, relatively poor memory for) irrelevant dimensions as these dimensions are ignored. In contrast, the latter way of learning may be slower and less efficient, but may result in attention allocated to all dimensions (and consequently, memory for both relevant and irrelevant dimensions).

Learning of similarity-based categories is a developmental default

It is hardly controversial that selective attention undergoes protracted development (see Hanania & Smith, 2010; Lane & Pearson, 1982; Plude, Enns, & Brodeur, 1994; for reviews), with infants and young children tending to distribute attention (Best, Yim, & Sloutsky, 2013; Deng & Sloutsky, 2015a; Plebanek & Sloutsky, 2017). If infants and young children tend to distribute attention rather than to attend selectively, how do they learn categories? One idea is that first categories that infants and young children learn are categories that do not require selective attention – these are sufficiently perceptually distinct and have enough within-category structure to be learned without selective attention (Sloutsky, 2010; Sloutsky & Fisher, 2004a).

In an attempt to examine the mechanism of category learning and its potential change with development, Deng and Sloutsky (2015b) presented 4-year-olds, 6-year-olds, and adults with a category learning task, in which participants learned two categories. The categories had a rule-plus-similarity structure, such that there was a single deterministic (or rule) feature and multiple probabilistic features (see Figure 1, for examples of stimuli). Therefore, this category structure allows examining what participants learn spontaneously: given the same set of stimuli, participants could learn either rule-based or similarity-based categories.

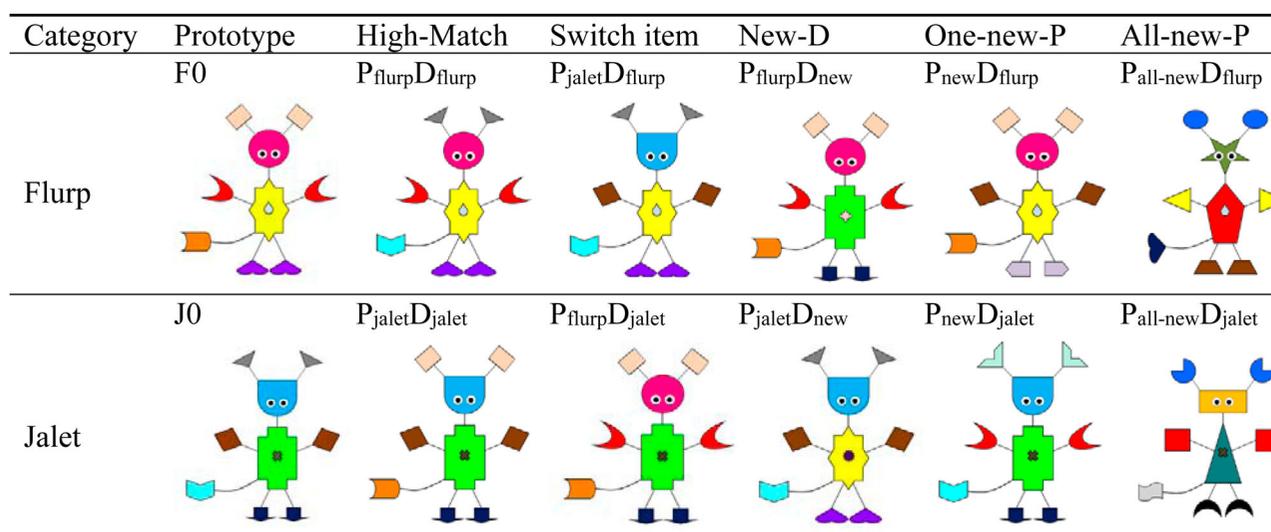


Figure 1. Examples of stimuli used in Deng and Sloutsky (2015b). There were two family resemblance categories, with each training item including a single deterministic feature D (which perfectly distinguished between the two categories) and multiple probabilistic features P (with each providing imperfect probabilistic information about category membership). The body mark (introduced as a body button) was the D feature, and all the other features – the head, body, hands, feet, antennae, and tail – were the P features. Each row depicts items within a category, whereas each column identified an item role (e.g. switch item) and item type (e.g. $P_{jalet}D_{flurp}$). The High-Match items were used in training and testing. The switch items, new-D, one-new-P, and all-new-P items were used only in testing. Neither prototype was shown in training or testing.

To establish what specifically was learned by participants, the authors presented them with categorisation and memory testing (various test items are presented in Figure 1). Categorisation trials included **High-Match** items (these were the items used in training), **Switch items** (these were the items that had the rule feature from one category and probabilistic features from another category), and **All-new-P** (these items had an old deterministic feature and all new probabilistic features). The goal of **High-Match** items was to test whether participants learned the category. The goal of **Switch** items was to examine what they learned about the category (i.e. whether they learned a rule-based or similarity-based category). Finally, the goal of **All-new-P** items was to examine whether participants could generalise on the basis of the rule features. There were also memory tests examining memory for the rule feature (i.e. New-D items) and each probabilistic feature (i.e. One-new-P items).

Results of categorisation testing (see Figure 2) indicated that whereas participants of all age groups ably learned the categories (as evidenced by high performance on High-Match items), only older participants relied in their categorisation on rule features (as evidenced by the above-chance performance on Switch items). In contrast, younger participants relied on the overall similarity (as evidenced by the below-chance performance on Switch items). In addition, younger participants were at chance for the All-new-P items, whereas

older participants were reliably above chance. These results suggested that whereas 4-year-olds learned similarity-based categories, 6-year-olds and adults learned rule-based categories.

Researchers also tested participants' memory for features. Memory results are important because they reflected what participants had learned about the categories. Whereas older children and adults remembered the rule features better than the probabilistic features, 4-year-olds remembered all the features equally well. Furthermore, follow-up studies indicated that 4-year-olds' memory for probabilistic features was (at least numerically) better than that of older participants (cf. Sloutsky & Fisher, 2004b). Taken together, these results indicated that whereas older children and adults spontaneously learned **rule-based** categories (if the structure supported such learning), young children spontaneously learned **similarity-based** categories.

