Generalization of novel names for relations in comparison settings: the role of conceptual distance during learning and at test.

Jean-Pierre Thibaut, Ella Stansbury and Arnaud Witt
(jean-pierre.thibaut@u-bourgogne.fr ; arnaud.witt@u-bourgogne.fr; ellastansbury@gmail.com)
LEAD-CNRS, UMR 5022, Université de Bourgogne Franche-Comté,
Pôle AAFE – Esplanade Erasme, 21065 Dijon, France

Abstract

Relational categories are notoriously difficult to learn. We studied the impact of comparison on relational concept learning with a novel word learning task in 3- and 4-year olds. We contrasted a no-comparison (single) condition and two comparison conditions. In the latter case, the set of learning pairs was composed of either close or far pairs (e.g., close pair: knife1- watermelon, knife2-orange; far pair: ax-evergreen tree, saw-log, for the "cutter for" relation). We also manipulated the transfer stimuli semantic distance (near or distant semantic domain, e.g., a scissor for a piece of paper in the close case, and a shaver for a face in the far domain case). The no-comparison condition led to random generalizations in the younger group only. Overall the close learning condition and the near transfer condition led to good performance. We discuss these results in terms of the role of semantic distance and how participants integrate stimuli depending on distance.

Key words: relational categories, relational language, comparisons, conceptual distance

Introduction

One can distinguish categories of objects from relational categories. Object categories are mostly defined by perceptual properties such as shape, texture, color among which shape commonalities play an important role (Jones & Smith, 1993). Relational categories refer to relations between objects and not to the objects' properties themselves. For example, there is no set of perceptually stable properties defining the entities we call "neighbors". Rather, "neighbor" refers to a relation that can be described as "something which is close to something else" and various objects, people or abstract entities such as events can be called "neighbors".

Observations and systematic studies show that object categories are acquired rapidly when the child begins to acquire a lexicon (Golinkoff, & Hirsh-Pasek, 2008) and, according to many authors, before names for relational categories are acquired. Descriptively, as noticed by Gentner, Anggoro, and Klibanoff (2011), the MacArthur Communicative Developmental Inventory database reveals that object nouns are frequent in the 8- to 16-month period whereas relational nouns appear in the 17-30-month range. Moreover, children might first misunderstand relational terms as referring to object categories (e.g., Hall & Waxman, 1993), because they focus on the object's properties at the expense of the relations connecting them. More generally, tasks relying on common relations such as analogical tasks reveal that younger children often prefer *object* matches over *relational* matches (e.g., Richland, Morrison, & Holyoak, 2006; Thibaut, French, & Vezneva, 2010).

Thus, understanding which factors promote the relation-based abstraction and generalization is essential. One of these factors is comparison. The benefits of comparison over the presentation of single exemplars have been obtained for several conceptual linguistic domains, such as objects (e.g., Gentner & Namy, 1999; Augier & Thibaut, 2013), adjectives (e.g., Waxman & Klibanoff, 2000), action verbs (e.g., Childers & Paik, 2009), names for parts (Gentner, Loewenstein, & Hung, 2007), relational nouns (Gentner, Anggoro, & Klibanoff, 2011; Thibaut & Witt, 2015) and perceptual categories (e.g., Hammer, Diesendruck, Weinshall, & Hochstein, 2009). In the case of relational categories, Gentner, et al. (2011) tested in which conditions relational nouns such as "*X is the dax for Y*" would give better generalization, in a single or comparison learning design. The comparison design was built around two familiar objects that were connected by a familiar relation (e.g., *"cutter for")*: the first one was the "operator" (e.g., a knife) and the other the "entity" (e.g., a watermelon). In all the experiments, at test, an entity (e.g., a sheet of paper) was introduced with three alternatives (i.e., a relational match-- a pair of scissors--, a taxonomic match-- a pile of sheets of paper--, and a thematic match-- a pencil--). Children had to point to the stimulus among the three alternatives that was the *dax* for the piece of paper (Relational Label condition). The comparison condition led to better results than the single stimulus condition. The authors also showed that starting with a simple case (two pairs that were semantically close), helped participants finding the common relation with a third, semantically less close, pair. The authors argued that this progressive alignment helped young children grasp and generalize new relational categories.

