Title
One versus many: Multiple examples in word learning.

Authors
Price, Gwendolyn F
Sandhofer, Catherine M

Publication Date
2021-09-01

DOI
10.1016/j.jecp.2021.105173

One Versus Many: Multiple Examples in Word Learning

Gwendolyn F. Price and Catherine M. Sandhofer

University of California, Los Angeles

Author Note

We are grateful to all of the families who participated in this study. We also thank all of the research assistants who coded the data, and our colleagues who provided feedback on previous drafts.

Correspondence concerning this article should be addressed to Gwendolyn F. Price, University of California, Los Angeles, Department of Psychology, 1285 Franz Hall, Box 951563, Los Angeles, CA 90095-1563. Email: gwenfp@gmail.com

Declarations of interest: none

Word Count: 5,175

One Versus Many: Multiple Examples in Word Learning

Abstract

A large body of research indicates that children can map words to categories and generalize the label to new instances of the category after hearing a single instance of the category labeled. Additional research demonstrates that word learning is enhanced when children are presented with multiple instances of a category through comparison or contrast. In this study, 3-year-old children participated in a novel noun generalization task in which a label was given for either 1) a single instance of a category, 2) multiple instances of a category, or 3) contrasting a category instance with non-category members. Children were asked to extend the label to new category at test either immediately (Study 1) or after a 10 second delay (Study 2). The results indicate that when tested immediately, children who heard a single instance labeled outperformed children who were presented with multiple instances. However, when tested after a brief delay, there was no difference between conditions.

Keywords: comparison, contrast, word learning, children

A large body of work demonstrates the benefits of comparing multiple objects (i.e., considering two or more instances simultaneously) versus labeling a single object for category learning and generalization (Gentner & Namy, 1999; Graham et al., 2010; Medin et al., 1993; Namy et al., 1997; Namy & Clepper, 2010a; Namy & Gentner, 2002; Spalding & Ross, 2000; Twomey et al., 2014; Vukatana et al., 2015). For example, providing learners with opportunities to compare multiple instances has been shown to result in: higher levels of exhaustive classification (Namy et al., 1997); increased ability to disregard irrelevant perceptual similarity in favor of relational matches (Gentner et al., 2011); accelerated verb learning (Childers et al., 2017); facilitated solving of complex relational problems (J. A. Dixon & Bangert, 2004) and increased likelihood or reasoning through analogy (Gick & Holyoak, 1983). In one study (Gentner & Namy, 1999), 4-year-old children were taught novel words. In the comparison condition, children were shown multiple examples (e.g., an apple, pear, watermelon, and grapes) and each was labeled with the same novel word. In the other condition, children were shown a single example (e.g., an apple) and it was labeled with a novel word. At test, children were shown a red balloon (a perceptual match to the apple) and a yellow banana (a taxonomic match to the apple) and were asked which one shared the same novel label. The results indicated that children in the comparison condition, who heard multiple category instances labeled, were more likely to select the taxonomic match over the perceptual match. Altogether, the robust finding in this body of work is that situations in which learners are encouraged to compare facilitate learning and generalization.

The ability to generalize a label to new instances of a category emerges over development. At 10 months of age, infants are able to successfully learn specific object-label pairings but fail to generalize a label to new instances of a category until around 2 years of age

(Taxitari et al., 2020). Around this same time, children's productive vocabulary becomes increasingly dominated by object names (Perry et al., 2010; Perry & Samuelson, 2011). In English and many other languages, object categories tend to be highly organized around similarity in shape. For example, the category of "spoons" includes objects that differ in color, material, and size, but largely share a similar spoon-shape. During this time children develop an attentional bias toward shape when generalizing new objects (Landau et al., 1988). That is, when shown a novel object and told, "See this, it's a lum," children will generalize the label lum to new objects that match the original lum in shape but not in color, texture, or material (Smith et al., 2002). This is commonly referred to as the shape bias.

Multiple studies indicate that children map object words to referents with either a single example or at the very least minimal exposure to a word (Behrend et al., 2001; Carey, 2010; Carey & Bartlett, 1978; Casasola, 2005; Goodman et al., 1998; Heibeck & Markman, 1987; Horst & Samuelson, 2008; Jaswal & Markman, 2001; Markson & Bloom, 1997; Wilkinson & Mazzitelli, 2003; Woodward et al., 1994). For example, Woodward, Markman, and Fitzsimmons (1994) presented 18-month old children with a single instance of a category and labeled it with a novel word. They then asked children to generalize the novel word to a new instance of the category. The results indicated that even at 18 months of age, children were able to successfully generalize a new word from viewing just a single example.

The literature seems to indicate that although children can appropriately generalize a category label from hearing a single example labeled in some studies (e.g., Woodward, Markman, and Fitzsimmons, 1994), children benefit from multiple instances of a category labeled in other studies (Gentner & Namy, 1999). On the surface this may seem contradictory, but the situations in which children are able to learn through a single exposure versus situations in which children seem to learn better through comparison are different. One possibility is that multiple examples seem to aid in learning when the to-be-learned information is more abstract (Gentner & Namy, 1999; Thibaut & Witt, 2015) or otherwise difficult (Graham et al., 2010; Medin et al., 1993; Waxman & Klibanoff, 2000). For example, Graham et al. (2010) taught 4-year-old children novel texture words, a category that may be more difficult for 4-year-old children to learn than object categories. Children were presented with either one example or two examples of the category. When asked to find another member of the category, children who viewed two examples showed higher performance than children who viewed a single example. In another study (Spencer et al., 2011), adult learners generalized food categories broadly (i.e., learners categorized items with more variation as being in the same category) when provided with a single example, but more narrowly (i.e., they were more restricted in what variation in items they allowed in the category) when provided with three similar examples simultaneously. It is worth noting that neither broad nor narrow generalization are indicative of more accurate generalizations in this study.

Multiple examples may also be more beneficial when learners must retain word-object pairings over a delay (Twomey et al., 2014). Recent research suggests that despite the fact that children can readily link a word to its referent in fast mapping studies, they appear to have difficulty retaining that link over short time delays. For example, Horst and Samuelson (2008) taught 24-month-old children the labels of novel objects using one example. When asked to select the referent immediately after learning the object label, children selected the correct object. However, when children were asked to select the referent after a 5-minute delay, they responded at chance levels, suggesting that children failed to retain the object-label link over the short delay. Similar work that included multiple exemplars of a category found that children were able to generalize over a delay when given multiple examples, but not when given a single example (Twomey et al., 2014). In addition, an eye-tracking study (Bion et al., 2013) found that children had difficulty retaining a label after a delay when training involved a single example. Although both 24, and 30-month-old children linked novel objects and labels when tested immediately, only the 30-month-old children showed fragile evidence of retention. Thus, although children may learn some word-object pairings quickly, they seem to exhibit difficulty retaining these same pairings over short delays.