To further examine mechanisms of early category learning, Deng and Sloutsky (2015b) attracted their attention to the deterministic feature by pointing to this feature on every training trial and commenting on the importance of this feature. Although in this condition 4-year-olds appeared to have learned a rule-based category (as evidenced by their categorisation and generalisation responses), they exhibited equivalently good memory for all features. Therefore, their categorisation decisions pointed to selective attention, whereas their pattern of memory did not. These findings suggest that

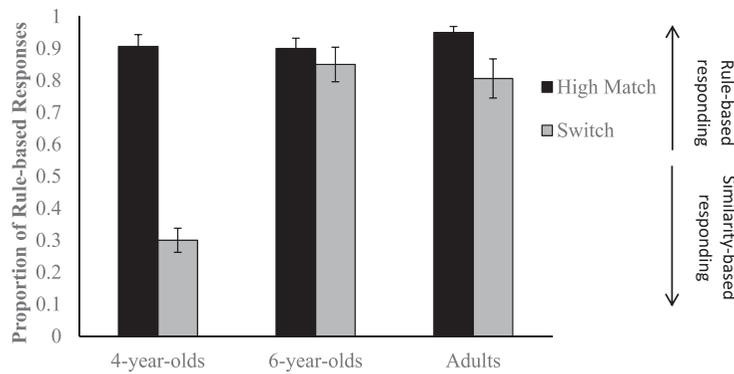


Figure 2. Categorisation performance: proportion of rule-based responses by trial type and training condition for 4-year-old children, 6-year-old children and adults (After Deng & Sloutsky, 2015b, Experiment 1).

Note: Error bars represent standard error of the mean.

category representations and category decisions may be decoupled early in development, but they become coupled in the course of development.

The role of category labels in category learning

Often times, children learn the category and the label, lexicalising this category, concurrently. This happens, for example, when a new object is shown to the child and is labelled (e.g. “look, a dax”). In this case, the child needs to figure out what are the other objects that are *daxes*, or learn the category of *dax*. In the lab studies, this type of learning is referred to as category learning by classification. First examples of categories are introduced and labelled. Then the participant needs to predict which of the novel items also belong to this category (i.e. have the same category label). Given that participants learn the category and the linguistic label at the same time, it is reasonable to ask: How does the label affect learning? And does this role change with development?

Several ideas have been proposed. Some have argued that from early in development, the category label is a category marker – an indicator that the items belong to the same category – and guides or supervises category learning (Gelman, 2003; Gelman & Markman, 1986; Waxman & Gelman, 2009; Waxman & Markow, 1995; Welder & Graham, 2001; Westermann & Mareschal, 2014). In contrast, others (Deng & Sloutsky, 2012, 2013, 2015a; Sloutsky & Fisher, 2004a; Sloutsky & Lo, 1999; Sloutsky, Lo, & Fisher, 2001) have argued that, at least early in development, labels are akin to other features of items, but their role may change in the course of development. How could these positions be tested and contrasted?

In an attempt to distinguish between labels being features and category markers, Yamauchi and Markman (1998, 2000) developed a paradigm potentially capable of settling the issue. The paradigm is based on the

following idea. Imagine two categories A (labelled “A”) and B (labelled “B”), each having five binary dimensions (e.g. Size: large vs. small, Colour: black vs. white, Shape: square vs. circle, Luminance: bright vs. dark, and Texture: smooth vs. rough). The prototype of Category A has all values denoted by “1” (i.e. “A”, 1, 1, 1, 1, 1) and the prototype of Category B has all values denoted by “0” (i.e. “B”, 0, 0, 0, 0, 0). There are two inter-related generalisation tasks – classification and induction. The goal of the classification task is to infer category membership (and hence the label) on the basis of presented features. For example, participants are presented with all the values for an item (e.g. ?, 0, 1, 1, 1, 1) and have to predict category label “A” or “B”. In contrast, the goal of the induction task is to infer a feature on the basis of category label and other presented features. For example, given an item (e.g. “A”, 1, ?, 1, 0, 1), participants have to predict the value of the missing feature. A critical manipulation that could illuminate the role of labels is the “low-match” condition. For low-match induction, participants were presented with an item “A”, ?, 0, 1, 0, 0 (which had the label of A, but more features in common with the prototype of Category B) and asked to infer the missing feature. For low-match classification, participants were presented with an item “?”, 1, 0, 1, 0, 0 (which again had more features in common with the prototype of Category B) and asked to infer the missing label.

These researchers argued that if the label is just a feature then performance on the classification and the induction tasks should be symmetrical. However, if the label is more than a feature and serve as a category marker, then inferring a label when features are provided (i.e. a classification task) should elicit different performance from a task of inferring a feature when the label is provided (i.e. an induction task). Specifically, category-consistent responding should be more likely in induction tasks (where participants could rely on the

category label) than in classification tasks (where participants had to infer the category label). This asymmetry should be particularly evident in the critical low-match condition: in the low-match classification task (when they predict the label and thus cannot rely on it), participants would be likely to identify low-match items (e.g. “?”, 1, 0, 1, 0, 0) as belonging to category B (because these items have more features in common with prototype B), whereas in the low-match induction (when they can rely on the label), participants would be likely to identify low-match items (e.g. “A”, ?, 0, 1, 0, 0) as belonging to category A.

Upon finding predicted asymmetries between the two conditions, these researchers concluded that category labels differed from other features in that adult participants were more likely to treat labels as category markers rather than as features. These findings have been replicated in a series of follow-up studies (Yamauchi & Yu, 2008; Yamauchi, Kohn, & Yu, 2007; see also Markman & Ross, 2003, for a review). However, when Deng and Sloutsky (2013) extended this paradigm to children, they found symmetric performance: regardless

of the condition (i.e. classification or induction), young children relied on multiple features rather than on the label (see Figure 3). It was concluded, therefore, that early in development category labels may function as features of objects, but they become more than features in the course of development.

This shift in the role of labels is related to the protracted development of selective attention discussed above: For a label to be used as a category marker, people should be able to selectively attend to relevant information and ignore irrelevant information. They should also have enough experience to realise that labels have higher cue validity than other features: even considering homonyms (which reduce cue validity), the probability that X belongs to category K, given that it has label “K” is very high.

