The benefits and costs of comparisons in a novel object categorization task: interactions with development

Augier Luc

Thibaut Jean-Pierre

Université de Bourgogne – LEAD – UMR CNRS 5022

Corresponding Authors

Luc Augier or Jean-Pierre Thibaut

Pôle AAFE – Esplanade Erasme -

Université de Bourgogne

21065 Dijon, France

Tel: +(0033)380 39 57 85

e-mail: Luc.Augier@u-bourgogne.fr

jean-pierre.thibaut@u-bourgogne.fr

Word count: 3971

Abstract

We investigated how children use within- or between-category comparisons to generalize novel names for novel objects on the basis of a non salient dimension (texture) rather than a salient one (shape). Previous studies have not experimentally addressed the costs associated with comparisons. We conjectured that increasing the number of stimuli to be compared (and thus converging evidence in favor of the target texture-based generalization), might not necessarily be beneficial, especially in young children (three-to-four-year-olds vs. five-to-six-year-olds). Our results showed that more evidence in favor of texture (i.e., more within-category exemplars sharing the same texture) did not linearly increase texture-based choices in the same way for younger and older children. They also revealed that between-category comparisons gave rise to texture-based generalizations in both age groups. Overall, our results show that even though within- and between-category comparisons contribute to generalizations based on texture, they also generate cognitive constraints which interact with age.

Introduction

One major challenge for children is to build a conceptual system that makes sense of the world. One difficulty is that generalization of novel concepts frequently means going beyond obvious and easily accessible perceptual appearances in favor of non-salient dimensions (conceptual or perceptual, see Rakison & Oakes, 2003).

What might promote generalizations based on non-salient dimensions? Several recent studies have suggested that the opportunity to compare items belonging to the same category provides an extremely powerful means for learning about the world (Namy & Gentner, 2002). Comparisons contribute to the discovery of unifying non-salient properties such as taxonomic commonalities (e.g., two objects belong to the category of furniture) or non-salient perceptual properties (e.g., object textures) that would remain unnoticed if the objects were processed in isolation (e.g., Gentner & Namy, 1999). This positive role of comparisons has been documented for a wide variety of stimuli and situations in both adults and children. In the case of children, this has been shown for object names (e.g., Gentner & Namy, 1999; Graham, Namy, Gentner, & Meagher, 2010; Namy & Clepper, 2010; Namy & Gentner, 2002) names for parts (Gentner, Loewenstein, & Hung, 2007), action verbs (Childers, 2011), adjectives (Waxman & Klibanoff, 2000), or perceptual categories (e.g., Hammer, Diesendruck, Weinshall, & Hochstein, 2009; Thibaut, 1991).

Typically, experiments involve comparison and no-comparison conditions, most often in the case of familiar stimuli. For example, in a novel name generalization task, Gentner and Namy (1999) pitted a taxonomic match against a perceptual match. In the comparison condition, children extended the novel name that was given to the exemplars (i.e. the standards, e.g., a bicycle and a tricycle) to a perceptually different taxonomic match (e.g., a skateboard), whereas in the no-comparison condition (e.g., a bicycle), they extended the name of the standards to the same-shape match (e.g. eyeglasses). Recently, Graham et al. (2010)

replicated this finding for categories of unfamiliar stimuli. The target unifying dimension was the non-salient dimension “texture”. In the comparison condition, preschoolers generalized the novel name to the stimulus with the same texture as the two standards rather than to the stimulus that shared the same shape as one of the two standards. In the “no comparison” condition, most four-year-olds generalized the novel name to the shape match.

In the present paper, we use the same novel name generalization task as in the studies referred to in the previous paragraph. However, our study specifically raises the issue of the cognitive costs of comparisons. Our key hypothesis is that increasing the number of stimuli involved in the task increases the number of comparisons which, in turn, will increase the cognitive costs of the task. Since young children have less developed executive functions (e.g. Anderson, 2002, see below), they might fail to integrate all the information available.

Within- and between-category comparisons.

We will also study the respective contribution of within- and between-category comparisons (hereafter “contrast-category” comparisons, that is comparisons involving stimuli from different categories) on categorizations of multidimensional stimuli. Both kinds of comparison increase the evidence in favor of the relevant dimension. According to Hammer, Bar-Hillel, Hertz, Weinshall and Hochstein (2008), within-category commonalities (e.g., all the *daxes* have property *i* and differ on their other properties) are clearly relevant in categorization in the sense that they contribute to the ability to distinguish categories. Hence, comparing stimuli belonging to the same category (e.g., several “daxes”) might highlight common properties.