Additionally, studies that indicate that viewing multiple instances simultaneously facilitates categorization often do not distinguish between multiple instances that provide learners with opportunities to compare or contrast. The opportunity to compare between multiple instances of a category can provide information about what matters for categorization and highlight the similarities between instances, but it can also indicate features that may be irrelevant to category membership. For example, seeing a red block and a blue block and hearing them both labeled as “block” should indicate that color red does not matter for category membership. Similarly, providing learners with the opportunity to contrast between category members and non-members can also inform about features that do not characterize the category. For example, hearing that a red block “is a block” and a red ball “is not a block” indicates that the color red is not indicative of category membership for blocks. Thus, comparison and contrast can provide, either implicitly or explicitly, information about the type of features that do not matter for category membership.

In two studies, we examined how presenting a single instance of a category versus multiple instances of the category affects young children’s performance on a generalization task. In both experiments, three-year-old children were presented with novel object categories in one of three learning conditions: a single example of the category (One-Example), three positive examples of the category (Comparison), or one example of the category compared to two non-category members (Contrast). Children then participated in a forced choice test which asked them to generalize to a novel instance of the category participated in a forced choice test. In Study 1, children participated in a forced choice test which asked them to generalize to a novel instance of the category immediately after the learning phase. In Study 2, we reversed the order of a distracter and learning phase such that children experienced a slight delay between learning and testing.

Study 1

Participants

Participants were 50 three-year-old children (*M* = 35.76 months, *SD* = 3.34 months), 24 girls and 26 boys. Prior to participating in the study, signed informed consent was obtained from each child's parent or guardian, and verbal assent was obtained from each child. Participants were recruited and tested in local preschool programs. All participants were learning English as a primary language and were fluent speakers of English but were not required to be monolingual English-speakers.

Stimuli

The stimuli in these studies were items the children were unfamiliar with, such as a toilet flapper, that were painted and texturized. No shape, texture, or color repeated between the 8 trials. These objects were 3-D objects the child was able to hold and play with. Figure 1 shows an example of the stimuli for a single trial, and all objects, textures, and colors used in the study are listed in Appendix A. The novel words were one to two syllables in length and followed English phonological conventions. The eight words used in the study were: wug, dac, fess, blick, modi, fep, tog, and lum. The novel words were randomly assigned to a novel category for each participant.

Procedure and Design

Children were randomly assigned to one of three between-subject conditions: the One-example condition (n=16), the Comparison condition (n=18), or the Contrast condition (n=16). There were no significant differences between the three conditions in children's age (F = 1.37, p = .264, η_p^2 = 0.06) nor gender (F = 1.28, p = .287, η_p^2 = 0.05).

There were eight trials. Each trial consisted of a learning phase, a distracter phase, and a test phase. Each trial began with a Distracter phase, which lasted for 10 seconds, then the Learning phase, which lasted for 30 seconds, and finally the Test phase, which lasted for approximately 30 seconds. In total, children learned eight different novel category-label pairings (i.e., one category-label pairing per trial).