This transition in the role of label is important because once a label denotes a category, much information about the category can be accumulated through conversations, reading, media, and education. This information can also be merged with what is accumulated through observation. As a result, linguistic labels may become “knowledge hubs” that afford non-trivial inferences that are impossible through observation, such as “cows and dolphins are mammals” or “plants and animals are alive.” Having these “knowledge hubs” in place is critical for semantic development – learning information about categories and using this information for linking the concepts together and forming conceptual networks and hierarchies. We focus on these issues in the next section.

Semantic development: from early concepts to conceptual networks and taxonomies

Language is not a necessary aspect of category learning: nonhuman animals and preverbal human infants can learn categories (Eimas & Quinn, 1994; Lazareva & Wasserman, 2008; Madole & Oakes, 1999; Smith et al., 2012, 2015; Younger & Cohen, 1985). However, lexicalisation of categories, or learning words for categories, is a critical step in integrating knowledge about objects, people, and events, and this integrated knowledge – semantic memory – is central to our ability to use concepts for planning, prediction, explanation, reasoning, and decision making. Furthermore, as discussed above, words may be a starting point for learning many new categories. For example, unobservable categories (such as *germs*, *heat*, or *energy*) can be only learned that way.

Importantly, regardless of how a concept is learned (i.e. from a category to lexicalisation or from a lexical entry to a category), words for categories eventually become part of category representation and they help connecting what is known about a given category and

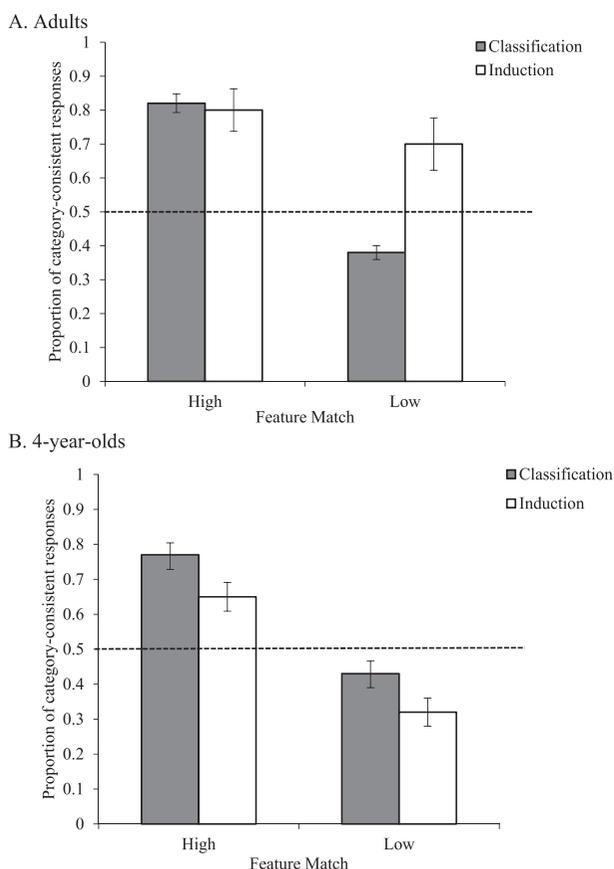


Figure 3. Proportion of category-consistent responses by feature match and testing condition (After Deng & Sloutsky, 2013, Experiment 1).

Note: Error bars represent standard error of the mean.

integrating it with newly learned information. For example, when learning a category (e.g. *cat*), some information can be acquired from observing the cats (e.g. shape, texture, and the pattern of locomotion), whereas other information cannot (e.g. the fact that cats have hearts, brains, and other internal organs). This latter information requires a verbal description and lexicalisation is critical here: the sentence *cats have brains* can only be expressed if one knows words for the cat and the brain. For the same reason, words are also important for forming conceptual hierarchies, such as *grey hound* → *dog* → *mammal* → *animal* → *living thing* → *thing*. All this is a product of development and in this section, we discuss how this development may occur.

From words to categories: learning words from context

Between birth and adulthood, a typical English-speaking child learns on average about 8–10 words per day: learning new words starts slowly, but accelerates dramatically during the second year of life (Bloom, 1973). Obviously, some of these words are learned by ostension (i.e. someone points to a putative referent and explicitly labels it), but many (if not most) are learned from context, including conversations and reading (see Goodman, McDonough, & Brown, 1998; Nagy, Herman, & Anderson, 1985), sometimes without a referent being present. When this happens, it is important to ask: How do people infer meanings of words, without having the referent present?

A number of ideas have been proposed to answer this question. Some have argued that the context provides the learner with syntactic, semantic, or social cues to meaning. Another possibility is that the context provides the learner with a rich network of associations that may help figuring out the meaning of the new word. Note that these ideas are not mutually exclusive and different sources may mutually strengthen each other.

One idea is that children use syntactic cues to disambiguate the meaning of a novel word. The idea goes back to Roger Brown (1957) who elegantly demonstrated that when presented with a novel word, 3–4 year-olds used the syntactic frame (e.g. “this is a sib” vs. “this one is sibbing”) to determine whether the word referred to an object, action, or property. These ideas generated much empirical support (see Bloom, 2000, for an extensive review). For example, Soja (1992) found that participants could use count noun/mass noun syntax to guide their learning of novel nouns. Syntactic cues have also been shown to facilitate learning of new verbs, the process known as syntactic bootstrapping (e.g. Fisher,

1996; Gleitman, 1990; Landau & Gleitman, 1985; Naigles, 1990). Another idea is that semantic cues can assist word learning. For example, Goodman et al. (1998) demonstrated that 2-year-olds could learn novel words when given sufficient semantic cues, such as familiar verbs (e.g. “Mommy feeds the ferret”) suggesting possible meanings of a novel noun. However, it is easy to see that syntactic and semantic cues provide only a rough guide as to what the meaning of the new word might be.