However, comparisons may also involve stimuli from different categories. In this latter case, properties which are common to stimuli from different categories (e.g., “*daxes* and *blickets* have property X”) do not differentiate between these categories. Indeed, being told “*X* is a *dax* and *Y* is not a *dax*” children can infer that featural commonalities between *X* and *Y* are not sufficient to define what a *dax* is. Hammer et al. (2008) claim that common properties of contrast categories are not very informative because knowing that one attribute does not discriminate between two categories does not tell which attribute might discriminate them except when the number of attributes defining the categories is very small. However, they might push them to further explore the stimuli in order to find properties that differentiate categories (see Hammer et al., 2008; see also Clark, 1993, for a discussion of the role of contrast in language learning). For example, in the context of novel adjective learning, Waxman and Klibanoff (2000) showed that three-year-olds were able to associate an object property (e.g., transparency) with an adjective (e.g., "*blickish*") when the object displaying that property (e.g., a transparent plate) was contrasted to an object from the same category without that property (i.e., an opaque plate).

Recently, Namy and Clepper (2010) studied the role of "contrast" items in the aforementioned novel name paradigm (Gentner & Namy, 1999). The comparison-plus-contrast condition featured two standards (i.e., two "*blickets*") and a contrast object (a "non *blicket*"). The contrast object and the two standards had the same, hence not distinctive, shape (e.g., the two standards were a bicycle and a tricycle; the contrast object was barbells). The authors replicated the main effect of comparison and found a limited effect of contrast, only when there were two standards. Most likely, knowing that shape was irrelevant did not tell participants which among the remaining dimension(s) was or were actually relevant (see Hammer et al., 2008).

The current study

Previous studies have demonstrated the positive effects of comparison on novel name generalizations. However, they neglected the cost of comparisons. Our central hypothesis is that these comparisons generate cognitive costs that will be handled more or less efficiently by executive functions, such as working memory, inhibition, and cognitive flexibility (e.g., Zelazo, Carter, Reznick, & Frye, 1997). In our comparison situations, increasing the number of standards belonging to the same category increases the number of comparisons that must be performed and held in working memory. Salient common perceptual attributes (such as shape) must be inhibited in order to find less salient unifying dimensions (such as texture). Shifting to a new dimension when a salient one, such as shape, is irrelevant for categorization requires cognitive flexibility. In this context, authors have defined the cognitive complexity of tasks. For example, Andrews and Halford (2002) defines it in terms of the number of sources of variation that are related and must be processed in parallel (see also Zelazo & Frye, 1998, for a similar cognitive metrics). Given that executive functions are known to develop (e.g., Anderson, 2002), we hypothesized that comparison costs might differentially influence different age groups, which motivated the inclusion of two age groups.

We used unfamiliar categories defined by the non-salient dimension “texture” as in Graham et al. (2010). Stimuli were constructed around two dimensions, texture and shape. In order to vary the processing load, we manipulated the number of standards (1, 2 and 4) and the presence of an item belonging to another category (No or 1 contrast item) which had the same shape as one standard stimulus. Adding more items, either standards or contrast items, increases texture relevance and decreases shape relevance, but increases cognitive costs. Thibaut, French, & Vezneva (2010) made a similar point in the case of analogies when they increased the number of related distractors which were supposed to increase the inhibition and flexibility costs.

We compared contrast and no-contrast conditions. As previously mentioned, a contrast object (e.g., “this is not a *dax*”) tells that the dimension value it shares with the *dax* category (e.g., a given shape) is not diagnostic. However, repetition of this non diagnostic dimension value (see Figure 1) also potentially increases its saliency, even though the comparison should lead subjects to discard it. Therefore children have to inhibit it, a process which might be harder in the younger age group. Recently, Hammer et al. (2009), working with multidimensional artificial categories, found that six-year-olds failed to integrate contrast information. This was not the case for adults (see also, Andrews, Livingston, & Kurtz, 2011, in adults). On these grounds, we predicted that younger children might not benefit from contrast conditions as older children would. In terms of inhibition and flexibility, the *within*- and the *contrast*-category comparisons are interesting because they have opposing interpretations. Within-category commonalities lead to choose the common dimension (here texture) whereas contrast-category similarities lead to the rejection of the common dimension (here shape).