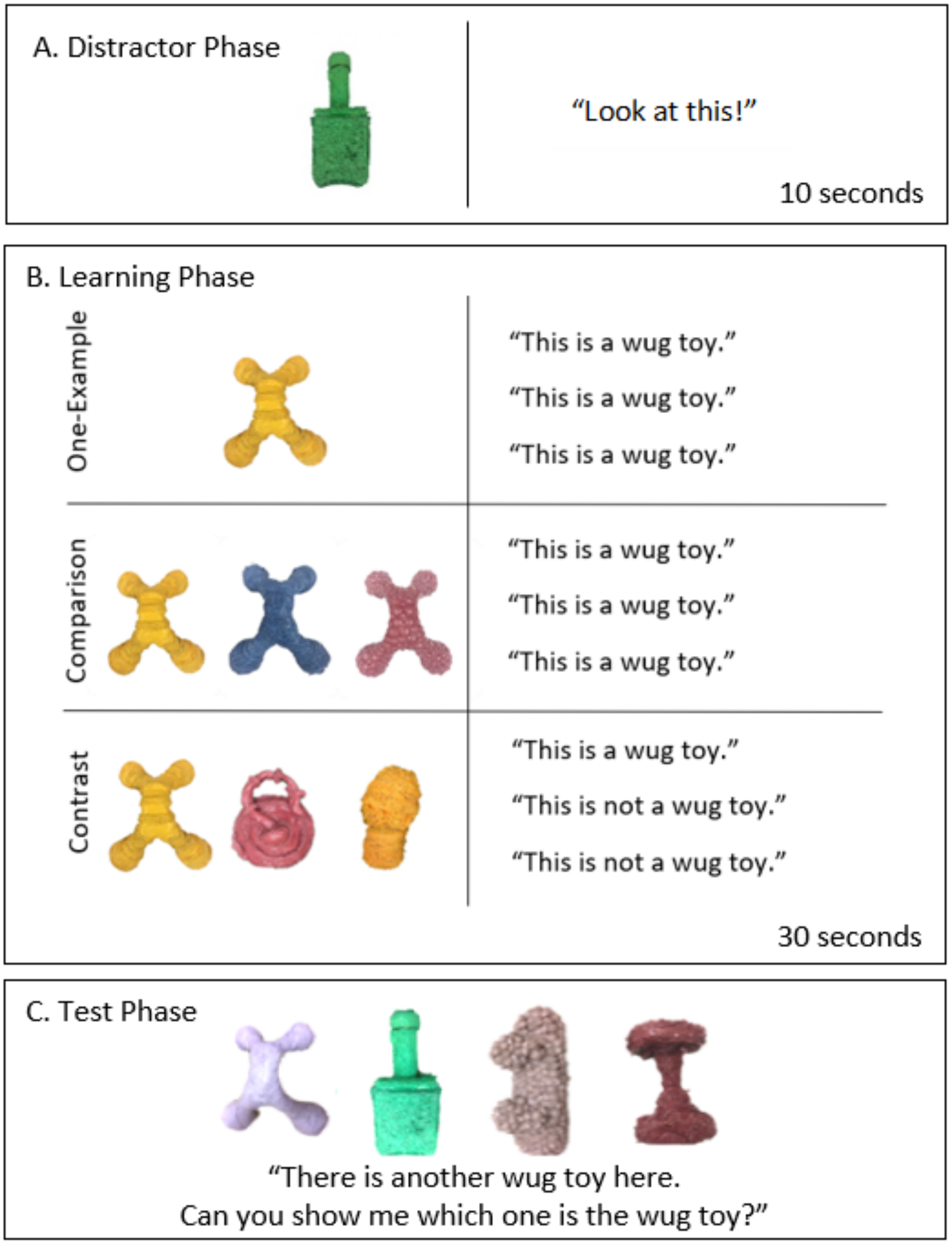


Figure 1. Experimental procedure. A) the Distracter phase B) the Learning phase, children participated in one of three conditions either One-Example, Comparison, or Contrast. C) the Test phase, children were asked to identify the target. In Study 1 children first completed the Distracter phase, then the Learning phase, and finally the Test phase. In Study 2 children first completed the Learning phase, then the Distracter phase, and finally the Test phase.

Distracter Phase

Each of the eight trials began with a distracter phase. The purpose of introducing a distracter object was to have an object present during testing that children had seen during the trial, but that was not the target object, thus ensuring that children were not responding based on the familiarity of objects during the test phase. The item was brought out of a bag and placed in the sight line of the child. The distracter object was presented for 10 seconds and introduced by the experimenter with a neutral phrase that did not contain a label (e.g., "Look at this!"). The child was also encouraged to play with the object (e.g., "You can play with it if you'd like."). After 10 seconds, the object was then removed from sight.

Learning Phase

The learning phase lasted 30 seconds and immediately followed the distracter phase. Children were either shown one or three objects, and all objects were 3-D objects that the children were able to hold and manipulate. An example of the objects presented during the learning phase in each condition are depicted in Figure 1 (Panel B). All labeling sentences were delivered by one experimenter in a neutral-happy tone of voice. All objects of the category children were learning, or target objects were labeled by the experimenter with a novel word (e.g., "This is a wug toy."), and all non-target objects were labeled by the experimenter with a negated novel word (e.g., "This is not a wug toy."). This phrasing was chosen intentionally to create ambiguity to what the novel word refers to (e.g., Slone & Sandhofer, 2017) and used adjectival syntax (Mintz & Gleitman, 2002), which can lead children to consider multiple properties for a novel word, rather than focusing only on shape (Hall et al., 1993; Waxman & Booth, 2001). For example, the syntax follows the same conventions as both "this is a square toy", indicating a shape, and "this is a blue toy," indicating an adjective.

In the One Example condition, children were shown a single target object labeled three times to match the number of times the novel word was heard in each condition (e.g., "This is a wug toy. This is a wug toy. This is a wug toy!"). In the Comparison condition, children were simultaneously shown three target objects from the target category. The objects were the same shape but different from each other in color and texture. Each object was labeled once with a novel word (e.g., "This is a wug toy."). In the Contrast condition, children were also simultaneously shown three objects, one target and two non-target objects. One non-target object was a different shape and texture than the target object but matched the target object in color. The other non-target object was a different shape and color than the target object but the same texture as the target object. The target object was labeled with the novel word once (e.g., "This is a wug toy."), the other two objects were each labeled once with a negation (e.g., "This is <u>not</u> a wug toy."). See supplemental material for additional condition information.

Test Phase

The test phase lasted approximately 30 seconds and immediately followed the learning phase. The objects and procedure for the test phase were identical for all conditions. Children were presented with four objects simultaneously: 1) the target object which shared the shape as the target objects in the learning phase but was a different texture and color, 2) the distracter object, 3) an unfamiliar object: an object the child had not seen before that differed from the target object in shape, texture, and color , and 4) a false match object: an object with a different shape from the target object, but with a texture and color that children had previously seen during the learning phase. Because comparison and contrast should provide information on features that do not matter for category membership, the false match object was created to test whether children were able to make use of this information. The false match object presented an amalgamation of features that the learning phase of Comparison and Contrast conditions should indicate are irrelevant for category membership. For children in the One-example condition, however, the false match object operated no differently than an additional unfamiliar object.

The experimenter prompted the child by saying “There is another wug toy here, can you show me the wug toy?” The test phase ended when children made a selection either by pointing to or picking up and handing over one of the objects. Children were given neutral feedback following their selection (e.g., “Thank you!”). Once the test phase ended, the procedure was immediately repeated for the next trial.

Results and Discussion

First, to examine how the learning condition affected test performance, we conducted a one-way Analysis of Variance (ANOVA). Figure 2 depicts the average performance by condition. The one-way ANOVA revealed a significant difference among conditions ($F = 4.998$, $p = .011$, $\eta_p^2 = 0.18$). A post-hoc power analysis was conducted and observed power for this ANOVA was .788.

We then examined specific condition differences through post-hoc independent samples t-tests and confirmed significant differences between the One-Example (M = 5.31, SD = 1.70) condition and the Comparison (M = 3.66, SD = 2.03) condition ($t = 2.54$, $p = .016$, $d = 0.88$) and the One-Example condition and the Contrast (M = 3.38, SD = 1.50) condition ($t = 3.25$, $p = .003$, $d = 1.15$). There was no significant difference between the Comparison and Contrast conditions ($t = 0.17$, $p = .867$, $d = 0.06$). Additionally, all conditions performed significantly above chance (Comparison: $t = 3.49$, $p = 0.003$, $d = 0.82$; Contrast: $t = 3.30$, p=0.005, $d = 0.82$; One-Example: $t = 7.79$, $p < 0.001$, $d = 1.95$).

One last set of analyses sought to determine whether the False Match object was chosen at levels above chance. One sample t-tests indicated that children in the One-Example condition chose the False Match significantly below chance levels (t = -8.348, p < 0.001, d = -1.84), but children chose the False Match object at chance levels for the Comparison (t = -0.838, p =.414, d = -0.198) and Contrast (t = -1.826, p = .088, d = -0.456) conditions.

