Title
Solid ground makes solid understandings: does simple comparison paves the way for more complex comparisons?

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 43(43)

ISSN
1069-7977

Authors
Lagarrigue, Yannick
Thibaut, Jean-Pierre

Publication Date
2021

Peer reviewed

Solid ground makes solid understandings: does simple comparison paves the way for more complex comparisons?

Yannick Lagarrigue (yannick_lagarrigue@etu.u-bourgogne.fr)
Jean-Pierre Thibaut (jean-pierre.thibaut@u-bourgogne.fr)
LEAD-CNRS, UMR 5022, Université de Bourgogne Franche-Comté,
Pôle AAFE –Esplanade Erasme, 21065 Dijon, France

Abstract

In this experiment, we investigated the role of dimensional distinctiveness on the generalization of novel names for unfamiliar objects. In a comparison design, we manipulated the sequence of trials difficulty, starting either with more difficult trials or with easier trials. To achieve this, we manipulated the dimensional distinctiveness of the first comparison trials and of the, later, transfer trials. Results showed that high-distinctiveness (easy) stimuli increased children's later performance in the low-distinctiveness (difficult) condition whereas low-distinctiveness early training led to no later improvement in easier trials. Last, a correct answer for the first trial in the first learning part predicted the level of performance in the second learning part. We interpret these findings in terms of differential costs of comparison for varying levels of distinctiveness and level of abstraction from one condition to another.

Keywords: Generalization; Word learning; Comparison; Transfer; Distinctiveness; Preschoolers.

Introduction

Children learn novel words in situations in which they encounter a target stimulus and a word. For example, a dog and its name, "Look, this is a dog". Later on, they will have to use this novel word despite broad dissimilarities (contextual and perceptual) between new dogs and the first one. When they learn novel nouns, children's challenge is to sample properties that are central for categories the nouns refer to (Murphy, 2002). Indeed, in many novel noun learning situations, irrelevant superficial and salient similarities or differences can be more cognitively prominent than variations of more relevant dimensions. The present paper compares several novel noun learning conditions in terms of their respective level of generalization.

Learning a novel word consists in finding out which features are relevant to decide whether or not a new object is also a member of a target category (Gentner & Namy, 1999; Jones & Smith, 1993). Former studies have shown that young children often extend novel object names according to their shape (e.g., Imai et al., 1994; Landau et al., 1988; Smith et al., 1996). Given that shape or other salient features can be conceptually irrelevant, understanding which situations facilitate generalizations based on relevant but non salient properties is a crucial issue for concept development.

In the context of novel noun generalization, it has been repeatedly shown that the opportunity to compare several learning examples (at least two) of a taxonomic category (e.g., apple and pear for fruits) associated with a common novel name, help children to generalize novel words taxonomically rather than perceptually (Augier & Thibaut, 2013; Gentner & Namy, 1999, Graham et al., 2010). In contrast, no-comparison situations tend to favor perceptually-based word extensions (e.g., Imai et al. 1994). For example, if children are presented with an apple introduced with a pseudoword (e.g., "This is a buxi") and asked which one between a banana (taxonomic match) and a balloon (perceptual match) could also be a "buxi", they point to the perceptual match (the balloon) beyond chance. In contrast, if they are presented with an apple and a pear introduced with the same pseudoword (e.g., "This is a buxi", "This is also a buxi"), children select the taxonomic match beyond chance (Gentner & Namy, 1999).

However, a recent meta-analysis on the effects of comparison showed that all comparison situations are not always associated with better performance (Alfieri et al., 2013). For instance, Thibaut and Witt (2015) manipulated the number of training items (two, three or four pairs illustrating the relation). Three training pairs gave the best performance whereas, surprisingly, four pairs significantly decreased performance (see also Augier & Thibaut, 2013; Simms et al., 2018, for similar results). The authors argued that the three-pair condition was the best compromise between informativeness and cognitive demands for the targeted age group. Augier et Thibaut (2013) also showed that the effects of comparison are modulated by age, with older children (five- to six-year-old) benefitting more from complex comparison situations than younger children (three- to four-year-old). In the present research, capitalizing on previous results, we introduced comparisons in sequences and manipulated the difficulty of the comparison trials in order to better understand which sequence of comparisons trials would lead to better results. In other words, the question of interest was to understand whether benefits in a first sequence of comparison trials would influence (positively or negatively) following trials.