We also compared two subclasses of within- and contrast-category comparisons which were defined around two stimuli: 1-1 (i.e., one standard and one contrast item) and the 2-0 (i.e., two standards and no contrast item) conditions. The question was whether, in the 1-1 condition, the contrast item would lead to increase the use of texture (against the irrelevant shape) in the same way as in the 2-0 condition. In this latter case, differences in shape within the category should highlight texture as the unifying feature.

Finally, in our context of executive functions, age is a crucial factor. Although previous studies have shown that four-year-olds can benefit from comparison, the question of how four- and six-year-olds will differentially integrate more comparisons remains open.

Methods

Participants

A total of 216 preschoolers were tested individually at school. Two age groups were recruited. The younger children (n= 108, mean age = 48.8m, SD = 6, range: 38-58m) and the older children (n= 108, mean age = 66.3m, SD = 5.5, range: 59-77m) were randomly assigned to one of the six experimental conditions with 18 children per condition. Informed consent was obtained from their school and their parents.

Design

Younger and older groups of children were presented with novel object sets in one of six between-subject experimental conditions resulting from the crossing of the variables Number of standards (1, 2, 4 standards) and Contrast (0 vs. 1 contrast) (See Figure 1).

Materials

Seven sets of seven artificial grey-scale objects pictured on cards were created. The dimensions of the cards were 12*9 cm (width*height) and the dimensions of the objects pictured on the cards were approximately 6*6 cm. Each participant completed two practice trials and five experimental trials. The objects' textures and shapes were chosen to be distinctive. Textures and shapes used in one trial differed from those used in other trials. There were two standards in the 2-0 and the 2-1 conditions and four in the 4-0 and the 4-1 conditions. The standards shared the same texture but had different shapes. The contrast object (in the 2-1 and 4-1 conditions) had the same shape as one of the standards. The first test object, the shape match, had the same shape as one of the two standards (and as the contrast object) but differed in terms of its texture. The second test object, the texture match, had the same texture as the standards (and a different texture than the contrast object). In 1-0 and 1-1 conditions, the standard had the same shape as the shape match (and the contrast object) and the same texture as the texture match. Each set was associated with one of seven two-syllable novel names: yo*uma*, *buxi*, *dajo*, *zatu*, *sepon*, *xanto*, *vira*.

****

Insert Figure 1 about here

****

Procedure

The experiment started with two practice trials which were followed by five test trials presented in a random order. Each standard was introduced with a novel count noun (e.g. “this is a *buxi*” and “this is a *buxi* TOO” for the other standards). In the contrast conditions, a contrast object was introduced below the standard(s) as a non-member of the category (e.g. “this is NOT a *buxi*”). The objects were presented sequentially and left in view. In the comparison conditions, 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 shape and the texture match) were introduced and the child was asked to point to the one which was also a member of the category (e.g., “Show me which one of these two is also a *buxi*”).

Results

Analyses were performed on the mean percentage of texture choices, and on consistency patterns (see below). A 3(Number of standards)* 2(Contrast)* 2(Age) ANOVA was conducted on the percentage of texture choices. Older children chose the texture match significantly more frequently (58%) than younger children (49%) (F (1,204) = 4.82, $p < .05$, = .023). The main effect of Number of standards was significant (F (2,204) = 54.97, $p <$ .001, = .35), with 26%, 58% and 77% of texture choices in the 1-, 2- and 4-standard conditions respectively, showing that more standards meant more texture choices. These two main effects were subsumed by a marginally significant Age by Number of standards interaction (F (2,217) = 2.94, p = .055, = .02). The REGWQ post hoc procedure (see Howell, 2007) revealed that, in both age groups, the 1-standard condition significantly differed from the 2- and the 4-standard conditions. It also revealed that the 4-standard condition was significantly higher than the two other conditions in the older group only. Furthermore, older children performed significantly better than younger children in the 4-

standard case ($p < 0.05$) (see Figure 2). Overall, this interaction shows that increasing the number of standards benefited younger children less it did older children.

