Journal of Experimental Child Psychology 234 (2023) 105704

Contents lists available at ScienceDirect

Journal of Experimental Child Psychology

journal homepage: www.elsevier.com/locate/jecp

Children's generalization of novel names in comparison settings: The role of semantic distance during learning and at test

Jean-Pierre Thibaut *, Arnaud Witt

Laboratory for Research on Learning and Development (LEAD), National Centre for Scientific Research (CNRS), UMR5022, Université de Bourgogne, 21065 Dijon Cedex, France

A R T I C L E   I N F O

Article history:
Received 31 December 2022
Revised 18 April 2023
Available online 7 June 2023

Keywords:
Cognitive development
Lexical development
Novel word learning
Object and relational nouns
Comparison design

A B S T R A C T

There is considerable evidence showing that, in novel noun learning and generalization tasks, comparisons of several learning stimuli lead to more taxonomically based generalizations of novel nouns than single stimulus presentations. This study investigated the role of semantic distance (close vs. far) between learning examples and between learning examples and transfer items (near vs. distant) in comparison designs. In two experiments, we investigated the case of object nouns (e.g., foods, Experiment 1) and relational nouns (e.g., is the cutter for, Experiment 2) in 4- to 6-year-old and 3- to 4-year-old children, respectively. As predicted, the comparison conditions led to better results than the no-comparison conditions. In comparison conditions, far training items and near generalization items gave the best performance. Semantic distance effects are discussed in terms of abstracted representations during learning as well as in terms of cognitive constraints on generalization. It is argued that both object nouns and relational nouns are construed in the light of the type of example used during learning (i.e., single or multiple). Depending on the distance between learning and generalization items, children build different categories and are more or less likely to accept distant referents.

* Corresponding author.
E-mail address: jean-pierre.thibaut@u-bourgogne.fr (J.-P. Thibaut).

https://doi.org/10.1016/j.jecp.2023.105704

Introduction

Both informal observations and controlled studies have long underlined that young children are fast and effective word learners (Carey, 2010; Carey & Bartlett, 1978; Heibeck & Markman, 1987; Markman, 1989). One remarkable feature of novel word learning is that it takes place following exposure to only a limited number of learning exemplars. Indeed, in many daily situations or experiments, one stimulus is introduced with its name and the child needs to generalize the new word to novel stimuli (e.g., Jones & Smith, 1993; Landau et al., 1988; Waxman, 1990). The challenge facing children therefore is to construct concepts for novel words that capture the dimensions that are relevant for the novel words while rejecting irrelevant dimensions (Gentner, 2010; Waxman et al., 2011).

To solve this challenge, it has been hypothesized that children *constrain novel word meaning* with lexical biases that lead them to select certain referents or favor certain possible meanings over others. Several biases have been described (Markman, 1989; Markman & Hutchinson, 1984; Waxman, 1990). One bias is the basic-level bias, according to which children spontaneously name a novel object at the basic level of categorization rather than at the other possible levels. For example, the word *apple* is generalized from one examplar of a Granny Smith apple to any other apple rather than to *Granny Smith apples* only or to any *fruit* (Markman, 1989; Waxman & Hatch, 1992; Waxman et al., 1991). The shape bias is another bias and refers to the generalization of novel object names to objects with the same shape as the learning object rather than to objects with the same color, texture, or size (Kucker et al., 2019). Several interpretations of these biases have been proposed (e.g., Diesendruck & Bloom, 2003; Jones & Smith, 1993; Landau et al., 1988; Markman, 1989). However, it is often thought that they enable children to find out which information is relevant in a given learning context (e.g., see discussion of the basic-level bias in Emberson et al., 2019; Murphy & Brownell, 1985).

However these biases are sometimes less appropriate or risk guiding the child toward incorrect information. For example, the basic-level bias is irrelevant when a subordinate level noun is required. In this respect, it has been shown that children may focus on the relevant categorization level when the communication situation and/or the available information are structured in a way that makes it more salient or relevant for communication (Waxman, 1990; Waxman & Hatch, 1992; Waxman et al., 1991). A priori, this differential focus on dimensions other than shape or any cognitively salient but semantically irrelevant dimension should be possible if the learning situation encourages the exploration of various dimensions. This is what we look at with reference to stimulus comparison designs.

The current study explored comparison designs in novel word learning. There is now extensive evidence suggesting that comparisons of two or more learning items can bias children toward more semantically based novel name generalizations (e.g., taxonomic choice rather than a shape choice) than single-stimulus learning situations (Hammer, 2015; Namy & Gentner, 2002; Thibaut, 1991; Tversky, 1977; see also Oakes et al., 2009, and Oakes & Ribar, 2005, in infants; Augier & Thibaut, 2013, Gentner & Namy, 1999, Graham et al., 2010, and Waxman & Klibanoff, 2000, in children; Gick & Holyoak, 1983, Lassaline & Murphy, 1998, and Markman & Gentner, 1993, in adults).

To illustrate, Gentner and Namy (1999) tested the extension of novel names in 4-year-old children with pictures of objects coming from familiar taxonomic categories (e.g., fruits). In the single-stimulus learning condition (no-comparison), a single stimulus was introduced as a *blicket*. Children extended the novel noun to a perceptually similar object (e.g., a balloon) rather than to a taxonomically related but perceptually dissimilar object (e.g., a banana). This preference was reversed in the comparison condition (two or four standards, e.g., an apple and an orange, also introduced as *blickets*). The single-stimulus results are consistent with a vast body of studies showing that young children generalize novel object names to perceptually similar objects, especially "shape-similar objects," in single-object designs (Jones & Smith, 1993; Landau et al., 1988). This type of forced-choice design is well-suited for studying factors biasing children's interpretations of novel names. Single-object learning designs have revealed important biases toward various dimensions. The comparison design presented here follows the same logic and focuses on how children might use additional information gained from comparing learning exemplars to extend novel names along dimensions other than those that would be chosen in a single design.

Comparison and semantic distance

The discussion above and the research conducted to date suggest that not all comparison conditions are equivalent. For example, providing a larger number of semantically consistent exemplars might lead to different proportions of semantic choices (Thibaut & Witt, 2015). In the current study, we explored another factor that could influence generalization performance, namely the semantic distance[1] between the learning items and the semantic distance between the learning items and the generalization items. Many previous studies have suggested that semantic distance affects novel noun generalization in both comparison and no-comparison contexts or in tasks where the alignment of semantic information is critical, such as analogy tasks.

One particular implementation of semantic distance is distance in a taxonomic hierarchy, as described by Rosch et al. (1978) and Rosch et al. (1976) (see Murphy, 2002, for a synthesis). Ever since Rosch et al. (1976), it has been generally accepted that both children and adults spontaneously categorize objects at what Rosch et al. called the basic level of categorization (e.g., dog, apple, chair) rather than at the subordinate level (e.g., Cox apple, poodle, kitchen chair) or at the superordinate level (e.g., mammal, fruit, furniture). Importantly, Rosch et al. showed that young children categorized two basic-level objects together (e.g., a dog with another type of dog) more easily than two objects from the same superordinate category (e.g., a dog with a cow or a fly). This result has been replicated many times (Mervis & Crisafi, 1982; Murphy, 2002). Here, we define semantic distance in terms of levels in the taxonomic hierarchy, that is, the distance between the learning items or the distance between the learning items and the generalization items separately. Members of the same basic-level category (e.g., two apples) are at a 0 distance in the taxonomy. Items from two different basic-level categories (e.g., banana and apple) are at a distance of 1. Two items are at a distance of 2 if they belong to two different but closely related superordinate categories (e.g., apple and steak belong to the same taxonomic category *food* and are at a distance of 2 because they have the superordinate-level categories of fruits and meat in between).

Superordinate categories lack semantic homogeneity, and features that unite them are difficult to find (e.g., features common to all pieces of furniture) (Murphy, 2002). Members of the same basic-level category are more similar to one another than members of a superordinate category, and they are easier to learn than superordinate-level categories (Markman, 1989; Murphy, 2002). The difficulty in learning superordinate categories can be interpreted as a difficulty in aligning their members on the basis of their commonalities and alignable differences (Markman & Gentner, 1993; Markman & Hutchinson, 1984; Waxman, 1990; Waxman & Hatch, 1992).

Studies involving other tasks have provided further support for this observation. It is more difficult to find semantic commonalities between entities belonging to distant semantic domains than between entities from close semantic domains. This has been shown for property or relation generalization tasks, novel name generalization, analogical reasoning, and problem solving (e.g., Alfieri et al., 2013; Barnett & Ceci, 2002; Gick & Holyoak, 1983; Green, 2016; Green et al., 2010; Klahr & Chen, 2011; Thibaut & French, 2016; Thibaut et al., 2010).

Taxonomic distance, as described above, has not been systematically manipulated in previous novel noun comparison experiments. Usually, learning stimuli come from the immediate (Level 1) superordinate category (e.g., apple and pear for fruits). Sometimes, other trial types have also been used, such as a caterpillar and a snake (i.e., animals from different immediate superordinate categories) or a derby hat and a baseball cap (arguably from the same basic-level category) in Namy and Gentner (2002). With regard to generalization items, the majority of them have been selected from the same immediate superordinate category as the training category (e.g., banana for apple

[1] Semantic distance is a widely used concept that nevertheless is difficult to define (see Kenett, 2019; Murphy, 2002). It has been left undefined in many articles, with the authors relying on intuition and illustrating it with examples but providing no working definition. Following on from Collins and Loftus (1975), the more semantic properties two concepts share, the more links there are between them. Alternatively, some authors have defined semantic distance as the "shortest path [direct or indirect] between two nodes" (Collins & Loftus, 1975, p. 412, Note 3). In latent semantic analysis (LSA), the semantic similarity between words in a given high-dimensional semantic space is defined as the probability of a given word co-occurring in a specific context. It should be noted that the two representations do not always converge (see Kenet, 2019). Another way to define semantic distance is to refer to psychological ratings collected from participants.

and pear, turnip for carrot), but there are exceptions to this. For example, a turtle has been used as the taxonomic choice for a learning item such as a caterpillar and belongs to the Level 2 superordinate category. In contrast, the generalization item *sombrero* belongs to the same basic-level category as a learning item such as derby hat. It should be noted that distance diversity was desirable in Namy and Gentner (2002) because it permitted generality.

More relevant for our research question is the study conducted by Callanan (1989), who manipulated semantic distance in a novel name learning paradigm for objects and contrasted single-object and comparison conditions in 3- and 5-year-old children. There were two single-object conditions, one in which children were told "This is a terval" and another incorporating an inclusion statement at the superordinate level such as "This is a wug. A wug is a kind of terval." In the corresponding comparison conditions, the children were told "This is a terval, and this is another terval" or the superordinate-level inclusion statement "This is a wug. A wug is a kind of terval. This is another terval" for a dog and a cat. The children then saw three basic-level, three superordinate-level, and three subordinate-level objects, introduced one by one. The single-object condition mostly elicited subordinate- and basic-level interpretations. The comparison condition with inclusion statements elicited more superordinate-level words and, in contrast to our comparison paradigm, was used by the author to teach superordinate-level nouns rather than basic-level nouns. However, taxonomically related items were more perceptually similar, and this might have contributed to the taxonomic choices.

