

10-17-2016

# Word and Concept Learning in Poor Comprehenders

Kayleigh Ryherd

*University of Connecticut - Storrs*, [kayleigh.ryherd@uconn.edu](mailto:kayleigh.ryherd@uconn.edu)

---

## Recommended Citation

Ryherd, Kayleigh, "Word and Concept Learning in Poor Comprehenders" (2016). *Master's Theses*. 1008.  
[http://digitalcommons.uconn.edu/gs\\_theses/1008](http://digitalcommons.uconn.edu/gs_theses/1008)

This work is brought to you for free and open access by the University of Connecticut Graduate School at DigitalCommons@UConn. It has been accepted for inclusion in Master's Theses by an authorized administrator of DigitalCommons@UConn. For more information, please contact [digitalcommons@uconn.edu](mailto:digitalcommons@uconn.edu).

Word and Concept Learning in Poor Comprehenders

Kayleigh Ryherd

B.A., The George Washington University, 2014

A Thesis

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

At the

University of Connecticut

2016

APPROVAL PAGE

Masters of Science Thesis

Word and Concept Learning in Poor Comprehenders

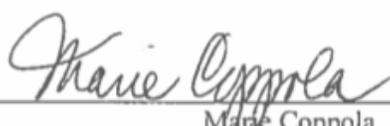
Presented by Kayleigh Ryherd, B.A.

Major Advisor



Nicole Landi

Associate Advisor



Marie Coppola

Associate Advisor



Eiling Yee

University of Connecticut  
2016

### **Acknowledgements**

I would like to thank my advisor, Dr. Landi, for her assistance on all stages of this project. I would also like to thank the UConn participants and all of the undergraduate RAs who ran the study for me. This research was funded by National Institutes of Health P01 HD001994-46, Nature and acquisition of the speech code and reading; Project 4: *Examinations of skilled and impaired spoken and written comprehension (PL, Nicole Landi; P01 PI, Jay G. Rueckl)*.

### **Abstract**

The current study aimed to examine how word and concept learning is related to comprehension ability beyond decoding skill. The study considered nonverbal concept construction and verbal label mapping in explicit and implicit paradigms to determine which aspects of concept learning were easiest for individuals with low comprehension skill. A sample of 28 undergraduate student participants completed two sets of word and concept learning, where they learned categories of novel items grouped on one nonverbal and one verbal feature. After learning, they were asked to indicate which items belonged to the same category. Accuracy, reaction time, and eye movements were measured during this test task. Using a continuous analysis, comprehension was not related to performance on the category learning task, but vocabulary and nonverbal IQ were. A later group analysis, splitting the subjects into better comprehenders (BC), poorer comprehenders with high vocabulary (PCHV) and poorer comprehenders with low vocabulary (PCLV) showed that PCLV subjects performed worse than their PCHV peers on all types of learning. Thus, the results suggest that vocabulary, but not comprehension, is related to word and concept learning. Limitations regarding the comprehension method used for the study are discussed. This study is one of few to consider novel word and concept learning in relation to comprehension.

## Word and Concept Learning in Poor Comprehenders

### Introduction

The simple view of reading posits that reading comprehension is the product of skilled decoding, or the translation of the sounds of speech into the letters of text, and listening comprehension (Gough & Tunmer, 1986). Indeed, difficulties with decoding often result in poor reading comprehension, a deficit most often classified as dyslexia. However, there also exists a population of individuals with impaired reading comprehension and adequate decoding skills. These individuals are known as poor comprehenders.

Poor comprehenders (PCs) exhibit poor reading comprehension skills despite intact decoding and general cognitive function. They make up as much as ten percent of adults and school-aged children. As a group, PCs show worse educational outcomes than their TD peers (Ricketts, Sperring, and Nation, 2014). Although the defining deficit in PCs lies in reading comprehension, they frequently exhibit difficulties in listening comprehension (Catts, Fey, & Zhang, 1999; Nation, Cocksey, Taylor, & Bishop, 2010), suggesting that their core weakness may be in language rather than reading *per se*. In addition to general comprehension difficulties, members of this population show many other weaknesses. Some are language-specific, such as lexical semantics (Nation & Snowling, 1998; Nation, Marshall, & Snowling, 2001; Landi & Perfetti, 2007; Henderson, Snowling, & Clarke, 2013), morphology (Tong, Deacon, Kirby, Cain, & Parrila, 2011; Adlof & Catts, 2015), narrative production (Cain & Oakhill, 1996; Cain, 2003; Cragg & Nation, 2006), and syntax (Yuill & Oakhill, 1988; Goff, Pratt, & Ong, 2005; Silva & Cain, 2015). PCs also show difficulty with higher-order skills, such as inference-making (Cain & Oakhill, 1999; Cain, Oakhill, & Lemmon, 2004; McMaster et al., 2012), and comprehension monitoring (Ehrlich, Remond, & Tardieu, 1999; Cain, Oakhill, & Bryant, 2004; van der Schoot,

Vasbinder, Horsley, Reijntjes, & van Lieshout, 2009). Finally, PCs show some more domain-general weaknesses, such as difficulties in working memory (Nation, Adams, Bowyer-Crane, & Snowling, 1999; Pimperton & Nation, 2010; 2014) and executive function (Cutting et al., 2009; Locascio, Mahone, Eason, & Cutting, 2010), although these weaknesses often appear to be limited to the verbal domain (Pimperton & Nation, 2010). Research over the past thirty years has provided this detailed description of PCs' various weaknesses. However, determining which weaknesses contribute to comprehension difficulties and which are just a result remains challenging.

One deficit consistently demonstrated in PCs is in lexical-semantics, or word meaning. Across studies, PCs perform worse than their TD peers on tasks tapping lexical-semantic knowledge. For example, PCs are slower and less accurate than TD individuals at making synonym judgments, which require meaning access, but they perform similarly to TD individuals when making rhyme judgments, which rely only on phonological skills. PCs are also slower and less accurate at reading low-frequency and irregular words aloud; these types of words require more support from semantic knowledge (Woollams et al. 2007; Nation & Snowling, 1998). PC children are also slower and less accurate at naming pictures with low-frequency names (Nation, Marshall, & Snowling, 2001). While PCs show typical priming for category coordinates when a pair co-occurs frequently in language (e.g. *cat* – *dog*), this priming disappears when the pair is not associated (e.g. *cow* – *goat*; Nation & Snowling, 1999), suggesting that PCs may not be sensitive to more abstract semantic relations. Further research demonstrates that PCs primarily show priming to the dominant meanings of words (e.g. *light* but not *flower* for *bulb*), with reduced or no priming for subordinate meanings, indicating difficulty inhibiting dominant meanings even when sentential context is biased towards subordinate ones (Henderson,

Snowling, & Clarke, 2013). In brain, TD individuals show more negative N200 and N400 amplitudes when a target is preceded by a semantically-unrelated prime than when it is preceded by a semantically-related prime, suggesting that they are sensitive to some semantic relationship between the prime and the target. PCs show a smaller difference in N200 and N400 amplitude between prime types, perhaps due to reduced sensitivity to semantic relationships (Landi & Perfetti, 2007).

Thus, while the lexical-semantic deficit in PCs has been reliably documented, we still do not know what contributes to this deficit. Put differently, we do not know how the lexical-semantic deficit in PCs came to be. An obvious way to look at the origin of these lexical-semantic weaknesses is to study younger children. Most studies of PCs are done on school-aged children between seven and eleven years old, but multiple studies have been looking younger, either by identifying PCs at an older age and using retrospective data or by focusing on listening comprehension (Silva & Cain, 2015; Kim, 2015; Florit, Roch, & Levorato, 2011; Catts et al., 2006). These types of studies reveal vocabulary difficulties in children who grow up to become poor comprehenders as early as kindergarten. Another way to investigate the origin of the lexical-semantic deficit in PCs is to study word learning in adolescents or adults. By focusing on word learning, researchers may be able to uncover more information about how PCs acquired a lexical-semantic deficit. While word learning in adolescents and adults surely does not perfectly match word learning in infants and children, some research has shown that it provides a helpful simulation.

To learn a new word, infants must map labels onto unnamed concepts. Thus, word learning can be divided into two parts: concept building and label mapping. Deficits in either process could lead to vocabulary difficulties later in life. Previous research has lent support to the

idea that early word learning is constrained not by concept-building mechanisms but instead by label mapping. Gleitman and colleagues (Gillette, Gleitman, Gleitman, & Lederer, 1999; Gleitman, Cassidy, Nappa, Papafragou, & Trueswell, 2005; Snedeker & Gleitman, 2004) demonstrate this by using a technique called the Human Simulation Paradigm (HSP). The HSP seeks to simulate infant word learning by modulating available information during word learning in adults. In HSP experiments, adult participants learn novel labels for early-learned words. All of the information given to participants in the HSP comes from actual mother-child interactions.

In the HSP, three types of information are presented. Some participants watch silent videos of interactions between mothers and their infants. Target words are replaced with a beep. These types of silent scenes provide participants with only the visual information about the context in which a word was produced. Other participants are given the nouns uttered in the same sentence as the target word, ultimately receiving lists of nouns that co-occurred with the target word. Finally, some participants receive the syntactic frames in which the mother had used the target word during these mother-child interactions. Syntactic frames preserve the function words and word order of a sentence while replacing content words with nonwords. These three types of information closely mirror what infants are able to use, because all information regardless of type comes directly from actual videos of mother-child interactions. Thus, this paradigm provides a reasonable simulation of infant word learning.

If early word learning is mainly constrained by concept knowledge, then adults should do well in this re-labeling task regardless of what information (scenes, nouns, or frames) is provided because their concepts should be fully developed. Target words come from a list of the earliest words babies learn, so this assumption is not unreasonable. However, despite adult-level knowledge about all target words, performance was indeed modulated by information type.

Participants receiving only the scenes or the nouns performed poorly, correctly identifying the target word only 15 percent of the time. Participants who received only the frames did better than only scenes or nouns, but the greatest success happened with multiple sources of information (e.g. scenes and nouns, or all three; Gillette et al., 1999). Thus, even though adults had fully-formed concepts for the target words, they still needed multiple types of information to correctly map a novel label onto these concepts.

While results from the HSP suggest that very early word learning is constrained by mapping rather than conceptual development, concepts do develop and word learning continues throughout life. Thus, later word learning, which inevitably involves more complex concepts, may be influenced by concept knowledge. One documented phenomenon showing conceptual development is known as the syntagmatic-paradigmatic shift. In a free association task, young children tend to provide a syntagmatic response, giving words that co-occur with the prompt in discourse (e.g. saying *fast* when provided with *run*). However, at around age seven, children shift toward paradigmatic responses, usually giving a word in the same grammatical form class that is semantically related to the prompt. For example, an older child might say *walk* when provided with *run* (Cronin, 2002). Nelson (1977) suggested that this shift might be more about reorganization of already-learned information rather than a change due to new information.