****

Insert Figure 2 about here

****

The Number of standards by Contrast interaction was significant (F (2,217) = 3.55, $p < .05$, = .03) (see Figure 3). The REGWQ post hoc procedure ($p < 0.05$) revealed the following hierarchy, 1-0 (17%) < 1-1 (34%) < 2-1 (54%) = 2-0 (62%), 2-1 < 4-0 (73%), 2-0 = 4-0, 2-0 < 4-1 (81%), 4-0 = 4-1. In sum, there are three important results. First, the comparison between 1-0 and 1-1 showed that children benefited from a contrast-category comparison; second, there was a difference between 1-1 and 2-0, suggesting that within-category comparisons are more powerful than contrast-category comparisons when the comparisons involve two stimuli; third, there was a significant difference between the 2-1 and the 4-1 conditions, and no difference between 2-0 and 4-0 (see discussion for an interpretation of this interaction).

****

Insert Figure 3 about here

****

Finally, we performed Chi-square tests of independence corrected for multiple comparisons (Bonferroni) ($p <.05$) on patterns of consistency (see Figure 4). Children were categorized as “texture-consistent” when they chose at least four texture matches (out of five), “shape-consistent” when the shape match was chosen at least four times, or “non consistent” in the other cases. Indeed, a given percentage (e.g., 50%) of shape choices can result from a majority of consistent participants (i.e., 50% shape consistent and 50% texture consistent participants) or from a majority of inconsistent participants.

****

Insert Figure 4 about here

****

Importantly, there were fewer texture-consistent children overall in the 1-0 (0% for both groups) than in the 1-1 condition (22% for younger and 28% for older children). This shows that contrast-category comparisons alone contribute to highlighting texture and/or decreasing the dominance of shape.

As shown in the analysis discussed above, only the older group benefited from “more comparison items”. There were significantly fewer texture-consistent older children in the 2-0 (44%) and 2-1 conditions (50%) than in the 4-0 (83%) and 4-1 (94%) conditions respectively. Comparisons between the two age groups showed that there were significantly fewer younger than older texture-consistent children in 2-1 (50% and 11% respectively), 4-0 (83% and 39% respectively) and 4-1 conditions (94% and 44% respectively).

Discussion

Previous studies have demonstrated the positive effects of comparison on novel name generalization. In the executive function framework, we hypothesized that the number of stimuli to be compared and the nature of the comparisons (within- or contrast-category) might matter. Indeed, presenting more stimuli to be integrated increases the number of comparisons and, thus, generates cognitive costs. More specifically, it was hypothesized that children of different age groups would not integrate the information in the same way because of differences in terms of executive function maturation (Zelazo et al., 1997). The results confirmed our predictions.

The discovery of a non salient dimension: role of the number of standards

The interaction between Age and Number of standards as well as the consistency patterns showed that the younger group, overall, obtained similar results in the 2- and 4-standard conditions whereas, for the older group, increasing the number of standards increased the percentage of texture-based generalizations (see Figures 2 and 4). This shows that younger children make sense of the comparisons but do not integrate the extra information in favor of texture: they might compare 2 or 3 stimuli but not all of them.

Within- and contrast-category comparisons

Regarding the two types of comparisons, two questions were raised. First, do contrast-category comparisons contribute to texture-based generalizations and, second, is the information coming from contrast exemplars more difficult to integrate than within-category comparisons?

The difference between the 1-0 and 1-1 conditions addresses the first question. It shows that the contrast condition (1-1) resulted in more texture-based choices and more texture consistency than the absence of contrast (1-0). To the best of our knowledge, this is the first evidence with young children that contrast-category comparisons *alone* can promote generalization based on a non-salient dimension. Indeed, with familiar categories, Namy and Clepper (2010) did not find any significant effect of contrast in their 1-standard condition.

The answer to the second question "How did within- and contrast-category comparisons influence generalizations?" is that contrast-category comparisons gave rise to less texture-based generalizations. *First,* and importantly, the interaction between contrast and number of standards shows that the 1-1 condition generated fewer texture choices than the 2-0 condition. *Second*, the consistency analysis shows that the 2-1 condition generated significantly more difficulties in the younger group than in the older group, whereas this difference between the two groups did not exist in the 2-0 condition. As shown in Figure 4, this is due to a decrease in the percentage of texture consistencies in the younger group. This suggests that the younger group found it harder to integrate the extra information provided by the contrast item (2-1).