Liu et al. (2001) also compared single-stimulus and comparison conditions in 3- to 5-year-old children. They manipulated semantic distance while controlling for perceptual similarity between the standard(s) and the response choices (taxonomically, thematically, and perceptually related). The learning stimuli came either from two different basic-level categories (e.g., an orange and a banana) or from the same basic-level category (e.g., two oranges). The *different* condition elicited significantly more taxonomic answers (70%) than perceptual choices (23%), whereas the results in the *same* basic-level conditions revealed no significant difference between the taxonomic and perceptual choices (53% vs. 44%, respectively). However, Liu et al. (2001) assessed the taxonomic status of the learning items (basic or superordinate) not as one factor in the same experiment but indirectly in a comparison between two experiments. They also did not control generalization distance. Interestingly, comparing the *same* and *different* basic-level conditions revealed no significant difference between them in terms of the number of taxonomic answers but instead indicated a higher proportion of perceptual items in the *same* basic-level condition, suggesting that perceptual similarity has a greater influence when the learning items are more homogeneous (basic level).

The current study: Aims and hypotheses

The current article reports two experiments that assessed the role of semantic distance between learning items as well as between learning items and generalization items. Experiment 1 introduced object nouns, whereas Experiment 2 dealt with relational nouns (see below). Semantic distance between learning items was manipulated in the comparison conditions, and the semantic distance between learning items and generalization items was manipulated in both the no-comparison and comparison conditions.

With regard to object nouns, previous research suggests a bias for basic-level objects in no-comparison conditions. This level would be favored because it is hypothesized to be the most cognitively accessible (Emberson et al., 2019; Waxman, 1990; Waxman & Hatch, 1992). In comparison designs, the available evidence regarding novel object nouns suggests that a greater learning distance elicits broader generalization (Liu et al., 2001), but only indirectly in a between-experiment comparison. Despite this, semantic distance between learning items and generalization items has not been systematically studied in connection with objects. The second experiment addressed the case of relational nouns, such as *neighbor* and *addition.* These nouns refer to relations between objects rather than to the objects themselves. There is also evidence suggesting that semantic distance between the connected domains plays a role (Gentner et al., 2011; Thibaut & Witt, 2015). However, as will become clearer in the introduction to Experiment 2, there is no systematic manipulation of distance between learning items themselves and between learning items and generalization items.

We believe that it is important to manipulate semantic distance at learning and between learning and generalization separately. As mentioned above in the Introduction, single-stimulus designs have led to the identification of a number of biases. Here, we used a forced-choice design to test whether providing comparison designs promotes conceptually based answers beyond the basic level and weakens the influence of perceptually similar distractors. We hypothesized that manipulating the distance between the learning items would have an effect on the constructed conceptual representation in the sense that a greater distance might lead to a broader conceptual space but also potentially to more conceptualization failures. Semantic distance at test is thought to capture what participants accept as being part of the category. Presentation conditions (no-comparison, close comparison, and far comparison) and generalization distance might also interact.

These are open questions. We addressed the role of semantic distance in terms of two views that make opposite predictions. One view of alignment suggests that starting with close comparisons provides a more robust basis for semantically based alignments. Close comparisons would pave the way for remote commonalities that would remain out of reach without these early, close accessible alignments. For example, Kotovsky and Gentner (1996) performed an analogical reasoning task showing that the use of close pairs in the base domain enabled children to apprehend very different instantiations of the relation in the target domain, which were not found without the presence of this early basis.

In contrast, the opposite hypothesis stresses the necessity for a greater distance between the learning stimuli in comparison designs, suggesting that variability at learning is important in order to enable a broad generalization space. This view is consistent with the available evidence from earlier experiments on comparisons (see Liu et al., 2001).

To summarize this debate, a small distance would lead to a firmer representation resulting from "close comparison" alignments, which would permit greater distances at test. In contrast, greater learning distances would emphasize the potential for a broader range of variations in the category. These earlier views and results, however, do not predict how learning distance interacts with generalization distance, which was a central question in the current study.

Beyond the previous debate on the two learning distances, close and far, we posit that not all comparisons are cognitively equal. Augier and Thibaut (2013) argued that they necessitate systematic attention to the learning stimuli, to salient as well as less salient dimensions, and require systematic comparisons of learning and generalization stimuli as well as the inhibition of salient irrelevant dimensions. Against this background, they showed that 4-year-old children did not benefit from a larger number of learning stimuli—that is, four versus two (i.e., more converging evidence)—in contrast to 6-year olds. Similarly, the authors demonstrated that three learning stimuli, rather than two or even four, resulted in the best performance in 3-year-olds. They argued that there is an optimal number of stimuli for comparison and that this might differ across ages.

As suggested by Andrews and Halford (2002), two comparison conditions might differ in the number of items of information that can be processed in parallel in working memory (see also Zelazo et al., 1997). Indeed, different levels of semantic distance differ in the number of explorations of the stimuli they require and therefore in the number of irrelevant salient dimensions that need to be processed. When semantic distance between learning and generalization items increases (i.e., far/distant comparisons), a larger number of features activated for the learning and generalization stimuli will not be alignable than is the case for close/near items and will need, as a result, to be inhibited. Conversely, starting with a larger number of irrelevant features might also require more flexibility in order to generate novel hypotheses. This reasoning was followed by Thibaut et al. (2010) who manipulated semantic distance between domains (i.e., between the base and target domains) by means of proportional analogies. Preschool children performed worse with analogies involving semantically distant domains than with those involving semantically close domains.

We used familiar learning stimuli in our two experiments in the same way as in many lexical learning experiments which have included classical single-stimulus designs. We then assessed the options that the children would select as a generalization stimulus. The set of options depended on the hypotheses. We further built on this logic and on biases previously studied with the single-object design. We tested children's capacity to depart from the choices they would have made in a single-object design when a second learning stimulus was provided and further tested how the addition of this second stimulus constrained the generalization of the word. Following a common practice in this literature, we used a puppet task and asked the children to learn the puppet's language rather than use their own language or to help the puppet understand this pretend language.

Experiment 1: Novel names for objects—The role of learning and generalization distance

This experiment examined the effect of learning and generalization distance in a comparison of familiar objects task. We studied which comparison condition would lead to better taxonomic generalization. The key question was whether the distance between semantic domains in the learning items and the distance between the learning items and the generalization items (i.e., the taxonomically related target) would differentially affect taxonomic choices.

Method

Participants

A total of 201 preschoolers were tested individually at school. Two age groups were recruited, most of them before the COVID-19 period. The younger children (n = 99; M_{age} = 4 years 10 months, SD = 3.7 months, range = 50–65 months) and the older children (n = 102; M_{age} = 6 years 7 months, SD = 3.3 months, range = 74– 87 months) were randomly assigned to one of the three experimental conditions with 33 or 34 children per condition. Informed consent was obtained from their school and their parents. It was realized in the context of a research convention between the laboratory, the university, and the academic authorities of the Côte d'Or ("Inspection Académique de Côte d'Or"). An a priori power analysis was conducted using G*Power (Faul et al., 2007) for sample size estimation, based on data from Gentner and Namy (1999) (N = 24), who compared children in the one-kind condition (n = 12; M = .69, SD = .24) with children in the two-kind condition (n = 12; M = .35, SD = .24). The effect size in Gentner and Namy was 1.42, considered to be extremely large using Cohen's (1988) criteria. With a significance criterion of α = .05 and power = .80, the minimum sample size needed with this effect size is n = 9 per group for a two-tailed Student's t test between two independent groups. Thus, the sample size per group (n = 33) (total sample of 201 divided by 2 [Age] $\times$ 3 [Comparison] conditions = 33.5) was more than adequate to test our hypotheses.

Design

Age (4-year-olds vs. 6-year-olds) was crossed with learning (no-comparison vs. close comparison vs. far comparison; between-participant factor) and generalization (near vs. distant; within-participant factor).

Materials

Seven sets of six objects pictured on cards were created for each condition of taxonomic distance (close or far; near or distant). Taxonomic distance can be defined, in terms of rank in a taxonomy, as the number of intermediate taxonomic steps leading to the nearest common superordinate category. For examples, two entities from the same basic-level categories (e.g., two apples) are at Level 0, and an apple and a banana are fruits, at Level 1, the nearest (familiar) superordinate category. An apple and a piece of beef are at Level 2, with food as a common superordinate category. In comparison, thematic relations refer to contextual semantic associations, such as dog and leash, and the strength of these associations can be assessed by adults' subjective ratings.

Each set corresponded to one object category (e.g., clothing accessories, food, tools) and was composed of a reference learning object, the standard (e.g., bracelet, apple, hammer) in the single-object condition. In the comparison conditions, we manipulated the semantic distance (close or far) between the two learning objects that were introduced in the learning pair. For each object category (e.g., food), the close comparison pairs were composed of two items from the same basic-level category (e.g., an apple–a second apple), whereas the far comparison pairs were composed of two learning items from a common superordinate category (e.g., fruit category for apple and cherry). There were two test pictures in both the close and far generalization conditions: first, a perceptual choice, which was the same in both distance conditions (e.g., a Christmas ball), and second, a taxonomic choice, which differed as a function of the generalization condition, near or distant (e.g., a banana or a piece of meat, respectively). Fig. 1 depicts the objects used to instantiate the close and far comparison conditions and the two generalization conditions for a food category (see Appendix A1 for the list of items). Semantic similarity ratings were collected from adults and confirmed that objects in the close comparison pairs were semantically more similar one to the other than objects in the far comparison pairs and that near generalization materials were more similar to the learning materials than were distant generalization materials (see Appendix A2 for the procedure and A3 for ratings). Perceptual similarity ratings revealed that perceptual choices were perceptually more similar to the learning items than were the taxonomic choices (see Appendix A3 for ratings). We forged seven bisyllabic labels (pseudowords) because bisyllabic labels have been shown to be more easily remembered than monosyllabic ones (Gathercole & Baddeley, 1993). Each set was associated with one of two-syllable novel names (e.g., yo*uma, buxi, dajo, zatu, sepon, xanto, vira*).

Procedure

The experiment started with two practice trials followed by 14 experimental trials presented in a random order. Each standard was introduced with a novel count noun (e.g. "This is a *buxi*") in the single-object condition. In the comparison conditions, both objects were named with the same label (e.g., "This is a *buxi*" for one standard and "This is a *buxi* TOO" for the other standard). A puppet named Yoshi was used in order to make the task more attractive for children and to make sense of the use of nonwords to designate object categories. The instructions were as follows: "In this game we are going to learn the language of Yoshi." The object(s) was (were) introduced sequentially and left in view. In the comparison conditions, the two stimuli were presented in a row and their location was determined randomly. The forced-choice test phase was identical in all conditions. The two test objects (i.e., the perceptual and taxonomic matches) were introduced, and children were asked to point to the one that was also a member of the category (e.g., "Show me which one of these two is ALSO a *buxi*").