Reorganization is a theme that runs through many theories of conceptual development. Bjorklund (1985) suggests that throughout development, the use of certain semantic relations becomes easier, which leads to a reorganization of semantic memory. For example, younger children tend to classify items based on thematic, or complementary, relations (e.g. *needle-thread*). Older children and adults tend to classify items based on taxonomic categories (e.g. *needle-pin*). However, when trained on taxonomic relationships, kindergarteners were able to

classify items based on taxonomic categories. In addition, when asked to classify more items a day after the taxonomic training, slightly under half of the children switched back to thematic classifications (Smiley & Brown, 1979). This suggests that the classification shift observed between younger and older children is one of tendency rather than simple ability, again pointing towards reorganization rather than acquisition of entirely new knowledge. That is, young children may be able to respond based on taxonomic relations, but their prepotent response involves thematic relations. Bjorklund (1985) also suggests that these taxonomic category relations may not be obvious to young children unless these relations are somehow made salient. Throughout development, taxonomic relations may become more obvious, but one can also to make relations salient through direct teaching.

Direct teaching may help children sort through information to pick out the most important aspects. Keil and Batterman (1984) demonstrated that over development, children naturally shift from using characteristic features to defining features when describing things. For example, while a kindergartener might describe twins using characteristic features, such as *they dress and look alike*, fourth graders tend to describe defining features, like *are born from the same mother at the same time*. However, Butler and Markman (2014) showed that children can flexibly use pedagogical cues provided by adults in order to organize objects by function. In their study, they showed that children frequently sort objects by visual features such as color or shape when experimenters demonstrate the function of the objects in an accidental manner (e.g. showing magnetism by accidentally dropping a paperclip on an object). When experimenters provide pedagogical cues while demonstrating the object's function, such as making eye contact and engaging in joint attention, children tend sort the objects by function. Butler and Markman also suggest that pedagogical cues may be interpreted as indicating defining features of an

artifact. Thus, over development, children may be using overt cues from adults to reorganize their semantic network and transition towards using relations similarly to adults. However, if an individual fails to notice those cues or use them correctly, they may continue to respond more like a younger child. Thus, it is important to study what information an individual takes from their environment, what cues they can pick up on, and how this attended-to information shapes their semantic network.

The framework of *preferential acquisition* highlights the role of the environment in word learning. Preferential acquisition is the hypothesis that words are learned in an order that corresponds to how well they are connected to other items in the learning environment. This framework is contrasted with another popular idea known as *preferential attachment*, which postulates that word learning proceeds according to how connected a given word is to the other words an individual already knows. Thus, while preferential attachment emphasizes the learner's own knowledge, preferential acquisition emphasizes the learner's interaction with his/her learning environment (Hills, Maouene, Maouene, Sheya, & Smith, 2009). One study by Beckage and colleagues (2011) demonstrates that while TD infants learn words in an order predicted by preferential acquisition, late talkers learn a completely different network of words. Investigation of word learning under the preferential acquisition framework allows for insights about how the infants are interacting with the world and what information they gather naturally.

To investigate the semantic networks of 15- to 36-month-old TD and late talking (LT) infants, Beckage and colleagues collected information on the items present in each infant's vocabulary using a parent report measure. In this study, LTs were defined as infants with atypically low vocabularies for their age. The authors aimed to see if LTs were just acquiring words more slowly or if they were actually learning words in a different way than TD infants. If

late talkers were acquiring words in the same manner as typically developing children, one would expect that the network of a typically developing child who knows 60 words to look similar to the network of a late talker who knows 60 words. The authors used co-occurrence statistics taken from the CHILDES corpus to connect the vocabulary items for each participant. Since the statistics were taken from a corpus, connections were the same for all infants; two infants with the exact same items in their vocabulary would have identical semantic networks. The authors found that TD individuals tended to have networks with lots of small-world structure, where individual items may not have a lot of direct connections, but most items can be reached through a small number of connections. Random networks created by randomly picking words out of the parent report measure also tended to exhibit small-world structure, suggesting that the learning environment itself largely has a small-world structure, at least for early-learned words. In contrast, late talkers showed much less small-world structure than both the typically developing children and the randomly generated networks (Beckage, Smith, & Hills, 2011). The authors concluded that LTs might be sampling the world in a fundamentally different way than their TD peers. When constructing their semantic networks, LTs may collect different information from their environment.

This insight could prove quite relevant for the study of PCs. The idea that PCs direct their attention in an abnormal way during learning matches well with a number of findings related to executive function, inhibitory control, and attention in PCs. For example, multiple studies have shown that PCs show weaker planning skills than their TD peers (Sesma, Mahone, Levine, Eason, and Cutting, 2009; Locascio, Mahone, Eason, and Cutting, 2010; Cutting, Materek, Cole, Levine, and Mahone, 2009). These planning deficits may be related to many findings indicating that PCs have trouble with inhibitory control. Henderson and colleagues (2013) found that PCs

have trouble inhibiting dominant meanings of words even with sentential context. Nation, Marshall, and Altmann (2003) also found that PCs make more fixations to target objects that are shorter in duration than skilled comprehenders, potentially due to weak inhibition skills. That is, they appear to be more distracted by distractor items. This interpretation is supported by marginally worse performance on an inhibitory control task found in the same study. Kieffer, Vukovic, and Berry (2013) found that both attention-shifting and inhibitory control contribute directly to reading comprehension ability. Although inhibition is often thought of as a domain-general skill, Pimperton and Nation (2010) demonstrated that in PCs, the inhibitory deficit appears to be only in the verbal domain. Finally, this profile of executive function deficits has been supported in brain as well. PCs have less gray matter volume than TDs in executive function regions (Bailey et al., 2016). In addition, better connectivity between executive and language regions has been shown to be positively correlated with reading comprehension ability (Aboud, Bailey, Petrill, & Cutting, 2016). Thus, PCs appear to have deficits in executive function, specifically in inhibition in the verbal domain.

The above findings suggest that PCs are not taking information from the world in a typical manner, just like the late talkers. If this is true, their word learning itself may be atypical. Most studies investigating word learning in PCs take place in a reading context. For example, Cain, Oakhill, & Lemmon (2004) tested PCs' ability to learn novel words from a story context as well as through direct instruction. In the story context task, participants read a sentence containing a novel word that was followed by a sentence providing the context necessary to determine the meaning of the novel word. They found that poor comprehenders with low vocabularies performed worse on this task than both better comprehenders and poor comprehenders with typical vocabularies. In comparison, when participants learned novel words

from sentences containing both the novel word and its definition, all poor comprehenders performed worse than better comprehenders regardless of their vocabulary abilities. Cain, Oakhill, and Elbro (2003) also gave the story context word inference task to poor comprehenders with typical vocabularies. They found that while poor comprehenders tended to make more errors than their more skilled peers, especially when more filler information was provided between the nonword and the informative context sentence, the pattern of errors was similar for both poor and better comprehenders. Both papers emphasize the possibility that poor comprehenders, especially those with poor vocabularies, may use inefficient strategies for word learning.

### **Summary & Prospectus**

Overall, more research needs to be done to better understand word learning in PCs. We do not know if it is simple label mapping, like in the HSP, or more complex concept structure that is impaired in PCs. If label mapping is PCs' main difficulty, then they should have little trouble building novel concepts. If the main deficit lies in conceptual organization, perhaps direct instruction will help guide PCs towards adult-like semantic structure. The current investigation aims to address both of these possibilities.

First, we consider what aspects of word learning are impaired in PCs. We test concept construction by nonverbally teaching participants groups of objects that move in the same manner, thus sharing a nonverbal feature. After concept construction, we teach participants to attach labels to these groups. By analyzing performance after nonverbal and after verbal blocks, we can break apart these two processes. Previous research showing that PCs' deficits tend to be specific to the verbal domain lead us to hypothesize that they will perform similarly to their TD

peers in nonverbal blocks but more poorly in verbal blocks (Pimperton & Nation, 2010; Nation, Adams, Bowyer-Crane, & Snowling, 1999).

Our second aim addresses the role of instruction type on word learning in PCs. During inference-making tasks, PCs make fewer inferences than their peers. However, when pointed towards relevant information, their performance rises to typical levels (Cain & Oakhill, 1999). This suggests that PCs have the ability to make inferences but do not do so spontaneously. This finding, in addition to the literature reviewed above pointing towards an executive function deficit in PCs, may indicate that PCs have trouble selecting important information. To investigate this possibility, we use two types of instruction in the word learning task: explicit and implicit. The explicit instruction type requires participants to respond based on category-relevant features during training, forcing their attention towards the most important features. In contrast, during implicit training blocks participants respond based on category-irrelevant features and category-relevant features are presented incidentally. This requires participants to direct their attention towards these features independently. We expect that both TD and PC participants will perform better in explicit blocks than in implicit blocks. However, if PCs' word-learning skills are primarily hampered by misdirected attention, they should show a greater benefit from explicit training than the TD participants.

One important caveat for the current investigation involves the use of the term "poor comprehender." PCs are often defined using standardized scores on reading comprehension and decoding assessments, indicating low comprehension in comparison to the population. The current study does not use this method; instead, we select better and poorer comprehenders from the collected sample for group analysis. While this group selection method may not allow for claims about canonically defined PCs, there still exists considerable variability in comprehension

skill in the undergraduate population. Thus, our investigation aims to comment more broadly on the relationship between word learning and comprehension skill rather than purely word learning in the PC population.

## Methods

### Participants

41 undergraduate students from the University of Connecticut participated in the experiment. One subject was excluded due to experimenter error. In order to focus on comprehension beyond decoding skill, all participants had to achieve at least a standardized score of 95 on the Woodcock-Johnson III Word Attack subtest (Woodcock, McGrew, Mather, & Schrank, 2001), a test of nonword decoding. Twelve subjects were excluded for Word Attack scores below 95 to ensure adequate decoding ability in all participants. The final sample consisted of 28 subjects (12 male; *M* age = 19, range = 18-22). All participants were native English speakers who were not native speakers of any other language. All participants also reported normal or corrected-to-normal vision and hearing.

### Behavioral Assessments

**Comprehension.** Comprehension was measured using the Comprehension subtest of the Nelson-Denny Reading Test (Brown, 1960). This task has been used in previous investigations of poor comprehenders (Landi, 2010; Landi & Perfetti, 2007). Participants had eight minutes to complete the test. In this task, participants read passages and answer comprehension questions. Some comprehension questions require literal answers that can be found in the text while others are more inferential, where participants must go beyond what is explicitly stated.

**Vocabulary.** Vocabulary was measured using the Vocabulary subtest of the Nelson-Denny Reading Test (Brown, 1960). Again, this task has been used in previous research (Landi,

2010) and was selected because it has been normed through college age. Participants had eight minutes to complete the test. In this task, participants select the best synonym for each item.

**Nonverbal IQ.** Nonverbal IQ was assessed using Raven's Advanced Progressive Matrices, Set II (Raven & Court, 1998). This assessment was selected because it has been used in multiple studies of PCs (Li & Kirby, 2014; Goff, Pratt, & Ong, 2005) and because it is difficult enough to provide variability even in a college-aged sample. Participants had ten minutes to complete the assessment. Raven's Advanced Progressive Matrices require participants to select the item that best completes a complex visual grid.

**Decoding.** Decoding ability was assessed using the Word Attack subtest of the Woodcock-Johnson III (Woodcock et al., 2001). Participants are asked to read nonwords aloud at their own pace.