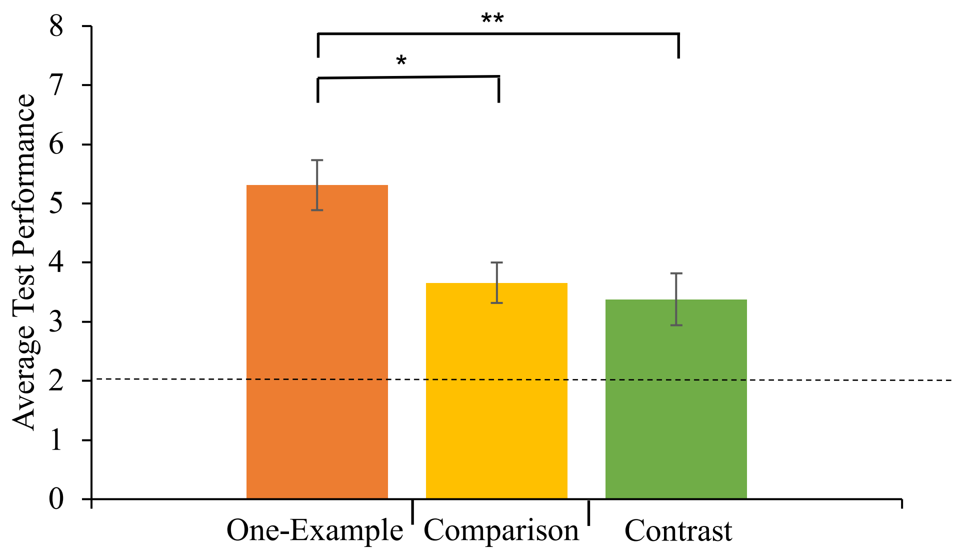


Figure 2. Mean number of target selections at test out of 8 trials in Study 1: Comparison (M=3.67), Contrast (M=3.38), and One-Example (M=5.31). Error bars depict standard error.

The results indicate that children were able to appropriately generalize to new instances of the category when they were presented a single example of the category. This finding is not surprising given the large number of studies showing that children generalize a category label from hearing a single example labeled (e.g., Behrend et al., 2001; Carey, 2010; Carey and Bartlett, 1978; Casasola, 2005; Goodman et al., 1998; Heibeck and Markman, 1987; Horst and Samuelson, 2008; Jaswal and Markman, 2001; Markson and Bloom, 1997; Wilkinson and

Mazzitelli, 2003; Woodward et al., 1994). Additionally, that the categories were defined by the shape of the objects may have contributed to children's ease of generalization. By two years of age, children typically attend to the shape of objects when they are labeled with syntax that is indicative of count nouns (e.g., Landau et al., 1988; Smith et al., 2002).

Surprisingly, however, children performed lower in the Comparison and Contrast conditions than in the One-Example condition. Further, despite the fact that the Comparison and Contrast conditions provided information that the features present in the false match object did not matter for category membership, children in these conditions did not choose the false match object at below chance levels, suggesting that children were not able to make use of the information indicating which features were irrelevant for category membership.

One possibility is that children were affected by the variation between objects which may have lowered attention to shape (Hanania & Smith, 2010). A non-significant difference between comparison and contrast ($t = 0.45$, $p = 0.65$, $d = 0.16$) may suggest there are no differences between these two types of presentation styles.

Because previous studies indicate that comparison facilitates learning in more difficult learning settings, introducing a small amount of difficulty, through the addition of a brief delay between learning and testing, may support greater learning in the comparison condition. Study 2 examined learning via one example, comparison, and contrast when children are asked to retain word-object pairings over a brief delay.

Study 2

Participants

Participants were 48 3-year-old children ($M = 35.10$ months, $SD = 3.69$ months), 23 girls and 25 boys. There were 16 participants in each condition. Participants were recruited from local preschool programs and through a birth records database. All participants were learning English as a primary language.

Stimuli Design and Procedure

The stimuli and design were the same as in Study 1. The procedure was the same as Study 1 with one exception: the three phases were ordered differently in Study 2. Children were first presented with the Learning phase, followed by the Distracter phase, then, and finally the Test phase. In moving the distracter phase in between the learning phase and the test, we introduced a small (10 second) delay between when children learned the words and when they were tested on the words.

Results and Discussion

Figure 3 depicts the average performance by condition. A one-way ANOVA was non-significant ($F = 0.23$, $p = 0.792$, $\eta_p^2 = 0.01$). Therefore, the short delay instituted in Study 2 did not show the same benefit for One-Example seen in Study 1. Children in the Comparison (M = 4.33, SD = 1.54), Contrast (M = 3.80, SD = 2.04), and One-Example (M = 4.13, SD = 1.94) conditions all performed above chance (Comparison: $t = 6.20$, $p < 0.001$, $d = 1.55$; Contrast: $t = 3.89$, $p = 0.001$ $d = 0.97$; One-Example: $t = 4.16$, $p = 0.001$, $d = 1.04$). Selection rates for each object type at test for both studies are reported in Table 1.

Table 1

	Study 1				Study 2			
	Target	Distracter	Novel	False Match	Target	Distracter	Novel	False Match

One-Example	66.41% (5.31)	14.84% (1.19)	11.72% (0.94)	6.25% (0.50)	50.78% (4.06)	10.94% (0.88)	17.97% (1.44)	20.31% (1.63)
Comparison	45.83% (3.67)	13.89% (1.11)	16.67% (1.33)	21.53% (1.72)	53.91% (4.31)	12.50% (1.00)	15.63% (1.25)	17.97% (1.44)
Contrast	42.19% (3.38)	19.53% (1.56)	21.88% (1.75)	18.75% (1.50)	48.44% (3.86)	8.59% (0.69)	21.88% (1.75)	19.53% (1.56)

Selection Rates for All Objects at Test

Note. The numbers in parentheses are the average number of times that item was selected at test out of 8. All numbers rounded to the nearest 100^{th}.

Additional one sample t-tests were conducted to determine whether the False Match object was chosen at levels above chance. Across all conditions, children chose the False Match object at chance levels (One-Example: t = -1.307, p = 0.211, d = -0.33; Comparison: t = -2.058, p = .057, d = -0.514; Contrast: t = -1.385, p = .186, d = -0.346).

Lastly, analyses were conducted to determine whether the implementation of a delay affected any condition type. The One-Example condition trended toward a significant decrease between immediate (Study 1) to delayed test (Study 2) (t = 1.914, p = 0.065, d = 0.68). On the other hand, neither the Comparison nor the Contrast conditions were significantly different between the two studies (Study 1: t = -1.045, p = 0.304 d = 0.36; Study 2: t = -0.784, p = 0.439, d = 0.28).