Results

We analyzed the proportions of taxonomic choices. Because proportions are not normally distributed, we conducted the analysis of variance (ANOVA) on arcsine transformations of the proportions (see Winer, 1971). We ran a 2 (Age: 4-year-olds vs. 6-year-olds) × 3 (Learning: no-comparison vs. close comparison vs. far comparison) × 2 (Generalization: near vs. distant) ANOVA on the proportion of taxonomic choices with age and learning as between-participant factors and generalization as a within-participant factor. Fig. 2 illustrates the results by condition. This analysis revealed a significant main effect of learning, $F(2, 195) = 14.44$, $p < .0001$, $\eta_p^2 = .13$, (no-comparison: $M = .61$, $SD = .35$; close comparison: $M = .80$, $SD = .35$; far comparison: $M = .93$, $SD = .35$), a significant effect of age, $F(1, 195) = 15.23$, $p = .0001$, $\eta_p^2 = .07$ (4-year-olds: $M = .68$, $SD = .35$; 6-year-olds: $M = .87$, $SD = .35$). The interaction between these two factors was not significant, $F(2, 195) = 0.43$, $p = .65$, $\eta_p^2 = .004$. Finally, there was a main effect of generalization; children performed significantly better in the near generalization condition ($M = .86$, $SD = .38$) than in the distant generalization condition ($M = .70$, $SD = .39$), $F(1, 195) = 49.40$, $p < .0001$, $\eta_p^2 = .20$. The triple interaction Generalization * Age * Learning, $F(2, 195) = 1.73$, $p = .18$, $\eta_p^2 = .018$, was not significant (see Fig. 2) (see Appendix A4 for results per item)

We also compared the proportions of correct responses with chance (50%), Student's t tests with a Bonferroni correction for multiple comparisons (significance at .0042, alpha threshold of .05 divided by 12, the number of independent comparisons with chance). These comparisons with chance tell us that even though performance in two conditions might significantly differ, even the best performance might not differ from chance. They revealed that 4-year-old children performed significantly above chance only in the far comparison–near generalization condition ($M = .62$), $t(32) = 3.10$, $p = .004$, and below chance in the no-comparison conditions (near: $M = .32$; distant: $M = .28$) and in the close comparison–distant transfer condition ($M = .34$) ($ps < .001$). The 6-year-olds were above chance in the far learning–near generalization condition ($M = .74$), $t(33) = 5.34$, $p < .001$, and

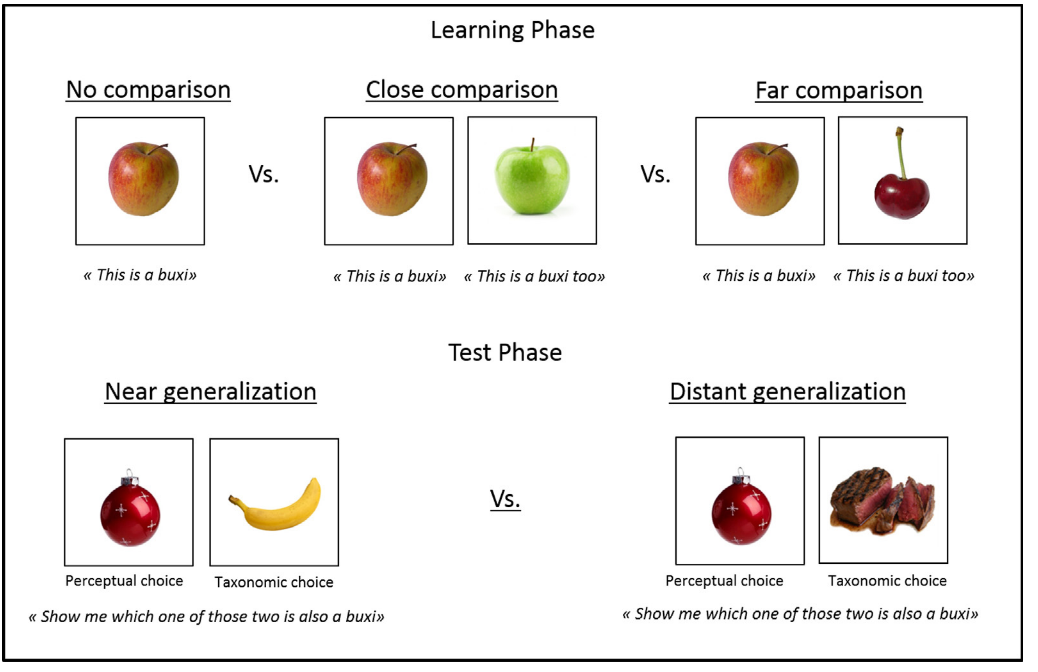

Fig. 1. Sample stimulus sets and instructions used in the six conditions crossing learning type (no-comparison vs. close comparison vs. far comparison) and generalization (near vs. distant).

significantly above chance in the far learning–distant generalization condition (M = .63), $t(33) = 3.07$, p = .004. They were at chance in the no-comparison and close comparison learning conditions whatever the generalization distance (ps > .0042). Overall, these results revealed, as expected, that the no-comparison conditions led to perceptually based generalization or, at best, chance performance in both age groups. In the other conditions, the hierarchy of conditions in the 4-year-olds is quite revealing. Near generalization was above chance in far comparison and at chance in close comparison, and distant generalization was driven by perception in close comparison.

A final analysis compared the individual profiles of responses across experimental conditions. We defined three profile types. Perceptually dominant children chose the perceptual distractor at least five times out of seven. Taxonomically dominant children chose the taxonomic choice at least five times out of seven trials. All the others were categorized as inconsistent. The results are given in Table 1. A chi-square test of independence was performed to examine the relation between learning type and choices at test, namely favoring perceptual choices, favoring taxonomic choices, or being inconsistent in the response mode (see Table 1 for results). At 4 years of age, children shifted from a perceptually based mode of categorization in the no-comparison condition to perceptual or inconsistent profiles in the close comparison condition and to inconsistent and taxonomically based profiles in the far comparison condition. This is true for the near generalization case, $\chi^2(4, N = 99) = 26.65$, $p < .0001$, and to a lesser extent in the distant generalization case, $\chi^2(4, N = 97) = 15.89$, $p < .01$, where inconsistent responses remained dominant in the far comparison condition. At 6 years of age, there was a shift from a scattered response mode in the no-comparison and taxonomically dominant participants in the near generalization case, $\chi^2(4, N = 102) = 10.45$, $p < .05$, and to a lesser extent in the distant generalization condition, $\chi^2(4, N = 102) = 9.81$, $p < .05$, where the number of inconsistent responses remained stable across conditions.

Together, these results confirm that close comparisons decrease the prominence of perceptually based answers observed in the absence of comparison, whereas far comparisons promote a taxonomically based generalization. However, the benefit of comparison is limited to the near transfer domains

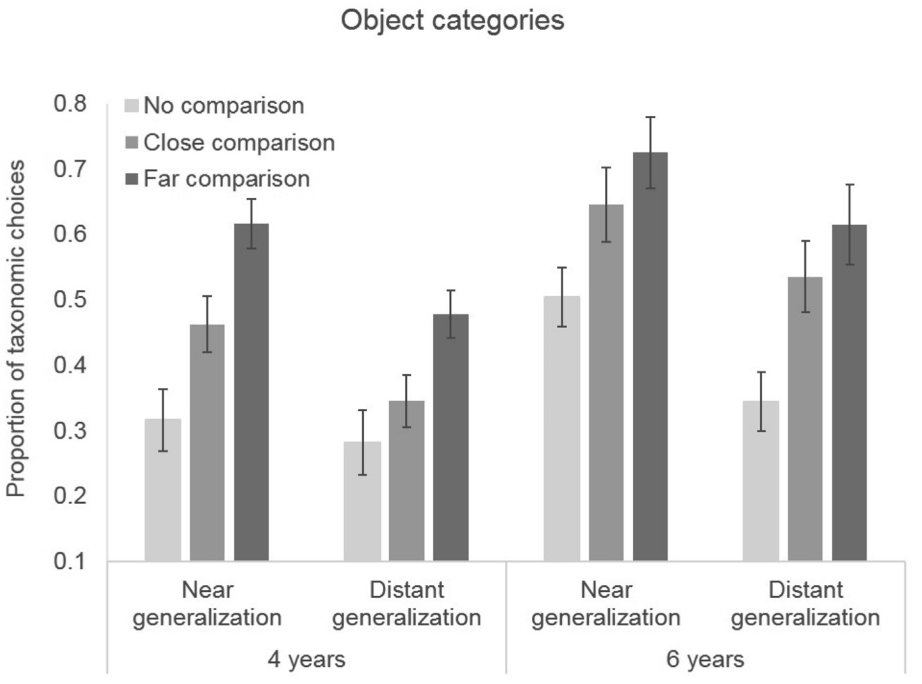

Fig. 2. Mean proportions of correct responses for the objects categories (taxonomic choices) as a function of age (2: 4-year-olds vs. 6-year-olds), learning (3: no-comparison vs. close comparison vs. far comparison), and generalization (2: near vs. distant). Error bars are standard errors of the means. Chance was .50.

Table 1
Distribution of individual profiles (perceptual, taxonomic, and inconsistent) as a function of experimental conditions, Learning (3: no-comparison vs. close comparison vs. far comparison) × Generalization (2: near vs. distant), for the two age groups (4-year-olds vs. 6-year-olds)

		No-comparison		Close comparison		Far comparison	
		Near	Distant	Near	Distant	Near	Distant
4-year-olds	Perceptual	23	23	12	17	3	8
	Inconsistent	4	5	13	13	16	18
	Taxonomic	6	5	8	3	14	7
6-year-olds	Perceptual	13	16	9	11	4	6
	Inconsistent	9	13	7	10	5	14
	Taxonomic	12	5	18	13	25	14

and to far transfer in older children, most likely because they are able to cope with the cognitive constraints of far comparison.

Discussion

As expected, comparison conditions led to more taxonomic choices than the no-comparison standard design. More important, semantic distance was a major determinant of comparison efficacy. Both distance between learning items and distance between learning and generalization items had an effect. As far as the learning conditions are concerned, the far comparison condition led to more taxonomic answers than the close comparison condition. Comparisons with chance and individual profiles confirm that taxonomic responding was favored in the far comparison condition. At test, near generalization (i.e., to close taxonomic categories) was easier than distant generalization (i.e., to more remote taxonomic categories) except for older children who could reliably choose distant generalization items in the far comparison condition. Confirming previous results, the no-comparison condition led to perceptually based generalization in most cases. Overall, this reveals a progression from no-comparison to far comparison learning that leads to broader semantic domains.

Augier and Thibaut (2013) argued that comparison situations generate cognitive costs that might prevent younger children from finding deep semantic regularities. Our results suggest that the effect of comparison on generalization is modulated by semantic distance, which in turn influences how similarities and differences are processed and aligned as relevant and irrelevant taxonomic dimensions. Because younger children might encounter more difficulties in capturing semantic similarities in the case of far learning items and/or applying them to distant domains, we hypothesized that semantic distance during learning and at test might differentially affect the benefits of comparison across age groups. However, near generalization was performed better than distant generalization. Taken together, these two results suggest that a broader learning range (i.e., items from the same superordinate category) leads to better near generalization.

The fact that only older children performed above chance on far comparisons in the distant generalization condition demonstrates a development from, first, accurate performance in the case of a broad learning distance and small generalization distance to, second, progressively better performance on wider generalization sets. One interpretation is that far comparison allowed both age groups to move the focus of attention away from perceptual similarities more than close comparison. However, for younger children, far comparison (i.e., items from the same superordinate category) led to generalization within the same superordinate category but not beyond it. In comparison, the same learning condition was sufficient to enable older children to understand higher-level properties and, hence, categories (i.e., a more remote superordinate category).

To conclude, our experiment suggests that the successful use of comparison depends on differences along the semantic dimensions of the compared objects as well as on executive capacities to process these more or less salient dimensions of stimuli. This finding has important implications regarding the role comparison plays in concept learning. Indeed, the executive constraints on comparison processing might explain the conditions under which comparisons can or cannot be successfully handled. The aim of Experiment 2 was to test whether these conclusions can be extended to learning and generalization of relational categories.