The role of sequences in learning has already been studied in previous studies but not in the framework of comparison. For example, Carvalho and Goldstone (2014) showed that blocked (i.e., the same stimuli repeated in a row) presentations led to more discoveries of commonalities among objects within the same category than interleaved (i.e., different stimuli interleaved in a sequence) presentations. Another conceptualization of sequences of learning trials is to think of them as a succession of trials, starting with concrete trials followed by less concrete later trials. Goldstone and Son (2005) showed that performance on pattern learning was significantly better in what they called a concreteness fading paradigm, which is supposed to promote more decontextualized representation, hence more transferable. Similarly, Kotovsky and Gentner (1996) showed that four year old who first saw easier instances of a relational similarity, were more able to abstract more difficult instances in later trials that would stay out of reach without the former easier trials. In Gentner et al. (2007), when children were first presented with high-similarity training stimuli, they were better to identify the common part than when they saw low-similarity training stimuli.

One popular interpretation of these experiments is the theory of progressive alignment (Gentner, 1983 ; Gentner & Rattermann, 1991). According to this theory, when children compare two objects, they first find perceptual similarities that ground the comparison and, then, discover less salient, deeper commonalities. Our proposal goes further and suggest that the process of comparison induce a change in the representation of the objects presented that could impact the result of a following process of comparison. More precisely, we examine preschool children's ability to learn and transfer across comparison situations that share a common underlying relevant feature (i.e., texture) which differ in in the way it is implemented. If feature alignment is an important step for successful learning from comparisons, then the difficulty of feature alignment in a first phase might influence identification and alignment of those features in a following comparison phase. In our study, we implemented a sequence of comparison trials and manipulated the easiness to identify the relevant feature. The stimuli were defined along two main dimensions, shape and texture, with shape being the salient bur irrelevant feature and texture the less salient but relevant feature.

In a noun generalization task with unfamiliar stimuli, Augier and Thibaut (2014) manipulated age (three- to four-year-old vs. five- to six-year-old) and the dimensions' distinctiveness (high vs. low) in comparison situations. According to Hammer and Diesendruck (2005, p.145) "the physical differences between stimuli, that is, their distinctiveness, affect people's (…) ability to discriminate between stimuli or to perceive them as similar". It means that a dimension (e.g., shape, texture, color) that would not be very different between two stimuli might be very difficult to identify and thus difficult to align. For example, the intuition tells us that a circle and a square are very different shapes, that is they are highly distinctive. In contrast, two oval shapes also differ in their shape, but the difference seem to be less pronounced, hence they have less distinctive shapes. In Augier and Thibaut study, both shape and texture were distinctive (i.e., very different) in the high-distinctive condition and both shape and texture were less distinctive (i.e., less different) in the low-distinctive condition. Their results showed that when the distinctiveness was high, both groups of age benefitted from comparison and achieved high categorization performance. When the distinctiveness was low, older children found the relevant dimension more often than chance in the comparison condition but not younger children. It means that the effects of comparison depend on dimensions' distinctiveness. The authors interpreted their finding in terms of complexity to integrate nonobvious commonalities when the differences were less easily noticed. However, even in the low distinctive case younger children still benefitted from comparison compared to a no-comparison situation. It means that they managed, sometimes, to find the relevant dimensions. In our study we manipulated the dimension's distinctiveness in the same way as Augier and Thibaut (2014) and considered high-distinctiveness comparison as easier comparison situation than low-distinctiveness comparison (see Figure 1 for an example of the stimuli used).