In a recent study with 42-month-old children, Thibaut and Witt (2015) manipulated the number of pairs of pictures of objects illustrating a target relation used in the training phase (2, 3 or 4 pairs such as an apple and a knife for "the knife is the *dax* for the apple") and the distance between the conceptual domains illustrating the target relation. For example, a knife with an apple and another knife with an orange belong to close conceptual domains whereas a knife with an orange and a tree with a saw belong to more distant domains. Results revealed that three learning pairs was the optimal number and that learning pairs from distant domains gave better generalization than ones from close domains whereas four led to worse results. In terms of alignment, this means that increasing the number of pairs, thus converging evidence, does not lead to a linear increase in performance.

Augier and Thibaut (2013) provide an interpretation of this type of non linear influence of the quantity of information in terms of executive functions, arguing that comparisons generate cognitive costs. They manipulated the number of items to-be-compared in 4-year-old and 6-year-old children. They speculated that introducing more evidence in favor of the target dimension (texture) also means more comparisons and more information to integrate, generating more executive costs. They showed that only the older group benefited from an increased number of standards (four training standards versus two standards). In the above study, Thibaut and Witt (2015) followed the same theoretical framework and argued that a larger number of learning pairs impedes young children's capturing of the relevant dimensions. There is an optimal number of amount of information that younger children can integrate in this comparison design and in their case, with their age group, it was three pairs. Recent studies on semantic analogies (Richland, Morrison, & Holyoak, 2006; Thibaut and French, 2016) or perceptual analogies (Thibaut, French, & Vezneva, 2010) (see also Richland & Burchinal, 2013) support this cognitive costs hypothesis. In these studies, irrelevant perceptual features or semantic distractors or the semantic distance between the stimuli used in the pairs of the A:B::C:D analogies they studied, explained part of children's performance because finding out nonobvious relevant relations requires inhibiting superficial irrelevant dimensions and integrating more difficult dimensions.

Goals of the present experiment

Gentner et al. (2011) Thibaut and Witt (2015) manipulated the semantic distance between items in the learning context for relational categories. However they did not manipulate the semantic distance between learning items and generalization items. Thibaut and Witt (2017) introduced this distance factor in a comparison setting involving object categories. They tested 4- and 6-year-olds and manipulated the semantic distance between the two learning items (e.g., two bracelets versus a bracelet and a watch), and the semantic distance between the learning items and the test items (e.g., a pendant –near domain versus a bow tie, remote domain). They tested whether smaller semantic distance between learning items would lead to more taxonomic (vs. perceptual) choices at test, than broader semantic distance during learning, especially in the case of distant test stimuli. Results revealed main effects of learning distance, of generalization distance and that only children aged 6 years benefited from broader semantic distance during learning. Four year-old children were less efficient in their generalizations to far test stimuli even with semantically distant learning items. The authors argued that the generalization depth depends on the learning exemplars. Indeed, there is a large body of literature showing to what extent generalization depends on the nature of the training items (Son, Landy, & Goldstone, 2008), on the one side, and factors affecting the generalization width on the other side (e.g., Klahr & Chen, 2011). Thus, knowing at which distance children generalize is a main issue in the study of the ontogeny of categories, subordinate, basic, and superordinate categories.

The following experiment dealt with relational categories. We manipulated the semantic distances between both the learning items and the test items (in the generalization phase). By contrast with Thibaut and Witt (2015), we introduced a control no-comparison condition in order to assess which comparison conditions would benefit the most from the possibility to compare stimuli. Further, we compared two age groups (3- and 4-year-olds) in order to better understand how cognitive resources might interact with semantic distances. Our hypothesis was that we should obtain an interaction between age and one (or both) semantic distance factors or, at least, different generalization patterns in each age group. As argued by Thibaut et al. (2010), this should result from the fact that older children would more systematically explore the semantic space in less obvious cases. Third, we manipulated the distance between the training items domain and the generalization items (near or distant). As suggested by Augier and Thibaut (2013) children of different ages might not benefit from comparison situations in the same way depending on the distance between learning instances and/or the distance between learning items and generalization instances. For example, it might well be that both age groups would generalize similarly in the close learning and close generalization case which are built around very similar exemplars in both the learning and the transfer stimuli. However, younger participants might encounter more difficulties to capture conceptual similarities in the case of more distant learning items or to apply what they understood with training items to more distant domains. It might also be that all the comparison conditions will not outperform the no-comparison conditions, either because they are difficult to unify or because participants might grasp the relevant relation in the no-comparison condition.