Experiment 2: The case of relations—Semantic distance during learning and at test

In the second experiment, we manipulated the same semantic distance factors as in Experiment 1—that is, between learning items and between learning items and generalization items—but in the case of relational nouns. Relational nouns refer to categories that are defined by relations between objects rather than by the intrinsic properties of the objects themselves (Childers, 2020; Gentner et al., 2011; Thibaut & Witt, 2015). For example, a pair of entities can be "neighbors" if they entertain a "proximity" relation. In the same way, "additions" involve many different objects. Relational nouns appear later than many object nouns, between 17 and 30 months of age (Fenson et al., 1994) because they refer to the relations between objects rather than to the objects themselves. This explains why children often misunderstand relational terms as referring to object categories (e.g., Hall & Waxman, 1993).

Gentner et al. (2011) contrasted comparisons and no-comparison learning conditions for relational nouns such as "X is the *dax* for Y." They introduced pairs built around two familiar objects connected by a familiar relation (e.g., "cutter of"), an operator (e.g., a knife). and an entity (e.g., a watermelon). At test, in all the experiments, a novel entity was introduced with three alternatives (e.g., a relational match: a pair of scissors; a taxonomic match: a pile of sheets of paper; and a thematic match: a pencil). Children were asked to show which stimulus among the alternatives would also be the *dax* for the piece of paper. The no-comparison condition (Experiment 1) gave fewer relational choices than the comparison condition (Experiment 2) in 4- and 6-year-old children. Thibaut and Witt (2015) extended these results, showing that far learning pairs led to better generalization than close learning pairs and that there was an optimal number of training pairs.

Gentner et al. (2011) and Thibaut and Witt (2015) manipulated the semantic distance between items in the learning pairs for relational categories, resulting in pairs made up of items from close categories (e.g., two different knives or two fruits) or far categories (e.g., a knife and an ax, a fruit and a log). However, they did not manipulate the semantic distance between the semantic domains instantiated by the learning items and the instantiated relation in the generalization items.

Here, we manipulated the semantic distance between the learning pairs and between the learning items and the generalization items. In contrast to Thibaut and Witt (2015), we introduced a no-comparison control condition to assess which learning or generalization condition would benefit the most from stimulus comparison. Furthermore, we compared two age groups (3- and 4-year-old children) to better understand how cognitive resources might interact with semantic distances. We selected these two age groups based on Gentner et al. (2011) and Thibaut and Witt (2015).

Relational and object nouns differ in the nature of their referents. Does this mean that object and relational nouns would behave differently in comparison settings? For object words, objects are assumed to be compared and aligned, starting with their perceptual similarities, which invite further explorations and, eventually, alignments on more semantically relevant properties. Comparisons, in the case of relational words, are like solving analogies. Children need to first map the relation between the objects in the two learning pairs and then map the resulting relation (if found) on the relational option. In terms of Fig. 3 below, it means, for example, "to find the relation between the knife and watermelon and the same for the other knife and the orange." Generalization requires understanding in what terms the two learning relations (the relation referred to by "being the dax for") can be translated to another domain (e.g., here "shaving"). The generalization requires uncovering a predicate–argument structure and mapping this relational structure on the generalization items, whereas perceptual similarities play no role.

One could predict that far comparisons should lead to broader generalization, as in Experiment 1. However, it could also be that once the relation is found in a pair, the pair becomes productive in the sense that children can apply the representation in many situations. Because this relation should be easy to find in close comparisons, this might mean excellent generalization results in this condition. This might also be the case in the no-comparison case; once children have discovered the relation between, say, the knife and the orange, they might generalize it in another context.

We also hypothesized an interaction between age and semantic distance factors. As argued by Thibaut et al. (2010), this interaction would result from the fact that older children more systematically explore the semantic space in less obvious cases. For example, both age groups would generalize similarly in close comparisons and in near generalization conditions because the stimuli are built around similar domains in both the learning and generalization stimuli. However, in far learning domains or in more distant generalization domains, younger participants might encounter greater difficulties in capturing relational similarities between learning pairs or between learning and generalization pairs.

Method

Participants

A total of 214 French-speaking preschoolers were tested individually in a quiet room at their school. Two age groups were recruited. The younger group was composed of 101 children (M_{age} = 3 years 10 months, range = 40–51 months), and the older group was composed of 113 children (M_{age} = 4 years 8 months, range = 53–60 months). All children were randomly assigned to one of the three experimental conditions with 63 (no-comparison), 76 (close comparison), and 75 (far comparison) children per condition. Informed consent was obtained from their school and their parents. An a priori power analysis was conducted using G*Power (Faul et al., 2007) for sample size estimation, based on data from Gentner et al. (2011) (N = 58), which compared performance for children who received progressive alignment (Experiment 3: n = 58; M = .44, SD = .26) with those who did not (Experiment 2: n = 32; M = .28, SD = .21). The effect size in Gentner et al. (2011) was .68, a medium (d' = .50) to large (d' = .80) effect size using Cohen's (1988) criteria. With a significance criterion of α = .05 and power = .80, the minimum sample size needed with this effect size is n = 36 per group for a two-tailed independent Student's t test, which was the sample size of our group (n = 36) for a

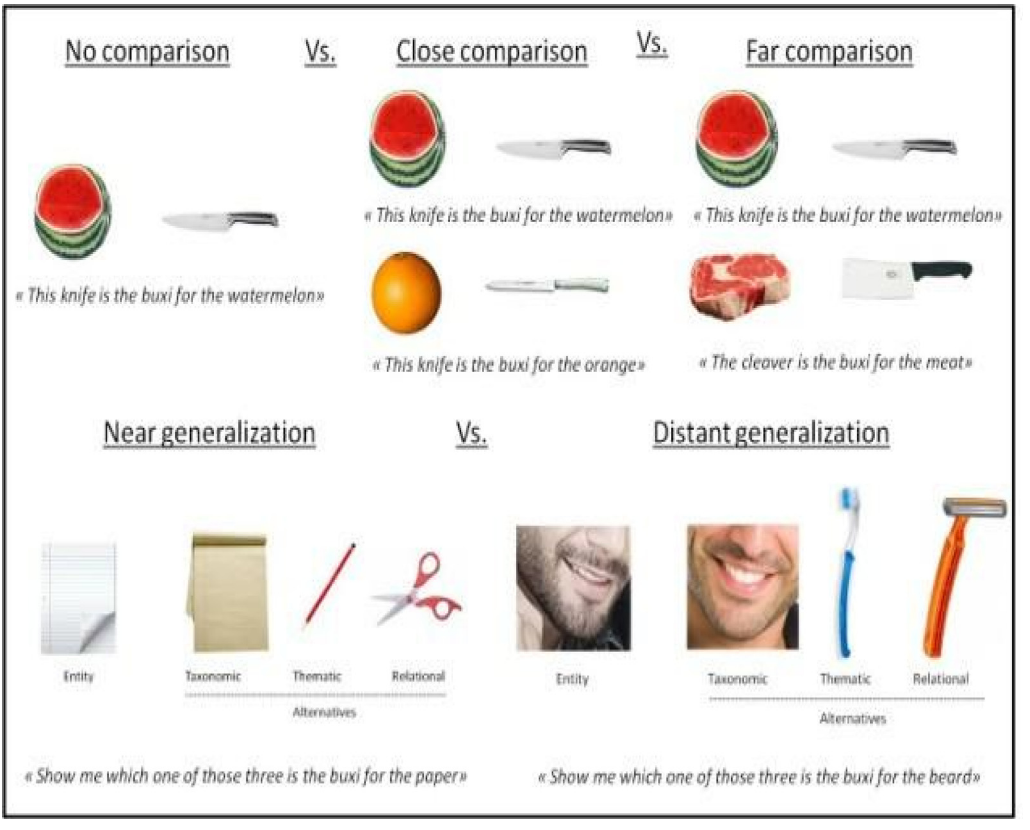

Fig. 3. Example of a stimulus set and instructions adapted for the six experimental conditions resulting from crossing learning type (no-comparison vs. close comparison vs. far comparison) and generalization distance (near vs. distant) factors.

two-tailed Student's t test between two independent groups. Thus, the mean sample size per group (n = 36) (total sample of 214 divided by 2 [Age] × 3 [Comparison] conditions = 35.7) was sufficient to test our hypothesis.

Design

Age was crossed with learning type (no-comparison vs. close comparison vs. far comparison; between-participant factor) and test distance (near vs. distant; within-participant factor).

Materials

Stimuli were adapted from Thibaut and Witt (2015, Experiment 1A). Fourteen sets of pictures were built. Each set corresponded to one of the seven relational categories used in this experiment (i.e., cutter for, home for, food of, baby of, container for, travel space for, cleaner for, and product of). All learning phase stimuli were organized around learning pairs, that is, one learning pair in the no-comparison case and two learning pairs in the comparison conditions (close and far comparison conditions). Each pair was composed of an operator and an entity, (e.g., a knife as an operator and a watermelon as an entity; see Fig. 3). The no-comparison condition pairs were composed of an operator-entity pair (e.g., either knife1–watermelon, knife2–orange, or cleaver–meat), the close comparison pairs condition was composed of semantically similar items (e.g., knife1–watermelon, knife2–orange), whereas the far comparison pairs were composed of less semantically similar pairs (e.g., knife1–watermelon, cleaver–meat). In this example knife1 and knife2 are semantically closer one to the other than knife1 and cleaver. The same is true for watermelon–orange (close) compared with watermelon–meat (far). Note that in both the close and far comparison conditions, each semantic relation was illustrated by two exemplars (e.g., there were two close "is the cutter for" learning exemplars each composed of two different pairs). Each was alternatively associated with a near generalization choice for half the participants (e.g., a pair of scissors for a piece of paper) or a distant generalization choice for the other half (e.g., shaver and beard). The test cards consisted of four pictures. The relationally correct answer was always the operator (e.g., a pair of scissors, a shaver, see Appendix B1 for the list of items). Independent participants (see Appendix B2 and B3) rated the solution operator(s) as semantically nearer to the learning operator(s) pictures in the near generalization condition than the solution operator(s) in the distant generalization solution. For example, in Fig. 3 the knife was rated as semantically closer to the pair of scissors than to the shaver. There was also a taxonomic card choice (e.g., pile of sheets of paper that was taxonomically related to the entity sheet of paper in Fig. 1) and a thematic card choice (e.g., pencil was thematically related to the entity sheet of paper in Fig. 3). The top part of Fig. 3 depicts the close and far comparison pairs for the "cutter for" relation during the learning phase in the no-comparison, close comparison, and far comparison conditions. The bottom part illustrates the four pictures introduced at test (entity plus taxonomic, thematic and relational choices) in the near and distant generalization conditions. The pictures used in our experiment were realistic pictures like those in Thibaut and Witt (2015). Independent similarity ratings confirmed that close comparison pairs were more semantically similar one to the other than far comparison pairs and that near generalization pairs were more similar to learning pairs than distant generalization pairs (see Appendix B3 for the similarity ratings per category.

Procedure

The stimuli were displayed on a laptop screen in a PowerPoint file. The sequence (order) of presentation of the learning pair(s) and the generalization stimuli was controlled, with four different orders of presentation per learning condition). During the learning and test phases, the experimenter kept the speech flow constant across items and experimental conditions. Prosodic emphasis was added for the pseudowords so that children noticed that the same label was used for each pair of stimuli instance (during the learning and test phases).