The present study

Our main research question was, whether comparisons not only support immediate taxonomic relation identification but also can confer insight that potentiates future comparisons. As in Augier et Thibaut (2014), we investigated this question in two group of age four-year-old and five-year-old. We compared three conditions. In the High-Low condition, high-distinctiveness objects were introduced before low-distinctiveness objects. In the Low-High condition, low distinctiveness objects were presented before high distinctiveness objects. In the Low-Low condition, low-distinctiveness objects were presented in both phases. In the High-Low condition, the purpose was to investigate if the understanding of the relevant dimension in the easier situation (high-distinctiveness) would be transferable to a less clear situation of comparison (with low-distinctive objects). The Low-Low condition was supposed to test whether there is a general practice effect, that is if improvements in the second – transfer phase could result from a larger number of practice trials. The purpose of the Low-High condition was to test if the level of abstraction obtained in the low distinctive situation thanks to comparison, compared to a no-comparison situation, would boost performance in easier cases of comparison.

In the first phase of experiment, the practice phase, we expected to replicate Augier and Thibaut (2014) results, i.e., (a) find more texture-based responses in the older group of children (b) In both age groups, better performance with high-distinctiveness objects than low distinctiveness case (c) less benefits of comparison with younger children.

Based on the progressive alignment theory suggesting that providing easily alignable dimensions first should help alignment of those features latter on, we predict better performance in the low condition after practice with high-similarity comparisons (High-Low condition). Indeed, if the level of abstraction achieved in the high-distinctive condition is sufficient to be applied in the low condition, comparison with distinctive objects may bootstrap transfer with less salient features' objects. In the Low-Low condition, predictions might differ for older and younger children. Given that only older children found the relevant dimension thanks to comparison in the low-distinctiveness condition, we hypothesized that only them could benefit from repetitive comparison with this level of distinctiveness. Then, children who fail to find the correct answer in the first trials might also fail in the transfer phase. In the low-high condition, the issue is more open: if comparison led to better abstraction, as shown by Augier and Thibaut, this abstraction might be sufficient to boost later performance in an easier condition. On the other hand, difficulties with this condition might lead to worse performance.

A last purpose, was to study whether participants who succeed in the second, transfer, phase found the correct answer early on or progressively. The first possibility would predict a positive correlation between the first trial and the performance in the transfer phase. On the contrary, an incremental view of learning would predict that given that children benefit from comparison, more trials should lead to better performance.

Methods

Participants

150 preschoolers participated in this study and were tested individually at school. They were split in two age groups: 90 younger (57 females, mean age = 52.83 ± 5.5m, range: 36-59m) and 60 older (32 females, mean age = 67.86 ± 4.9m, range: 60-76m). They were randomly assigned to one of the three experimental conditions with 20 children per condition for the older and 30 children per condition for the younger. We recruited more older children because even in the low-distinctiveness condition, they already have good performances so there is not much room for improvement. Informed consent was obtained from their school and their parents. The procedure was in accordance with the Declaration of Helsinki and followed institutional ethics board guidelines for research on humans.

Design

We used a forced-choice categorization task, in which children had to decide which of two simultaneously presented objects was of the same kind as the standard(s). Each child saw two familiarization trials followed by five practice trials and five transfer trials. No feedbacks were given. Younger and older groups of children participated in one of the three between-subject experimental conditions: High-Low, Low-Low or Low-High. In the High-Low condition, high-distinctiveness objects were introduced in the first phase (i.e., practice phase) and low-distinctiveness objects in the transfer phase. In the Low-Low condition, low-distinctiveness objects were presented in both practice and transfer phases. In the Low-High condition, low-distinctiveness objects were presented in the practice phase and high-distinctiveness objects in the transfer phase.

