Mapping the similarity space of children and adults’ artifact categories

Gil Diesendruck*, Rubi Hammer, Or Catz

Department of Psychology, Bar-Ilan University, Ramat-Gan 52900, Bar-Ilan, Israel

Received 1 July 2002; received in revised form 1 February 2003; accepted 1 February 2003

Abstract

Most theorists agree that people’s categorization of artifacts is influenced by two factors: (a) in what respect objects are similar, and (b) to what degree objects are similar. Research with children has identified two respects — physical appearance and function — which are important for the categorization of artifacts. Previous work, however, has not mapped decisively the relative contribution of each of these respects to children’s categorization decisions. In the present study, children and adults were shown a novel standard object and were asked to categorize 10 test objects of varying degrees of physical and functional similarity to the standard. We found that while adults included as category members objects of high functional similarity even if of low physical similarity, children accepted as members only objects that were highly similar to the standard both functionally and physically. In other words, children and adults weighted functional similarity comparably, but children gave more weight than adults to physical similarity. These results may help illuminate the conflicting findings and theories on the development of artifact categories.

© 2003 Elsevier Science Inc. All rights reserved.

Keywords: Categorization; Artifacts; Conceptual development

1. Introduction

Common sense tells us that the reason why two objects belong to the same category is because they are similar. A few decades of philosophical (e.g., Goodman, 1972) and psychological (e.g., Rosch & Mervis, 1975; Tversky, 1977) analysis...
of this issue has taught us that this intuition is too vague. Similarity needs to be constrained both in quality and degree. Namely, to determine that two objects belong to the same category, at the least one needs first to define in what respects the objects are similar, and then compute the degree to which the objects are similar in each relevant respect (Medin, Goldstone, & Gentner, 1993). The goal of the present study is to investigate these two characteristics in children and adults’ representations of artifact categories.

According to a “similarity-view” of categorization (e.g., Goldstone & Barsalou, 1998), decisions about the category membership of objects are made based almost entirely on computations of the respects and degrees of similarity between objects. Hampton (1998), for instance, defines three fundamental aspects of categories: an intensional representation of attributes (e.g., where the respects get defined), a metric of similarity, and a threshold criterion for categorization. Smith and Heise (1992) suggest that the similarity between objects can be computed based on their “distance in a psychological space” (p. 243). Moreover, they note that similarity is dynamic: in different contexts and with development, the weights of the different respects upon which objects vary may change, thus changing also the psychological metric of similarity (see also Jones & Smith, 1993).

Scholars often associated with a “core-view” of categorization differ from a similarity-view primarily in their attribution of a causal role in categorization to people’s conceptual beliefs about the entities to be categorized (e.g., Bloom, 1996; Rips, 2001). Nonetheless, also according to some proponents of a core-view, respects and gradients of similarity may be particularly important in accounting for the development of children’s artifact categories. Keil, Smith, Simons, and Levin (1998), for instance, propose a hybrid model of categories in which causal understanding defines, but is at the same time enriched by, notions of the kinds of properties central in a domain (the respects), and by computations of similarity. Along similar lines, Gelman and Diesendruck (1999) claim that while essentialist beliefs are crucial for determining category membership, a similarity-based prototype assists in the quick recognition of category exemplars, especially in a domain such as artifacts wherein categories may be less essentialized.

Motivated by these divergent views regarding categorization, researchers have been debating to what extent children’s artifact categories are determined by “superficial” properties, such as objects’ physical appearance, as opposed to conceptually “deeper” properties, such as objects’ function. To address this question head-on, researchers have often confronted children with a direct choice between categorizing objects based on their physical or functional similarity. The results of these studies have been mixed. Some studies found that children as old as 5 years of age, but not adults, extend labels to objects based on physical similarity, even at the expense of functional similarity (Graham, Williams, & Huber, 1999; Landau, Smith, & Jones, 1998; Merriman, Scott, & Marazita, 1993; Smith, Jones, & Landau, 1996; Tomikawa & Dodd, 1980). In contrast, a number of other studies found that children as young as 2 years of age extend labels to objects based on functional similarity, irrespective of physical similarity (Deak, Ray, & Pick,
Scholars have proposed, and investigated, a number of factors that varied across studies and that could therefore partly account for the disparity in the findings. For instance, it has been shown that the plausibility of the function-appearance relation, the time allowed children to examine the objects, the availability of the functional information, and the intentional character of objects’ functions, all affect children’s categorization strategy (Deak et al., 2002; Diesendruck et al., 2003; Kemler Nelson, Russell, Duke, & Jones, 2000).