Overall, the present results show that contrast-category comparisons provide relevant information for texture based generalization (1-0 vs 1-1), even though it is more difficult to integrate than within-category comparisons (1-1 vs 2-0, and the difference between the age groups for 2-1). As mentioned in the introduction (see Hammer et al., 2008), within-category comparisons provide cumulative evidence in favor of the common dimension value (here, same texture) and tell that “non-repeated” dimension values (here, same shape) are irrelevant. By contrast, the contrast-category comparisons demonstrate that contrast-category similarities (i.e., common shape) are irrelevant and thus another dimension should be looked for, but does not tell which one. Another difficulty with contrast-category similarities is that the repetition of a common dimension value (same shape) increases its saliency but tells that it is irrelevant and thus must be inhibited, as in the 1-1 condition. The 2-1 condition is more difficult, in this respect because there are two types of commonalities (contrast-category for shape and within-category for texture) that children must interpret in opposite ways (i.e., “reject shape” vs. “favor texture”). In our case, this requires both inhibition capacities (“inhibit the salient common shape”) and cognitive flexibility (“search for another dimension unifying the category”). The younger group’s difficulties in 2-1, together with the fact that they benefit less from “more standards”, are consistent with the general observation that cognitive monitoring (and thus executive functions) is less developed in younger children, (e.g., Zelazo et al., 1997). This explanation also fits with Hammer et al.’s (2009) results showing that even six-to-nine year olds’ failed to integrate contrast items in a category learning task.

The dynamic of comparisons

Gentner and colleagues have argued that comparisons lead to generalization by structural alignment which increases the saliency of less obvious “within-category” dimensions (e.g., Gentner & Namy, 1999). Incorporating contrast-category comparisons in alignment processes as we did in the present experiment, shows that comparisons not only highlight alignable common within-category features and/or downplay alignable within-category differences (shape) but also downplay alignable common contrast-category features (shape) and/or highlight alignable distinctive contrast-category features (texture).

In this context, increasing the number of standards decreases the relevance of alignments based on the a priori salient but irrelevant dimension (shape). Since decreasing shape saliency does not guarantee that unifying but less salient dimensions will be found, it becomes necessary to search for other dimensions on which the stimuli might be aligned. This is easier when more stimuli with the same dimension value are provided. However, as we said before, providing more items means coordinating more comparisons within the set and also between each standard and both matches, which was more difficult for the younger group.

To conclude, our main question focused on whether children would be able to integrate information from a growing set of comparisons involving within- and contrast-category information. In general, our data suggest that, even though both groups benefited from comparisons, younger children were less able to integrate all the available information. This suggests that comparisons generate their own cognitive costs which must be taken into account when a comparison situation is devised.

References

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology, 8*(2), 71-82.

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

Andrews, J. K., Livingston, K. R., & Kurtz, K. J. (2011). Category learning in the context of co-presented items. *Cognitive Processing, 12*(2), 161-175.

Childers, J. B. (2011). Attention to multiple events helps two-and-a-half-year-olds extend new verbs. *First Language, 31*(1), 3-22.

Clark, E. V. (1993). *The lexicon in acquisition*. New York, NY, US: Cambridge University Press.

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.

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., Bar-Hillel, A., Hertz, T., Weinshall, D., & Hochstein, S. (2008). Comparison processes in category learning: From theory to behavior. *Brain Research, 1225*, 102-118.

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

Howell, D.C. (2007). Statistical methods for psychology (6th Ed.). Belmont, CA: Wadsworth Publishing.

Namy, L. L., & Clepper, L. E. (2010). The differing roles of comparison and contrast in children's categorization. *Journal of Experimental Child Psychology, 107*(3), 291-305.

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.

Rakison, D. H., & Oakes, L. M. (2003). *Early category and concept development : making sense of the blooming, buzzing confusion*. Oxford: Oxford University Press.

Thibaut, J. P., French, R., & 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. (1991). Récurrence et variations des attributs dans la formation des concepts. Unpublished doctoral thesis, University of Liège, Liège.

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

Zelazo, P. D., & Frye, D. (1998). Cognitive complexity and control: II. The development of executive function in childhood. *Current Directions in Psychological Science, 7*, 121-126.

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.

Acknowledgments

The authors wish to thank the Conseil Régional de Bourgogne for its financial support (PhD grant to the first author), Faber Program. We thank the children who participated in the study, their parents and the school staff for their assistance, and Laurent Bergerot for his help in producing the stimuli.

Figure Captions

Figure 1. Sample stimulus set and instructions used in the six experimental conditions which crossed the two factors: Number of standards (1, 2 or 4 standards) and Contrast (0 or 1 contrast).

Figure 2. Mean percentage of texture-based responses (+SEM) as a function of Age and Number of standards. Braces show significant differences between columns ($p <.05$).