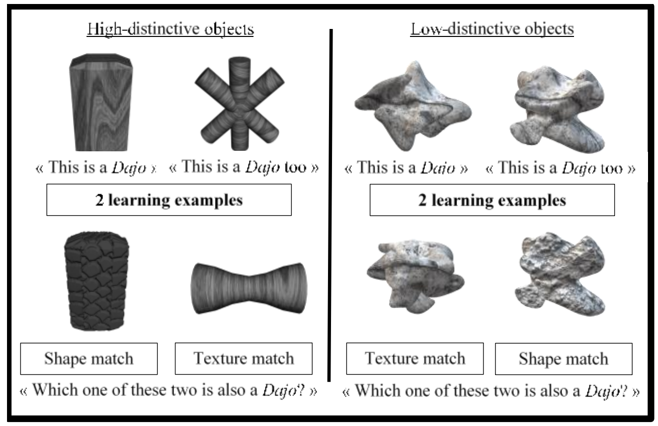

Figure 1: Example of sample stimulus sets and instructions used in the two levels of distinctiveness.

Materials

Twenty-one sets of four artificial grey-scale objects depicted on laminated cards were adapted from Augier and Thibaut (2013), fourteen for the low-distinctiveness condition and seven for the high-distinctiveness condition. Each set was composed of two standards and two test objects. The two standards had the same texture but different shapes (see Figure 1). The first test object, the shape match, had the same shape as one of the two standards but differed in its texture. The other test object, the texture match, had the same texture as both standards but had a different shape (see Figure 1).

Each object was printed on a 12 cm by 9 cm laminated card. The objects' textures and shapes in the low distinctiveness condition were created to be less distinctive (Figure 1, left panel) than in the high distinctiveness condition (Figure 1, right panel). Seventy-five adults rated the distinctiveness of pairs of objects on a 7-point Likert-like scale (ranging from not similar at all to extremely similar). Objects presented in pairs of different-texture-but-same-shape objects were judged to be significantly less distinctive in the low-distinctiveness condition (Mean = 3.48 ± 1.29) than in the high-distinctiveness condition (Mean = 1.67 ± 1.19), $t(74) = 15.85$, $p < .001$. Objects presented in pairs of different-shape-but-same-texture pairs of objects were judged to be significantly more similar in the low distinctiveness set (M = 3.47, SD = 1.19) than in the high-distinctiveness set (M = 1.81 ± 1.10), $t(74) = 13.82$, $p < .001$. Textures and shapes that were used in one set differed from those used in the other sets. The order in which the sets were presented was pseudo-randomized within and across participants.

We also created twelve different bi-syllabic labels (pseudo-words). Each of these novel names (Youma, Buxi, Dajo, Zatu, Sepon, Xanto, Vira, Loupo, Sampi, Loga, Kufa, and Budan) was randomly associated with one of the twelve sets performed by each child. The names were pseudo-randomized in a counterbalanced design across sets and participants.

Procedure

Children were introduced to the puppet “Yoshi” who “lives very far away”. They were told that Yoshi had some unknown objects with strange names that they would have to learn. The experimenter introduced the first standard with a novel-name (e.g., “This is a Dajo”) and children were asked to repeat the novel word. Then, a second standard was introduced with the same label as the first one (e.g., “This is a Dajo too. They both are Dajo.”). The two standards were presented in a row and their location was determined randomly. Children were asked to look carefully at the objects. In the forced-choice test phase, the two test objects (i.e., the shape and the texture match) were introduced simultaneously and children were asked to point at the one that was also a member of the category (e.g., “Show me which one of these two is also a Dajo.”). The test objects were presented in a row and their location was determined pseudo-randomly.

Results

Anova analysis We first conducted an ANOVA with 2 Age (young vs. old) and 3 Conditions (High-Low vs. Low-Low vs. Low-High) as between subject factors on the percentage of texture match choices at the transfer phase. Older children made significantly more texture choices (77% ± 32.69) than younger children (61% ± 35.30) (F (1, 144) = 12.42, p < .001, η^2_P = .08, BF_{10} = 39.56). The ANOVA also revealed a significant effect of Condition (F (2, 144) = 3.67 p = .003, η^2_P = .05, BF_{10} = 1.54). All groups of post-hoc were computed using separate paired-samples comparison and the sequentially acceptive step-up Bonferroni procedure, with an initial alpha level of .05 (Hochberg, 1988).