These studies have been crucial for our conceptualization of physical appearance and function as important respects by which artifacts are categorized, and more so, for our understanding of the circumstances under which one or the other factor dominates children’s categorization. Nonetheless, these studies do not provide a definitive answer about the unbiased relative contribution of physical and functional similarity to children’s artifact categories. This assertion is especially valid if we take a closer look at the methodologies employed by most previous studies on this issue.

In most studies, the categorization procedure involved at least one of the following two characteristics (with slight variations, see Deak et al., 2002; Diesendruck et al., 2003; Graham et al., 1999; Kemler Nelson, Frankenfield, et al., 2000; Kemler Nelson, Russell, et al., 2000; Landau et al., 1998; Merriman et al., 1993; Smith et al., 1996). First, children were typically shown a novel target artifact, and were asked to extend its label either to a functionally similar physically dissimilar novel object, or to a functionally dissimilar physically similar novel object (though see Kemler Nelson et al., 1995, for the use of gradients of physical similarity). In other words, the relative weight of physical or functional similarity was presumably inferred by children’s dichotomous absolute choice. This procedural characteristic not only leads to the above logically problematic inference, but it also runs counter to the theoretical importance of assessing gradients of similarity in artifact categorization (Diesendruck & Gelman, 1999; Kalish, 1995; Malt & Johnson, 1992), and to the arguably natural correlation between objects’ form and function (Gentner & Ratterman, 1991; Gibson, 1969).

The second methodological characteristic of most previous studies has to do with potential imbalances in the cognitive effect of physical and functional information. Specifically, in most studies, either functional information about the objects was not explicitly available at the time children had to make a categorization decision (e.g., Smith et al., 1996), or it was made available by having the experimenter demonstrate and/or verbally describe it (e.g., Deak et al., 2002; Diesendruck et al., 2003; Kemler Nelson, Frankenfield, et al., 2000; but see Kemler Nelson, Russell, et al., 2000, Experiment 2). The lack of simultaneity in the availability of information is known to affect adults’ categorization or similarity judgments (Medin et al., 1993; Palmer, 1978). Moreover, as Deak et al. (2002) note, “if an adult takes pains to make one attribute available or salient (e.g., repeatedly demonstrating an object’s function), this might indicate the adult’s intentions or
The purpose of the present study was to circumvent these methodological issues in order to provide a more balanced assessment of the relative — and unbiased — weight of physical and functional similarity in children and adults’ artifact categories. In a way, the study was meant to provide a “baseline” assessment of children and adults’ artifact categories. For that purpose, first, instead of forcing participants to decide between these two dimensions when categorizing artifacts, the present study gave participants the freedom to choose objects of varying degrees of similarity on these two dimensions. Even though this procedure still required an absolute decision as to whether or not each object was a member of a certain category, the gradient similarity composition of the objects enabled us to assess how the degree of physical and/or functional similarity between objects affected children and adults’ categorization decisions.

Second, instead of presenting actual three-dimensional objects in which the physical appearance information is made available visually but the functional information verbally, we used as stimuli computer-animated objects. In this way, both the appearance and function information were made available to participants through the same medium — visually, and were available simultaneously for the same amount of time. In other words, participants were left to observe and evaluate on their own the degree of physical and functional similarity between objects, without the direction or interference of another adult.

2. Method

2.1. Participants

Thirty-nine adults (mean age = 27.5 years, range = 19–48 years) and 40 children (mean age = 5 years and 4 months, range = 4 years and 6 months–6 years) participated in this study. Approximately the same number of males and females participated. Twenty additional adults provided ratings of similarity on the stimuli. All participants were Hebrew speakers.