Figure 3. Mean percentage of texture-based responses (+SEM) as a function of Contrast and Number of standards. Braces show significant differences between columns ($p < .05$).

Figure 4. Percentage of children, for each experimental condition and for each group of age, who were either texture-consistent, shape-consistent, or non-consistent.

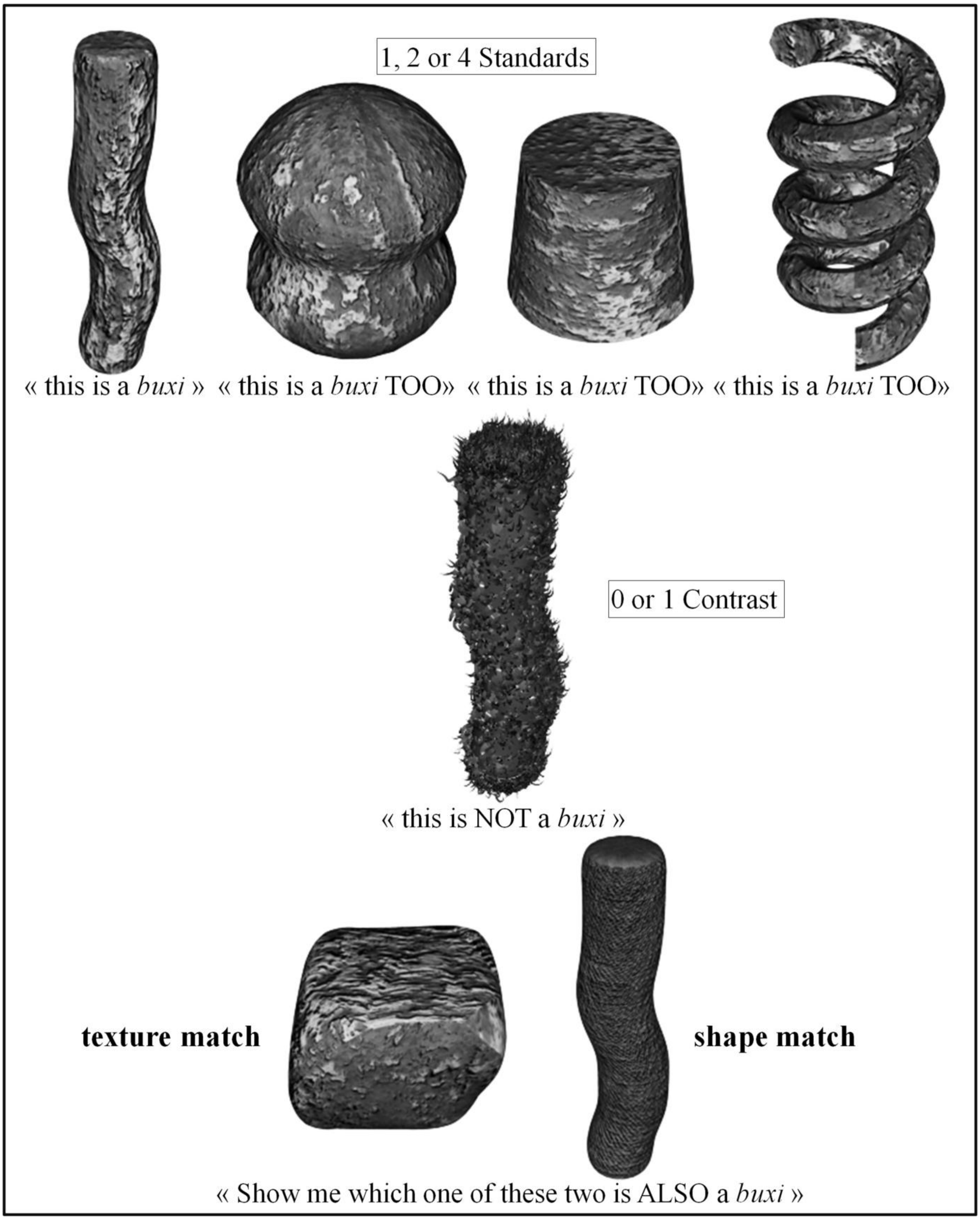
1, 2 or 4 Standards
« this is a *buxi* »
« this is a *buxi* TOO»
« this is a *buxi* TOO»
« this is a *buxi* TOO»
0 or 1 Contrast
« this is NOT a *buxi* »
texture match
shape match
« Show me which one of these two is ALSO a *buxi* »

Figure 1.