The third idea is that word learning could be assisted by the “theory of mind.” In particular, solving a social problem of what the speaker is calling attention to may guide word learning (e.g. Akhtar & Tomasello, 2000). This social problem could be solved by focusing on the speaker’s gaze direction, facial expression (Baldwin, 1991 1993), or on some relevant aspects of a social situation (Akhtar, 2002). For example, Akhtar (2002) presented 2-year-olds with a word learning task. In one condition, participants’ attention was attracted to the texture of objects (“this is a smooth one and this is a fuzzy one”), and in another condition their attention was attracted to the shape of the objects. Participants were then shown a triad of novel objects, and one of the novel objects was labelled (“this is a dacky one”). The remaining objects matched the labelled object either in shape or in texture. The results indicated that the context affected inferred meaning of the new adjective – participants were more likely to infer that “dacky” was a shape word in the shape-relevant context. However, while social problem solving may offer some assistance in figuring out the meaning, its assistance is rather limited when entities have multiple feature dimensions. For example, what would be the set of relevant features to communicate that the word *cat* refers to all cats and only to cats and how could these features be communicated?

Finally, a more recent proposal (Sloutsky, Yim, Yao, & Dennis, 2017) suggests that the context in which words are presented provides associative cues that trigger a candidate meaning of a novel word. Two types of associative cues are of particular importance – syntagmatic and paradigmatic (Brown & Berko, 1960; Dennis, 2005; Ervin-Tripp, 1970; Nelson, 1977).

Syntagmatic associations refer to words that co-occur in close temporal proximity (e.g. The *dog* was *barking* at the car). Words associated syntagmatically tend to be thematically related. Paradigmatic associations refer to words playing the same role in sentences and appearing in similar sentential contexts (e.g. “He went home to feed the *cat*” and “She drove home to feed the *dog*”). Words associated paradigmatically also tend to be taxonomically related.

While syntagmatic associations exhibit an early onset, paradigmatic associations tend to emerge later in development appearing around 6-years of age (McNeill, 1963; Nelson, 1977). Under this construal, early in development (i.e. before paradigmatic associations come online), only syntagmatic associations provide cues to the meaning of a novel word. At the same time, later in development, both syntagmatic and paradigmatic associations provide cues to novel words. If this is the case, then early in development syntagmatic associates should be better cues than paradigmatic ones.

To illustrate, imagine that the learner is presented by a sequence "... , furry, dax" or by a sequence "... , cat, dax". When syntagmatic associations predominate (which is believed to be the case for young children), the word *furry* is associated with the word *animal* (strengths of such associations can be measured by a free association task (Nelson, McEvoy, & Schreiber, 2004) allowing to calculate the probability that the word *animal* will be recalled in a free association task, given the word *furry*). In this case, *furry* will activate the word *animal*, and this may result in *dax* being interpreted as a kind of animal. At the same time, if paradigmatic associations are not formed yet, the word *cat* will not be associated with the word *animal*. As a result, the word *furry* will be a better cue to the meaning of *dax* (i.e. suggesting that *dax* is an animal) than the word *cat*.

In the case of emergence of paradigmatic associations (which we believe are a product of development), the word *furry* remains associated with the word *animal* syntagmatically, whereas the word *cat* become associated with the word *animal* paradigmatically. As a result, both words may activate the word *animal*: *furry* via the syntagmatic route, whereas *cat* via the paradigmatic route. Under this construal, the probability that the word *dax* would be considered as referring to some sort of an animal, would be proportional to the forward associative strength between the words accompanying *dax* and the word *animal* (or, perhaps, features of animacy or the set of animals). The initial meaning of a word learned from context is broad and imprecise and additional information would be needed to zero in on a precise meaning of the word. This initially broad meaning may become more precise with additional experience with the word.

If syntagmatic and paradigmatic associations play a role in inferring meanings of novel words, young children whose associative repertoire is limited to syntagmatic associations (Nelson, 1977) should learn words from contexts that involve syntagmatic associations. At the same time, adults, whose repertoire includes both syntagmatic and paradigmatic associations should learn words from contexts that have associations of

either type. These predictions have been supported in new research coming out of our lab (Sloutsky et al., 2017). Specifically, in a series of experiments, Sloutsky et al. presented 4-year-olds and adults with sets of words that included a single nonsense word (e.g. *dax*) and asked them to indicate whether the nonsense word was an animal or an artifact. Across experiments, adults reliably identified the appropriate category of the nonsense word when lists contained associatively (all members of the list were semantic associates of the word *animal*, e.g. *zoo*, *farm*, *furry*, *creature*, *giraffe*, *hamster*, *bear*, and *feeding*) or taxonomically (all members of the list were animals referred to by count nouns, e.g. *cat*, *dog*, *fish*, *bird*, *horse*, *squirrel*, *cow*, and *rabbit*) related items, whereas children could only identify the appropriate category when lists contained associatively related items. Furthermore, a computational model developed by these researchers indicated that only the syntagmatic network initially affected the model performance, which was sufficient to account for the child data. In contrast, to capture the adult data, additional learning in the paradigmatic network was needed, which developed later in training. These results suggest a syntagmatic to paradigmatic shift in development and provide a mechanistic account for the shift: while word co-occurrence (which appears in linguistic environment early in development) gives rise to syntagmatic associations, experience with language (which accumulates with development) gives rise to paradigmatic associations.

Linking words and categories: the development of semantic knowledge

With the dramatic growth in language, children start to learn words for both known and unknown categories. There is much evidence that these emerging concepts (as well as fully developed ones) get organised into some form of semantic (or conceptual) network. This evidence includes studies of semantic priming, the development and decline of semantic memory, as well as of property and picture verification tasks (see Rogers & McClelland, 2004, for a review). In short, language allows the development of a conceptual network (also referred to as semantic knowledge or semantic memory) that represents one's knowledge about the world.

In subsequent studies, researchers attempted to examine whether priming effects in adults stem from thematic relations (or syntagmatic associations) or from taxonomic relations (or paradigmatic associations). These researchers found evidence for priming effects due taxonomic relations in the absence of thematic relations (e.g. Ferrand & New, 2003; Thompson-Schill,

Kurtz, & Gabrieli, 1998) and for priming effects due thematic relations in the absence of taxonomic relations (e.g. Ferrand & New, 2003). However, priming effects were more reliable for words that were both taxonomically and thematically related (e.g. dog and cat) than for words that were related only taxonomically (e.g. dog and cow) or only thematically (e.g. dog and bone) (McRae & Boisvert, 1998; Moss, Ostrin, Tyler, & Marslen-Wilson, 1995; Perea & Rosa, 2002).