**Nonword fluency.** Nonword fluency was assessed using the Test of Word Reading Efficiency Phonemic Decoding Efficiency subtest (Torgesen, Wagner, & Rashotte, 1999). Participants are instructed to read as many nonwords aloud as they can in 45 seconds.

### **Stimuli**

The novel items participants learned were cartoon robots, used with permission from artist Andy Martin. Visual similarity ratings were obtained from 9 undergraduate subjects at the University of Connecticut who did not participate in the main task. Participants doing this visual similarity rating task were shown two robots and asked to indicate how similar they looked on a scale from 1 to 5, where 1 corresponded to "very similar" and 5 to "very dissimilar." Three robots were then chosen for each category using these visual similarity ratings. Each category consisted of two robots that were rated as similar and one robot that was rated as dissimilar to the other two. This method ensured that visual similarity was an unreliable cue to category

membership. Once the robots were put into categories, each category was assigned a motion pattern and a name. The motion patterns were created using Microsoft PowerPoint and included paths such as zig-zags, arcs, and loop-de-loops. Names were CVC constructions. To reduce phonological interference, each category name had a unique onset and none of the names rhymed.

### Apparatus

The category-learning task was presented using E-Prime 2.0 (Psychology Software Tools, Pittsburgh, PA). Eye movements were measured using an Eyelink-1000 Plus desktop-mounted eyetracker from SR Research in remote mode using a 16mm lens and a sampling rate of 500Hz. Five-point calibration occurred at the beginning of each block and as needed, indexed by performance on drift corrections that were completed at every trial. We utilized eyetracking methodologies to allow finer-grained measurement of participants' choosing behavior during test trials.

### Category<sup>1</sup> learning task – training

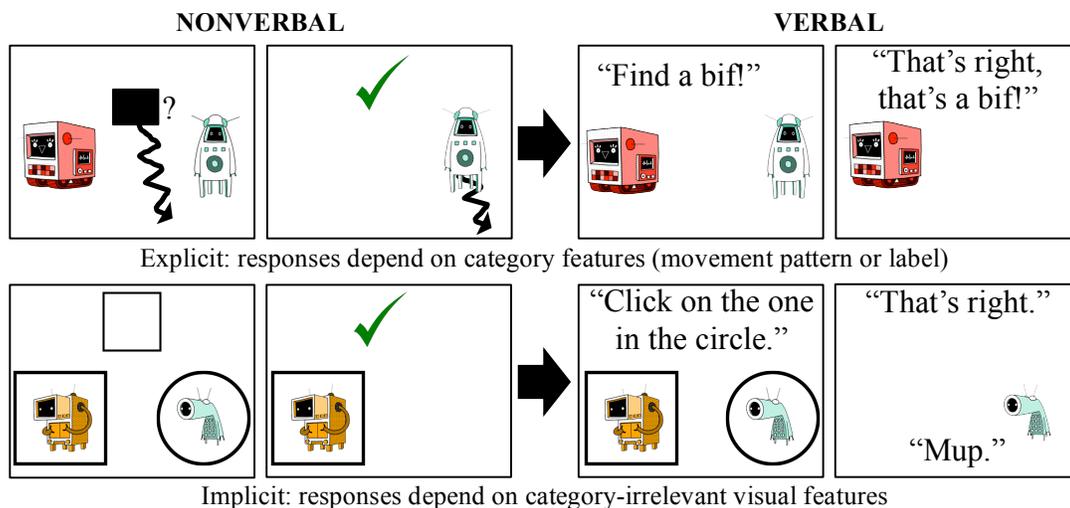


Fig. 1. Example trials for all four training conditions.

<sup>1</sup> We do not intend to comment on the difference between categories and concepts. This paper uses the phrase "category learning" to refer to learning in our task.

In this task, participants learned four novel categories. There were two training types: nonverbal and verbal. In nonverbal training, participants learned the motion patterns of each robots. In verbal training, participants learned each robot's label. Since motion patterns and labels were common among robots in the same category, this task intended to teach category membership. Nonverbal training always preceded verbal training. This was done primarily to attempt to simulate development, where infants might construct concepts before learning their labels. There were also two instruction types. In explicit training, participants were required to respond based on category-relevant features. In implicit training, participants responded to category-irrelevant features and saw or heard category-relevant features incidentally. Some participants received implicit training first while others received explicit training first.

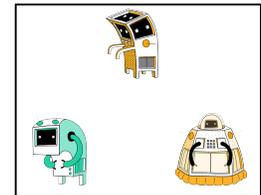
In all types of training, trials began with a drift correction. Then, participants clicked on a fixation cross in the center of the screen. In explicit nonverbal training, participants saw a black box make a motion between two robots. This motion pattern was identical to the motion pattern of one of the robots' categories. After the motion finished, a question mark appeared on screen and participants were allowed to respond by clicking on either robot. If their selection was incorrect, a red X appeared on-screen and they were permitted to try again. If their selection was correct, a green checkmark appeared on-screen. Then, the distractor robot disappeared from the screen and the correct robot went through its motion pattern. In explicit verbal training, participants saw two robots and heard, "Find a *[name]*." They then clicked on a robot. For incorrect responses, participants heard "Try again!" and made another selection. After a correct response, participants heard "That's right! That's a *[name]*." In implicit nonverbal training, participants saw two robots on the screen. One robot was inside a circle and the other was inside a square. In addition, an empty shape (either a square or circle) appeared in the center of the

screen. Participants were supposed to click on the robot inside the shape that matched the empty center shape. For incorrect responses, participants saw a red X. For correct responses, participants saw a green checkmark and then the robot went through its motion pattern. In implicit verbal training, participants saw the same type of robots-in-shapes display. Participants heard, “Click on the one in the *circle/square*.” If participants clicked on the robot in the wrong shape, they heard “Try again!” and were allowed to make another response. After clicking on the robot in the correct shape, participants heard “That’s right!” followed by the category name. See Figure 1 for example training trials. Participant completed 54 trials of training in each condition. Feedback was given after each trial.

Besides procedural instructions about how to operate trials (such as where to click), instructions during training were relatively minimal. Participants were told that they would be learning “families” of robots, but they were not told anything about the robot “families.”

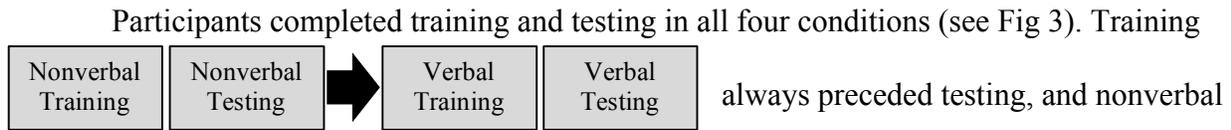
### Category learning task – testing

The testing block was identical for all types of training. At test, participants saw three robots arranged in a triangle on the screen, with two on the bottom of the screen and one on top. Two of the robots were from the same category. Participants were asked to indicate which of the robots on the bottom of the screen was from the same “family” as the robot on the top by clicking. See Figure 2 for an example testing trial. Participants completed 108 test trials. The total number of possible combinations for testing is 216. Due to time constraints, these 216 combinations were split in half and placed into two separate pseudorandomized lists. Lists were counterbalanced across subjects.



*Fig. 2. Example trial for test. Participants are instructed to select the robot on the bottom that comes from the same “family” as the robot on the top.*

## Procedure



*Fig. 3. Experimental design. Participants complete this sequence twice: once with explicit tasks and once with implicit tasks.*

Instruction type order (implicit or explicit

first) was counterbalanced across subjects.

## Results

In this results section, we will first examine task differences without considering individual differences to investigate our manipulations. Next, we will report continuous analyses looking at the relationships between our behavioral measures and the experimental task. Finally, we will perform a group analysis to further explore these relationships.

### Behavioral Measures

See Table 1 for descriptive statistics of each of the behavioral measures. See Table 2 for correlations between behavioral measures. Four subjects did not complete the Raven's Advanced Matrices due to time constraints; any analyses presented below that include Raven's are done with these subjects excluded.

### Accuracy (see Fig. 4)

Accuracy for one subject in the explicit verbal condition was removed. This subject had an accuracy of zero, demonstrating misunderstanding of the testing task rather than difficulty learning the concepts. Another subject showed extremely low accuracy (4%) in the explicit nonverbal condition. His/her accuracy for this condition was removed for the same reason.

One-sample t-tests confirmed that accuracy at test for all conditions was significantly greater than zero, indicating that learning did occur in all conditions (Table 3).

To investigate the effect of condition on accuracy, we ran linear mixed-effect models. Condition had a significant effect on accuracy,  $\chi^2(3) = 56.26, p < 0.001$ . Orthogonal contrasts revealed that explicit conditions were more accurate than implicit conditions,  $b = 0.10, t(79) = 7.62, p < 0.001$ . They also showed that verbal conditions were more accurate than nonverbal conditions,  $b = 0.05, t(79) = 3.92, p < 0.001$ . Follow-up pairwise *t*-tests showed that the implicit nonverbal condition was less accurate than all other conditions (all *ps* less than 0.001, Bonferroni-corrected). In addition, the explicit verbal condition was more accurate than the implicit verbal condition ( $p = 0.003$ , Bonferroni-corrected). Finally, accuracy in the explicit nonverbal condition was not significantly higher than accuracy in the implicit verbal condition ( $p = 0.08$ , Bonferroni-corrected) or the explicit verbal condition ( $p = 0.25$ , Bonferroni-corrected).

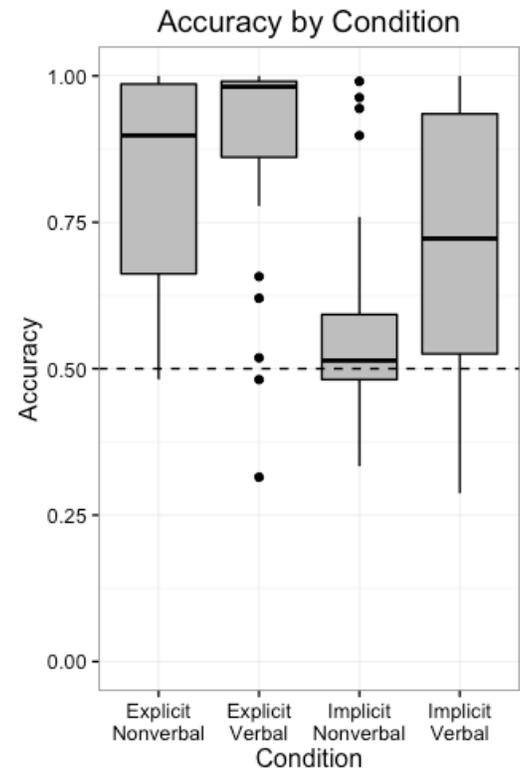


Fig. 4. Accuracy by condition. The dashed line indicates chance level.