To explore more precisely this result on the transfer phase with low-distinctiveness objects, we ran t-tests as post-hoc analysis for each age group. In the younger group, there were more texture-based categorizations in the transfer phase in the High-Low condition (M = 67%, SD = 33.60) than in the Low-Low condition (M = 47%, SD = 30.00), t (58) = 2.42, p = .019, d = .6, BF_{10} = 2.88). This means that in the transfer phase with low-distinctiveness objects, younger children significantly benefited from practice with high-distinctiveness objects (see Figure 2a) which was not the case with earlier practice with low-distinctiveness objects. In the older group, there were not more texture-based categorizations in the transfer phase in the High-Low condition (M = 87%, SD = 24.52) than in the Low-Low condition (M = 70%, SD = 32.12), t (38) = 1.88, p = .07, d = .60, BF_{10} = 1.22).

In order to explore the effect of a practice phase with high-distinctiveness objects on the performance with low-distinctiveness objects, we ran independent sample t-tests (two-tailed) to compare the practice phase in the Low-High condition and the transfer phase in the High-Low condition in the two-age group. There were not significant differences for younger children (L-H practice phase M = 51%, SD = 37.39; H-L transfer phase M = 67%, SD = 33.7=60), t (58) = 0.67, p = .507, d = .17, BF_{10} = 0.32) and for older children (L-H practice phase M = 74%, SD = 33.78; H-L transfer phase M = 87%, SD = 24.52), t (38) = 1.39, p = .172, d = .44, BF_{10} = 0.66).

In order to explore the effect of a practice phase with low-distinctiveness objects on the performance with low-distinctiveness objects, we ran independent sample t-tests (two-tailed) to compare the practice phase in the Low-High condition and the transfer phase in the Low-Low condition in the two-age group. There were not significant differences for younger children (L-H practice phase M = 51%, SD = 37.39; L-L transfer phase M = 47%, SD = 30.00), t (58) = 0.45, p = .651, d = .12, BF_{10} = 0.29) and older children (L-H practice phase M = 74%, SD = 33.78; L-L transfer phase M = 70%, SD = 32.12), t (38) = 0.38, p = .703, d = .12, BF_{10} = 0.33).

In order to compare the effects of a first practice phase with low-distinctiveness objects on the performance with high-distinctiveness objects, we used an independent sample t-test (two-tailed) to compare the transfer phase in the Low-High condition and the practice phase in the High-Low condition in the two-age group. There were not significant differences for older (H-L practice phase M = 92%, SD = 18.80; L-H transfer phase M = 77%, SD = 41.69), t (38) = 1.47, p = .151, d = .46, BF_{10} = 0.72, or younger children (H-L practice phase M = 76%, SD = 33.80; L-H transfer phase M = 61%, SD = 35.81), t (58) = 1.71, p = .093, d = .44, BF_{10} = 0.88). It suggests that a practice with low-distinctiveness objects did not benefit generalizations with high-distinctiveness objects.

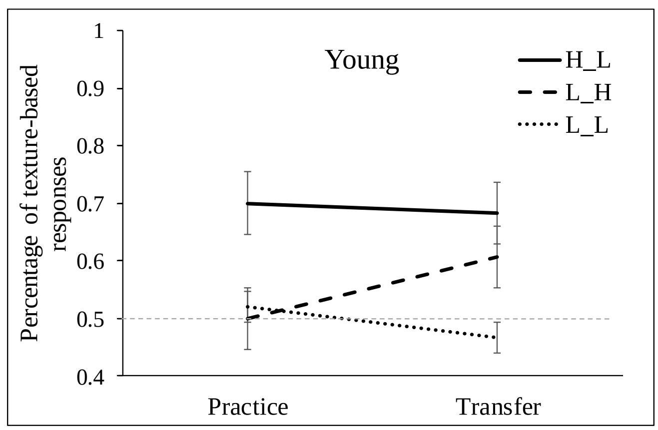

Figure 2a: Young children mean percentage of texture based-categorization in the practice and the transfer phase for the three conditions: High-Low (H-L), Low-Low (L-L) and Low-High (L-H). Vertical bars represent standard error of the mean.