Although considerable progress has been made in characterising the structure of adults' semantic knowledge, origins of this knowledge and how it develops remain unclear. Findings from studies of word associations in children are controversial. Some researchers argue that the ability to represent semantic relations transpires very early in development, with infants as young as 24 months of age exhibiting evidence of semantic priming. Others argue for protracted development, presenting evidence of semantic (or taxonomic) sensitivity emerging during elementary school years.

For example, Arias-Trejo and Plunkett (2009) used an intermodal preferential looking paradigm to examine 18- and 21-month-old infants' responses to related prime-target pairs such as cat-dog, compared with unrelated pairs such as plate-dog. The related pairs were strongly associated according to adult associative norms. Infants first heard a phrase such as "I saw a cat", followed by a target word (dog). They then concurrently saw two images, one related (e.g. a dog) and one unrelated (e.g. a door). The critical manipulation was whether the initial phrase contained a prime word related to the target word and picture (e.g. "I saw a cat ... dog") or an unrelated word (e.g. "I saw a swing ... dog"). The researchers found that 18-month-olds looked significantly longer to the picture named by the target word, regardless of whether it followed a related or an unrelated prime phrase. In contrast, 21-month-olds looked significantly longer to the named picture in the related prime-target condition ("I saw a cat ... dog") but not in the unrelated prime-target condition ("I saw a swing ... dog"). The older infants' sensitivity to the word relatedness provides evidence for lexical organisation in infants by 21 months of age.

To further identify the type of lexical relatedness, Arias-Trejo and Plunkett (2013) used the same paradigm with 21- and 24-month-olds, with prime-target word pairs that were either thematically or taxonomically related or were unrelated. Twenty-four-month-olds, but not 21-month-olds, exhibited a priming effect for words that were either thematically or taxonomically related, and the age-related differences suggest that semantic relationships between words developed between 21 and 24 months of age. Similarly, Willits,

Wojcik, Seidenberg, and Saffran (2013) also provide convergent evidence showing that, even in the absence of visual referents, infants by 24 months of age are able to represent semantic relations between words when processing language.

However, a recent study by Unger, Fisher, Nugent, Ventura, and MacLellan (2016) found that even preschoolers were unable to represent relations that were only taxonomic: these participants appear to recognise only links between concepts that are related along multiple dimensions (both thematic and taxonomic). Older children increasingly recognise links between concepts that are related along one dimension (either thematic or taxonomic), and starting at approximately the age of second grade, taxonomic relations are prioritised over thematic relations. This finding is in contrast to the infant literature where semantic priming is found in 24-month infants, which could be possibly due to the methodological differences between studies with infants and with older children (see also Sloutsky, Deng, Fisher, & Kloos, 2015, for related evidence).

Although the precise structuring principles of children's semantic network remains unclear, most researchers would agree that there is a dramatic development in children's semantic knowledge: children learn links between words and categories and form connections among concepts. Importantly, the development of semantic knowledge is critical and fundamental to the development of conceptual networks and hierarchies.

Development of conceptual hierarchies

One critical step in conceptual development is establishing a structure for concepts. Among various possible conceptual structures (Kemp, Shafto, & Tenenbaum, 2012), taxonomic hierarchy is the most general and well-studied one. An example of taxonomic hierarchy is grey hound → dog → mammal → animal → living thing → thing. This kind of hierarchy is based on class-inclusion relations, with lower-level categories being exhaustive with respect to a higher-level category, and sub-classes of higher-level category being mutually exclusive, with no common members in it.

In order to form a taxonomic hierarchy of concepts, some researchers (Inhelder & Piaget, 1964) argued that the ability to understand the class-inclusion relations, or the logical constraints, is of critical importance. The idea of the logic of classes is that multidimensional sets of stimuli can be divided into proper subsets focusing on one dimension at a time. As proposed by Inhelder and Piaget (1964), the development of conceptual hierarchies is a function of the development of the logic of classes. According to this view, fundamental

developmental changes occur with respect to understanding of class-inclusion relations, which is closely tied to the ability of understanding quantifiers, such as *all*, *some*, *some are not*, and *none*. Once these are mastered, a classification scheme based on these relations can be applied to any domain of knowledge. However, logic alone is not sufficient for building the taxonomic hierarchies of concepts. One also needs to know dimensions that distinguish sub-categories, to learn words denoting different level of classes, and to acquire knowledge of a domain in which a taxonomy is to be built. Therefore, most contemporary theories consider domain knowledge as a necessary component of the development of conceptual hierarchies (e.g. Carey, 1985; Chi, Hutchinson, & Robin, 1989; Inagaki & Hatano, 2002; Keil, 1981) and argue that a taxonomic hierarchy of concepts may result from knowledge of a domain. According to this view, knowledge of a domain reveals class-inclusion relations, which gives rise to one's hierarchical representations of concepts. For example, knowledge of mammals may help a child understand that all dogs are mammals, but not all mammals are dogs. In this case, the child does not necessarily need to understand logic and apply class-inclusion relations.

An alternative view, has been proposed by Rogers and McClelland (2004), according to which conceptual hierarchies can be formed implicitly, on the basis of shared predicates. For example, *salmon* and *trout* share most of the predicates (e.g. can swim, had gills, has scales, has skin, has bones, etc.), whereas *salmon* and *eagle* share only some of the predicates (e.g. has skin, has bones). Therefore, *salmon* and *eagle* are members of a broader class (i.e. vertebrates) than *salmon* and *trout*. Although this is a promising approach, it runs into some obvious difficulties. First, in order to bypass the problem of class inclusion, one need to have a representative sample of predicates associated with a given category. However, given vast differences in experience, there is no guarantee that this would happen. As a result, there may be more individual variability in conceptual hierarchies than is currently observed.

Another potential problem is that simply adding the predicates may result in a wrong taxonomy. For example, animals sharing the habitat (e.g. *dolphin* and *tuna* or *bat* and *hawk*) may share more predicates with members of a broader class than with members of a narrow class (e.g. *dolphin* and *cow*), yet people eventually learn to classify such animal correctly.