Learning phase. A puppet, Yoshi, was used in order to make the task more attractive for children. The experimenter introduced the game with the following instructions (the example is for the close comparison condition; the instructions were the same for the far and no-comparison learning conditions): "Hello, we are going to play a game together. In this game we are going to teach Yoshi the word *buxi.* We are going to show him what a *buxi* is. Look [the knife1 and the watermelon appeared at the top of the screen], this knife is the *buxi* for the watermelon. Now look, [the knife2 and the orange appeared below the knife1–watermelon pair], this knife is the *buxi* for the orange." In the no-comparison condition, the experimenter introduced one pair during the learning phase. In all conditions, the learning pair(s) remained in view during the entire trial until children gave an answer (pointing).

Test phase. The test started with these instructions: "Now let's look at this (these) [gesturing across the learning pair(s)]. You see how this (these) [gesturing across the operator(s)] is (are) a *buxi (buxies)* for this (these) [gesturing across the entity (entries)]? Now, your turn. Which one of these [the test cards—entity: paper; taxonomic: pieces of paper; thematic: pencil; relational: scissors—appeared at the bottom of the screen] is the *buxi* for the paper?" To avoid answers before children analyzed the three test cards, we asked them to refrain from answering before a picture of Yoshi appeared on the screen (5 s after the test cards appeared). Children chose among the three test cards by pointing to the one on the screen that was the *buxi* for the paper. This procedure was repeated for the 14 experimental relational categories. The presentation order of the relational categories and the position of the three choices (left, middle, or right) were counterbalanced, and the labels were interchanged among pairs across participants.

Coding and analysis of the data

We computed the proportion of relational choices made at test as well as the proportions of alternative choices (taxonomic and thematic choices) for the 7 near generalization trials and the 7 distant generalization trials (total of 14 trials). A Kolmogorov–Smirnov test was used to test for normality on the main dependent variable, namely the proportion of relational choices. The proportion of relational choices for the 3-year-old group, $D(101) = .118$, $p > .10$, and the proportion of relational choices for the 4-year-old group, $D(113) = .077$, $p > .20$, both were normal, indicating that the data were normally distributed in both groups.

Results

As in Experiment 1, we analyzed the proportions of taxonomic choices. Again, because proportions are not normally distributed, we conducted our ANOVA on arcsine transformations of the proportions (see Winer, 1971). We ran a three-way ANOVA with age (3-year-olds or 4-year-olds) and learning type (no-comparison, close comparison, or far comparison) as between-participant factors and test distance (near or distant) as a within-participant factor. In particular, within each age group we assessed which comparison condition significantly differed from the corresponding no-comparison condition. Finally, we compared the proportion of relational answers with chance and analyzed individual profiles of answers.

The three-way ANOVA revealed a significant main effect of age, $F(1, 208) = 22.57$, $p < .00001$, η^2_p = .10 (younger = .70; older = .83), learning type, $F(2, 208) = 33.93$, $p < .0001$, $\eta^2_p = .25$ (no-comparison = .60; close comparison = .88; far comparison = .82). A posteriori comparisons (Tukey HSD [honestly significant difference]) revealed that the no-comparison condition was significantly lower than the two comparison conditions (*ps* < .001), which did not differ one from the other ($p = .17$). Learning type did not interact with age, $F(2, 208) = 1.60$, $p = .20$, $\eta^2_p = .01$. There was no main effect of test distance, $F(1, 208) = 1.97$, $p = .16$, $\eta^2_p = .01$. Test distance interacted with learning type, $F(2, 208) = 11.234$, $p < .0001$, $\eta^2_p = .10$, but not with age, $F(1, 208) = 0.09$, $p = .76$, $\eta^2_p = .0004$. As depicted in Fig. 4, Tukey HSD revealed that the close and far comparisons gave significantly higher relational choices than the no-comparison in the near generalization case (*ps* < .0001), whereas in the distant generalization case only close comparisons led to a higher level of relational choices ($p = .004$; far vs. no-comparison: $p = .67$). There was no significant difference between close and far comparisons regardless of generalization distance. The interaction among age, learning type, and test distance was not significant, $F(2, 208) = 0.33$, $p = .72$, η^2_p = .003. (see Appendix B4 for the proportions of relational choices per item for each learning condition)

We compared each condition against chance as a function of age, comparison, and generalization distance. Children performed significantly above chance in all the comparison conditions (*ps* < .0042, Bonferroni-corrected *p*-value threshold) except for the younger children in the far comparison–distant generalization condition ($p = .026$). In the two no-comparison conditions, the younger group performed below chance in the near case ($p < .0042$) and at chance in the distant case ($p = .22$), and older children were at chance in the no-comparison–near generalization case ($p = .23$) but beyond chance in the no-comparison–distant generalization condition ($p < .001$).

Finally, we compared the individual profiles of responses between experimental conditions. As in Experiment 1, we defined three profile types. Two were taxonomic/thematic profiles for children who chose the taxonomic or thematic distractor five to seven of seven times and relational profiles for children who chose the correct relational choice on five to seven of seven trials. The other children were categorized as inconsistent. The results are given in Table 2.

A chi-square test of independence was performed to examine the relation between learning type and choices at test, namely favoring distractor choices, favoring relational choices, or being inconsistent in their response mode. Most 3-*year*-olds were inconsistent in the close comparison conditions as well as in the far comparison–distant generalization condition, where they were inconsistent or chose the distractors. In the no-comparison condition, there were more distractor-consistent participants than in the near generalization case, $\chi^2(4, N = 101) = 33.07$, $p < .00001$, which is consistent with the previous result (i.e., fewer relational choices in the no-comparison–near generalization case than in the distant generalization case), $\chi^2(4, N = 101) = 10.69$, $p < .05$. The no-comparison condition also had a substantial number of inconsistent participants. At 4 years of age, there was a switch from inconsistency or distractor-based consistency in the no-comparison condition to inconsistency or relational consistency in the comparison conditions in both the near generalization condition, $\chi^2(4, N = 113) = 41.52$, $p < .00001$, and the distant generalization condition, $\chi^2(4, N = 113) = 15.79$, $p < .01$. In this latter condition, relational choices increased in comparison conditions, whereas the number of inconsistent participants remained stable.

Discussion

In this study, we compared no-comparison and comparison conditions in two age groups in a relational noun learning paradigm. As predicted, no-comparison conditions led to fewer relational choices in both age groups, whereas both groups performed significantly better in the close comparison conditions than in the no-comparison conditions (see Fig. 4). Despite this, although the optimal condition for the younger children was the close comparison condition, the case of far comparison–near generalization also differed significantly from chance. The 4-year-olds significantly differed from chance in all comparison conditions, showing that they understood the target relation, including the distant generalization.

These results extend Gentner et al. (2011) by showing, first, that relational language is important for understanding relations and, second, that comparison situations contribute to this understanding. They also extend Thibaut and Witt (2015), who found that both the number of training relations and the semantic distance between training items had a significant effect on performance. Our results show not only that both age groups had similar patterns of performance but also that semantic distance is an important factor during both learning and test. Taken together, these results are important because they dissociate the role of semantic distance between learning items from semantic distance between learning and generalization items (see General Discussion).

General discussion

In two experiments, we investigated the role of semantic distance in the generalization of novel object nouns and novel relational nouns. We systematically manipulated, as a function of age, semantic distances for learning and generalization items or the distance between semantic domains for relational nouns. Although previous studies of comparisons revealed an effect of distance between learning items, they did not simultaneously separate the distance between learning and transfer items and investigate possible common patterns of results for object nouns and relational nouns. The main point was to assess whether close comparison items would lead to better encoding of the targeted common features or relations and to better generalization than far comparison items (Kotovsky & Gentner, 1996). Finally, age was hypothesized to interact with taxonomic distance because longer distances were expected to require more cognitive control.

First, although our results revealed a positive role of comparison in generalization, this role was modulated by the semantic distance between learning items and, to some extent, by the type of noun. Far learning items produced better transfer in the case of object nouns, whereas the two learning distances were equivalent for relational nouns, but only in the near generalization case. However, interactions between learning and generalization distance revealed a slightly different pattern across types of nouns. Far learning led to better results for both generalization distances, with better results being observed for near generalization items than for distant generalization items in the case of object names. For relational nouns, both generalization distances led to above-chance results, with an advantage for the close learning condition over the far learning condition, but only in the case of distant generalization. As shown by profile analyses, this was due to the younger group. Finally, and contrary to expectations, the observed pattern of results was similar for both age groups.

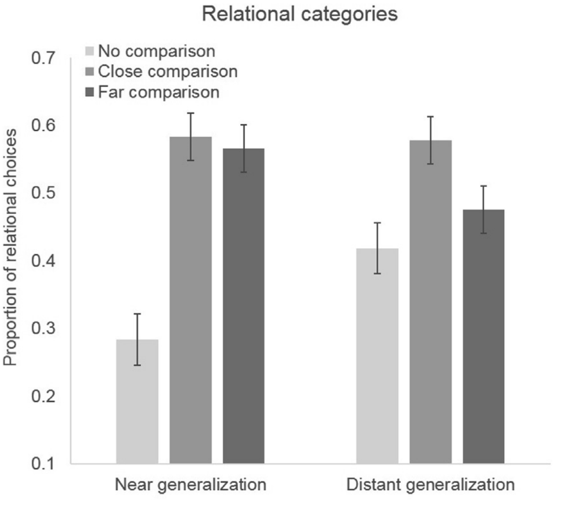

Fig. 4. Mean proportions of correct responses for the relational categories (relational choices) as a function of learning type (no-comparison vs. close comparison vs. far comparison) and generalization (near vs. distant). The error bars are standard errors of the means.

Table 2
Distribution of individual profiles (taxonomic/thematic, relational, and inconsistent pickers) as a function of experimental conditions, Learning (3: no-comparison vs. close comparison vs. far comparison) × Generalization (2: near vs. distant), for the two age groups (3-year-olds vs. 4-year-olds)

		No-comparison		Close comparison		Far comparison	
		Near	Distant	Near	Distant	Near	Distant
3-year-olds	Taxonomic/thematic	22	11	7	7	8	15
	Inconsistent	4	14	19	21	20	19
	Relational	1	2	12	10	8	2
4-year-olds	Taxonomic/thematic	25	11	7	5	3	3
	Inconsistent	9	20	14	14	16	23
	Relational	2	5	17	19	20	13

No comparison versus comparison in learning

As expected, the no-comparison conditions gave rise to fewer taxonomic choices in Experiment 1 and fewer relational choices in Experiment 2, especially in the case of distant generalization compared with near generalization.

In the case of *object nouns,* the two learning comparison conditions had the same effect in both age groups, with far comparison being better than close comparison and near generalization being better than distant generalization. This was confirmed by analyses that compared the results with chance, a more stringent test of performance, and consistency profiles, with above-chance performance and a lower level of perceptual consistency being observed in children in the two comparison conditions.

Recall that in Experiment 1 both perceptual and taxonomic matches were *a priori* viable solutions given that the perceptual match always shared a common shape with the learning items. Therefore, we wanted to identify the distance condition in which children would successfully align the stimuli on the basis of taxonomically or relationally relevant properties compared with a baseline single no-comparison condition, which was expected to favor perceptual matches (Diesendruck et al., 1998; Gentner & Namy, 1999; Imai et al., 1994; Landau et al., 1988).

Close semantic distance between learning items (i.e., less variability) led to more perceptual matches than greater distance, probably because the learning items were perceptually more similar (two items from the same basic-level category are, on average, perceptually quite similar; Archambault et al., 2000; Murphy & Smith, 1982; Xu & Tenenbaum, 2007). This was true for the distant generalization condition but also for the near generalization condition, albeit to a lesser extent.