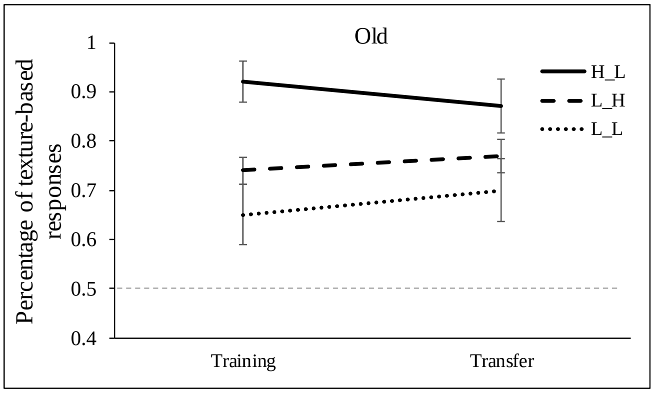

Figure 2b: Old children mean percentage of texture based-categorization in the training and the transfer phase for the three conditions: High-Low (H-L), Low-Low (L-L) and Low-High (L-H). Vertical bars represent standard error of the mean.

Chance comparisons We compared the percentage of texture match choices with chance (i.e., 50%) in each phase of learning in the three condition for the two age groups.

For older children all t-tests were significant except in the training phase in the Low-Low condition (M = 65%, SD = 37.80), *t* (19) = 1.78, *p* = .09, *d* = .40, BF_{10} = 0.87).

For younger children, in the Low-Low condition, t-tests were not significant in both the practice (M = 55%, SD = 30.50), *t* (29) = 0.96, *p* = .346, *d* = .175, BF_{10} = 0.30) and the transfer phases (M = 47%, SD = 29.90), *t* (29) = -0.47, *p* = .644, *d* = -.08, BF_{10} = 0.22). In the High-Low condition, the t-tests were significant in both the practice (M = 76%, SD = 33.80), *t* (29) = 4.21, *p* < .001, *d* = .77, BF_{10} = 125.98) and the transfer phase (M = 67%, SD = 33.60), *t* (29) = 2.83, *p* = .008, *d* = .52, BF_{10} = 5.19). In the Low-High condition, t-tests were not significant for both the practice (M = 51%, SD = 37.39), *t* (29) = 0.19, *p* = .847, *d* = .036, BF_{10} = 0.19) and the transfer phase (M = 61%, SD = 35.81), *t* (29) = 1.63, *p* = .113, *d* = .30, BF_{10} = 0.63).

Pattern analysis In order to provide a more precise picture of participants' distribution of answers, individual patterns of answers analysis were also performed. Indeed, a mean score can result either from a bimodal distribution of consistency or from a large number of inconsistent participants plus a small number of consistent participants. Children were categorized as texture-consistent when they chose at least four texture matches (out of five trials) and shape-consistent when they chose at least four times the shape matches (out of five). They were classified as inconsistent in other cases. Chi-square tests of independence were then run on the patterns of consistency of the transfer phase between conditions.

In the younger group, the distribution of the pattern of responses in the transfer phase between the conditions was marginally significant. They were more texture-consistent and less inconsistent in the High-Low condition than in the Low-Low condition (χ^2 (2, 60) = 5.71, *p* = 0.057) (see Figure 3).

In the older group, the distribution of the pattern of responses in the transfer phase between the conditions was significant. Indeed, they were more texture-consistent and less inconsistent in the High-Low condition than in the Low-Low condition (χ^2 (2,40) = 7.20, *p* = 0.027).

Correlations analysis Pearson's **c**orrelations between the first training trial and the mean of the transfer phase was conducted for each condition, to examine whether the first trial would predict the performance in the transfer phase. This analysis was motivated by the contrast between a cumulative-progressive view of learning and an all-or none view of learning. An all-or-none view would predict the first trial to be correlated with the transfer trials scores.