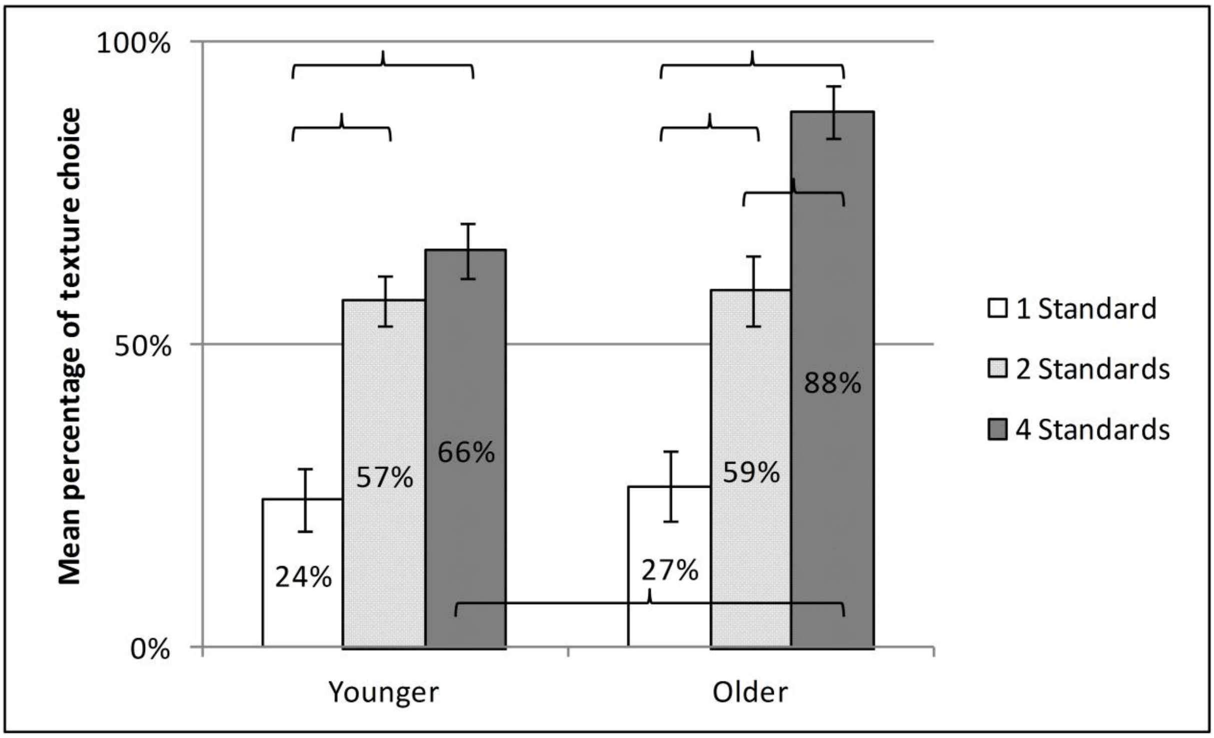

Figure 2.