Methods

Participants. One hundred and fifty-four french speaking preschoolers were tested individually in a quiet room at their school. Two age groups were recruited. The younger group was composed of 77 children (mean age = 3 years, 10 months; range: 40 - 51 months) and the older group was composed of 77 children (mean age = 4 years, 8 months, range: 53 – 60 months). All children were randomly assigned to one of the three experimental conditions with 52 (*no* comparison) 50 (*close* comparison) or 52 (*far* comparison) children per condition. Informed consent was obtained from their school and their parents.

Design. Three and four-year-old children were compared. This factor was crossed with Learning type (*no* vs. *close* vs. *far* comparison, between-subject factor) and Test distance (Near vs. Distant, within-subject 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 7 relational categories used in this experiment (i.e., *cutter for*, *home for*, *food of*, *baby of*, *container for, travel space for,* *cleaner for, product of)*. All training phase stimuli were organized around training pairs, that is one training pair in the no-comparison case and two training pairs in the comparison conditions (*close* or *far* training condition). Each pair was composed of an operator and an entity, (e.g., a knife as an operator and a watermelon as an entity, see Figure 1). The *no-comparison* condition pairs were composed of an operator-entity pair (e.g., either knife1-watermelon or knife2-orange or cleaver-meat), the *close* training pairs condition was composed of conceptually similar items (e.g., knife1-watermelon, knife2–orange), while the *far* pairs were composed of less conceptually similar pairs (e.g., knife1-watermelon, cleaver-meat). In this example knife1 and knife2 are conceptually closer one to the other than knife1 and cleaver. The same is true for watermelon-orange (close) compared to watermelon-meat (far). Note that in both the close and the far training conditions, each semantic relation was illustrated by two exemplars (e.g., there were two close "is the cutter for" training exemplars each composed of two different pairs), each being alternatively associated with a "nearer" generalization choice for half of the participants (e.g., a scissor for a piece of paper) or a "remote" generalization choice (shaver and beard) for the other half. The test cards consisted of four pictures. The relationally related correct answer was always the operator (e.g., a pair of scissors or a shaver). Independent participants (see below) rated the solution operator(s) as conceptually nearer to the training operator(s) pictures in the "near" generalization condition than the solution operator(s) in the distant generalization solution. For example, in Figure 1, the knife was rated as conceptually closer to the pair of scissors than to the shaver. There was also a taxonomic card choice (e.g., pile of sheets of paper which was taxonomically related to the entity "sheet of paper" in Figure 1), and a thematic card choice (e.g., pencil was thematically related with the entity sheet of paper in Figure 1). The top part of Figure 1 depicts the *close* and *far* training pairs for the "*cutter for*" relation in the learning phase, as a function of the learning condition (*no comparison* vs. *close* comparison vs. *far* comparison), and the 4 pictures introduced at test (entity and taxonomic, thematic or relational choice), as a function of the generalization distance (near vs. distant).

We forged 14 different bisyllabic labels (pseudo-words) which are, as shown by Gathercole and Baddeley (1993), easier to remember than monosyllabic pseudo-words (e.g., buxi, dajo, zatu, xanto, vira). Syllables were of the CV type which is the dominant word structure in French (from the French database, Lexique.org, New, Pallier, Brysbaert, & Ferrand, 2004). The pictures used in our experiment were realistic pictures like in Thibaut & Witt (2015). Independent similarity ratings from 61 students confirmed that *close* learning pairs were more conceptually similar one to the other than *far* learning pairs (similarity ratings: *close* pairs ($M = 6.15$, $SD = .68$), *far* pairs ($M = 4.98$, $SD = 0.69$), $t(26) = 4.47$, $p < .001$) and that *close* generalization pairs were more similar to learning pairs than *far* generalization pairs (similarity ratings: *close* pairs, $M = 4.51$, $SD = .71$; *far* pairs, $M = 3.59$, $SD = 1.10$, $t(26) = 2.64$, $p < .02$).