For the younger children, the correlation was significant in the three conditions, High-Low (*r* (53) = 0.36, *p* = .007, BF_{10} = 6.60), Low-High (*r* (58) = 0.50, *p* < .001, BF_{10} = 514) and Low condition (*r* (58) = .51, *p* < .001, BF_{10} = 417). We interpret this result in terms of executive function in the general discussion. We did not look at the older group, because the majority of them (17 participants out of 20) were correct on the first training trial.

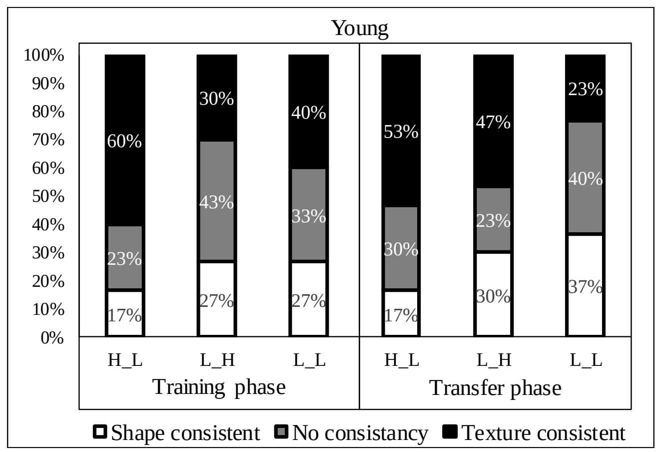

Figure 3: Percentage of young children who were either texture-consistent, shape-consistent, or inconsistent for each phase in each condition.

Discussion

Former evidence regarding the positive role of comparison in novel name learning has been obtained with comparisons situations that did not assess how a first phase in one condition would influence performance in another condition. In this experiment we examined whether learning to categorize objects at a given distinctiveness level (either low or high) would influence later categorization of high- or low-distinctiveness objects. Overall, our results showed that training with high-distinctiveness objects led to better scores in low-distinctiveness object categorization than training with low-distinctiveness objects. Correlation analyses showed that the first training trial predicted transfer performance, providing evidence in favor of an all-or-none view of learning.

In both age groups, we found more texture based-responses in the transfer phase after training with high distinctiveness objects than after training with low distinctiveness objects. This result is in line with the notion of progressive alignment suggesting that comparisons children make in a simple domain highlight the relational structure in this domain, which, in turn makes possible even more abstract comparisons (Kotovsky and Gentner, 1996). In our study, children's first exposure to the high-distinctiveness objects allowed them to align texture with texture even in the case of more complex learning objects. This is important because high-distinctiveness objects are likely to be compared and even young children might benefit from this comparison to gains deeper knowledge about them. Indeed, our findings suggest that these high-distinctive objects potentiate following less obvious comparison situations. It suggests that participants encoded texture across trials as a relevant type of feature in this situation, rather than as disconnected tokens. Exemplars of texture might have been stored as a result of stimulus processing, and the code in which these easy cases were stored was applied to new, less obvious, cases.

This result stands in sharp contrast with the Low-Low condition, in which participants could not use the trials of the practice part as additional opportunities to learn texture before the second part occurred. This result suggests that less distinctiveness objects did not allow any progressive learning that might occur across repeated presentations of the same learning situations (e.g., Scott & Dienes, 2010). This result is surprising because children, especially the older, benefited from comparison with low-distinctiveness objects but they could not use the gained insight for the following trials. It suggests there is a limit of the benefits from repeating comparison situations with the same material.

We also examined the course of learning thanks to the correlation between the first trial and the global performance. A first hypothesis was that once a dimension has been found by participants, they have difficulties to inhibit this first hypothesis or if they were able to inhibit it, they were unable to generate a new description of the objects (i.e., a lack of cognitive flexibility). On the contrary it is possible that learning is composed of trial-and-error testing in which participants try different hypotheses until they converge on the correct one. The significant correlation between the first trial in the practice phase and the mean performance of the transfer part of the experiment speaks in favor of the first hypothesis, that is an explanation in terms of executive functions (see Augier & Thibaut, 2013; Simms et al. 2018). These results suggest a link between the comparison process and cognitive flexibility necessary to apprehend objects from a novel point of view. Moreover, it seems that repeating comparisons situations does not allow for a progressive improvement and everything is settled after the first trial.