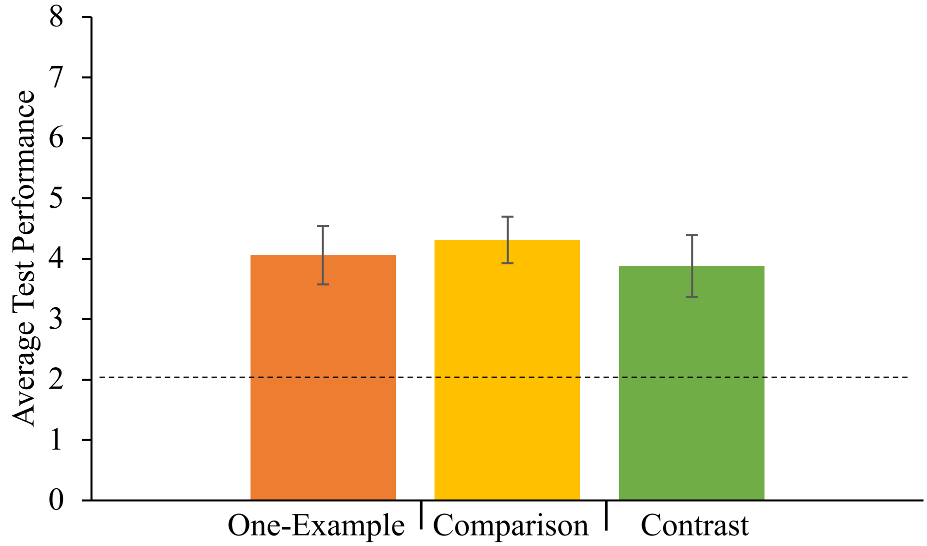


Figure 3. Average selection of target at test out of 8 trials for Study 2: Comparison (4.31), Contrast (3.80), and One-Example (4.06). Error bars represent standard error.

General Discussion

In these experiments, we examined two sets of findings: both a single example *and* multiple examples can lead children to generalize words to new exemplars. We hypothesized that there could be two possibilities: first, that the One Example condition would show lower performance because although children can pick up shape words quickly, viewing a single example may not provide enough information to generalize to a new object compared to conditions that provide children with multiple examples or explicit information about the features that do not matter for category membership. The second possibility was that multiple examples in the Comparison or Contrast conditions would lead to better performance because multiple examples have supported children's category formation and generalization in multiple studies. We found that for shape categories, when children were tested immediately, hearing a single example labeled led to higher performance than hearing multiple examples labeled (both for comparative and contrastive examples). Thus, 3-year-old children who have a history of learning and generalizing shape words can readily generalize new shape labels from a single example. However, when tested after a 10 second delay, performance no longer differed between children who heard a single example labeled and children who heard multiple examples labeled.

These results prompt an open question: why was there a difference when children had to remember the word for 10 seconds? First, although, a 10-second delay is a very brief delay, and admittedly shorter than delays used in other retention studies, short-term memory is also very brief (Cowan, 2008) and when items are not rehearsed or actively maintained, short term memory persists for mere seconds. Because children were actively engaging with new information during the 10 second delay (i.e., the introduction of the distracter object in Study 2) children were unlikely to be rehearsing the previously learned information during the delay. Second, some forgetting can be useful for learning if the memory trace is strong enough to be recalled (Bjork, 2011). However, too much forgetting inhibits recall (Storm, 2011; Vlach, 2014). The memory formed through experience with a single example may not be a strong enough memory trace to survive past a 10 second retention interval. Third, although no label was provided during the distracter phase, children were invited to engage with the object and were told "Look at this!" and, "You can play with it if you'd like." Thus, the interaction between the distracter phase and the learning phase shared similarity in experience – in both cases children were perceiving and interacting with novel 3-D objects.

In this way, the distracter phase may have provided some interference with the memories formed during the learning phase. Distractions in the environment lower word learning performance in toddlers (W. E. Dixon et al., 2006), so it is possible that the distracter object may have reduced performance for the One-Example condition in Study 2 because it introduced distracting material rather than simply being a delay. As such, the effect of differences between the One-Example conditions from Study 1 and 2 might then have more to do with the recency of the distracter than a decay of memory. Future work needs to investigate whether the delay increased forgetting, retrieval, or acted as retroactive interference.

An additional intriguing possibility is that the difference between performance in Study 1 and Study 2 could be due to order effects. The order in which items are presented in can impact learning and generalization (Kovack-Lesh & Oakes, 2007; Samuelson & Horst, 2007; Sandhofer & Doumas, 2008). Beyond acting as a delay, the change in the presentation order of the learning and distracter phases of each trial between Study 1 and 2 may have impacted children's ability to generalize to new instances of the category. Order effects are particularly intriguing because one likely explanation for order effects involves delays between when particular instances are presented in sequence. Although the current study wasn't designed to distinguish between the differential effects of delays and order effects, future research should seek to disentangle the contributions of each to learning.

The current results are supported by findings that show that delays between learning and test show differential impacts depending on how the information is presented during learning. For example, a study by Vlach, Ankowski, and Sandhofer (2012) taught children novel words for shape categories by giving them four examples of the category either simultaneously (i.e., all examples presented at the same time), massed (i.e., examples shown one after another), or spaced (i.e., examples shown one after another with a delay in between each instance). Children were tested either immediately afterwards, or after a 15-minute delay. Children who saw all the examples simultaneously during learning performed best when tested immediately but showed reduced performance after a delay. However, children who saw the examples spaced out in time did not show declines in performance after a delay. The current studies expand upon these findings by showing when children learn words via a single example, even a very brief 10 second delay can reduce generalization performance; however, comparison and contrast did not show similar reductions.

Comparison conditions have been found to be helpful in many word learning studies (Ankowski et al., 2013; Graham et al., 2010; Namy & Clepper, 2010a), particularly when contrasted with conditions in which children are provided with a single example. (Gentner & Namy, 1999; Oakes et al., 2009). Interestingly in the studies reported here, a single example resulted in higher performance versus when children were given the opportunity to compare between multiple objects. In Namy and Gentner's (1999) study, when provided with a single example, children matched based on perceptual features. In their study, comparison likely helped children look beyond perceptual similarities. Thus, it is not surprising that in the current study, children succeeded in making a perceptual match with a single example. What is surprising is that children performed lower in the Comparison condition than they did in the One Example condition.