### Reaction Time (see Fig. 5)

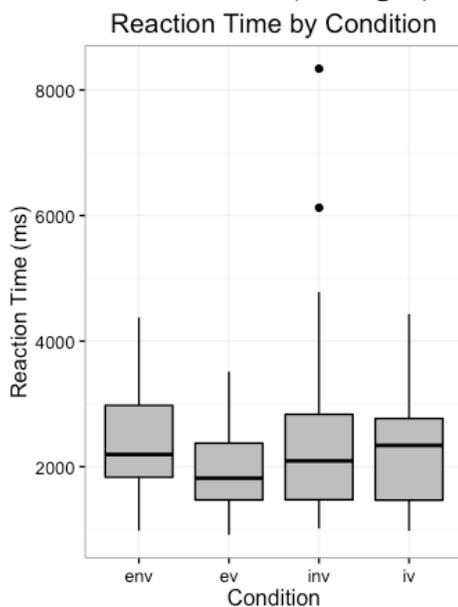


Fig. 5. Reaction time by condition.

Due to an error in the presentation software, any trials where participants double-clicked the mouse had erroneous reaction times. Thus, all double-click trials were removed from RT analyses. Double-click trials accounted for no more than two percent of all trials in any given condition. In addition, only accurate trials were used for RT analysis. Thus, the participant with zero accuracy in the explicit verbal condition was not

included for that condition. In addition, RTs for the participant with four percent accuracy in the explicit nonverbal condition were removed, since the low accuracy meant that RTs from only two trials in that condition were kept. For descriptive statistics on RTs, see table 4.

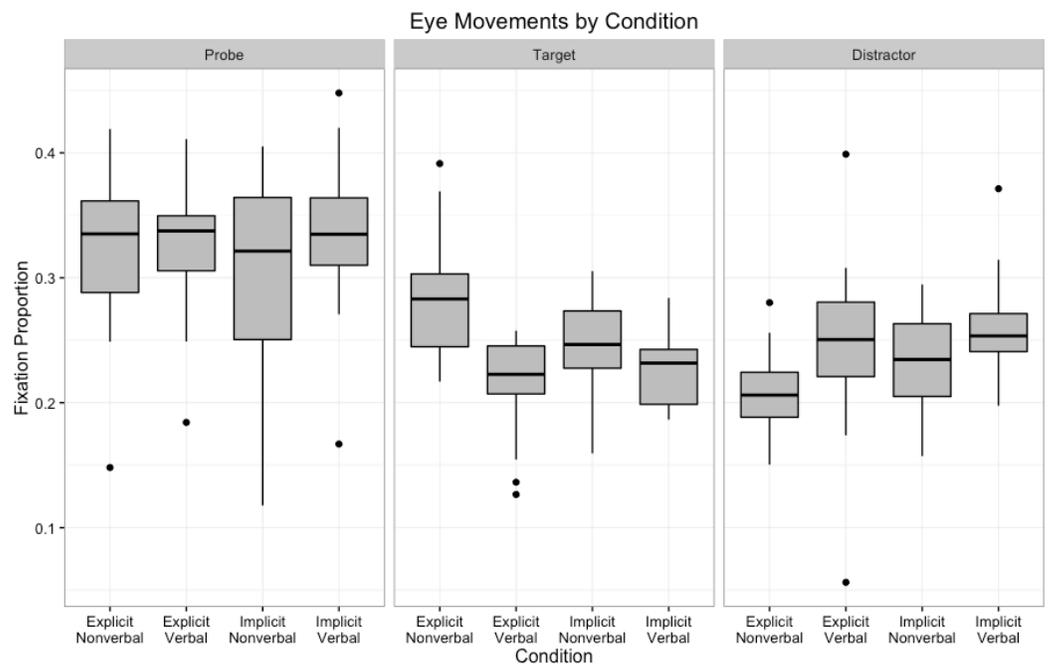
Linear mixed-effects models were used to investigate the effect of condition on RTs. Condition did not have a significant effect on RTs,  $\chi^2(3) = 6.36, p = 0.09$ . Thus, RTs did not vary by condition. Both planned orthogonal contrasts and pairwise t-tests showed that RTs were not significantly different between any conditions. None of the behavioral measures or age had significant effects on reaction time.

### **Eye Movements (see Fig. 6)**

For descriptive statistics on eye movements, see Table 5. We used linear mixed-effects models to see if eye movements were different depending on the area of interest and condition. Eye movements are reported here as fixation proportions taken over the entire test trial. Note that the analyses here only take into account test trials, where participants indicated which robots belonged to the same “family.” Test trials lasted until participants made a response. The choice to use fixation proportion over the same trial follows analyses of other papers using the visual world paradigm (Magnuson, Tanenhaus, Aslin, & Dahan, 2003). Condition did not have a significant effect on fixation proportion ( $\chi^2(3) = 2.51, p = 0.47$ ). Interest area did have a significant effect on fixation proportion  $\chi^2(3) = 67.73, p < 0.0001$ . Contrasts revealed that across conditions, participants looked at the probe robot more than the other two robots,  $b = 0.043, t(222) = 16.41, p < 0.0001$ . Participants also looked more at the target robot more than the distractor robot across conditions,  $b = 0.046, t(222) = 11.38, p < 0.0001$ . Adding the interaction between condition and interest area significantly improved the fit of the model,  $\chi^2(3) = 67.73, p < 0.0001$ . Orthogonal contrasts showed that three interaction contrasts were significant. First, the

difference between looking at the probe versus looking at target or distractor was larger for verbal conditions than it was for nonverbal conditions,  $b = 0.055$ ,  $t(216) = 2.32$ ,  $p = 0.02$ . Next, the difference between fixation proportion to the target and to the distractor was different for explicit and implicit conditions,  $b = 0.099$ ,  $t(216) = 2.73$ ,  $p = 0.007$ . Inspection of means showed that participants looked more at the target than the distractor in the explicit conditions but this trend was reversed for implicit conditions. Similarly, the difference between fixation proportion to the target and to the distractor was different for nonverbal and verbal conditions,  $b = -0.013$ ,  $t(216) = -3.56$ ,  $p =$

0.0005. In the verbal conditions, participants looked more at the distractor than the target, and this trend was reversed



for the nonverbal conditions.

Fig. 6. Eye movements by condition.

### Order Effects

To see whether getting the implicit or explicit tasks first affected performance on the category learning task, we ran linear mixed effects models.

**Accuracy (see Fig. 7).** One-sample t-tests showed that performance in all conditions was significantly above chance regardless of order, save one exception. Participants receiving

implicit tasks first did not demonstrate above-chance performance in the implicit nonverbal condition (see Table 6). Adding order to the intercept model did not

significantly improve fit

( $\chi^2(1) = 3.17, p = 0.21$ ). In addition, adding order to the model predicting accuracy from condition also did not improve fit ( $\chi^2(1) = 3.10, p = 0.08$ ). However, adding the interaction between condition and order did significantly improve fit ( $\chi^2(4) = 13.83, p = 0.008$ ). Contrasts revealed one significant interaction ( $b = 0.079, t(76) = 3.18, p = 0.0021$ ). While participants tended to be more accurate in the explicit tasks relative to the implicit tasks, participants who got the explicit tasks first showed a smaller difference in accuracy between explicit and implicit tasks than those who received the implicit tasks first.

**Reaction time.** Similar to accuracy, adding order to the intercept model predicting reaction time did not significantly improve fit ( $\chi^2(1) = 2.18, p = 0.14$ ). The same was true for adding order to the model predicting reaction time from condition ( $\chi^2(1) = 2.13, p = 0.14$ ). Adding the interaction between order and condition did significantly improve fit ( $\chi^2(4) = 17.59, p = 0.002$ ). Participants who received the explicit tasks first were faster in explicit tasks than implicit tasks, but this trend was reversed for those who completed implicit tasks first ( $b = -603.74, t(76) = -3.72, p = 0.0004$ )

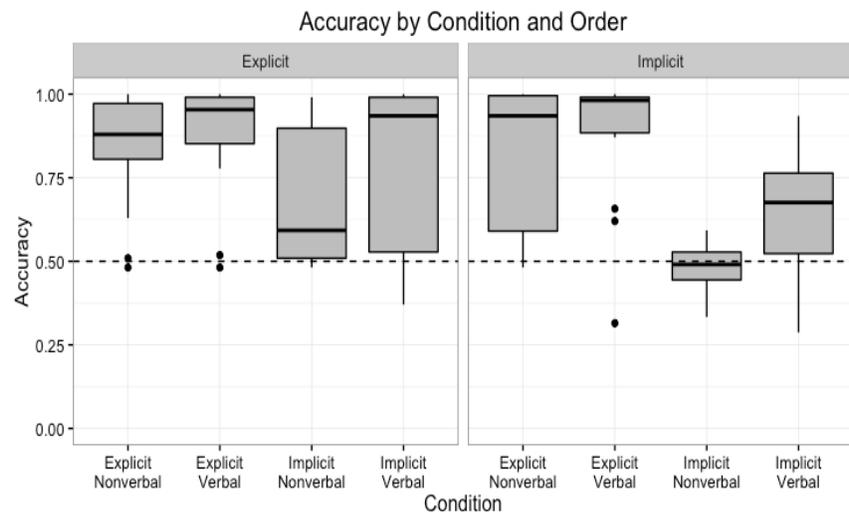


Fig. 7. Accuracy by condition and order.

**Eye movements.** Adding order to the intercept model predicting fixation proportion did not improve model fit ( $\chi^2(1) = 0.03, p = 0.86$ ). Since previous analyses showed that the model including the interaction of condition and interest area provided the best fit, the next step was to add the order interaction term to this model. Adding this term did not improve fit ( $\chi^2(12) = 5.45, p = 0.94$ ).

### Visual Similarity

To see whether visual similarity between test items affected performance on the category-learning task, we ran linear mixed-effects models. Three types of visual similarity were considered: probe-target (PT), probe-distractor (PD), and target-distractor (TD).

**Accuracy.** All three types of visual similarity, when entered separately into different models, improved fit over the intercept model (PT:  $\chi^2(1) = 9.09, p = 0.003$ ; PD:  $\chi^2(1) = 33.03, p < 0.0001$ ; TD:  $\chi^2(1) = 30.88, p < 0.0001$ ). Adding different visual similarities to the model (e.g. PT to PD model) only improved fit when the base model was PT, suggesting that the best predictors of accuracy were PD and TD visual similarity. However, adding any type of visual similarity or the interaction term to the model predicting accuracy from condition did not improve model fit. Thus, visual similarity affects accuracy similarly across conditions.

**Reaction Time.** Adding any type of visual similarity to either the intercept model or the model predicting RT from condition did not improve model fit.

**Eye Movements.** Adding any type of visual similarity to the model predicting fixation proportion from interest area did not improve model fit.

### Individual Differences

To investigate the effect of comprehension ability on performance in the category-learning task, we utilized a novel method inspired by a recent technique used to define PCs. This

technique predicts reading comprehension from a variety of low-level skills, such as nonverbal IQ, decoding, and vocabulary. Then, the predicted comprehension score is compared to the individual's measured comprehension score. Participants with an actual comprehension score much lower than their predicted comprehension score are labeled "unexpected poor comprehenders" (Li & Kirby, 2014; Tong, Deacon, & Cain, 2013). Our analysis is similar; however, instead of creating groups, we extracted the residuals from the model to provide a continuous measure of comprehension beyond low-level skills.