Conclusion

To the best of our knowledge, our data were the first to contrast different sequences of learning comparison conditions in a novel noun learning task. In our case, the first, high-distinctiveness, comparisons allow some form of abstraction of the texture dimension which are usable for more difficult, demanding cases later on. However, it does not tell us how the alignment was performed: transfer of an abstract notion of texture, learning of shape irrelevance. Further studies involving objects composed of different kinds of dimensions are required to explore more clearly how transfer occurs between trials.

Acknowledgments

The authors wish to thank the National Agency of Research (ANR) for their financial support (COMPARE project). They also would like to thank Marie Revellin for her help in running the experiment.

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.

Augier, L., & Thibaut, J. P. (2014). Dimensional Distinctiveness Constraints on Comparison Processing Across Development. In M. Bello P., Guarini M., McShane M. & Scasselati B. (Eds.) *Proceedings of the 36th Annual Conference of the Cognitive Science Society* (pp. 1874-1879) Austin TX: Cognitive Science Society.

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.

Carvalho, P. F., & Goldstone, R. L. (2014). Putting category learning in order: Category structure and temporal arrangement affect the benefit of interleaved over blocked study. *Memory & Cognition*, *42*(3), 481-495.

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.

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

Gentner, D., & Rattermann, M. J. (1991). Language and the career of similarity. Center for the Study of Reading Technical Report; no. 533.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. Cognitive science, 7(2), 155-170.

Gershkoff-Stowe, L., & Rakison, D. H. (Eds.). (2005). Building object categories in developmental time. Psychology Press.

Goldstone, R. L., & Son, J. Y. (2005). The Transfer of Scientific Principles Using Concrete and Idealized Simulations. *Journal of the Learning Sciences*, *14*(1), 69-110.

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.

Hammer, R., & Diesendruck, G. (2005). The role of dimensional distinctiveness in children's and adults' artifact categorization. *Psychological Science*, *16*(2), 137-144.

Hochberg, Y. (1988). A sharper Bonferroni procedure for multiple tests of significance. *Biometrika*, 75(4), 800-802.

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.

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

Kotovsky, L., & Gentner, D. (1996). Comparison and Categorization in the Development of Relational Similarity. *Child Development*, *67*(6), 2797-2822.

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

Ly, A., Verhagen, A. J. & Wagenmakers, E.-J. (2016). Harold Jeffreys. *Journal of Mathematical Psychology*, 72, 19-32.

Morey, R. D., & Rouder, J. N. (2018). BayesFactor: Computation of Bayes Factors for Common Designs. [R package]. Retrieved from https://cran.r-project.org/package=BayesFactor.

Murphy, G. L. (2002). The big book of concepts. MIT Press.

R Core Team (2021). R: A Language and environment for statistical computing. (Version 4.0) [Computer software]. Retrieved from https://cran.r-project.org. (R packages retrieved from MRAN snapshot 2021-04-01).

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, 56, 356-374.

Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian t tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, 16, 225-237.

Scott, R. B., & Dienes, Z. (2010). Knowledge applied to new domains: The unconscious succeeds where the conscious fails. *Consciousness and Cognition*, *19*(1), 391-398.

Simms, N. K., Frausel, R. R., & Richland, L. E. (2018). Working memory predicts children's analogical reasoning. *Journal of Experimental Child Psychology*, *166*, 160-177.

Smith, L. B., Jones, S. S., & Landau, B. (1996). Naming in young children: A dumb attentional mechanism? *Cognition*, *60*(2), 143-171.

The jamovi project (2021). jamovi. (Version 1.8) [Computer Software]. Retrieved from https://www.jamovi.org.

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