The results for *relational nouns* also revealed a consistent scale of difficulty. Our results are consistent with Gentner et al. (2011), who showed that a relational Label condition elicited more relational matches than a No Label condition (4-year-olds or 3-year olds in a progressive alignment relational condition) (see also Thibaut & Witt, 2015, for a discussion of the optimal learning input). Our systematic manipulation of distance extended these results. Importantly, both age groups performed above chance in the two learning comparison conditions. This was not the case in the no-comparison condition. Interestingly, the younger group of children experienced more difficulties in the case of far learning and distant generalization than in the case of close learning and distant generalization, as shown by the profile analysis and the younger group's performance, which was not significantly above chance. This result is interesting when seen in the light of the one obtained in Experiment 1, which showed better results for the far learning condition in the younger group. These younger children's difficulties were most likely located at the learning level rather than at the generalization stage. We believe that the younger group found it more difficult to extract the targeted common relation in the far condition because these two different instantiations of the same relation were difficult to unify conceptually. As a result, younger children would generalize more poorly in the far learning condition than in a condition where they encoded the relation more efficiently. In other words, when the relation has once been discovered, it can be applied in many contexts.

With regard to alignments during comparisons (Markman & Gentner, 1993), we believe that these may occur because close learning comparisons result in item representations that are built more around surface similarities or because semantic similarities remain embedded in perceptual similarities. Therefore, children are unable to apply these local similarities to new stimuli that do not contain them. These accessible similarities during learning might have prevented children from seeing deeper semantic similarities and might have pushed them toward the perceptual distractor (Namy et al., 2007).

Alignment, taxonomic, and relational distance

Another main question addressed by this study was the generalization scope as a function of learning distance. Results showed better results for near generalization than for distant generalization. Once children discovered the object commonalities or the relational commonalities during comparison, they applied them to generalization items more efficiently in the case of near generalization. Two interpretations are possible. First, the abstracted representations lacked generality, and participants failed to see how they might apply to new cases (i.e., they did not see how the generalization options might afford the same object property or relation). In terms of alignment, therefore, when distant generalizations are required, perceptual similarities and/or local similarities cannot underpin the generalization items or are not compatible with the encoding/representation of distant generalization items (see also Xu & Tenenbaum, 2007).

The second interpretation is that although the representation that the children built was sufficiently general, they did not see how the object option might instantiate it or how a pair of objects could instantiate the targeted relation (e.g., "I understand that this relation applies to the objects, but I do not know how to apply it"). As hypothesized in the Introduction, it might also be the case that distant generalization conditions involve stimuli that might activate a larger number of irrelevant properties that need to be inhibited or that might obscure the translation of the encoded learning stimuli to the generalization stimuli (see Thibaut et al., 2010, for similar reasoning with semantic analogies).

The course of novel noun generalization

Combining our results with results obtained with the single-object design allowed us to build a broader understanding of the role of semantic distance and levels of classification in novel noun learning. First, in the case of object nouns, we have seen that children's early nouns refer to the basic level of categorization—the so-called basic-level bias (Emberson et al., 2019; Waxman, 1990). In Experiment 1, given that there was no "same basic-level object" option that would be the "natural" selection in the no-comparison learning condition, children did not go for the taxonomic choice (same superordinate-level category) because it did not correspond to the basic-level bias. The great majority of them chose the perceptually similar options, a choice that is more in line with the shape bias (see Kucker et al., 2019, for discussion). What our data add is the finding that, despite the basic-level bias or the shape bias, comparison designs allow children to access broader superordinate levels. Our results also show that the breadth of a category depends on the semantic distance between items, with broader learning giving the best generalization results. In other words, our data reveal a hierarchy of generalization difficulty as a function of learning and generalization distance.

Relations to other works and free-choice designs

In our experiments, we used a forced-choice design to test whether children understood a particular targeted common feature (taxonomic or relational) in the presence of other options, which could be more or less salient (here, a perceptual distractor in the object noun of Experiment 1). This paradigm is well-suited for studying children's understanding of targeted relations or how they resist an a priori bias (see above). Other experimental approaches are possible. For example, a free-choice task might yield different results if children are asked whether there are other "daxes" or to find other daxes in a group of stimuli. They can then decide that no stimulus has the same name as the learning stimulus or that several stimuli could have the same name. In a forced-choice paradigm, it is possible that children might select one stimulus while considering that the other option is also possible. For example, recent studies have used a free-choice task within a Bayesian perspective. The reasoning regarding generalization was as follows. In a multiple-instance learning condition (usually three instances in the tested conditions, equivalent to our comparison condition), children might refer to what the authors call a "suspicious coincidence" (Xu & Tenenbaum, 2007) in the input. For example, if the three perceptually very similar instances have the same shape and same color, then there is good reason for children to conclude that these strong similarities mean that the items should belong to the same basic-level category, or even to the same subordinate category (Xu & Tenenbaum, 2007), rather than to the same superordinate category. With more diverse items (from different basic categories), participants have been found to select more diverse generalization items, that is, from the superordinate (and thus broader) category. Even though these results have led to various interpretations (see Spencer et al., 2011; Jenkins et al., 2015), this free-choice paradigm, which *is less constraining* in terms of what children can or might include, assesses which items might be included or "accepted" in the same set. The results seem to show that more variability leads to a more diverse set of selected stimuli. These results are consistent with our results showing that children choose more distant items with far learning examples in a forced-choice paradigm.

Final thoughts and limitations

There are probably other ways to manipulate semantic distance, especially in terms of the steps along the taxonomy scale (e.g., same basic-level categories, basic-level categories from the immediate superordinate-level category, basic-level categories from remote superordinate categories). We could have divided the scale into three or more distances. It would also be possible to manipulate the response format. In the current study, participants needed to choose among three options, to some extent making the task resemble a reasoning task. They might have spontaneously chosen none of the stimuli or more than one stimulus, whereas the task required them to choose only one. In reality, however, choices are open and children can decide not to include an item under a term and to use broader superordinate or nonspecific terms such as "stuff" and "thing." This kind of task might lead to differences in category extension. Therefore, it would be interesting to contrast an open task of this type with the current one in order to study the extension of novel terms in more detail. Finally, one might question our use of familiar stimuli given that children might have used their existing knowledge of nouns and their extension to categorize our novel nouns. Although this is a plausible objection, we actually capitalized on this existing knowledge. It would be impossible to build new stimuli or artifacts given that children would know nothing about their relations (if any). Using unfamiliar but recognizable stimuli, such as unfamiliar fruits, would be a possible strategy as long as participants recognized them as fruits or as specific but unknown types of basic-level fruit. However, this would be a viable strategy only provided that participants recognized them. Nevertheless, it is true that relying on familiar stimuli might limit the scope of our results. Importantly, one should recall that we worked with a forced-choice design, and the results showed that children differentially broadened their taxonomic and relational choices as a function of condition. It would be interesting to extend the current work by working with free-choice designs, younger children, or unfamiliar stimuli.

Data availability

Data will be made available on request.

Acknowledgments

The authors wish to thank the Agence Nationale de la Recherche (France) (ANR-18-CE28-0019-01 COMPARE, ANR Grant) and the Conseil Régional de Bourgogne Franche-Comté (PARI program) for their financial support, and to Marion Mallot, Jessica Miles and Eva Maire for designing the stimuli and/or collecting the data.

Appendix A

A1

List of materials for the close versus far comparison conditions, and the near versus far generalization conditions.

	Learning phase			Generalization phase		
	Standard item	Close item	Far item	Perceptual	Close taxonomic	Far taxonomic
Set 1						
Clothing accessories	bracelet1	bracelet2	watch	tyre	pendant	bow tie
Tools	hammer1	hammer2	axe	ostrich head	pincers	chainsaw
Clothing	sock1	sock2	jeans	pipe	sweater	hat
Food	apple1	apple2	cherry	bulb	banana	beefsteak
Animals	ladybird1	ladybird2	beetle	ball	butterfly	duck
Music player	guitar1	guitar2	cello	bottle	keyboard	Hi-Fi
Game/toy	ball1	ball2	cuddly toy	orange	Lego	video game
Set 2						
Food	pear1	pear2	strawberry	candle	pineapple	fry
Food2	pumpkin1	pumpkin2	tomato	ball	cucumber	grilled chicken
House tools/ appliances	broom1	broom2	feather duster	bush	vacuum	blender
Animals	snake1	snake2	lizard	rope	alligator	bird
Vehicle	bike1	bike2	scootering	glasses	rollerblade	boat
Office items	pencil1	pencil2	ruler	candy cane	scissors	laptop
Clothing accessories	knit cap1	knit cap2	hardhat	turtle	crown	bootsA

A2. Semantic and perceptual similarity ratings for object categories

Independent similarity ratings from 54 students confirmed that the close comparison object condition was semantically closer one to the other than the objects composing the far comparison pairs (M_{Close} = 5.52, SD = 0.39; M_{Far} = 4.44, SD = 1.06), $t(13) = 3.31$, $p < .006$ (Bonferroni-corrected p-value threshold .05/8 = .00625), and that close generalization stimuli were semantically more similar to the two learning stimuli than far generalization stimuli (M_{Near} = 3.75, SD = 0.58; $M_{Distant}$ = 2.48, SD = 0.80), $t(13) = 4.41$, $p < .006$. For the purpose of our experiment, it is also crucial that semantically related generalization items are perceptually less similar to the learning items than the perceptually similar lures. In this perspective, perceptual similarity ratings revealed that the perceptual choices were perceptually more similar to the learning items than the semantically related choices (taxonomic choices) in both the close and far conditions ($M_{Perceptual}$ = 4.77, SD = 0.54; M_{Near} = 2.13, SD = 0.62; $M_{Distant}$ = 1.86, SD = 0.69), $t(13) = 10.59$, $p < .001$ and $t(13) = 11.63$, $p < .001$, respectively. Importantly, we also performed perceptual similarity and semantic similarity ratings between the close comparison stimuli (e.g., two apples) and the far comparison stimuli (e.g., an apple and a cherry) on the one hand and the taxonomically related generalization item on the other hand. These ratings showed that the generalization stimuli were equally distant to both types of learning items. This was true for both types of generalization items (a) near generalization items: perceptual distance (M_{Close} = 2.18, SD = 0.68; M_{Far} = 2.02, SD = 0.59), $t(13) = 1.58$, $p = .14$; semantic distance (M_{Close} = 3.09, SD = 0.52; M_{Far} = 3.36, SD = 0.60), $t(13) = -2.96$, $p = .01$ (>.006); (b) distant generalization items: perceptual distance (M_{Close} = 1.88, SD = 0.61; M_{Far} = 1.86, SD = 0.74), $t(13) = 0.24$, $p = .81$; semantic distance (M_{Close} = 2.18, SD = 0.68; M_{Far} = 2.02, SD = 0.59), $t(13) = -0.19$, $p = .85$. This is central because we want to avoid performance differences between near and distant generalization items being due to perceptual but also semantic similarity differences between learning items. We included semantic similarity differences in order to keep only taxonomic distance influence. For example, if we get a difference between close and far generalization items (e.g., between jewel pendent and bow tie), we do not want it to be due to other semantic information (e.g., the fact that the jewel pendent would be more thematically related to bracelet than to the bow tie) than the taxonomic distance.

A3
Details of semantic and perceptual similarity ratings by categories.