In this analysis, vocabulary, nonword decoding, and nonverbal IQ were centered and scaled. Age was also centered but not scaled. Then, reading comprehension was predicted from vocabulary, nonword decoding, nonverbal IQ, and age. Together, these predictors accounted for 33.5 percent of the variance in reading comprehension. Then, residuals were extracted from the model. Linear mixed-effects models were used to see if the residuals (i.e. comprehension beyond low-level skills) had a significant effect on performance in the task. Residuals did not have a significant effect on accuracy ( $\chi^2(1) = 1.41, p = 0.23$ ) or reaction time ( $\chi^2(1) = 0.81, p = 0.37$ ) across conditions.

### **Behavioral Measures & Performance: Continuous Analysis**

**Accuracy.** First, we added the behavioral measures to the model predicting accuracy from condition. Each behavioral measure was added separately to the accuracy-condition model. Adding vocabulary to the model significantly increased the fit ( $\chi^2(1) = 6.30, p = 0.01$ ). Adding nonverbal IQ to the model also significantly increased the fit ( $\chi^2(1) = 17.43, p < 0.001$ ). Including the interaction term did not increase fit for vocabulary ( $\chi^2(3) = 1.29, p = 0.73$ ) or for nonverbal IQ ( $\chi^2(3) = 6.54, p = 0.09$ ). Adding comprehension to the accuracy-condition model did not result in better fit ( $\chi^2(1) = 0.63, p = 0.42$ ). The same was true for age ( $\chi^2(1) = 0.06, p =$

0.80), decoding ( $\chi^2(1) = 0.83, p = 0.36$ ) and nonword fluency ( $\chi^2(1) = 1.93, p = 0.16$ ). Thus, vocabulary and nonverbal IQ are significant predictors of accuracy at test, but they do not interact with condition.

**Reaction Time.** None of the behavioral measures or age had significant effects on reaction time.

**Eye Movements.** Next, we added the different behavioral measures separately to the model predicting eye movements from interest area. None of the behavioral measures significantly increased the fit of the model (comprehension:  $\chi^2(1) = 1.74, p = 0.19$ ; vocabulary:  $\chi^2(1) = 0.68, p = 0.41$ ; nonword decoding:  $\chi^2(1) = 1.14, p = 0.29$ ; nonverbal IQ:  $\chi^2(1) = 0.25, p = 0.61$ ; age:  $\chi^2(1) = 0.036, p = 0.84$ ).

### Behavioral Measures & Performance: Group Analysis

Typical analyses of poor comprehension split participants into groups of better and poorer comprehenders who are matched on decoding (Henderson et al., 2013; Nation & Snowling, 1998). We followed this model to look more at the effect of comprehension and vocabulary on performance in the category-learning task, adding further group membership based on vocabulary. We split the participants into three groups: better comprehenders (BC), poor comprehenders with high vocabulary (PCHV) and poor comprehenders with low vocabulary (PCLV). This split was motivated by

the method of Cain et al. (2004), who investigated

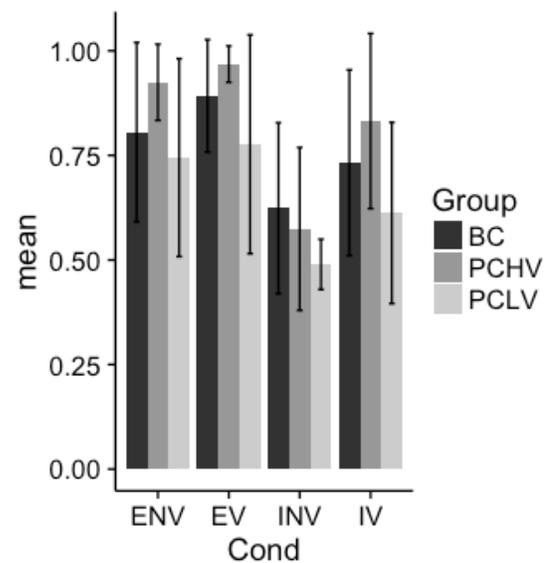


Fig. 8. Accuracy by group.

word meaning inference in groups split by both comprehension and vocabulary, as well as our

continuous findings indicating the influence of vocabulary on performance in this task. All three groups were matched on nonverbal IQ, nonword decoding, and nonword fluency. The BC and PCHV groups were matched on vocabulary. The PCLV and PCHV groups were matched on comprehension. Each group had 8 subjects, for a total sample size of 24. Four subjects were not selected to allow for better matching. See Table 7 for more description of the groups.

**Accuracy (see Fig. 8).** One-sample t-tests showed that learning occurred for all groups in all conditions except for the implicit nonverbal condition, where no group performed significantly above chance (see Table 8). In addition, the PCLV group did not show above-chance performance in the implicit verbal condition. Linear mixed-effects models showed that adding group to the intercept model did not significantly improve fit ( $\chi^2(2) = 5.19, p = 0.07$ ). However, follow-up Bonferroni t-tests showed that while BC did not differ significantly from PCHV ( $p = 0.92$ ) or PCLV ( $p = 0.14$ ), PCHV exhibited higher accuracy across conditions than PCLV ( $p = 0.01$ ). Adding group to the model predicting accuracy from condition also did not significantly improve fit ( $\chi^2(2) = 5.49, p = 0.06$ ). Note that while adding group to the accuracy-condition model may marginally improve fit, the only significant effects are from condition rather than group.

Also considered was the improvement in accuracy after the verbal block, where participants have both more information and more exposure to the categories. In these models, the dependent variable is accuracy in nonverbal blocks subtracted from accuracy in verbal blocks. Adding group to the intercept model did not significantly improve fit ( $\chi^2(2) = 0.69, p = 0.70$ ). Adding training condition (implicit vs. explicit) marginally improved fit ( $\chi^2(1) = 3.39, p = 0.07$ ). Finally, adding the interaction between group and condition did not improve fit over the condition model ( $\chi^2(4) = 4.92, p = 0.30$ ). Inspection of means showed that while the difference

between verbal and nonverbal blocks was mostly similar across groups in the explicit conditions, the PCHV group had a numerically greater increase in accuracy after verbal training than either the PCLV or BC groups. Follow-up Bonferroni t-tests show that none of the groups differed significantly, even when only considering the implicit training condition.

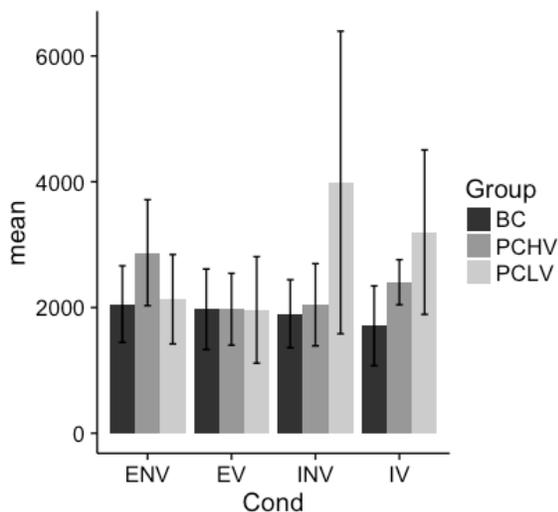


Fig. 9. Reaction time by group.

pairwise Bonferroni t-tests confirmed this finding ( $p = 0.003$ ). The PCHV and PCLV groups did not differ ( $p = 0.21$ ); neither did the BC and PCHV groups ( $p = 0.37$ ). Adding condition to the model did not significantly improve fit ( $\chi^2(3) = 5.93$ ,  $p = 0.11$ ). However, adding the interaction term did significantly improve fit ( $\chi^2(9) = 32.68$ ,  $p = 0.0002$ ). This improvement in fit was driven by a single interaction contrast. While the BC group was slower in the explicit conditions, the PCLV group was slower in the implicit conditions ( $b = -878.64$ ,  $t(61) = -4.14$ ,  $p = 0.001$ ).

### Eye movements

(see Fig. 10.). Linear mixed-effects models showed that there was no main effect of group on

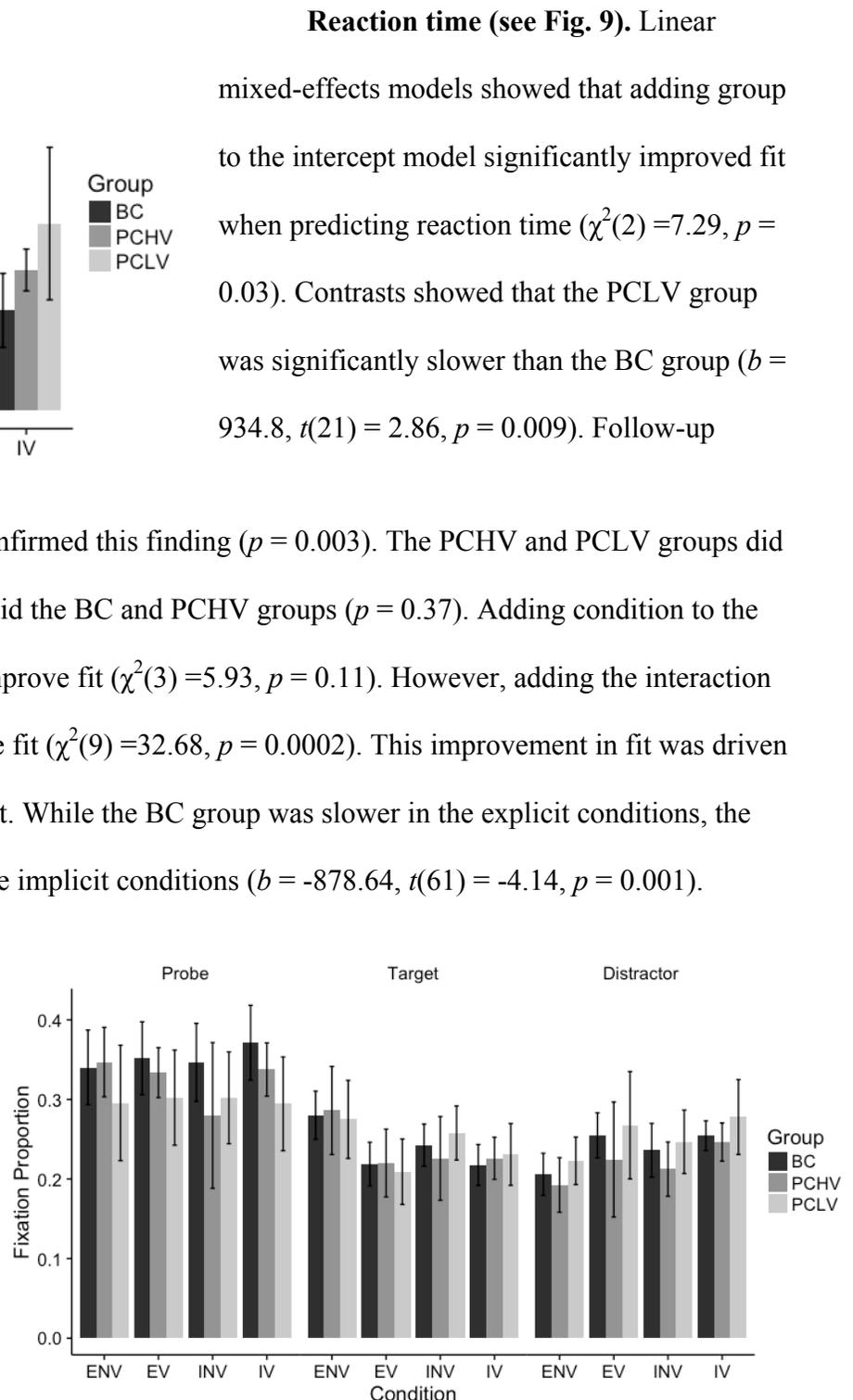


Fig. 10. Fixation proportion by group, condition, and interest area.

fixation proportion across interest areas ( $\chi^2(2) = 1.89, p = 0.39$ ). In addition, adding group to the model predicting eye movements from interest area did not improve fit ( $\chi^2(2) = 1.89, p = 0.39$ ). However, adding the interaction significantly improved fit ( $\chi^2(6) = 28.54, p = 0.0001$ ). This improvement was driven by one interaction contrast. The difference between probe fixation proportion and fixation proportion to others (target and distractor together) was larger for the BC group than for the PCLV group ( $b = 0.02, t(186) = 3.03, p = 0.003$ ). Thus, the BC group's looks to the probe relative to other interest areas were greater than that of the PCLV group, across conditions. Finally, adding condition to the model did not significantly improve fit ( $\chi^2(3) = 3.64, p = 0.30$ ), but adding the interaction of condition did ( $\chi^2(27) = 78.85, p < 0.0001$ ). However, none of the three-way interaction contrasts were significant.