One possibility is that children's bias to attend to shape and generalize new labels for objects on the basis of shape (Arias-Trejo, 2010; Landau et al., 1988) may have benefitted the generalizing shape-based objects from a single example. For other kinds of categories, such as texture and relational categories, one example may not be enough to generalize the category structure (Gentner & Namy, 1999; Graham et al., 2010). However, presumably younger children, who have not yet developed a bias to attend to the shape of objects, may benefit more from comparison because generalization by shape is difficult at earlier ages (Taxitari et al., 2020).

There are other potential explanations of why we might find a difference between the One Example condition and the Comparison and Contrast conditions. One possibility is that the syntax used (i.e., "This is an X toy.") while ambiguous to property type, as in, it could be a "square toy" or "blue toy" could make the task more difficult for children because children might interpret the syntax to indicate a non-shape adjective (Hall et al., 1993; Waxman et al., 1997; Waxman & Booth, 2001). Perhaps then children are less likely to consider a shape match to be a correct choice. As a result, children in the One-Example condition may have learned the novel word as a color or texture, but instead defaulted to shape when those features did not appear at test. Second, the differences between the conditions include different number of objects and resultingly differential object variability displayed which may affect competition for attention. Further research may be able to disentangle these possibilities.

Interestingly, the Contrast condition did not seem to help children learn new words differently from the Comparison condition. The kind of multiple examples, that is, comparative or contrastive examples, did not seem to make a difference in children's word learning in both studies. Additionally, despite there being only one target example like in the One-Example condition, children in the Contrast condition performed significantly lower at test than children in the One-Example condition. So instead of helping children determine what did not belong in the category, the contrasting examples in the Contrast condition hurt performance at test and did so to the same degree as the Comparison condition. While the Comparison and Contrast conditions provided different kinds of information to the learner and may then affect learning differently despite their similar results on these tasks, it is also possible that they both served as superfluous or even distracting material from the children's natural inclination to focus on the shape of the object due to the shape bias. Some research shows the learning benefits of multiple examples, and how they differ based on whether the examples are comparative or contrastive (Gentner & Namy, 1999; Graham et al., 2010; Namy & Clepper, 2010). In these studies, the learning benefits of Comparison and Contrast could have been stifled due to the short or nonexistent delays between learning and test, and the shape bias making the additional examples unnecessary. Future research should examine these learning conditions with longer delays and other category types in order to fully understand their potential for learning.

These studies sought to ask how presenting a single instance of a category versus multiple instances of the category affects young children's performance on a generalization task. In summary, the current studies further the literature on comparison and contrast and indicates that multiple examples do not facilitate categorization in all learning situations. These studies also found an instance in which comparing examples of the same category (i.e., Comparison condition) and examples of different categories (i.e., Contrast condition) equally contributed to children's word learning, showing that in at least some instances whether the example is of the same category or not may not be as important as how many examples there are. Further, the current studies highlight the role of even a very brief delay during learning and testing and how variations in how information is presented may differentially support learning and retention.

In sum, we found that when given an immediate test, children in the One Example condition outperformed children who were given multiple examples to compare or contrast, but after a short delay, the difference in performance was no longer present. This study adds to the literature on generalization by presenting an instance in which performance extending a novel label is increased when provided fewer examples rather than more examples. These results also add to an increasing number of studies indicating that a short delay can affect performance in generalization tasks and in doing so both sheds new light on children's word learning and lays the groundwork for future research.

References

Ankowski, A. A., Vlach, H. A., & Sandhofer, C. M. (2013). Comparison Versus Contrast: Task Specifics Affect Category Acquisition. *Infant and Child Development*, *3*, 1–23. https://doi.org/10.1002/icd

Arias-Trejo, N. (2010). Young childrens extension of novel labels to novel animate items in three testing conditions. *International Journal of Behavioral Development*, *34*(3), 206–217. https://doi.org/10.1177/0165025409350951

Behrend, D. A., Scofield, J., & Kleinknecht, E. E. (2001). Beyond fast mapping: Young children's extensions of novel words and novel facts. *Developmental Psychology*, *37*(5), 698–705. https://doi.org/10.1037/0012-1649.37.5.698

Bion, R. A. H., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word – object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*(1), 39–53. https://doi.org/10.1016/j.cognition.2012.08.008

Bjork, R. A. (2011). On the symbiosis of learning, remembering, and forgetting. In *Successful Remembering and Successful Forgetting: A Festschrift in Honor of Robert A. Bjork* (pp. 1–21). Psychology Press.

Carey, S. (2010). Beyond fast mapping. *Language Learning and Development*, *6*(3), 184–205. https://doi.org/10.1080/15475441.2010.484379

Carey, S., & Bartlett, E. (1978). Acquiring a Single New Word. *Papers and Reports on Child Language Development*, *15*, 17–29.

Casasola, M. (2005). When less is more: How infants learn to form an abstract categorical representation of support. *Child Development*, *76*(1), 279–290. https://doi.org/10.1111/j.1467-8624.2005.00844.x

Childers, J. B., Paik, J. H., Flores, M., & Lai, G. (2017). Does Variability Across Events Affect Verb Learning in English, Mandarin, and Korean? *Cognitive Science*, *41*, 808–830. https://doi.org/10.1111/cogs.12398

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Prog Brain Res*, *169*, 323–338. https://doi.org/10.1016/S0079-6123(07)00020-9.What

Dixon, J. A., & Bangert, A. S. (2004). On the spontaneous discovery of a mathematical relation during problem solving. *Cognitive Science*, *28*(3), 433–449. https://doi.org/10.1016/j.cogsci.2003.12.004

Dixon, W. E., Salley, B. J., & Clements, A. D. (2006). Temperament, distraction, and learning in toddlerhood. *Infant Behavior and Development*, *29*(3), 342–357. https://doi.org/10.1016/j.infbeh.2006.01.002

Gentner, D., Anggoro, F. K., & Klibanoff, R. S. (2011). Structure Mapping and Relational Language Support Children's Learning of Relational Categories. *Child Development*, *82*(4), 1173–1188. https://doi.org/10.1111/j.1467-8624.2011.01599.x

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

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology*, *15*, 1–38. https://doi.org/0010-0285183/010001-38