2.2. Materials

Two sets of computer-animated objects were created for the experiment. The animations were created using 3D Studio Max 3R® software, and were presented on a 13.1 in. IBM Thinkpad® laptop computer screen. An especially developed program written in Visual Basic enabled the simultaneous presentation of video
animations, and subsequent recording of participants’ responses. Each animation displayed an object with a novel shape performing a novel function (e.g., burning logs or transforming billiard balls into smaller ones) for 10 s (see Fig. 1). Each set of objects consisted of a “standard” and 10 test objects. The standard object occupied 390 × 290 pixels on the computer screen, whereas the test objects occupied 210 × 160 pixels. Five of the test objects were highly similar in terms of their function to the standard, but had graded physical similarity to it. We refer to them as “functional matches.” The other five objects were highly similar in terms of their physical appearance to the standard, but had graded functional similarity to it. We refer to them as “physical matches.”

Twenty adults rated the degree of similarity between each of the test objects to the standard of each set. As the participants sat in front of a computer screen, the experimenter activated all of the objects in one set for 1 min. The experimenter then stopped the animation of 9 of the 10 test objects, leaving active only 1 test object at a time and the standard. Participants were asked to estimate the percentage (0–100) of physical, functional, and global similarity between each test object and the standard. As can be seen in Table 1, functional matches were more similar functionally, t(18) = 7.40, P < .001, and less similar physically, t(18) = −3.93, P < .005, to the standard than were physical matches. The two types of test objects did not differ significantly in terms of global similarity.

2.3. Procedure

The procedure for adults and children were identical. Adults were tested individually in a university laboratory. Children were tested individually in a quiet area of their preschools. In order to best capture participants’ categorization behavior, we employed two different instructions. In particular, adults and children were randomly divided into a Kind and a Naming condition. The only difference in between the conditions was in the specific generalization task participants performed.

In both conditions, participants sat in front of a laptop computer screen (approximately 2 ft away), and the experimenter told them that they were going to see a number of objects and would then be asked some questions about them. The experimenter first activated the standard of the first set for 20 s. In the Kind condition, participants were simply told to look at the object. In the Naming con-

<table>
<thead>
<tr>
<th>Type of test object</th>
<th>Physical similarity</th>
<th>Functional similarity</th>
<th>Global similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Physical matches</td>
<td>78.3</td>
<td>34.8</td>
<td>56.4</td>
</tr>
<tr>
<td>Functional matches</td>
<td>49.7</td>
<td>88.5</td>
<td>67.6</td>
</tr>
</tbody>
</table>
Fig. 1. Gray scale examples of stimuli used in the experiment.
dition, the experimenter labeled the standard repeatedly with a novel name (e.g., “This is a Bargovan” — the other meaningless Hebrew-sounding novel name used was “Tirpal”). The experimenter then activated the 10 test objects for about 60 s, with each object appearing on a small window around the standard. While the participants looked at the animations of the test objects, the experimenter gave participants the generalization instructions. In the Kind condition, the participants were asked to evaluate which objects were of “the same kind” as the standard. In the Naming condition, the participants were asked to evaluate which were also “Bargovans.” Participants were asked not to express their decisions at this point.

After the participants had watched all the objects playing simultaneously, the experimenter told them they were now going to see each of the test objects separately. The animations of 9 of the 10 test objects were paused, leaving only the standard and 1 test object at a time active for about 20 s. The test object window was increased to the size of the standard, and the experimenter pointed to both objects on the screen and told participants, “Look at these.” After this stage, all test objects were again activated, together with the standard, and the experimenter pointed to each test object and asked participants whether it was “the same kind” as the standard (Kind condition), or whether it was a “Bargovan/Tirpal” (Naming condition). Participants were free to choose as many test objects as they wished. Placement on the screen of the 10 test objects around the standard was constant across participants, but the order in which they were activated (from top left or right) was counterbalanced. After the participants made their selections, the experimenter removed all objects from the screen and presented the second set in the exact same way.