We also compared the similarity between entities (e.g., similarity between watermelon and orange - close case- or between watermelon and meat -far case-) and between operators in the close and the far learning pairs (e.g., similarity between knife1 and knife2 (close case), or between knife1 and cleaver (far case)). *Close* entities ($M = 5.72$, $SD = .58$) were significantly more similar one to the others than *far* entities ($M = 3.82$, $SD = 1.21$), $t(26) = 5.29$, $p < .0001$, and *close* operators ($M = 5.78$, $SD = .81$) were significantly more similar one to the others than *far* operators ($M = 2.74$, $SD = .94$), $t(26) = 3.99$, $p < .001$).

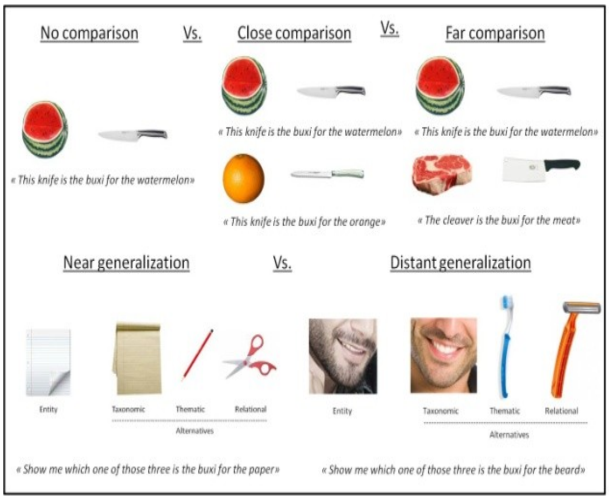

Figure 1: Example of a stimulus set and instructions adapted for the six experimental conditions resulting from crossing Learning Type (No comparison vs. Close vs. Far comparison) and Test distance (Near vs. Distant generalization) factors.

Similarity ratings also confirmed that entities or operators in the close generalization pairs (e.g., sheet of paper and scissors, respectively) were more similar to the entities (e.g., watermelon, orange and meat) or operators (e.g., knife1, knife2 and cleaver) in the learning pairs than entities and operators in the far generalization pairs (e.g., bearded face and shaver, respectively). Entities in the *close* generalization pairs were significantly more similar to entities in the learning pairs ($M = 2.90$, $SD = 1.42$) than entities in the *far* generalization pairs ($M = 1.65$, $SD = 1.25$), $t(26) = 2.48$, $p < .02$, and operators in the *close generalization pairs* were significantly more similar to operators in the learning pairs ($M = 3.61$, $SD = .86$) than operators in the *far* generalization pairs ($M = 2.74$, $SD = .94$), $t(26) = 2.56$, $p < .02$).

Procedure. Our procedure was close to Thibaut and Witt (2015) procedure. The stimuli were displayed on a laptop screen. The rate of presentation of the learning pair (or pairs) and the generalization stimuli was manually controlled. We illustrate it with the "*cutter for*" relational category. During the initial and test phases, the experimenter kept the speech flow constant, across items and experimental conditions (i.e. participants). Prosodic emphasis was added for the pseudo-words so that children noticed that the same label was used for each instance pair of stimuli (during the learning and the 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 learning 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* ***buxy***. *We are going to show him what buxy means.*" *"Look! (the knife1 and the watermelon appeared at the top of screen) this knife is the* ***buxy*** *for the watermelon. Now look, (the knife2 and the orange appeared below the knife1-watermelon pair) this knife is the* ***buxy*** *for the orange"*. In the no-comparison condition (only one pair was presented during the learning phase) the learning phase stopped after the first pair. The learning pair(s) remained in view during the entire trial until the child pointed to an answer.

Test phase: The test started with these instructions: "*Now let's look at this (at these) (gesturing across the learning pair(s)). You see how this(these) (gesturing across the operator(s)) is(are)* ***buxy(ies)*** *for this(these) (gesturing across the entity(ies))? Now it's 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 buxy for the paper?*" In order to avoid answers before children analyzed the three test cards, we asked them to refrain themselves from answering before a picture of Yoshi appeared on the screen (4-5 seconds after the test cards appeared).

Children chose among the three test cards by pointing to the one on the screen that was the *buxy* 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. The extent to which children learned relational categories during the initial phase was assessed by coding the proportion of relational choices made at test, as well as the proportions of alternative choices (taxonomic and thematic choices). For each participant, the number of relational choices was calculated and the proportion of relational choices was computed for the 7 near transfer trials and the 7 distant transfer trials (7 + 7 trials for a total of 14 trials).