Goodman, J. C., McDonough, L., & Brown, N. B. (1998). The Role of Semantic Context and Memory in the Acquisition of Novel Nouns. *Child Development*, *69*(5), 1330–1344.

https://doi.org/10.1111/j.1467-8624.1998.tb06215.x

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

Hall, D. G., Waxman, S. R., & Hurwitz, W. M. (1993). How two- and four-year-old children interpret adjectives and count nouns. *Child Development*, *64*(6), 1651–1664. https://doi.org/https://www.jstor.org/stable/1131461

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

Horst, J. S., & Samuelson, L. K. (2008). Fast mapping but poor retention by 24-month-old infants. *Infancy*, *13*(2), 128–157. https://doi.org/10.1080/15250000701795598

Jaswal, V. K., & Markman, E. M. (2001). Learning proper and common names in inferential versus ostensive contexts. *Child Development*, *72*(3), 768–786. https://doi.org/10.1111/1467-8624.00314

Kovack-Lesh, K. A., & Oakes, L. M. (2007). Hold your horses: How exposure to different items influences infant categorization. *Journal of Experimental Child Psychology*, *98*(2), 69–93.

Landau, B., Smith, L. B., & Jones, S. S. (1988). The Importance of Shape in Early Lexical Learning. *Cognitive Development*, *3*, 299–321. https://doi.org/https://doi.org/10.1016/0885-2014(88)90014-7

Markson, L., & Bloom, P. (1997). Evidence against a dedicated system for word learning in children. *Nature*, *385*, 813–815.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993). Respects for Similarity. *Psychological Review*, *100*(2), 254–278. https://doi.org/10.1037/0033-295X.100.2.254

Mintz, T. H., & Gleitman, L. R. (2002). Adjectives really do modify nouns: The incremental and restricted nature of early adjective acquisition. *Cognition*, *84*(3), 267–293. https://doi.org/10.1016/S0010-0277(02)00047-1

Namy, L. L., & Clepper, L. E. (2010a). The differing roles of comparison and contrast in children's categorization. *Journal of Experimental Child Psychology*, *107*(3), 291–305. https://doi.org/10.1016/j.jecp.2010.05.013

Namy, L. L., & Clepper, L. E. (2010b). The differing roles of comparison and contrast in children's categorization. *Journal of Experimental Child Psychology*, *107*(3), 291–305. https://doi.org/10.1016/j.jecp.2010.05.013

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. https://doi.org/https://doi.org/10.1037/0096-3445.131.1.5

Namy, L. L., Smith, L. B., & Gershkoff-Stowe, L. (1997). Young children's discovery of spatial classification. *Early Childhood Research Quarterly*, *12*(2), 163–184.

Oakes, L. M., Kovack-Lesh, K. A., & Horst, J. S. (2009). Two are better than one: Comparison influences infants' visual recognition memory. *Journal of Experimental Child Psychology*, *104*(1), 124–131. https://doi.org/10.1016/j.jecp.2008.09.001

Perry, L. K., & Samuelson, L. K. (2011). The shape of the vocabulary predicts the shape of the bias. *Frontiers in Psychology*, *2*, 1–12. https://doi.org/10.3389/fpsyg.2011.00345

Perry, L. K., Samuelson, L. K., Malloy, L. M., & Schiffer, R. N. (2010). Learn locally, think globally: Exemplar variability supports higher-order generalization and word learning. *Psychological Science*, *21*(12), 1894–1902. https://doi.org/10.1177/0956797610389189.Learn

Samuelson, L. K., & Horst, J. S. (2007). Dynamic noun generalization: Moment-to-moment interactions shape children's naming biases. *Infancy*, *11*(1), 97–110. https://doi.org/10.1207/s15327078in1101_5

Sandhofer, C. M., & Doumas, L. A. A. (2008). Order of Presentation Effects in Learning Color Categories. *Journal of Cognition and Development*, *9*(2), 194–221. https://doi.org/10.1080/15248370802022639

Slone, L. K., & Sandhofer, C. M. (2017). Consider the category: The effect of spacing depends on individual learning histories. *Journal of Experimental Child Psychology*, *159*, 34–49. https://doi.org/10.1016/j.jecp.2017.01.010

Smith, L. B., Jones, S. S., Landau, B., Gershkoff-Stowe, L., & Samuelson, L. (2002). Object name learning provides on-the-job training for attention. *Psychological Science*, *13*(1), 13–19. https://doi.org/10.1111/1467-9280.00403

Spalding, T. L., & Ross, B. H. (2000). Concept learning and feature interpretation. *Memory and Cognition*, *28*(3), 439–451. https://doi.org/10.3758/BF03198559

Spencer, J. P., Perone, S., Smith, L. B., & Samuelson, L. K. (2011). Learning Words in Space and Time: Probing the Mechanisms Behind the Suspicious-Coincidence Effect. *Psychological Science*, *22*(8), 1049–1057. https://doi.org/10.1177/0956797611413934

Storm, B. C. (2011). The Benefit of Forgetting in Thinking and Remembering. *Current Directions in Psychological Science*, *20*(5), 291–295. https://doi.org/10.1177/0963721411418469

Taxitari, L., Twomey, K. E., Westermann, G., & Mani, N. (2020). The Limits of Infants' Early Word Learning. *Language Learning and Development*, *16*(1), 1–21. https://doi.org/10.1080/15475441.2019.1670184

Thibaut, J., & Witt, A. (2015). Young children ' s learning of relational categories : multiple comparisons and their cognitive constraints. *Frontiers in Psychology*, *6*(May), 1–10. https://doi.org/10.3389/fpsyg.2015.00643

Twomey, K. E., Ranson, S. L., & Horst, J. S. (2014). That's more like it: Multiple exemplars facilitate word learning. *Infant and Child Development*, *122*(October 2013), 105–122. https://doi.org/10.1002/icd.1824

Vlach, H. A. (2014). The Spacing Effect in Children ' s Generalization of Knowledge : Allowing Children Time to Forget Promotes Their Ability to Learn. *Child Development Perspectives*, *8*(3), 163–168. https://doi.org/10.1111/cdep.12079

Vlach, H. A., Ankowski, A. A., & Sandhofer, C. M. (2012). At the same time or apart in time? The role of presentation timing and retrieval dynamics in generalization. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, *38*(1), 246–254. https://doi.org/10.1037/a0025260