**Order effects.** Because participants completed the behavioral tests and category-learning task in the same session, we were unable to ensure that equal numbers of each group completed each order. By chance, only one PCLV participant did the experiment with the implicit-first order. As such, we will not be presenting order-by-group analyses.

## Discussion

The current study investigated how performance in a novel category learning task relates to comprehension ability beyond the level of decoding, with two main aims. First, we explored the role of the label. Can individuals with comprehension difficulties learn nonverbal concepts just as well as their typical peers? Next, we introduced two types of training, explicit and implicit, to see if direct instruction would affect performance on the word learning task. Our category learning task expanded upon previous research by teaching individuals novel concepts with novel labels. Previous research has primarily focused on teaching new labels for known concepts or, in one study, combining known concepts to create novel ones. While we did not find a

significant effect of comprehension, both vocabulary and nonverbal IQ affected performance in our category learning task. We discuss the two main manipulations and the effect of these skills below.

Before we discuss the main questions, we will consider basic effects. One interesting result concerning information type was shown with eye movements. Across comprehension skill levels, participants looked more at the target than the distractor during nonverbal conditions. This pattern was reversed for the verbal conditions. Despite these patterns of looking, participants in general were more accurate in verbal conditions than nonverbal conditions. Potentially, participants were more unsure during nonverbal conditions, prompting them to examine the targets more carefully but still leading to a lower accuracy. In contrast, in verbal conditions it may have been easier to retrieve names than retrieving motion patterns was in nonverbal conditions. Once the name of the probe was retrieved, it may have been easy to attach that same label to the target, and the longer looking times on the distractor may reflect extra effort in retrieving the distractor's label. These explanations are purely speculative, however, and further study is required to confirm or deny them.

Our first main question was concerned with whether explicit training on category-relevant features would affect learning. Overall, explicit training resulted in higher accuracy at test than implicit training. Participants were more accurate and looked at the target more than the distractor after explicit training. However, comprehension skill, either continuously or by group, did not significantly predict how performance was modulated by instruction type. We hypothesized that PCs would show an increased benefit from explicit training, perhaps performing closer to typical individuals in explicit rather than implicit tasks. However, because our analysis did not find differences due to comprehension either continuously or by group, we

were unable to test this hypothesis. In addition, even though we did find a significant effect of vocabulary on performance, there was no interaction between vocabulary and condition. This suggests that individuals with low vocabularies do not show an increased benefit from explicit training. Similarly, Cain, Oakhill, and Lemmon (2004) showed that even after directed learning, where definitions were provided explicitly in simple sentence frames, PCs with low vocabularies needed more repetitions to successfully learn than better comprehenders and PCs with typical vocabularies. Thus, in both children and adult populations, PCs with low vocabularies show slower and/or reduced word learning than TD peers, even with explicit instruction. Multiple factors may explain why PCs, especially those with low vocabularies, show difficulty learning words even with the most explicit training. First, their working memory may not have enough capacity to learn new words at the same rate as their TD peers. Indeed, studies of PCs have demonstrated working memory deficits, especially in the verbal domain (Pimperton & Nation, 2010; Nation et al., 1999). Thus, working memory may restrict any additional benefits or insights provided by an explicit teaching method. In addition, PCs with low vocabularies may have different standards for word learning than their higher-vocabulary peers. Previous findings in comprehension monitoring suggest that PCs have lower standards for coherence than their TD peers; incoming information does not have to fit perfectly with previously-read text to be accepted (Van Dyke, Matsuki, & Landi, under revision). Perhaps these lowered standards for coherence are also present in word learning in PCs, especially those with lower vocabularies. They may not carefully evaluate new information to determine how it fits with the lexical item, leading to disorganized or poorly constructed word meanings.

The second hypothesis centered on the role of information type during word and concept learning in PCs. Performance after verbal training was greater than after nonverbal training

across participants; however, this difference may be due to the fact that participants had double the amount of exposure to the novel categories after verbal training. Because PCs frequently show domain-specific verbal deficits, we hypothesized that they might perform similarly to TD participants during nonverbal learning and show a deficit after verbal label mapping. This was not what we found. In the continuous analysis, comprehension skill did not interact with condition. In addition, group analysis did not show a significant interaction between group and the nonverbal-verbal contrast. Thus, while PCs with low vocabularies overall perform worse than their peers, it appears that they perform the same when comparing nonverbal and verbal learning. This could indicate multiple things. First, PCs with low vocabularies may have trouble with both concept construction and verbal mapping, leading to reduced performance after both nonverbal and verbal training. This could suggest that concept construction actually includes verbal mapping, such that both processes are impaired when concept learning ability is reduced. Similar to what has been discussed above, working memory may also constrain both processes.

Overall, vocabulary, not comprehension, was one of the strongest predictors of performance on the category learning task. At the surface level, this makes sense; vocabulary should be a relatively useful index of word-learning ability, which this task was designed to simulate. In the continuous analysis, vocabulary predicted accuracy over and above the role of condition. At the group level, PCLV individuals performed significantly worse than their PCHV peers in all conditions and numerically worse than better comprehenders in all conditions. In contrast, PCHV participants performed similarly to the BC participants, with PCHV even showing numerically higher accuracy than BC in all conditions except implicit nonverbal. Because of the small sample size, numeric trends may be indicative of significant trends, although these should be interpreted cautiously. Thus, it was not comprehension skill that distinguished performance on the word-

learning task but vocabulary ability. Since no significant group-by-condition interaction emerged, the current investigation does not shed much light onto what parts of category learning were especially difficult for PCLV individuals.

A study by McGregor and colleagues (2013) may provide more insight. In this experiment, participants were tested directly after word learning, to investigate encoding abilities, and at later follow-up visits, to look at remembering skills. The authors studied two groups: a language-impaired (LI) group, comprised of members with some diagnosed language impairment, and a typically-developing (TD) group. The LI and TD groups were matched on age, education level, and IQ. The LI group showed lower written and spoken language skills and lower vocabularies than the TD group. During training, participants learned novel items with novel names. The items were combinations of two already-known objects or animals (e.g. pony-snake, banjo-pizza), and the names were nonwords. This allowed for semantic as well as form-based foils during testing. The authors found that individuals with LI showed less encoding of forms, meanings, and the links between them than TD individuals. In contrast, the LI group performed similarly to the TD group during remembering meanings and form-meaning links, only showing worse performance at the follow-up visits on tasks tapping form memory. Finally, scores on vocabulary measures were related to performance on word form encoding tasks. Overall, LI individuals seem to have a deficit in encoding both form and meaning, but only their memory for form is worse over time. While the LI group does not have exactly the same profile as PCs, they still share some similarities (i.e. oral and written language weaknesses, vocabulary deficits). Perhaps PCs, especially those with poor vocabularies, also have difficulty encoding new forms and meanings. This encoding difficulty may again tie into the working memory hypothesis discussed above or it may be the result of poor attentional skills.

Another strong predictor of performance in the word-learning task across participants was nonverbal IQ. The measure of nonverbal IQ used in the current study was Raven's Advanced Progressive Matrices, which require an individual to determine the rules by which items in a matrix are arranged and to select a missing item based on those rules. Some studies have suggested that Raven's Matrices increase in the amount of abstraction needed throughout the task (Carpenter, Just, & Shell, 1990). Early matrices can be solved on simple perceptual features alone, whereas later matrices require abstraction across rows and columns. In the test blocks of the word-learning task, participants must recall the name and movement-pattern features in order to respond based on "family," overcoming the more obvious visual similarities between robots. The category-relevant features, while grounded in perceptual experience, are nonetheless more abstract at test than visual similarity. Thus, participants with higher Raven's scores may be better at abstracting category rules and responding accurately according to them. In addition, previous research suggests that the items in Raven's Advanced Matrices can be split into those that can be solved by visual pattern matching and those that require relational reasoning (Baldo, Bunge, Wilson, & Dronkers, 2010). Further investigation into this split using these data could further illuminate the relationship between visual pattern matching, relational reasoning, and performance on this task. While all three of the groups had statistically equal nonverbal IQ scores, the PCLV group had a numerically lower mean. The observed statistical similarity may have been driven by the PCLV group's larger variance. Thus, with a larger sample size, the PCLV group may indeed show lower scores on Raven's Advanced Progressive Matrices, allowing us to make a claim about the abstraction abilities of this group.

Results from the Human Simulation Paradigm suggest that the main constraint on early word learning is sound-to-meaning mapping rather than concept development. However,

research on phenomena such as the syntagmatic-paradigmatic shift show that semantic networks do change through reorganization over development. In our experiment, PCs with low vocabularies showed numerically lower accuracy at test in the category-learning task than both the better comprehenders and PCs with typical vocabularies. However, their increase in accuracy after a verbal training block was both numerically and statistically similar to the increase in accuracy experienced by better comprehenders. PCs with low vocabularies are receiving the same amount of benefit from verbal blocks as the better comprehenders, suggesting that their label mapping skills are intact. In contrast, their achievement after nonverbal blocks is numerically lower than better comprehenders. This pattern of results, while not statistically significant, could with a larger sample size indicate that PCs have trouble with novel concept formation but not label mapping.

This interpretation is somewhat in contrast with results from Hudson, Landi, and Cutting (2016). They showed that PCs have trouble attaching new labels to known items, showing lower accuracy both during learning and during a word application task, in which the novel word appeared in a sentence and participants were asked to judge the acceptability of the sentence. However, weaker concept knowledge could lead to poorer form-to-meaning mappings. That is, PC participants may have performed more poorly than TD controls on this task even if nonwords were not used.