Results. The purpose of this experiment was to assess which of the three learning conditions (no-comparison, close comparison, far comparison) would give the best transfer performance as a function of the transfer item conceptual distance and age. We ran a three-way analysis of variance (ANOVA) with Age (*3* and *4 years*) and Learning type (*no comparison*, *close comparison*, *far comparison*) as between factors and Test distance (*Near*, *Distant*) as a within factor. We also wanted to identify *within* each age group, which comparison condition significantly differed from the corresponding no-comparison condition, and which comparison conditions differed from one another. Finally, we compared the proportion of relational answers to chance.

The ANOVA revealed significant main effects of Age, $F(1,148) = 14.31$, $p < .001$, $\eta^2_p = .09$ (Younger = .46; Older = .57); Learning Type, $F(2,148) = 10.38$, $p < .001$, $\eta^2_p = .12$ (no-comparison = .43; close training = .58, far training = .53); a posteriori comparisons (Tukey HSD) revealed that the No-comparison condition was significantly lower than the two comparison conditions ($p < .001$), which did not differ one from the other (p = .10). There was also a main effect of Test distance, $F(1, 148) = 7.63$, $p < .01$, $\eta^2_p = .04$ (near generalization = .55, distant generalization = .49). No interaction reached significance. Differences between each comparison condition (2 learning) and the corresponding no-comparison condition within each age group was assessed with planned comparisons for each test distance. Indeed, here we were interested by differences between training conditions. In the four-year-old group, the analyses revealed a significant difference between the no-comparison condition and close learning for both near ($F(1, 148) = 11$, $p < .01$) and distant ($F(1, 148) = 10.93$, $p < .01$) generalizations. There was no difference between the no-comparison and the far learning condition for the two generalization conditions (near, $F(1, 148) = 2.34$, $p = .13$, and distant, $F(1, 148) = 1.46$, $p = .23$). In the older group, the no-comparison group was significantly lower than in the close-learning-distant-generalization case, $F(1,148) = 9.97$, $p < .01$, and the far learning-near-generalization condition, $F(1,148) = 4.89$, $p < .05$. We also compared the performance between learning comparison conditions (close versus far learning) within each generalization condition (near versus distant generalization). In younger children, in both the close and the far learning case, there was no difference between near and distant generalization.

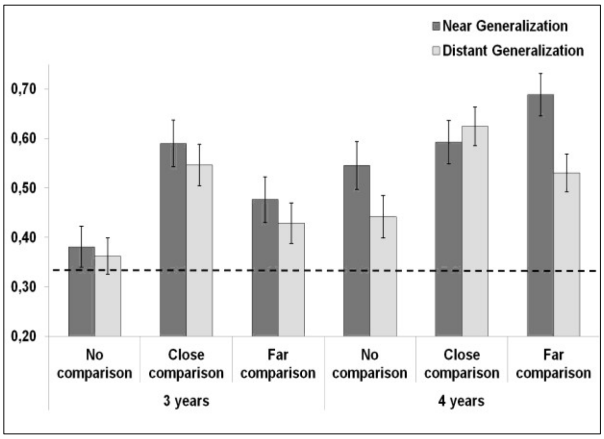

Figure 2. Proportions of relational choices as a function of the learning (no-comparison vs. close comparison vs. far comparison) and the generalization conditions (near generalization vs. distant generalization). The error bars correspond to one standard error and the dashed line represents chance level (33.33%) between the taxonomic, the thematic and the relational choices.

As for the comparison between close and far learning, they differed in the distant generalization case, $F(1,148) = 4.03$, $p < .05$, but not in the near generalization case, $F(1,148) = 2.94$, $p = .09$. In the older group, in the close learning case, there was no difference between near and distant generalization, $F < 1$, whereas in the far learning case, near generalization was significantly better than the distant generalization, $F(1,148) = 10.91$, $p < .01$. As for the comparison between close and far learning, they did not differ statistically in both generalization cases, respectively $F(1,148) = 2.46$, $p = .12$ and $F(1,148) = 2.97$, $p = .09$.

Finally, all the conditions significantly differed from chance ($p_s < .01$) except the two no-comparison conditions in the younger group ($p_s > .20$). Thus, the older group captured part of the relational meaning in the no comparison case (see general discussion) (see Figure 2).