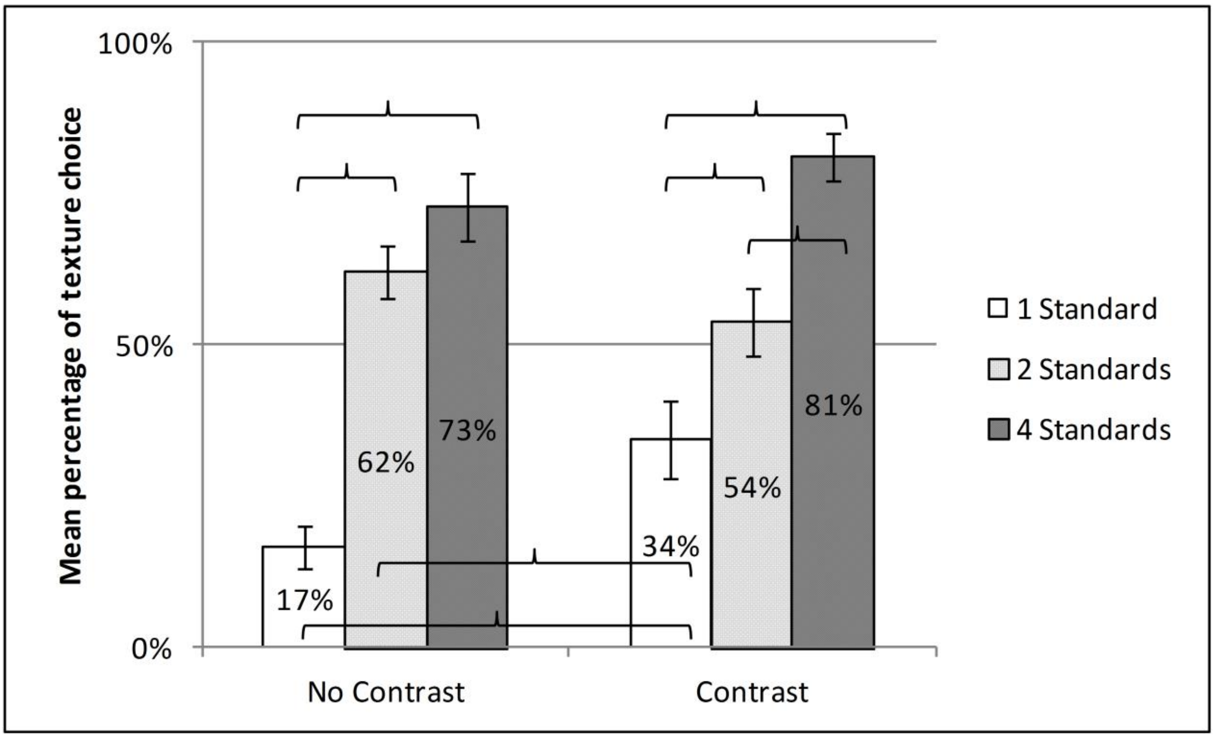

Figure 3.

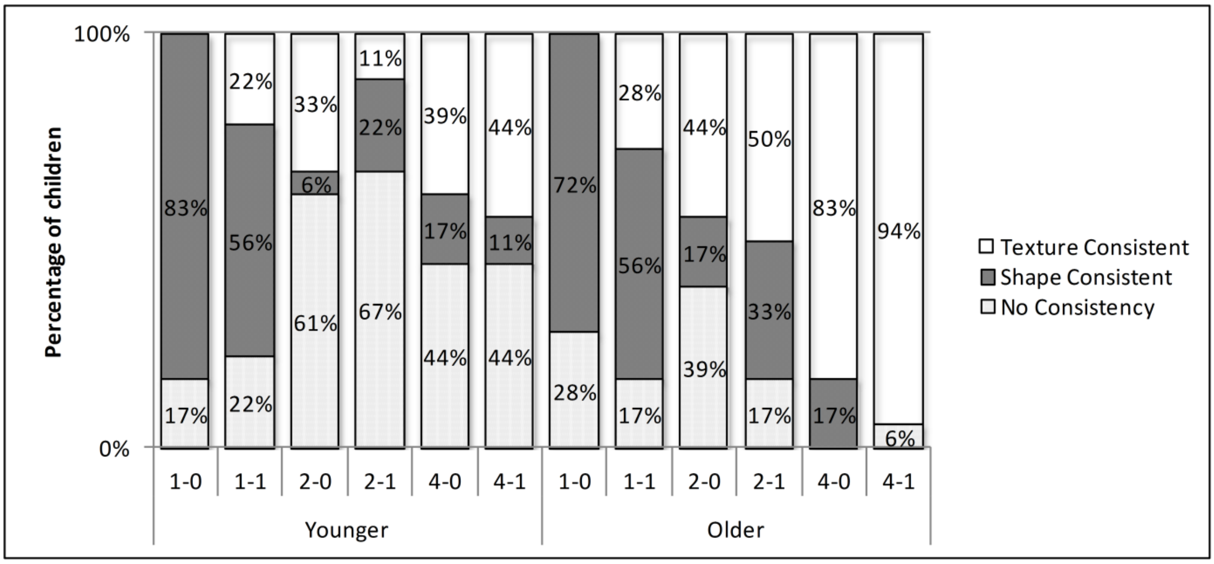
100%
0%
Percentage of children
83%
17%
22%
56%
22%
33%
6%
61%
11%
22%
67%
39%
17%
44%
44%
11%
44%
72%
28%
28%
56%
17%
44%
17%
39%
50%
33%
17%
83%
17%
94%
6%
1-0
1-1
2-0
2-1
4-0
4-1
1-0
1-1
2-0
2-1
4-0
4-1
Younger
Older
Texture Consistent
Shape Consistent
No Consistency

Figure 4.