Set	Category	Name	Semantic similarity ratings				Perceptual similarity ratings				
			Learning		Generalization		Learning		Generalization		
			Close	Far	Near	Distant	Close	Far	Near	Distant	Distractor
1	**1**	**"clothing accessories"**	5.93	4.77	3.57	2.93	6.77	5.15	2.73	1.98	5.21
	2	**"tools"**	5.87	3.36	3.48	2.86	6.92	5.77	2.22	1.48	4.77
	3	**"clothing"**	5.67	4.15	4.13	2.89	6.92	5.15	1.60	1.10	4.33
	4	**"food"**	6.00	4.92	4.31	2.27	6.85	4.23	1.92	1.69	4.28
	5	**"animals"**	5.80	5.07	3.72	1.71	6.62	5.46	2.85	1.23	4.73
	6	**"music player"**	5.53	5.38	4.86	3.16	6.85	5.77	1.56	1.50	5.69
	7	**"game/toy"**	5.73	1.41	3.51	2.38	6.08	4.77	2.10	1.83	5.22
2	**8**	**"food"**	5.46	4.93	4.44	2.59	6.69	5.31	2.10	1.75	4.52
	9	**"food2"**	5.62	4.87	4.11	1.92	6.23	5.23	1.96	2.28	5.40
	10	**"house tools/appliances"**	5.23	4.73	3.69	0.79	6.77	5.23	1.85	2.30	5.42
	11	**"animals"**	4.92	4.87	3.67	2.58	5.92	4.08	1.60	1.79	4.75
	12	**"vehicle"**	4.62	5.13	3.08	2.88	6.69	5.08	2.06	1.70	4.08
	13	**"office items"**	5.38	3.50	3.24	1.70	6.46	3.62	1.47	1.47	4.40
	14	**"clothing accesorries"**	5.54	5.13	2.64	4.13	6.85	4.62	3.75	3.94	3.96

A4
Proportion of taxonomic choices for each of the 14 object categories used in the two sets.

	Object category													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
No-comparison	.31	.40	.43	.40	.37	.50	.29	.34	.30	.36	.39	.28	.22	.45
Close comparison	.54	.57	.48	.41	.44	.62	.48	.48	.48	.30	.37	.41	.32	.55
Far comparison	.41	.58	.60	.57	.62	.78	.48	.66	.68	.47	.64	.71	.49	.67

Appendix B

B1. List of materials for Experiment 2 (relational nouns) for the close versus far comparison and the near versus far generalization conditions

Materials for Experiment 2 (relational nouns) for the close versus far comparison and the near versus far generalization conditions.

	Test choices				
	Entity	Operator	Taxonomic	Thematic	Relational
Set 1					
Learning phase “Cutter for”					
Standard	watermelon	knife			
Close comparison	orange	knife2			
Far comparison	meat	Cleaver			
Test phase					
Near generalization	sheet of paper		sheet of paper2	pencil	scissors
Distant generalization	beard	beard2	toothbrush	razor	
Learning phase “Home for”					
Standard	man1	house			
Close comparison	man2	house2			
Far comparison	woman	flat			
Test phase					

(*continued*)

	Test choices				
	Entity	Operator	Taxonomic	Thematic	Relational
Near generalization	American Indian		American Indian	American Indian2 bow	tepee
Distant generalization	dog		dog2	bone	doghouse
Learning phase "Baby of"					
Standard	cat	kitten			
Close comparison	dog	puppy			
Far comparison	deer	fawn			
Test phase					
Near generalization		bear	bear2	fish	bear cub
Distant generalization		woman	woman2	dress	baby
Learning phase "Clothes for"					
Standard	hand	glove			
Close comparison	hand2	mitten			
Far comparison	head	woolly hat			
Test phase					
Near generalization		foot	foot2	soccer ball	sock
Distant generalization		racket	racket2	tennis ball	racket cover
Learning phase "Product of"					
Standard	cow	milk			
Close comparison	goat	milk			
Far comparison	chicken	eggs			
Test phase					
Near generalization		bee	bee2	flower	honey
Distant generalization		tree	tree2	axe	apple
Learning phase "Wash for"					
Standard	teeth	toothbrush			
Close comparison	teeth2	toothbrush2			
Far comparison	ear	cotton bud			
Test phase					
Near generalization		hand	hand2	ring	soap
Distant generalization		car	car2	wheel	car wash

(*continued on next page*)

(*continued*)

	Test choices				
	Entity	Operator	Taxonomic	Thematic	Relational
Learning phase "Travel space for"					
Standard	car	road			
Close comparison	quad bike	country lane			
Far comparison	train	railway			
Test phase					
Near generalization		boat	boat2	sailor	sea
Distant generalization		rocket	rocket2	astronaut	space
Set 2					
Learning phase "Cutter for"					
Standard	log	saw			
Close comparison	plank	saw2			
Far comparison	hedge	hedge trimmer			
Test phase					
Near generalization		grass	grass2	flowers	lawnmower
Distant generalization		nail	nail2	nail varnish	nail clippers
Learning phase "House for"					
Standard	dog	doghouse			
Close comparison	dog2	doghouse2			
Far comparison	cat	basket			
Test phase					
Near generalization		bird	bird2	grain	nest
Distant generalization		man	man2	briefcase	flat
Learning phase "Baby of"					
Standard	woman	baby			
Close comparison	woman2	baby2			
Far comparison	dog	puppy			
Test phase					
Near generalization		lion	lion2	savanna	lion cub
Distant generalization		whale	whale2	bear cub	calf
Learning phase "Clothes for"					
Standard	tennis racket	tennis racket cover			

(*continued*)

	Test choices				
	Entity	Operator	Taxonomic	Thematic	Relational
Close comparison	ping-pong racket	ping-pong racket cover			
Far comparison	guitar	guitar case			
Test phase					
Near generalization		phone	phone2	phone charger	phone shell
Distant generalization		torso	torso2	legs	shirt
Learning phase “Product of”					
Standard	wheat	flour			
Close comparison	corn	popcorn			
Far comparison	coconut tree	coconut			
Test phase					
Near generalization		potato	potato2	peeler	fry
Distant generalization		fish	fish2	boat	fishstick
Learning phase “Wash for”					
Standard	plates	sponge			
Close comparison	glace	sponge2			
Far comparison	floor	mop			
Test phases					
Near generalization		patio	patio2	garden table	high-pressure cleaner
Distant generalization		laundry	laundry2	shoe	washer
Learning phase “Baby of”					
Standard	dog	puppy			
Close comparison	dog2	puppy2			
Far comparison	cat	kitten			
Test phase					
Near generalization		bear	bear2	fish	bear cub
Distant generalization		doll	doll2	doll clothes	baby dollB

B2. Semantic similarity ratings for relational categories

A total of 61 undergraduate students (49 female) enrolled in a psychology course (M_{age} = 25.15 years) were divided into four subgroups and completed online surveys. Two groups rated, on a 7-point scale, semantic similarity between learning pairs (operator–entity pairs between each other) and between learning items (entities or operators between each other). In the first task, participants judged the similarity between close and distant pairs. In the second task, participants rated the similarity between entities and between operators in close pairs (e.g., watermelon, orange; knife1, knife2) and distant pairs (e.g., watermelon, meat; knife1, knife3). Categories and sets were randomly assigned. Results revealed that close comparison pairs (M = 6.15, SD = 0.68) are perceived as more similar than far comparison pairs (M = 4.98, SD = 0.69), $t(13) = 6.28$, $p < .0001$. In addition, the entities and operators of the close pairs ($M_{Entities}$ = 5.72, SD = 0.58; $M_{Operators}$ = 5.78, SD = 0.81) are found to be more similar than those of the far pairs ($M_{Entities}$ = 3.82, SD = 1.21; $M_{Operators}$ = 4.49, SD = 0.89), $t(13) = 6.16$, $p > .0001$ and $t(13) = 6.43$, $p < .001$, respectively. The two other groups of participants rated, on a 7-point scale, semantic similarity between learning and generalization pairs and between entities or operators from learning and generalization items (entities or operators). In each group, participants saw 63 near pairs and 63 distant pairs so that all comparisons between pairs and items (near and distant) were made. Results revealed that near generalization pairs (M = 4.51, SD = 0.71) were judged to be semantically more similar to learning pairs than distant generalization pairs (M = 3.59, SD = 1.10), $t(13) = 4.74$, $p < .001$. Similarity ratings also confirmed that entities or operators in the near generalization pairs (e.g., sheet of paper and scissors) were more similar to the entities (e.g., watermelon, orange, meat) or operators (e.g., knife1, knife2, cleaver) in the close comparison pairs than in the far generalization pairs (e.g., bearded face and shaver). Entities in the near generalization pairs were significantly more similar to entities in the learning pairs (M = 2.90, SD = 1.42) than entities in the distant generalization pairs (M = 1.65, SD = 1.25), $t(13) = 3.98$, $p = .001$ (<.0083, Bonferroni-corrected p-value threshold), and operators in the near generalization pairs were significantly more similar to operators in the learning pairs (M = 3.61, SD = 0.86) than operators in the distant generalization pairs (M = 2.74, SD = 0.94), $t(13) = 3.99$, $p = .001$.

B3
Details of semantic similarity ratings by categories

Set	Category	Name	Pairs ratings				Item Ratings							
							Operators				Entities			
			Learning		Generalization		Learning		Generalization		Learning		Generalization	
			Close	Far	Near	Distant	Close	Far	Near	Distant	Close	Far	Near	Distant
1	**1**	Cutter for	6.25	4.875	3.41672794	2.08780637	6.375	5.6875	4.45654762	3.62470238	5.25	3.1875	0.56875	0.41666667
	2	Home for	6.5	4	4.92383578	4.31574755	6.3125	5.3125	4.80193452	4.06279762	6.3125	5.5	4.9953869	2.34002976
	3	Baby of	6.6875	5.625	5.34558824	4.82941176	5.125	3.9375	3.82782738	3.5016369	5.0625	3.625	3.2796131	2.56979167
	4	Clothe for	6.3125	5.5625	5.39644608	2.53327206	6.0625	5.5	4.91235119	1.97827381	6.5	4.3125	4.71116071	1.15550595
	5	Product of	5.9375	5.4375	4.73633578	3.70177696	5.5	4.375	1.9296131	1.28705357	5.25	4.5	4.17827381	4.24181548
	6	Whash for	5.25	5.375	4.59859069	3.71354167	5.375	4.5625	3.44598214	2.67053571	5	4.125	2.82068452	1.18720238
	7	Travel space for	6.625	5.375	4.40931373	3.00839461	5.9375	4	3.10699405	2.53511905	6.5625	4.4375	3.9047619	0.40327381
2	**8**	Cutter for	6.4375	5.1875	4.18229167	3.99816176	5.875	5.625	3.84285714	2.57529762	5.3125	5.125	0.81309524	0.4547619
	9	Home for	6.75	5.3125	4.63860294	4.03578431	6.375	4.4375	4.06622024	2.94181548	6.25	4.875	3.44419643	2.59553571
	10	Baby of	6.375	4.5625	5.04681373	4.78756127	6.1875	3.625	3.3578869	2.42053571	6.1875	3	2.528125	1.72931548
	11	Clothe for	6.75	4.5625	3.68063725	2.28958333	6	4.5625	3.02931548	1.16889881	5.625	1.6875	1.45342262	0.45431548
	12	Product of	4.375	3.3125	3.63106618	2.61973039	3.3125	2.625	2.70550595	2.04270833	5.0625	2.9375	2.87619048	1.56547619
	13	Whash for	5.4375	4.75	3.63823529	2.57230392	6.3125	3.5625	2.77321429	3.3735119	5.5	1.5625	1.30104167	0.43675595
	14	Baby of	6.375	5.8125	5.55128676	5.72322304	6.1875	5.0625	4.31919643	4.18452381	6.1875	4.625	3.80282738	3.54761905
Mean			6.14732143	4.98214286	4.51398372	3.5868785	5.78125	4.49107143	3.61253189	2.74052934	5.71875	3.82142857	2.90553784	1.64986182
SD			0.68448076	0.69485946	0.71447719	1.10315355	0.80999985	0.89464793	0.86241342	0.93903121	0.58411978	1.20890554	1.41838761	1.2530718

B4
Proportion of relational choices for each of the 14 relational categories used in the two sets.