3. Results

The main dependent measure was the mean number of objects selected by participants. Preliminary analyses revealed no effect of gender or stimulus set on this measure, and therefore these factors were excluded from subsequent analyses.

In order to evaluate whether the categorization instructions mattered, we conducted a mixed three-way ANOVA with age group (children versus adults) and condition (naming versus kind) as between-subjects factors, and type of test object (physical matches versus functional matches) as within-subjects factor. The analysis revealed no difference between the overall number of test objects chosen by participants in the Naming condition \((M = 3.52, \text{S.D.} = 2.60)\) and in the Kind condition \((M = 3.28, \text{S.D.} = 2.46)\), \(F(1, 75) = 0.34, P = \text{n.s.}\) (see Table 2). There was also no interaction between condition and age or between condition and type of test object. Children and adults alike manifested a similar pattern of categorization whether they were asked to extend a name or to choose objects of the same kind. This variable was not included in the remaining analyses.

The main question of this study was whether or not children and adults’ artifact categories differed in terms of the relative weights of physical and functional
Table 2
Mean number of test objects chosen by type of test object, age group, and condition

<table>
<thead>
<tr>
<th>Age group</th>
<th>Condition</th>
<th>n</th>
<th>Functional matches</th>
<th>Physical matches</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>M</td>
<td>S.D.</td>
</tr>
<tr>
<td>Children</td>
<td>Naming</td>
<td>22</td>
<td>1.27</td>
<td>1.56</td>
</tr>
<tr>
<td></td>
<td>Kind</td>
<td>18</td>
<td>1.36</td>
<td>1.59</td>
</tr>
<tr>
<td>Adults</td>
<td>Naming</td>
<td>20</td>
<td>3.80</td>
<td>1.29</td>
</tr>
<tr>
<td></td>
<td>Kind</td>
<td>19</td>
<td>3.24</td>
<td>1.66</td>
</tr>
</tbody>
</table>

similarity. Our analyses revealed that the answer to this question was affirmative regarding physical similarity but negative regarding functional similarity. First, the ANOVA described above showed that participants overall chose more functional matches ($M = 2.41$, S.D. = 1.88) than physical matches ($M = 1.01$, S.D. = 1.19), $F(1, 75) = 82.89, P < .001$, and that adults chose more test objects overall ($M = 4.36$, S.D. = 2.04) than did children ($M = 2.48$, S.D. = 2.62), $F(1, 75) = 12.11, P < .001$.

More to the point, the analyses revealed a significant interaction between type of test object and age group, $F(1, 75) = 66.51, P < .001$. The $t$-tests revealed that adults chose more functional matches ($M = 3.53$, S.D. = 1.49) than children did ($M = 1.31$, S.D. = 1.55), $t(77) = -6.46, P < .001$, but adults ($M = 0.83$, S.D. = 0.89) did not differ from children ($M = 1.17$, S.D. = 1.41) in terms of the number of physical matches chosen, $t(77) = 1.29, P = n.s$. In other words, adults were more likely than children to include in their artifact categories objects that were physically highly dissimilar from the standard, as long as they were functionally highly similar to it. Adults and children, nonetheless, were equally likely to exclude from their categories objects that were functionally highly dissimilar from the standard, even if they were physically highly similar to it.

The above interaction was also revealed in an analysis of how the ratings of similarity predicted participants’ categorization. Specifically, the ratings of global, functional, and physical similarity provided by the adult judges were entered into a hierarchical multiple regression of the frequency with which each of the 20 test objects (across the two sets) were selected by the participants. In this analysis, global similarity ratings and age group (adult or child) were entered in the first step, functional and physical similarity were entered together in the second step, and the interactions between each of the three similarity ratings and age group were entered in the third step. As can be seen in Table 3, in the final step the analysis revealed a significant contribution of age group ($\beta = 0.37, P < .001$), of functional similarity ($\beta = 0.80, P < .05$), and of the interaction between age group and functional similarity ($\beta = 0.82, P < .05$) to participants’ frequency of selection. In fact, the overall explanatory power of the regression model was quite substantial ($R^2 = .82, P < .001$).
Table 3
Hierarchical multiple regression results for all participants, regressing similarity variables (global, functional and physical) on frequency of stimuli selection