In sum, although people often behave as if they have conceptual hierarchies, it is not fully understood *if*, *when*, and *how* such conceptual hierarchies are developed. Evidence for the early onset of conceptual hierarchies is limited. Even if a child exhibits the ability to classify

items at a superordinate level or draws inductive inferences on the basis of a superordinate class, this ability does not necessarily indicate the presence of a conceptual hierarchy – because the child can rely on similarity. Although it is clear that the ability to understand class-inclusion relations (along with the ability to use quantifiers) and the knowledge of how these relations can be applied in a particular domain can greatly contribute to the development of conceptual hierarchies, it remains controversial as to whether the former is really necessary for such development (cf. Rogers & McClelland, 2004).

Conclusions

Conceptual development supports many uniquely human behaviours ranging from accumulation of knowledge that is not directly observable to multiple ways of using this knowledge. Importantly, conceptual development has humble origins – it is based on the ability to form categories, the ability that humans share with many non-human animals. However, this ability is greatly amplified (if not transformed) by language: (a) lexicalisation helps turning categories into knowledge hubs as well as to mark to-be-learned categories and (b) language is an important source of knowledge about the concepts. Knowledge acquired through perceptual experience coupled with knowledge acquired through language (including reading) leads to the formation of conceptual networks and hierarchies.

However, much is unknown about each constituent part of conceptual development (i.e. the development of category learning, lexicalisation, and semantic organisation) as well as about their interaction. Therefore, a challenge for future research is to establish precise details of conceptual development, including the way the constituent components change and interact in the course of development.

Acknowledgements

We thank members of the Cognitive Development Lab for helpful comments.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Institute for Educational Sciences (IES) [grant number R305A140214] and National Institutes of Health (NIH) [grant number R01HD078545] and [grant number P01HD080679] to Vladimir Sloutsky.

References

- Akhtar, N. (2002). Relevance and early word learning. *Journal of Child Language*, 29, 677–686. doi:10.1017/S0305000902005214
- Akhtar, N., & Tomasello, M. (2000). The social nature of words and word learning. In R. M. Golinkoff, & K. Hirsh-Pasek (Eds.), *Becoming a word learner: A debate on lexical acquisition* (pp. 115–135). Oxford: Oxford University Press.
- Arias-Trejo, N., & Plunkett, K. (2009). Lexical-semantic priming effects during infancy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1536), 3633–3647. doi:10.1098/rstb.2009.0146
- Arias-Trejo, N., & Plunkett, K. (2013). What's in a link: Associative and taxonomic priming effects in the infant lexicon. *Cognition*, 128, 214–227. doi:10.1016/j.cognition.2013.03.008
- Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development*, 62, 874–890. doi:10.1111/j.1467-8624.1991.tb01577.x
- Baldwin, D. A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology*, 29, 832–843. doi:10.1037/0012-1649.29.5.832
- Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences*, 22, 577–609. doi:10.1017/S0140525X99002149
- Barton, M. E., & Komatsu, L. K. (1989). Defining features of natural kinds and artifacts. *Journal of Psycholinguistic Research*, 18, 433–447. doi:10.1007/BF01067309
- Best, C. A., Yim, H., & Sloutsky, V. M. (2013). The cost of selective attention in category learning: Developmental differences between adults and infants. *Journal of Experimental Child Psychology*, 116, 105–119. doi:10.1016/j.jecp.2013.05.002
- Bhatt, R. S., & Quinn, P. C. (2010). How does learning impact development in infancy? The case of perceptual organization. *Infancy*, 16, 2–38. doi:10.1111/j.1532-7078.2010.00048.x
- Bloom, L. (1973). *One word at a time: The use of single word utterances before syntax*. The Hague: Mouton.
- Bloom, P. (2000). *How children learn the meanings of words*. Cambridge, MA: MIT Press.
- Brown, R. (1957). Linguistic determinism and the part of speech. *The Journal of Abnormal and Social Psychology*, 55, 1–5. doi:10.1037/h0041199
- Brown, R., & Berko, J. (1960). Word association and the acquisition of grammar. *Child Development*, 31, 1–14. doi:10.2307/1126377
- Carey, S. (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.
- Chi, M. T. H., Hutchinson, J. E., & Robin, A. F. (1989). How inferences about novel domain-related concepts can be constrained by structured knowledge. *Merrill-Palmer Quarterly*, 35, 27–62.
- Deng, W., & Sloutsky, V. M. (2012). Carrot eaters or moving heads: Inductive inference is better supported by salient features than by category labels. *Psychological Science*, 23, 178–186. doi:10.1177/0956797611429133
- Deng, W., & Sloutsky, V. M. (2013). The role of linguistic labels in inductive generalization. *Journal of Experimental Child Psychology*, 114, 432–455. doi:10.1016/j.jecp.2012.10.011
- Deng, W., & Sloutsky, V. M. (2015a). Linguistic labels, dynamic visual features, and attention in infant category learning. *Journal of Experimental Child Psychology*, 134, 62–77. doi:10.1016/j.jecp.2015.01.012
- Deng, W., & Sloutsky, V. M. (2015b). The development of categorization: Effects of classification and inference training on category representation. *Developmental Psychology*, 51, 392–405. doi:10.1037/a0038749
- Dennis, S. (2005). A memory-based theory of verbal cognition. *Cognitive Science*, 29, 145–193. doi:10.1207/s15516709cog0000_9
- Eimas, P. D. (1994). Categorization in early infancy and the continuity of development. *Cognition*, 50, 83–93. doi:10.1016/0010-0277(94)90022-1
- Eimas, P. D., & Quinn, P. C. (1994). Studies on the formation of perceptually based basic-level categories in young infants. *Child Development*, 65, 903–917. doi:10.1111/j.1467-8624.1994.tb00792.x
- Ervin-Tripp, S. M. (1970). Substitution, context and association. In L. Postman & G. Keppel (Eds.), *Norms of word association* (pp. 383–467). New York: Academic Press.
- Ferrand, L., & New, B. (2003). Semantic and associative priming in the mental lexicon. In P. Bonin (Ed.), *Mental lexicon: Some words to talk about words* (pp. 25–43). Hauppauge, NY: Nova Science Publisher.
- Fisher, C. (1996). Structural limits on verb mapping: The role of analogy in children's interpretations of sentences. *Cognitive Psychology*, 31, 41–81. doi:10.1006/cogp.1996.0012
- Fodor, J. (1972). Some reflections on LS Vygotsky's "Thought and Language". *Cognition*, 1, 83–95. doi:10.1016/0010-0277(72)90046-7
- Gelman, S. A. (2003). *The essential child: Origins of essentialism in everyday thought*. New York: Oxford University Press.
- Gelman, S. A., & Markman, E. (1986). Categories and induction in young children. *Cognition*, 23, 183–209. doi:10.1016/0010-0277(86)90034-X
- Gentner, D. (1981). Some interesting differences between nouns and verbs. *Cognition and Brain Theory*, 4, 161–178.
- Gentner, D., & Kurtz, K. J. (2005). Relational categories. In W. K. Ahn, R. L. Goldstone, B. C. Love, A. B. Markman, & P. W. Wolff (Eds.), *Categorization inside and outside the lab* (pp. 151–175). Washington, DC: APA.
- Gleitman, L. R. (1990). The structural sources of verb meanings. *Language Acquisition*, 1, 3–55. doi:10.1207/s15327817la0101_2
- Goodman, J. C., McDonough, L., & Brown, N. B. (1998). The role of semantic context and memory in the acquisition of novel nouns. *Child Development*, 69, 1330–1344. doi:10.1111/j.1467-8624.1998.tb06215.x
- Hanania, R., & Smith, L. B. (2010). Selective attention and attention switching: Towards a unified developmental approach. *Developmental Science*, 13, 622–635. doi:10.1111/j.1467-7687.2009.00921.x
- Homa, D., Rhoades, D., & Chambliss, D. (1979). Evolution of conceptual structure. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 11–23. doi:10.1037/0278-7393.5.1.11
- Inagaki, K., & Hatano, G. (2002). *Young children's naive thinking about biological world*. New York, NY: Psychology Press.
- Inhelder, B., & Piaget, J. (1964). *The early growth of logic in the child: Classification and seriation*. New York: Harper and Row.
- James, W. (1983/1890). *The principles of psychology (originally published in 1890)*. Cambridge, MA: Harvard University Press.