	Relational category													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
No-comparison	.57	.49	.49	.37	.43	.49	.41	.52	.43	.44	.41	.30	.44	.43
Close comparison	.71	.54	.55	.61	.51	.61	.55	.63	.59	.59	.50	.47	.58	.54
Far comparison	.51	.57	.45	.51	.60	.53	.45	.51	.61	.48	.67	.52	.49	.47

References

Alfieri, L., Nokes-Malach, T. J., & Schunn, C. D. (2013). Learning through case comparisons: A meta-analytic review. *Educational Psychologist, 48*(2), 87–113. https://doi.org/10.1080/00461520.2013.775712.

Andrews, G., & Halford, G. S. (2002). A cognitive complexity metric applied to cognitive development. *Cognitive Psychology, 45*(2), 153–219.

Archambault, A., Gosselin, F., & Schyns, P. G. (2000). *A natural bias for the basic level?* Psychology Press.

Augier, L., & Thibaut, J.-P. (2013). The benefits and costs of comparisons in a novel object categorization task: Interactions with development. *Psychonomic Bulletin & Review, 20*(6), 1126–1132.

Barnett, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy for far transfer. *Psychological Bulletin, 128*(4), 612–637.

Callanan, M. A. (1989). Development of object categories and inclusion relations: Preschoolers' hypotheses about word meanings. *Developmental Psychology, 25*(2), 207–216.

Carey, S. (2010). Beyond fast mapping. *Language Learning and Development, 6*(3), 184–205.

Carey, S., & Bartlett, E. (1978). Acquiring a single new word. *Proceedings of the Stanford Child Language Conference, 15*, 17–29.

Childers, J. B. (2020). *Language and concept acquisition from infancy through childhood: Learning from multiple exemplars*. Springer International. https://doi.org/10.1007/978-3-030-35594-4.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum.

Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review, 82*(6), 407–428. https://doi.org/10.1037/0033-295X.82.6.407.

Diesendruck, G., & Bloom, P. (2003). How specific is the shape bias? *Child Development, 74*(1), 168–178.

Diesendruck, G., Gelman, S. A., & Lebowitz, K. (1998). Conceptual and linguistic biases in children's word learning. *Developmental Psychology, 34*(5), 823–839.

Emberson, L. L., Loncar, N., Mazzei, C., Treves, I., & Goldberg, A. E. (2019). The blowfish effect: Children and adults use atypical exemplars to infer more narrow categories during word learning. *Journal of Child Language, 46*(5), 938–954.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191.

Fenson, L., Dale, P. S., Reznick, J. S., Bates, E., Thal, D. J., Pethick, S. J., Tomasello, M., Mervis, C. B., & Stiles, J. (1994). Variability in early communicative development. *Monographs of the Society for Research in Child Development, 59(5, Serial No. 242)*.

Gathercole, S. E., & Baddeley, A. D. (1993). Phonological working memory: A critical building block for reading development and vocabulary acquisition? *European Journal of Psychology of Education, 8*(3), 259. https://doi.org/10.1007/BF03174081.

Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science, 34*(5), 752–775. https://doi.org/10.1111/j.1551-6709.2010.01114.x.

Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure mapping and relational language support children's learning of relational categories. *Child Development, 82*(4), 1173–1188. https://doi.org/10.1111/j.1467-8624.2011.01599.x.

Gentner, D., & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14*(4), 487–513. https://doi.org/10.1016/S0885-2014(99)00016-7.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*(1), 1–38.

Graham, S. A., Namy, L. L., Gentner, D., & Meagher, K. (2010). The role of comparison in preschoolers' novel object categorization. *Journal of Experimental Child Psychology, 107*(3), 280–290. https://doi.org/10.1016/j.jecp.2010.04.017.

Green, A. E. (2016). Creativity, within reason: Semantic distance and dynamic state creativity in relational thinking and reasoning. *Current Directions in Psychological Science, 25*(1), 28–35. https://doi.org/10.1177/0963721415618485.

Green, A. E., Kraemer, D. J. M., Fugelsang, J. A., Gray, J. R., & Dunbar, K. N. (2010). Connecting long distance: Semantic distance in analogical reasoning modulates frontopolar cortex activity. *Cerebral Cortex, 20*(1), 70–76. https://doi.org/10.1093/cercor/bhp081.

Hall, D. G., & Waxman, S. R. (1993). Assumptions about word meaning: Individuation and basic-level kinds. *Child Development, 64*, 1550–1570. https://doi.org/10.1111/j.1467-8624.1993.tb02970.x.

Hammer, R. (2015). Impact of feature saliency on visual category learning. *Frontiers in Psychology, 6* 451.

Heibeck, T. H., & Markman, E. M. (1987). Word learning in children: An examination of fast mapping. *Child Development, 58*(4), 1021–1034.

Imai, M., Gentner, D., & Uchida, N. (1994). Children's theories of word meaning: The role of shape similarity in early acquisition. *Cognitive Development, 9*(1), 45–75. https://doi.org/10.1016/0885-2014(94)90019-1.

Jenkins, G. W., Samuelson, L. K., Smith, J. R., & Spencer, J. P. (2015). Non-Bayesian noun generalization in 3-to 5-year-old children: Probing the role of prior knowledge in the suspicious coincidence effect. *Cognitive science, 39*(2), 268–306.
Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development, 8*(2), 113–139.
Kenett, Y. N. (2019). What can quantitative measures of semantic distance tell us about creativity? *Current Opinion in Behavioral Sciences, 27*, 11–16. https://doi.org/10.1016/j.cobeha.2018.08.010.
Klahr, D., & Chen, Z. (2011). Finding one's place in transfer space. *Child Development Perspectives, 5*(3), 196–204.
Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67*(6), 2797–2822. https://doi.org/10.1111/j.1467-8624.1996.tb01889.x.
Kucker, S. C., Samuelson, L. K., Perry, L. K., Yoshida, H., Colunga, E., Lorenz, M. G., & Smith, L. B. (2019). Reproducibility and a unifying explanation: Lessons from the shape bias. *Infant Behavior and Development, 54*, 156–165.
Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive Development, 3*(3), 299–321.
Lassaline, M. E., & Murphy, G. L. (1998). Alignment and category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 24*(1), 144–160.
Liu, J., Golinkoff, R. M., & Sak, K. (2001). One cow does not an animal make: Young children can extend novel words at the superordinate level. *Child Development, 72*(6), 1674–1694. https://doi.org/10.1111/1467-8624.00372.
Markman, A., & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology, 25*(4), 431–467. https://doi.org/10.1006/cogp.1993.1011.
Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. MIT Press.
Markman, E. M., & Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology, 16*(1), 1–27. https://doi.org/10.1016/0010-0285(84)90002-1.
Mervis, C. B., & Crisafi, M. A. (1982). Order of acquisition of subordinate-, basic-, and superordinate-level categories. *Child Development, 53*(1), 258–266.
Murphy, G. L. (2002). *The big book of concepts*. MIT Press.
Murphy, G. L., & Brownell, H. H. (1985). Category differentiation in object recognition: Typicality constraints on the basic category advantage. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*(1), 70–84.
Murphy, G. L., & Smith, E. E. (1982). Basic-level superiority in picture categorization. *Journal of Verbal Learning and Verbal Behavior, 21*(1), 1–20.
Namy, L. 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*(1), 5–15.
Namy, L. L., Gentner, D., & Clepper, L. E. (2007). How close is too close? Alignment and perceptual similarity in children's categorization. *Cognitie, Creier, Comportament/Cognition, Brain, Behavior, 11*(4), 647–651. 654–659.
Oakes, L. M., Kovack-Lesh, K. A., & Horst, J. S. (2009). Two are better than one: Comparison influences infants' visual recognition memory. *Journal of Experimental Child Psychology, 104*(1), 124–131. https://doi.org/10.1016/j.jecp.2008.09.001.
Oakes, L. M., & Ribar, R. J. (2005). A comparison of infants' categorization in paired and successive presentation familiarization tasks. *Infancy, 7*(1), 85–98.
Rosch, E. (1978). Principles of categorization. In E. Rosch & B. Lloyd (Eds.), *Cognition and categorization* (pp. 27–48). Lawrence Erlbaum.
Rosch, E., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8*(3), 382–439.
Spencer, J. P., Perone, S., Smith, L. B., & Samuelson, L. K. (2011). Learning words in space and time: Probing the mechanisms behind the suspicious-coincidence effect. *Psychological science, 22*(8), 1049–1057.
Thibaut, J.-P. (1991). *Récurrence et variations des attributs dans la formation des concepts*. University of Liège. Unpublished doctoral thesis.
Thibaut, J.-P., & French, R. M. (2016). Analogical reasoning, control and executive functions: A developmental investigation with eye-tracking. *Cognitive Development, 38*, 10–26. https://doi.org/10.1016/j.cogdev.2015.12.002.
Thibaut, J.-P., French, R., & Vezneva, M. (2010). Cognitive load and semantic analogies: Searching semantic space. *Psychonomic Bulletin & Review, 17*(4), 569–574. https://doi.org/10.3758/PBR.17.4.569.
Thibaut, J.-P., & Witt, A. (2015). Young children's learning of relational categories: Multiple comparisons and their cognitive constraints. *Frontiers in Psychology, 6*. https://doi.org/10.3389/fpsyg.2015.00643 643.
Tversky, A. (1977). Features of similarity. *Psychological Review, 84*(4), 327–352.
Waxman, S. R. (1990). Linguistic biases and the establishment of conceptual hierarchies: Evidence from preschool children. *Cognitive Development, 5*(2), 123–150.
Waxman, S. R., & Hatch, T. (1992). Beyond the basics: Preschool children label objects flexibly at multiple hierarchical levels. *Journal of Child Language, 19*(1), 153–166.
Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology, 36*(5), 571–581.
Waxman, S. R., & Leddon, E. M. (2011). Early word-learning and conceptual development: Everything had a name, and each name gave birth to a new thought. In U. Goswami (Ed.), *The Wiley-Blackwell handbook of childhood cognitive development* (pp. 180–208). Wiley-Blackwell.
Waxman, S. R., Shipley, E. F., & Shepperson, B. (1991). Establishing new subcategories: The role of category labels and existing knowledge. *Child Development, 62*(1), 127–138.
Winer, B. J. (1971). *Statistical principles in experimental design* (Vol. 2). McGraw–Hill.
Xu, F., & Tenenbaum, J. B. (2007). Sensitivity to sampling in Bayesian word learning. *Developmental Science, 10*(3), 288–297.
Zelazo, P. D., Carter, A., Reznick, J. S., & Frye, D. (1997). Early development of executive function: A problem-solving framework. *Review of General Psychology, 1*(2), 198–226. https://doi.org/10.1037/1089-2680.1.2.198.