While the current study is able to add some interesting insights to the study of word and category learning in PCs, it is not without limitations. One major limitation was the small variance in comprehension skill observed in our sample. 17 of our 28 subjects received a raw score of 11, 12, or 13 on the comprehension test. Because most of our subjects were of similar ages, scaled scores retained this limited variability. It is possible that the college-aged sample we

collected data from indeed had small variability in comprehension scores. However, it is also possible that our choice of comprehension measure contributed to the variability problem. The Nelson-Denny Reading Test comprehension subtest begins with its longest passage, which is nearly a page long. This fact, combined with the eight-minute time limit we imposed in order to reduce total testing time, meant that most subjects completed only one or two passages in the test, often performing at ceiling on both passages. Thus, the lack of a significant relationship between comprehension ability and performance on the category learning task may be due in part to the limited variability we had on the comprehension measure. However, our group is currently re-running the study on carefully selected PC and TD groups, with promising preliminary results.

Another limitation of the current study is that we had no training of verbal mapping without previous nonverbal concept construction. This was in part due to our research interests; we wanted to see how well PCs could map novel labels onto entirely novel concepts, rather than mapping new labels on to old concepts (e.g. Cain, Oakhill, & Lemmon, 2004; Hudson, Landi, & Cutting, 2016) or onto novel concepts that are dependent on old concepts (McGregor et al., 2013). Still, this design does not allow us to look at pure label mapping abilities. However, related data from another project suggests that comprehension skill is related to label mapping over and above decoding ability (see Supplemental Data). Thus, PCs may indeed have pure label-mapping difficulties. Another potential weakness was present in the implicit conditions. The presence of the circle or square may have added an additional feature to the categories, perhaps making the task more difficult. This would lead to the lower accuracy seen in the implicit blocks. However, shape was randomized across trials such that any given robot could be seen in both a circle and a square during different trials of the same block. The lack of consistent association between robot and shape hopefully mitigated this concern at least somewhat.

Overall, this study has shown that vocabulary is a strong predictor of performance on a verbal and nonverbal category learning task. Participants with poor comprehension and poor vocabulary show less word learning than participants with poor comprehension and typical vocabulary. In addition, comparison of the verbal and nonverbal conditions shows that PCs may have more difficulty constructing concepts than mapping labels onto those concepts. Even explicitly directing attention towards category-relevant features did not help PCs perform similarly to better comprehenders, suggesting that attentional differences are not the only factor differentiating word learning in PCs from that of their TD peers. Future research should investigate the relationship between working memory capacity, attention, and performance on this type of novel word learning task. In addition, we are currently re-running this paradigm on carefully-selected groups of adolescents. These individuals fit the more canonically defined profile of PCs. This will allow us to address the limitation of our comprehension measure, since the measure used to select these adolescents shows a much greater variability. We also plan to analyze the eye movement data from only accurate trials to ensure that this analysis only involves those trials where participants are completing the task. In addition, we plan to modify our comprehension measure and re-run the task on college-aged students. These replications will allow us to see if, given a better measure, comprehension ability beyond decoding is related to word and concept learning or if these types of learning are primarily tied to vocabulary.

### References

- About, K. S., Bailey, S. K., Petrill, S. A., & Cutting, L. E. (2016). Comprehending text versus reading words in young readers with varying reading ability: Distinct patterns of functional connectivity from common processing hubs. *Developmental Science*, 1–25.  
doi:10.1111/desc.12422
- Adlof, S. M., & Catts, H. W. (2015). Morphosyntax in poor comprehenders. *Reading and Writing*, 28(7), 1051–1070. doi:10.1007/s11145-015-9562-3
- Bailey, S., Hoeft, F., About, K., & Cutting, L. (2016). Anomalous gray matter patterns in specific reading comprehension deficit are independent of dyslexia. *Annals of Dyslexia*.  
doi:10.1007/s11881-015-0114-y
- Baldo, J. V., Bunge, S. A., Wilson, S. M., & Dronkers, N. F. (2010). Is relational reasoning dependent on language? A voxel-based lesion symptom mapping study. *Brain and Language*, 113(2), 59–64. <http://doi.org/10.1016/j.bandl.2010.01.004>
- Beckage, N., Smith, L., & Hills, T. (2011). Small worlds and semantic network growth in typical and late talkers. *PLoS ONE*, 6(5). doi:10.1371/journal.pone.0019348
- Bjorklund, D. F. (1985). The role of conceptual knowledge in the development of organization in children's memory. In C. J. Brainerd & M. Pressley (Eds.), *Basic Processes in Memory Development* (pp. 103–142). New York: Springer.
- Brown, J. I. (1960). The Nelson-Denny Reading Test.
- Butler, L. P., & Markman, E. M. (2014). Preschoolers use pedagogical cues to guide radical reorganization of category knowledge. *Cognition*, 130(1), 116–127.  
doi:10.1016/j.cognition.2013.10.002

Cain, K. E. (2003). Text comprehension and its relation to coherence and cohesion in children's fictional narratives. *British Journal of Developmental Psychology, 21*, 335–351.

doi:10.1348/026151003322277739

Cain, K. E., & Oakhill, J. V. (1996). The nature of the relationship between comprehension skill and the ability to tell a story. *British Journal of Developmental Psychology, 14*(2), 187–201.

doi:10.1111/j.2044-835X.1996.tb00701.x

Cain, K. E., & Oakhill, J. V. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and Writing, 11*(5/6), 489–503.

doi:10.1023/A:1008084120205

Cain, K. E., Oakhill, J. V., & Bryant, P. E. (2004). Children's Reading Comprehension Ability: Concurrent Prediction by Working Memory, Verbal Ability, and Component Skills. *Journal of Educational Psychology, 96*(1), 31–42. doi:10.1037/0022-0663.96.1.31

Cain, K. E., Oakhill, J. V., & Elbro, C. (2003). The ability to learn new word meanings from context by school-age children with and without language comprehension difficulties.

*Journal of Child Language, 30*(3), 681–94. doi:10.1017/S0305000903005713

Cain, K. E., Oakhill, J. V., & Lemmon, K. (2004). Individual Differences in the Inference of Word Meanings from Context: The Influence of Reading Comprehension, Vocabulary Knowledge, and Memory Capacity. *Journal of Educational Psychology, 96*(4), 671–681.

doi:10.1037/0022-0663.96.4.671

Carpenter, P. a, Just, M. a, & Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. *Psychological Review, 97*(3), 404–431. doi:10.1037/0033-295X.97.3.404

- Catts, H. W., Fey, M. E., & Zhang, X. (1999). Language Basis of Reading and Reading Disabilities: Evidence From a Longitudinal Investigation. *Scientific Studies of Reading*, 3(4), 331–361. doi:10.1207/s1532799xssr0304
- Catts, H. W., Adlof, S. M., & Weismer, S. E. (2006). Language deficits in poor comprehenders: a case for the simple view of reading. *Journal of Speech, Language, and Hearing Research*, 49(2), 278–293. doi:10.1044/1092-4388(2006/023)
- Cragg, L., & Nation, K. (2006). Exploring Written Narrative in Children with Poor Reading Comprehension. *Educational Psychology*, 26(1), 55–72. doi:10.1080/01443410500340991
- Cronin, V. S. (2002). The syntagmatic-paradigmatic shift and reading development. *Journal of Child Language*, 29(April 2002), 189–204. doi:10.1017/S0305000901004998
- Cutting, L. E., Materek, A., Cole, C. A. S., Levine, T. M., & Mahone, E. M. (2009). Effects of fluency, oral language, and executive function on reading comprehension performance. *Annals of Dyslexia*, 59(1), 34–54. doi:10.1007/s11881-009-0022-0
- Ehrlich, M.-F., Remond, M., & Tardieu, H. (1999). Processing of anaphoric devices in young skilled and less skilled comprehenders: Differences in metacognitive monitoring. *Reading and Writing: An Interdisciplinary Journal*, 11(1), 29–63. doi:10.1023/A:1007996502372
- Florit, E., Roch, M., & Levorato, M. C. (2011). The relationship between listening comprehension of text and sentences in preschoolers: Specific or mediated by lower and higher level components? *Applied Psycholinguistics*, 34(2), 395–415. doi:10.1017/S0142716411000749
- Gillette, J., Gleitman, H., Gleitman, L., & Lederer, a. (1999). Human simulation of vocabulary learning. *Cognition*, 73, 135–176.

Gleitman, L., Cassidy, K., Nappa, R., Papafragou, A., & Trueswell, J. (2005). Hard words.

*Language Learning and Development, 1*(1), 23–64. doi:10.1038/scientificamerican0491-138

Goff, D. A., Pratt, C., & Ong, B. (2005). The Relations Between Children's Reading

Comprehension, Working Memory, Language Skills and Components of Reading Decoding in a Normal Sample. *Reading and Writing, 18*(7-9), 583–616. doi:10.1007/s11145-004-7109-0

Henderson, L., Snowling, M. J., & Clarke, P. (2013). Accessing, Integrating, and Inhibiting

Word Meaning in Poor Comprehenders. *Scientific Studies of Reading, 17*(3), 177–198. doi:10.1080/10888438.2011.652721

Hills, T. T., Maouene, M., Maouene, J., Sheya, A., & Smith, L. B. (2009). Longitudinal Analysis

of Early Semantic Networks. *Psychological Science, 20*(6), 729. Retrieved from <http://pss.sagepub.com/content/20/6/729.abstract>

Hudson, N.S., Landi, N., & Cutting, L.E. (2016, July). Word Learning in Good and Poor

Comprehenders: Evidence of a Specific Reading Comprehension Deficit Profile. Poster session presented at the 23rd Annual Society for the Scientific Study of Reading Meeting, Association,

Keil, F. C., & Batterman, N. (1984). A characteristic-to-defining shift in the development of

word meaning. *Journal of Verbal Learning and Verbal Behavior, 23*(2), 221–236. doi:10.1016/S0022-5371(84)90148-8

Kieffer, M. J., Vukovic, R. K., & Berry, D. (2013). Roles of attention shifting and inhibitory

control in fourth-grade reading comprehension. *Reading Research Quarterly, 48*(4), 333–348. doi:10.1002/rrq.54

- Kim, Y.-S. (2015). Language and Cognitive Predictors of Text Comprehension: Evidence From Multivariate Analysis. *Child Development, 86*(1), 128–144. doi:10.1111/cdev.12293
- Landi, N. (2010). An examination of the relationship between reading comprehension, higher-level and lower-level reading sub-skills in adults. *Reading and Writing, 23*(6), 701–717. <http://doi.org/10.1007/s11145-009-9180-z>
- Landi, N., & Perfetti, C. A. (2007). An electrophysiological investigation of semantic and phonological processing in skilled and less-skilled comprehenders. *Brain and Language, 102*(1), 30–45. doi:10.1016/j.bandl.2006.11.001
- Li, M., & Kirby, J. R. (2014). Unexpected Poor Comprehenders Among Adolescent ESL Students. *Scientific Studies of Reading, 18*(2), 75–93. doi:10.1080/10888438.2013.775130
- Locascio, G., Mahone, E. M., Eason, S. H., & Cutting, L. E. (2010). Executive dysfunction among children with reading comprehension deficits. *Journal of Learning Disabilities, 43*, 441–454. doi:10.1177/0022219409355476
- Magnuson, J. S., Tanenhaus, M. K., Aslin, R. N., & Dahan, D. (2003). The time course of spoken word learning and recognition: studies with artificial lexicons. *Journal of Experimental Psychology. General, 132*(2), 202–227. <http://doi.org/10.1037/0096-3445.132.2.202>
- McGregor, K. K., Licandro, U., Arenas, R., Eden, N., Stiles, D., Bean, A., & Walker, E. (2013). Why Words Are Hard for Adults With Developmental Language Impairments. *Journal of Speech Language and Hearing Research, 56*(6), 1845. doi:10.1044/1092-4388(2013)12-0233)

McMaster, K. L., van den Broek, P., Espin, C. A., White, M. J., Rapp, D. N., Kendeou, P., ...