- Keil, F. C. (1981). Constraints on knowledge and cognitive development. *Psychological Review*, 88, 197–227. doi:10.1037/0033-295X.88.3.197
- Kemp, C., Shafto, P., & Tenenbaum, J. B. (2012). An integrated account of generalization across objects and features. *Cognitive Psychology*, 64, 35–73. doi:10.1016/j.cogpsych.2011.10.001
- Kloos, H., & Sloutsky, V. M. (2008). What's behind different kinds of kinds: Effects of statistical density on learning and representation of categories. *Journal of Experimental Psychology: General*, 137, 52–72. doi:10.1037/0096-3445.137.1.52
- Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, 99, 22–44. doi:10.1037/0033-295X.99.1.22
- Landau, B., & Gleitman, L. R. (1985). *Language and experience: Evidence from the blind child*. Cambridge, MA: Harvard University Press.
- Lane, D. M., & Pearson, D. A. (1982). The development of selective attention. *Merrill-Palmer Quarterly*, 28, 317–337.
- Lazareva, O. F., & Wasserman, E. A. (2008). Categories and concepts in animals. In R. Menzel, & J. Byrne (Eds.), *Learning theory and behavior*. Vol. [1] of *learning and memory: A comprehensive reference* (Vol. 4, pp. 197–226). Oxford: Elsevier.
- Love, B. C., Medin, D. L., & Gureckis, T. M. (2004). SUSTAIN: A network model of category learning. *Psychological Review*, 111, 309–332. doi:10.1037/0033-295X.111.2.309
- Madole, K. L., & Oakes, L. M. (1999). Making sense of infant categorization: Stable processes and changing representations. *Developmental Review*, 19, 263–296. doi:10.1006/drev.1998.0481
- Markman, A. B., & Ross, B. H. (2003). Category use and category learning. *Psychological Bulletin*, 129, 592–613. doi:10.1037/0033-2909.129.4.592
- McNeill, D. (1963). The origin of associations within the same grammatical class. *Journal of Verbal Learning and Verbal Behavior*, 2, 250–262. doi:10.1016/S0022-5371(63)80091-2
- McRae, K., & Boisvert, S. (1998). Automatic semantic similarity priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24, 558–572. doi:10.1037/0278-7393.24.3.558
- Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology*, 51, 121–147. doi:10.1146/annurev.psych.51.1.121
- Moss, H. E., Ostrin, R. K., Tyler, L. K., & Marslen-Wilson, W. D. (1995). Accessing different types of lexical semantic information: Evidence from priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21, 863–883. doi:10.1037/0278-7393.21.4.863
- Nagy, W. E., Herman, P. A., & Anderson, R. C. (1985). Learning words from context. *Reading Research Quarterly*, 20, 233–253. doi:10.2307/747758
- Naigles, L. (1990). Children use syntax to learn verb meanings. *Journal of Child Language*, 17, 357–374. doi:10.1017/S0305000900013817
- Nelson, D. L., McEvoy, C. L., & Schreiber, T. A. (2004). The University of South Florida free association, rhyme, and word fragment norms. *Behavior Research Methods, Instruments, & Computers*, 36, 402–407. doi:10.3758/BF03195588
- Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: A review of research and theory. *Psychological Bulletin*, 84, 93–116. doi:10.1037/0033-2909.84.1.93
- Nosofsky, R. M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, 115, 39–57. doi:10.1037/0096-3445.115.1.39
- Perea, M., & Rosa, E. (2002). The effects of associative and semantic priming in the lexical decision task. *Psychological Research*, 66, 180–194. doi:10.1007/s00426-002-0086-5
- Plebanek, D. J., & Sloutsky, V. M. (2017). Costs of selective attention: When children notice what adults miss. *Psychological Science*, 28, 723–732. doi:10.1177/0956797617693005
- Plude, D. J., Enns, J. T., & Brodeur, D. (1994). The development of selective attention: A life-span overview. *Acta Psychologica*, 86, 227–272. doi:10.1016/0001-6918(94)90004-3
- Rogers, T. T., & McClelland, J. L. (2004). *Semantic cognition: A parallel distributed processing approach*. Cambridge, MA: MIT Press.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology*, 7, 573–605. doi:10.1016/0010-0285(75)90024-9
- Rouder, J. N., & Ratcliff, R. (2004). Comparing categorization models. *Journal of Experimental Psychology: General*, 133, 63–82. doi:10.1037/0096-3445.133.1.63
- Shepard, R. N., Hovland, C. I., & Jenkins, H. M. (1961). Learning and memorization of classifications. *Psychological Monographs: General and Applied*, 75, 1–42. doi:10.1037/h0093825
- Sloutsky, V. M. (2010). From perceptual categories to concepts: What develops? *Cognitive Science*, 34, 1244–1286. doi:10.1111/j.1551-6709.2010.01129.x
- Sloutsky, V. M., Deng, W., Fisher, A. V., & Kloos, H. (2015). Conceptual influences on induction: A case for a late onset. *Cognitive Psychology*, 82, 1–31. doi:10.1016/j.cogpsych.2015.08.005
- Sloutsky, V. M., & Fisher, A. V. (2004a). Induction and categorization in young children: A similarity-based model. *Journal of Experimental Psychology: General*, 133, 166–188. doi:10.1037/0096-3445.133.2.166
- Sloutsky, V. M., & Fisher, A. V. (2004b). When development and learning decrease memory: Evidence against category-based induction in children. *Psychological Science*, 15, 553–558. doi:10.1111/j.0956-7976.2004.00718.x
- Sloutsky, V. M., & Lo, Y.-F. (1999). How much does a shared name make things similar? Part 1. Linguistic labels and the development of similarity judgment. *Developmental Psychology*, 35, 1478–1492. doi:10.1037/0012-1649.35.6.1478
- Sloutsky, V. M., Lo, Y.-F., & Fisher, A. V. (2001). How much does a shared name make things similar? Linguistic labels, similarity, and the development of inductive inference. *Child Development*, 72, 1695–1709. doi:10.1111/1467-8624.00373
- Sloutsky, V. M., Yim, H., Yao, X., & Dennis, S. (2017). An associative account of the development of word learning. *Cognitive Psychology*, 97, 1–30. doi:10.1016/j.cogpsych.2017.06.001
- Smith, J. D., Berg, M. E., Cook, R. G., Murphy, M. S., Crossley, M. J., Boomer, J., ... Grace, R. C. (2012). Implicit and explicit categorization: A tale of four species. *Neuroscience and Biobehavioral Reviews*, 36, 2355–2369. doi:10.1016/j.neubiorev.2012.09.003
- Smith, J. D., & Minda, J. P. (2000). Thirty categorization results in search of a model. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 3–27. doi:10.1037/0278-7393.26.1.3
- Smith, J. D., Zakrzewski, A. C., Johnston, J. J. R., Roeder, J. L., Boomer, J., Ashby, F. G., & Church, B. A. (2015).