<table>
<thead>
<tr>
<th>Variables</th>
<th>Adults</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>$\beta$</td>
<td>$R^2$</td>
<td>$R^2$ change</td>
</tr>
<tr>
<td>Step 1</td>
<td>.45**</td>
<td>.45**</td>
<td>.45**</td>
<td>.45**</td>
</tr>
<tr>
<td>Age group</td>
<td>0.187</td>
<td>0.37**</td>
<td>0.01</td>
<td>0.56**</td>
</tr>
<tr>
<td>Global similarity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>.57**</td>
<td>.12**</td>
<td>.57**</td>
<td>.12**</td>
</tr>
<tr>
<td>Age group</td>
<td>0.187</td>
<td>0.37**</td>
<td>0.001</td>
<td>0.06</td>
</tr>
<tr>
<td>Global similarity</td>
<td>0.007</td>
<td>0.80</td>
<td>0.003</td>
<td>0.22</td>
</tr>
<tr>
<td>Functional similarity</td>
<td>0.001</td>
<td>0.80</td>
<td>0.003</td>
<td>0.22</td>
</tr>
<tr>
<td>Physical similarity</td>
<td>0.007</td>
<td>0.80</td>
<td>0.003</td>
<td>0.22</td>
</tr>
<tr>
<td>Step 3</td>
<td>.82**</td>
<td>.25**</td>
<td>.82**</td>
<td>.25**</td>
</tr>
<tr>
<td>Age group</td>
<td>0.187</td>
<td>0.37**</td>
<td>0.007</td>
<td>0.80*</td>
</tr>
<tr>
<td>Global similarity</td>
<td>0.001</td>
<td>0.06</td>
<td>0.003</td>
<td>0.22</td>
</tr>
<tr>
<td>Functional similarity</td>
<td>0.011</td>
<td>0.82*</td>
<td>0.003</td>
<td>0.22</td>
</tr>
<tr>
<td>Physical similarity</td>
<td>0.007</td>
<td>0.80</td>
<td>0.003</td>
<td>0.22</td>
</tr>
</tbody>
</table>

* $P < .05$.
** $P < .001$.

In order to better capture the directionality of the interaction between age group and functional similarity, we conducted two separate multiple hierarchical regressions for each age group. In these analyses, global similarity ratings were entered in the first step, and functional and physical similarity ratings were entered together in the second step. As can be seen in Table 4, adults’ object selections were mostly explained by functional similarity ($\beta = 1.34$, $P < .05$), with physical similarity

Table 4
Hierarchical multiple regression results for children and adults separately, regressing similarity variables (global, functional and physical) on frequency of stimuli selection

<table>
<thead>
<tr>
<th>Variables</th>
<th>Adults</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>$\beta$</td>
<td>$R^2$</td>
<td>$R^2$ change</td>
<td>B</td>
<td>$\beta$</td>
<td>$R^2$</td>
<td>$R^2$ change</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 1</td>
<td>.42**</td>
<td>.42**</td>
<td>.44**</td>
<td>.44**</td>
<td>.45**</td>
<td>.45**</td>
<td>.45**</td>
<td>.45**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global similarity</td>
<td>0.0152</td>
<td>0.65**</td>
<td>0.0066</td>
<td>0.66**</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Step 2</td>
<td>.85**</td>
<td>.85**</td>
<td>.85**</td>
<td>.85**</td>
<td>.85**</td>
<td>.85**</td>
<td>.85**</td>
<td>.85**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global similarity</td>
<td>0.0089</td>
<td>−0.38</td>
<td>0.0066</td>
<td>0.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional similarity</td>
<td>0.0133</td>
<td>1.34*</td>
<td>−0.0002</td>
<td>−0.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical similarity</td>
<td>0.0041</td>
<td>0.28</td>
<td>0.0013</td>
<td>0.21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* $P < .05$.
** $P < .001$. 
Fig. 2. Three-dimensional plots of adults’ (a) and children’s (b) frequency of object selection by physical and functional similarity.
not making a significant contribution. In turn, children’s object selections were primarily explained by global similarity, with neither functional ($\beta = -0.04$) nor physical similarity ($\beta = 0.21$) making a unique significant contribution. In fact, global similarity was only significant in the first step of the regression, suggesting that for children, categorization decisions were made based on an overall impression of the objects that included somewhat balanced considerations of functional and physical similarity. This general pattern of age differences was also manifested in the pattern of correlations. Specifically, for adults, the correlation between functional similarity and frequency of stimuli selection was quite high ($r = .92$, $P < .001$). For children, the correlations between physical and functional similarity to frequency of selection were moderate, and almost identical ($r = .34$ and .33, respectively, $Ps < .05$).