Discussion

We contrasted no-comparison and comparison conditions in two groups of children. The purpose was to assess which conditions would give the best generalization results as a function of conceptual distance between learning items and between learning items and generalization items. Our results show that the two age groups did not benefit from comparison in the same way, but also that comparison conditions did not always outperform the no-comparison conditions. Younger children were significantly better in the close training comparison conditions than in the no-comparison conditions which did not differ significantly from chance (see Figure 2). The latter result suggests that younger children could not make sense of the target relation in the no-comparison case and that the optimal condition was the close training condition. The far training case seems to be in between, since it did not differ significantly from any of the two conditions (no-comparison and close training), despite being significantly different from chance, which tells us that these children could make sense of the target relations in the far training case. Older children had more contrasted results. They significantly differed from chance in the two generalization conditions, which suggests that they were able to understand the target relation despite the scarcity of information. The far training condition provides an interesting case: older children (4-year olds) were better in the near-transfer-far-training than in the remote-transfer-far-training condition, whereas younger children (3-year olds) had similar in these two far training conditions. As shown by figure 2, this is due to the fact that the near condition in the far training case was the easiest condition at age four whereas they both remained low in the younger age group. This pattern makes sense. It can be argued that the far training condition is more difficult to unify than the close training condition because it is more difficult to grasp commonalities between items from far domains, and it is still a step further to apply it to remote objects. On the other hand, when you understand the relation between far training domains, applying it to a near domain is rather straightforward whereas going beyond this already difficult first task involves extra comparisons and a need to unify very dissimilar cases. For younger kids, the far training case seemed to be very difficult, which explains why younger kids had low performance in both generalization cases (no difference with the no-comparison case).

Taken together, these results are important because they show that conceptual distance between learning items and between learning and generalization items have separate contributions. Gentner et al. (2011) showed that relational language compared to no relational language is important for conceptualization of relations. Thibaut and Witt (2015) found that both the number of training relations and conceptual distance between training items had a significant effect on performance. Here we extend these results and show how, depending on age, conceptual distance within training stimuli and between training and generalization items are important and impact performance. Younger children increased their performance only for close training items, suggesting difficulties to capture relational regularities for far training items. The next step, for older children, was to capture relational regularities for far cases but only in near generalization cases, the distant generalization case being beyond their reach (in our paradigm). In sum, an analogy between the present results and Thibaut and Witt (2015) can be drawn. In Thibaut and Witt (2015) participants' best results were obtained for the three-pair case and declined in the four-pair case. Here the older children's best performance was obtained for far-training-close transfer. As it has been argued earlier, going beyond that point would require unifying more disparate cases, which probably involves more comparisons, deeper encoding, and inhibition of a larger number of irrelevant properties that are spontaneously activated during the comparisons.

The fact that in older (4-y-o) children, the no-comparison conditions led to beyond chance results suggests that 4-year-old children could reliably use the information conveyed by a single exemplar of the target relation to make sensible hypotheses about the relation. Note that our results are consistent with Gentner et al. (2011), since older children were also beyond chance in their no-comparison condition. Figure 2 suggests that this can be achieved for near generalizations particularly easily. To some extent, the comparison taking place at test between the single stimulus and the transfer object is analogous to the one taking place between training items in the close training case which explains the similar performance. This means that the test stimuli can be considered as a opportunity to compare the training stimulus with other stimuli and to refine the dimensions one abstract from the training pair, including the target relation which, at first glance, might not have been salient for participants when the training pair was initially considered.

Overall these results underline a complex but consistent scale of difficulty. Younger children needed to compare to find out the target relation, but had their best performance in the close training case. Their performance decreased in the far training case suggesting that conceptually distant training cases make relational extraction difficult. Do their difficulties arise because they fail to align the entities and the operators in terms of role, between pairs, or do they fail to understand that the relation illustrated in each pair is the same? Or this understanding might be so fragile that they cannot apply it to novel instances at test.

The older group could reliably extract the relation from one pair which suggests that once they found the relation, the depth of its conceptual implementation was sufficient to allow a generalization towards novel instances. The fact that their best game seemed to be in the far training-close generalization case suggests that they could handle more variability but that this possibility seemed to remain limited to close categories.