Carlson, S. (2012). Making the right connections: Differential effects of reading intervention for subgroups of comprehenders. *Learning and Individual Differences, 22*(1), 100–111. doi:10.1016/j.lindif.2011.11.017

Nation, K., Adams, J. W., Bowyer-Crane, C. A., & Snowling, M. J. (1999). Working memory deficits in poor comprehenders reflect underlying language impairments. *Journal of Experimental Child Psychology, 73*(2), 139–58. doi:10.1006/jecp.1999.2498

Nation, K., Cocksey, J., Taylor, J. S. H., & Bishop, D. V. M. (2010). A longitudinal investigation of early reading and language skills in children with poor reading comprehension. *Journal of Child Psychology and Psychiatry, 51*(9), 1031–1039. doi:10.1111/j.1469-7610.2010.02254.x

Nation, K., Marshall, C. M., & Altmann, G. T. M. (2003). Investigating individual differences in children's real-time sentence comprehension using language-mediated eye movements. *Journal of Experimental Child Psychology, 86*(4), 314–329. doi:10.1016/j.jecp.2003.09.001

Nation, K., Marshall, C. M., & Snowling, M. J. (2001). Phonological and semantic contributions to children's picture naming skill: Evidence from children with developmental reading disorders. *Language and Cognitive Processes, 16*, 241–259.

doi:10.1080/01690960042000003

Nation, K., & Snowling, M. J. (1998). Semantic Processing and the Development of Word-

Recognition Skills: Evidence from Children with Reading Comprehension. Nation, K., & Snowling,

M. J. (1998). Semantic Processing and the Development of Word-Recognition Skills:

Evidence from Children with Reading Co. *Journal of Memory and Language, 39*(1), 85–

101. doi:10.1006/jmla.1998.2564

- Nation, K., & Snowling, M. J. (1999). Developmental differences in sensitivity to semantic relations among good and poor comprehenders: Evidence from semantic priming. *Cognition*, *70*(1), 4–9. doi:10.1016/S0010-0277(99)00004-9
- Nelson, K. (1977). The syntagmatic-paradigmatic shift revisited: a review of research and theory. *Psychological Bulletin*, *84*(1), 93–116. doi:10.1037/0033-2909.84.1.93
- Pimperton, H., & Nation, K. (2010). Suppressing irrelevant information from working memory: Evidence for domain-specific deficits in poor comprehenders. *Journal of Memory and Language*, *62*(4), 380–391. doi:10.1016/j.jml.2010.02.005
- Pimperton, H., & Nation, K. (2014). Poor Comprehenders in the Classroom: Teacher Ratings of Behavior in Children With Poor Reading Comprehension and Its Relationship With Individual Differences in Working Memory. *Journal of Learning Disabilities*, *47*(3), 199–207. doi:10.1177/0022219412454172
- Raven, J. C., & John Hugh Court. (1998). *Raven's progressive matrices and vocabulary scales*. Oxford, UK: Oxford Psychologists Press.
- Ricketts, J., Sperring, R., & Nation, K. (2014). Educational attainment in poor comprehenders. *Frontiers in Psychology*, *5*, 1–12. doi:10.3389/fpsyg.2014.00445
- Schneider, W., Eschman, A., and Zuccolotto, A. (2012). *E-Prime User's Guide*. Pittsburgh: Psychology Software Tools, Inc.
- Sesma, H. W., Mahone, E. M., Levine, T., Eason, S. H., & Cutting, L. E. (2009). The contribution of executive skills to reading comprehension. *Child Neuropsychology*, *15*(3), 232–246. doi:10.1080/09297040802220029

- Silva, M., & Cain, K. E. (2015). The relations between lower and higher level comprehension skills and their role in prediction of early reading comprehension. *Journal of Educational Psychology, 107*(2), 321–331. doi:10.1037/a0037769
- Smiley, S. S., & Brown, A. L. (1979). Conceptual preference for thematic or taxonomic relations: A nonmonotonic age trend from preschool to old age. *Journal of Experimental Child Psychology, 28*(2), 249–257. doi:10.1016/0022-0965(79)90087-0
- Snedeker, J., & Gleitman, L. (2004). Why It Is Hard To Label Our Concepts. *Weaving a Lexicon, 427*(February), 257–294. doi:10.1038/nature02339.1.
- Tong, X., Deacon, S. H., & Cain, K. E. (2013). Morphological and syntactic awareness in poor comprehenders: another piece of the puzzle. *Journal of Learning Disabilities, 47*(1), 22–33. doi:10.1177/0022219413509971
- Tong, X., Deacon, S. H., Kirby, J. R., Cain, K. E., & Parrila, R. (2011). Morphological awareness: A key to understanding poor reading comprehension in English. *Journal of Educational Psychology, 103*(3), 523–534. doi:10.1037/a0023495
- Torgesen, J. K., Wagner, R., & Rashotte, C. (1999). TOWRE–2 test of word reading efficiency. *Austin, TX: Pro-Ed.*
- van der Schoot, M., Vasbinder, A. L., Horsley, T. M., Reijntjes, A., & van Lieshout, E. C. D. M. (2009). Lexical ambiguity resolution in good and poor comprehenders: An eye fixation and self-paced reading study in primary school children. *Journal of Educational Psychology, 101*(1), 21–36. doi:10.1037/a0013382
- Woodcock, R. W., McGrew, K. S., Mather, N., & Schrank, F. (2001). Woodcock-Johnson III NU tests of achievement. *Rolling Meadows, IL: Riverside Publishing.*

- Woollams, A. M., Ralph, M. A. L., Plaut, D. C., & Patterson, K. (2007). SD-squared: On the association between semantic dementia and surface dyslexia. *Psychological Review*, *114*(2), 316–339. doi:2007-05396-004 [pii]\n10.1037/0033-295X.114.2.316 [doi]
- Yuill, N., & Oakhill, J. V. (1988). Understanding of anaphoric relations in skilled and less skilled comprehenders. *British Journal of Psychology*.

	Mean	SD	Range
<b>Reading comprehension<sup>a</sup></b>	174.0	13.7	157-210
<b>Vocabulary<sup>a</sup></b>	224.4	18.2	183-255
<b>Nonverbal IQ<sup>b</sup></b>	15.8	4.6	3-22
<b>Decoding<sup>c</sup></b>	102.5	5.5	95-113
<b>Word fluency<sup>c</sup></b>	102.5	10.3	84-120

Table 1. Descriptive statistics for sample.

<sup>a</sup>Scaled scores<sup>b</sup>Raw scores<sup>c</sup>Standardized scores

	1	2	3	4	5
<b>1. Reading comprehension</b>	-	0.53*	0.39*	0.25	0.23
<b>2. Vocabulary</b>		-	0.52*	0.45*	0.31*
<b>3. Nonverbal IQ</b>			-	0.29	0.24
<b>4. Decoding</b>				-	0.58*
<b>5. Word fluency</b>					-

Table 2. Correlations between behavioral measures.

\* $p < .02$ 

Condition	Mean	SD	$t(39)$	p-value
Implicit Nonverbal	0.56	0.17	2.17	.02
Explicit Nonverbal	0.74	0.26	5.77	< .0001
Implicit Verbal	0.70	0.22	5.75	< .0001
Explicit Verbal	0.87	0.17	13.65	< .0001

Table 3. Statistics describing accuracy at test for each condition of the category learning experiment. T-tests indicate whether the mean accuracy is significantly greater than chance (0.5).

Condition	Mean	SD
Implicit Nonverbal	2537	1631
Explicit Nonverbal	2380	828
Implicit Verbal	2348	1011
Explicit Verbal	1938	651

Table 4. Descriptive statistics for RTs by condition.

	Probe		Target		Distractor	
	Mean	SD	Mean	SD	Mean	SD
Implicit Nonverbal	0.305	0.033	0.243	0.039	0.231	0.036
Explicit Nonverbal	0.326	0.072	0.280	0.043	0.207	0.030
Implicit Verbal	0.337	0.053	0.226	0.028	0.257	0.034
Explicit Verbal	0.331	0.048	0.216	0.037	0.248	0.057

Table 5. Descriptive statistics for fixation proportions by condition and interest area.

Condition	Order	M Accuracy	P-value
Implicit nonverbal	Implicit first	0.48	0.84
	Explicit first	0.68	<b>0.005</b>
Explicit nonverbal	Implicit first	0.81	<b>&lt;0.001</b>
	Explicit first	0.84	<b>&lt;0.001</b>
Explicit verbal	Implicit first	0.88	<b>&lt;0.001</b>
	Explicit first	0.87	<b>&lt;0.001</b>
Implicit verbal	Implicit first	0.64	<b>0.004</b>
	Explicit first	0.81	<b>0.0003</b>

Table 6. T-tests comparing accuracy in each condition by order to chance ( $\mu = 0.5$ ).

	BC (n = 8)		PCHV (n = 8)		PCLV (n = 8)	
	Mean	SD	Mean	SD	Mean	SD
<b>Reading comprehension</b>	191.6	13.5	165.6	3.4	167.1	4.8
<b>Vocabulary</b>	238.4	12.2	232.9	10.6	202.1	11.1
<b>Nonverbal IQ</b>	17.6	2.6	18.2	3.1	11.8	5.6
<b>Decoding</b>	102.3	5.7	104.6	5.7	100.1	4.7
<b>Word fluency</b>	101.2	5.5	103.1	12.5	99.1	9.4

Table 7. Behavioral scores for the three groups.

Condition	Group	M Accuracy	P-value
Implicit nonverbal	PCLV	0.49	0.68
	PCHV	0.57	0.16
	BC	0.62	<i>0.065</i>
Explicit nonverbal	PCLV	0.74	<b>0.02</b>
	PCHV	0.92	<b>&lt;0.001</b>
	BC	0.81	<b>0.002</b>
Explicit verbal	PCLV	0.78	<b>0.01</b>
	PCHV	0.97	<b>&lt;0.001</b>
	BC	0.89	<b>&lt;0.001</b>
Implicit verbal	PCLV	0.61	0.09
	PCHV	0.83	<b>0.001</b>
	BC	0.73	<b>0.01</b>

Table 8. T-tests comparing accuracy in each condition by group to chance ( $\mu = 0.5$ ).