To better illustrate these relations, we plotted the mean rating of functional and physical similarity of each of the 20 test objects with the object’s corresponding proportion of selection. Fig. 2a and b present the resulting plots for adults and children, respectively. In effect, these figures may be taken to represent the map of children and adults’ artifact categories along the dimensions of functional and physical similarity.

4. Discussion

A vast literature on children and adults’ artifact categorization has identified two important respects that affect categorization decisions — namely physical appearance and function — and discussed some of the factors that may sway children’s decisions towards one or the other respect. The goal of the present study was to try and assess the relative contribution of each of these respects to children and adults’ artifact categorization, and do so in as unbiased a fashion as possible. The methodological changes provided us more complete and valid maps of the physical and functional similarity space of children and adults’ artifact categories.

Consistent with earlier findings (Graham et al., 1999; Kemler Nelson, Frankenfield, et al., 2000; Landau et al., 1998; Smith et al., 1996), we found that adults were highly sensitive to functional similarity. Adults chose more objects that were functionally similar to a category standard than did children, and independent ratings of the functional similarity between objects were the best predictor of adults’ category membership decisions.

In contrast, children’s artifact categorization was neither exclusively determined by physical nor by functional similarity, but instead seemed equally sensitive to the two dimensions. In fact, children’s categorization decisions were most strongly related to global judgments of similarity. In other words, when children were left on their own to observe functioning artifacts, and the physical and functional information about the artifacts were equally available, children ended up constructing categories in which members were highly similar on both dimensions. This “baseline” categorization strategy may be a manifestation of children’s sensitiv-

...to the “natural” correlation between object form and object function (Gentner & Ratterman, 1991; Gibson, 1969; see also Kemler Nelson, Frankenfield, et al., 2000, for findings consistent with this conclusion).

Using Jones and Smith’s (1993) terms, the present findings indicate that the two-dimensional similarity space of adults’ artifact categories is quite stretched in the function dimension, but very narrow in the appearance dimension. That is, marginal changes in functional similarity lead to substantial changes in category membership. In turn, the similarity space of children’s artifact categories is equally and quite stretched on both the function and appearance dimensions. Marginal changes in either one of these dimensions lead to substantial changes in category membership.

This analysis intimates that the disparity in the findings of the earlier studies, in which children were given a dichotomous categorization task, might have derived from differences in stimulus composition. It could be that in studies where physical appearance prevailed (e.g., Landau et al., 1998; Smith et al., 1996), the functionally similar objects were not perceived as so clearly similar and thus children were not strongly inclined to include them in the target’s category. In studies where function prevailed (e.g., Diesendruck et al., 2003; Kemler Nelson, Russell, et al., 2000), perhaps the functionally dissimilar test objects were perceived as highly dissimilar and thus children rejected them as members of the category. For the sake of illustration, we can see this pattern of object selection with our stimuli as well. Consider the actual two objects displayed in Fig. 2b with the following “coordinates” of rated percentage of physical and functional similarity, respectively: Object A (88.25, 48.5) and Object B (18.25, 77). Children selected Object A more than twice as often than they did Object B. In contrast, consider this other pair of objects: Object C (77.25, 10) and Object D (35.5, 90.5). In this case, children selected Object D almost twice more often than they did Object C.