Final thoughts: there are probably other ways to manipulate the semantic distance, especially in terms of the steps along the taxonomy scale (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 3 or more distances. Also it is possible to manipulate the response format. In the present case, participants had to choose between three options, which gives the task the flavor of a reasoning task. They might spontaneously choose none of the stimuli but the task requests them to choose one, as in any forced choice task. By contrast, in the real world, choices are open and children can decide to not include an item under a term and to use broader superordinates or non-specific terms like “stuff”, “thing”. This kind of task might lead to differences in category extension. Thus it would be interesting to contrast such an open task to the present one to more thoroughly study the extension of a novel term.

Acknowledgments

The authors wish to thank the Conseil Regional of the Bourgogne Franche-Comté (PARI program) for their financial support, and to Marion Mallot, Jessica Miles and Eva Maire for collecting the data.

References

Augier, L., & Thibaut, J.P (2013). The benefits and costs of comparisons in a novel object categorization task: Interactions with development. *Psychonomic Bulletin & Review*, 2, 1126-1132.

Childers, J. B., & Paik, J. H. (2009). Korean-and English-speaking children use cross-situational information to learn novel predicate terms. *Journal of Child Language*, *36*(1), 201-224

Gentner, D, & Namy, L. L. (1999). Comparison in the development of categories. *Cognitive Development, 14*(4), 487–513.

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), 1173–88.

Gentner, D., Loewenstein, J., & Hung, B. (2007). Comparison facilitates children’s learning of names for parts. *Journal of Cognition and Development, 8*(3), 285–307.

Goldstone, R. L., Day, S., & Son, J. Y. (2010). *Comparison*. In B. Glatzeder, V. Goel, & A. von Müller (Eds.) On thinking: Volume II, towards a theory of thinking. Heidelberg, Germany: Springer Verlag GmbH. (pp. 103-122).

Golinkoff, R. M., & Hirsh-Pasek, K. (2008). How toddlers begin to learn verbs. *Trends in cognitive sciences*, *12*(10), 397-403.

Hall, D. G., & Waxman, S. R. (1993). Assumptions about Word Meaning: Individuation and Basic-Level Kinds. *Child Development*, *64*(5), 1550-1570.

Hammer, R., Diesendruck, G., Weinshall, D., & Hochstein, S. (2009). The development of category learning strategies: What makes the difference?. *Cognition*, *112*(1), 105-119.

Jones, S. S., & Smith, L. B. (1993). The place of perception in children's concepts. *Cognitive Development*, *8*(2), 113-139.

Klahr, D., & Chen, Z. (2011). Finding one’s place in transfer space. *Child Development Perspectives, 5*(3), 196-204.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The importance of shape in early lexical learning. *Cognitive development, 3*(3), 299-321.

New, B., Pallier, C., Brysbaert, M., & Ferrand, L. (2004). Lexique 2: A new French lexical database. *Behavior Research Methods, Instruments, & Computers*, *36*(3), 516-524.

Richland, L. E., & Burchinal, M. R. (2013). Early executive function predicts reasoning development. *Psychological science, 24*(1), 87-92.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children’s development of analogical reasoning: Insights from scene analogy problems. *Journal of experimental child psychology, 94*(3), 249-273.

Son, J. Y., Smith, L. B., & Goldstone, R. L. (2008). Simplicity and generalization: Short-cutting abstraction in children’s object categorizations. *Cognition, 108*(3), 626-638.

Thibaut, J.P, French, R.M., & Vezneva, M. (2010). The development of analogy making in children: Cognitive load and executive functions. *Journal of experimental child psychology, 106*(1), 1-19.

Thibaut, J. P., & Witt, A. (2015). Young children's learning of relational categories: multiple comparisons and their cognitive constraints. *Frontiers in psychology, 6* : 643.

Thibaut, J. P., & French, R. M. (2016). Analogical reasoning, control and executive functions: a developmental investigation with eye-tracking. *Cognitive Development, 38*, 10-26.

Thibaut, J. P., & Witt, A. (2017). Generalizing novel names in comparison settings: Role of conceptual distance during learning and at test. In Proceedings of the 39th Annual Meeting of the Cognitive Science Society (pp. 3314-3319).

Waxman, S. R., & Klibanoff, R. S. (2000). The role of comparison in the extension of novel adjectives. *Developmental Psychology, 36*(5), 571–581.