- Generalization of category knowledge and dimensional categorization in humans (*Homo sapiens*) and nonhuman primates (*Macaca mulatta*). *Journal of Experimental Psychology: Animal Learning and Cognition*, 41, 322–335. doi:10.1037/xan0000071
- Soja, N. N. (1992). Inferences about the meanings of nouns: The relationship between perception and syntax. *Cognitive Development*, 7, 29–45. doi:10.1016/0885-2014(92)90003-A
- Thompson-Schill, S. L., Kurtz, K. J., & Gabrieli, J. D. E. (1998). Effects of semantic and associative relatedness on automatic priming. *Journal of Memory and Language*, 38, 440–458. doi:10.1006/jmla.1997.2559
- Unger, L., Fisher, A. V., Nugent, R., Ventura, S. L., & MacLellan, C. J. (2016). Developmental changes in semantic knowledge organization. *Journal of Experimental Child Psychology*, 146, 202–222. doi:10.1016/j.jecp.2016.01.005
- Waxman, S. R., & Gelman, S. A. (2009). Early word-learning entails reference, not merely associations. *Trends in Cognitive Sciences*, 13, 258–263. doi:10.1016/j.tics.2009.03.006
- Waxman, S. R., & Markow, D. B. (1995). Words as invitations to form categories: Evidence from 12-to 13-month-old infants. *Cognitive Psychology*, 29, 257–302. doi:10.1006/cogp.1995.1016
- Welder, A. N., & Graham, S. A. (2001). The influences of shape similarity and shared labels on infants' inductive inferences about nonobvious object properties. *Child Development*, 72, 1653–1673. doi:10.1111/1467-8624.00371
- Westermann, G., & Mareschal, D. (2014). From perceptual to language-mediated categorization. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369, 20120391. doi:10.1098/rstb.2012.0391
- Willits, J. A., Wojcik, E. H., Seidenberg, M. S., & Saffran, J. R. (2013). Toddlers activate lexical semantic knowledge in the absence of visual referents: Evidence from auditory priming. *Infancy*, 18, 1053–1075. doi:10.1111/infa.12026
- Yamauchi, T., Kohn, N., & Yu, N. Y. (2007). Tracking mouse movement in feature inference: Category labels are different from feature labels. *Memory & Cognition*, 35, 852–863. doi:10.3758/BF03193460
- Yamauchi, T., & Markman, A. B. (1998). Category learning by inference and classification. *Journal of Memory and Language*, 39, 124–148. doi:10.1006/jmla.1998.2566
- Yamauchi, T., & Markman, A. B. (2000). Inference using categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 26, 776–795. doi:10.1037/0278-7393.26.3.776
- Yamauchi, T., & Yu, N. (2008). Category labels versus feature labels: Category labels polarize inferential predictions. *Memory & Cognition*, 36, 544–553. doi:10.3758/MC.36.3.544
- Younger, B. A., & Cohen, L. B. (1985). How infants form categories. *Psychology of Learning and Motivation*, 19, 211–247. doi:10.1016/S0079-7421(08)60528-9
- Yu, C., & Smith, L. B. (2012). Embodied attention and word learning by toddlers. *Cognition*, 125, 244–262. doi:10.1016/j.cognition.2012.06.016