It is important to remember that the notion of “similarity” is a psychological construct, not entirely derivable from mathematical computations. Consequently, there are many psychological factors that may make two objects appear more or less similar on any dimension. Thus one may conceptualize many of the factors identified as influencing children’s categorization decisions (e.g., whether or not the functions are demonstrated) as carrying their effect by changing the psychological map of similarity. The present methodology allowed us to see this possibility even under circumstances with little directive influence by other people. In fact, current findings suggest that in dichotomous categorization tasks, children are influenced by characteristics of the stimulus pairings per se, such as the relative distinctiveness of function versus physical appearance (Hammer & Diesendruck, 2002). In a way, these findings demonstrate how the process of comparing features shared by novel categories — undertaken here autonomously by children and adults — plays a fundamental role in conceptual understanding (Namy & Gentner, 2002).

An additional point of theoretical significance coming out of the present study has to do with the null-finding regarding categorization task. Earlier studies found differences in children’s responses to tasks involving name extension and similarity...
judgments (Gentner, 1978; Landau et al., 1998; Smith et al., 1996). Based on these findings, Smith et al. (1996) suggested that naming might recruit special mechanisms of attention and memory. In the present study, we found that adults and children categorized artifacts similarly when asked to extend a standard’s name and when asked to select objects of the same kind as the standard (see Diesendruck & Bloom, 2003, for compatible findings). These findings intimate that naming and kind categorization in children may be governed by the same mechanisms, a conclusion consistent with the notion that children’s categorization are driven by general conceptual beliefs about object kind (Bloom, 2000; Gelman & Diesendruck, 1999; Keil, 1989).

Before we conclude, a number of qualifications about the generalizability of the present findings have to be made. First, it is an interesting question whether the modality by which people get to know artifacts affects their conception of the artifacts. In the present study, participants had only visual information about the artifacts, whereas in others (e.g., Kemler Nelson’s) children often had the opportunity to manipulate the objects. It is plausible that an object’s function appears more salient if children are allowed to physically interact with objects than if they can only watch the objects functioning. In this sense, it is all the more remarkable that children in the present study were as sensitive to the functional information as they were to the physical information. Second and related, due to our theoretical goals, the objects in the present study had to be displayed as computer animations. While it is possible that these objects were conceptualized differently than actual three-dimensional objects (see for instance, Deak & Bauer, 1996, for the effect of stimulus type on children’s categorization), we cannot envision a study with three-dimensional objects that would replicate precisely the present one. Finally, as with every psychological study, it is unclear to what extent the results would be replicated with a different population of children and adults. More to the point, it would be valuable to examine in more detail the developmental changes in the similarity map of artifact categories that occur from preschool to young adulthood.

In conclusion, the present findings set on firmer ground the current theoretical debate between a “similarity” and a “core” view of artifact categorization. Specifically, the study showed that when functional and appearance information about artifacts are simultaneously available to children for the same length of time, through the same medium, and without adult direction, children weigh these two respects equally and highly. Thus an appropriate account should not define children’s artifact categories in terms of a single dimension, but instead must allow for some flexibility. Moreover, the findings revealed that compared to adults, children are over sensitive to the physical similarity between objects, but are adequately sensitive to the functional similarity between objects. With development, children learn to give more weight to what adults believe are relevant for determining artifact category membership — namely, functional similarity — and learn to lower the weight of what is considered less relevant — namely, physical similarity.
Acknowledgments

This research was supported by a grant from The Israel Foundations Trustees (Special Fund for Post-Doctoral Research Grants) to the first author. We thank Paul Bloom for comments on a draft of this paper, Matar Ferera for help with data collection, and the teachers, parents, and children from the following preschools, for participating in the study: Gan Haya — Holon, Gan Rimon — Shfaym, Gan Havatzelet — Kiriat Ono, Ganei Ilan — Kiriat Ono.

References