Vukatana, E., Graham, S. A., Curtin, S., & Zepeda, M. S. (2015). One is Not Enough: Multiple Exemplars Facilitate Infants' Generalizations of Novel Properties. *Infancy*, *20*(5), 548–575. https://doi.org/10.1111/infa.12092

Waxman, S. R., & Booth, A. E. (2001). Seeing pink elephants: Fourteen-month-olds' interpretations of novel nouns and adjectives. *Cognitive Psychology*, *43*(3), 217–242. https://doi.org/10.1006/cogp.2001.0764

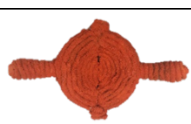
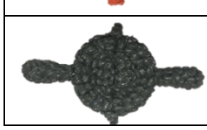
Waxman, S. R., & Klibanoff, R. S. (2000). The Role of Comparison in the Extension of Novel Adjectives. *Developmental Psychology*, *36*(5), 571–581. https://doi.org/10.1037//O012-1649.36.5.571

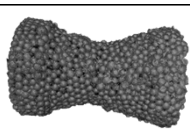
Waxman, S. R., Senghas, A., & Benveniste, S. (1997). A cross-linguistic examination of the noun-category bias: Its existence and specificity in French- and Spanish-speaking preschool-aged children. *Cognitive Psychology*, *32*(3), 183–218. https://doi.org/10.1006/cogp.1997.0650

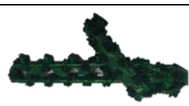
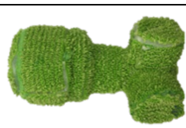
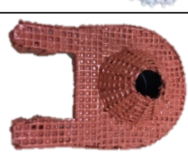
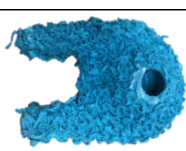
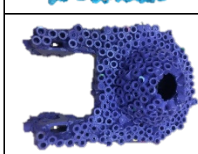
Wilkinson, K. M., & Mazzitelli, K. (2003). The effect of “missing” information on children’s retention of fast-mapped labels. *Journal of Child Language*, *30*(1), 47–73. https://doi.org/10.1017/S0305000902005469

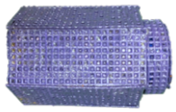
Woodward, A. L., Markman, E. M., & Fitzsimmons, C. M. (1994). Rapid word learning in 13- and 18-month-olds. *Developmental Psychology*, *30*(4), 553–566. https://doi.org/10.1037//0012-1649.30.4.553

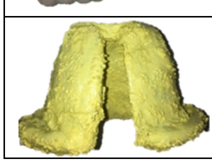
Appendix A

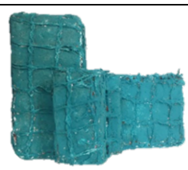
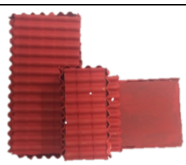
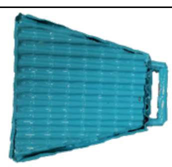
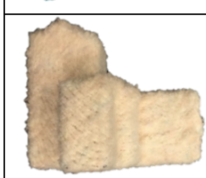
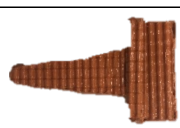
Object	Texture	Color
TRIAL A		
	Pony Beads	Pyrrole Red
	Bordette	Cerulean Blue
	Couscous	Moss Green
	Regular Yarn	Satin Granite Gray
	Bordette	Moss Green
	Regular Yarn	Cerulean Blue
	Aluminum Foil	Lavender
	Regular Yarn	Moss Green
	Velcro	Raw Sienna
TRIAL B		
	Giraffe Print Bumpy Cloth	Cranberry
	Large Yarn	Orange Spray
	White Beans	Dark Moss Green

	Burlap	Yellow
	Large Yarn	Dark Moss Green
	Burlap	Orange Spray
	Grip Liner	Light Bright Blue
	Burlap	Dark Moss Green
	Webbing	Flat Red-Brown
TRIAL C		
	Perler Beads (lying flat)	Ocean Breeze
	Pipe Cleaners	Silver
	Elbow Pasta	Golden Brown
	Bean Bag Filler	Fuchsia
	Pipe Cleaners	Golden Brown
	Bean Bag Filler	Silver
	Scour Pads	Hunter Green
	Bean Bag Filler	Golden Brown
	Scratchy Fabric	Tangerine
TRIAL D		

	Wide Mesh	Neon Yellow
	Duct Tape	Ivory
	Thin Fabric	Violet
	Large Grain Sand	Cobalt Turquoise
	Duct Tape	Violet
	Large Grain Sand	Ivory
	Pom Pom Trim	Green Isle
	Large Grain Sand	Violet
	Mop Pad	Light Green
TRIAL E		
	Hot Glue Dots	Neutral Gray
	Hard Plastic Grid	Terra Cotta
	Crocheted Fabric	Seaside Blue
	Perler Beads (standing up)	Purple

	Hard Plastic Grid	Purple
	Crocheted Fabric	Terra Cotta
	Cosmetic Sponges	Pink Blast
	Crocheted Fabric	Purple
	Plastic Spikes	Black
TRIAL F		
	Sponge	Mint Green
	Regular Yarn (spaced out 1 in)	Dark Yellow
	Loofah Mesh	Satin Wildflower Blue
	Faceted Jewels	Burgundy
	Regular Yarn (spaced out 1 in)	Burgundy

	Loofah Mesh	Dark Yellow
	Bandage	Purple Pearl
	Loofah Mesh	Burgundy
	Pom Poms	Latte
TRIAL G		
	Metal Scrubber	Permanent Green Light
	Round Foam Strips	Sea Mist Pearl
	Cellophane	Pink
	Classic Loft Batting	Light Grey
	Cellophane	Sea Mist Pearl
	Round Foam Strips	Light Grey
	Washcloth	Light Yellow

	Cellophane	Light Grey
	Small Sand	Dark Orange
TRIAL H		
	None – smooth	Dark Peach
	Wide Cloth & Wire Netting	Aqua
	Felt	Orange
	Adobe Style Plastic Roof Tiles	Bright Red
	Wide Cloth & Wire Netting	Orange
	Adobe Style Plastic Roof Tiles	Aqua
	Fuzzy Socks	Beige
	Adobe Style Plastic Roof Tiles	Orange

	Long Bumpy Mop Cloth	Light Pink