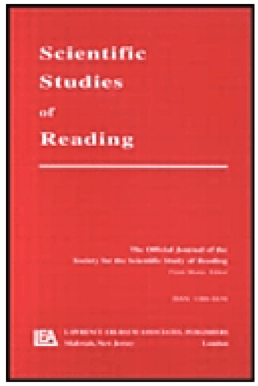


Category learning in Poor Comprehenders

Journal:	*Scientific Studies of Reading*
Manuscript ID	HSSR-2017-0128.R2
Manuscript Type:	Research Article
Keywords:	reading comprehension, category learning, lexical processing, lexical semantics

Category learning in Poor Comprehenders

Kayleigh Ryherd[ac*]

Nicole Landi [abc]

[a]University of Connecticut

[b]Haskins Laboratories

[c]CT Institute for the Brain and Cognitive Sciences

*Corresponding Author
kayleigh.ryherd@uconn.edu
406 Babbidge Road, Unit 1020
Storrs, CT 06269
603-438-3107

Abstract

Poor comprehenders (PCs) are characterized by poor reading comprehension despite intact decoding and general cognitive ability. Poor word meaning knowledge is one of the earliest deficits associated with a PC profile. We examined processes underpinning word learning in PCs using a category learning paradigm. Adolescent participants (20 TD, 19 PC, ages 13-18) learned novel categories with two key manipulations: information type (verbal vs. nonverbal) and training type (directed vs. undirected). We found that PCs showed more benefit from directed training than TD individuals overall; however, both groups performed similarly when receiving directed blocks first. Moreover, when undirected training was received first, TD individuals showed better performance in directed as compared to undirected blocks, whereas PCs who receive undirected training first showed no significant difference between training types. Our investigation indicates that PCs may have different strategies for learning, especially when their attention is not externally directed towards relevant features.

Keywords: reading comprehension; category learning; lexical semantics

Introduction

Poor comprehenders (PCs) have poor reading comprehension despite intact word reading ability and typical general cognitive functioning. PCs also exhibit a range of sub-clinical deficits in many language-related skills (see Landi & Ryherd, 2017 for a recent review). This range of deficits reflects the complex nature of reading comprehension, which involves multiple component processes. One of the earliest and most consistent weaknesses found in PCs is in vocabulary, reflecting a deficit in word-level meaning knowledge (Catts, Adlof, & Weismer, 2006; Kim, 2015; Nation & Snowling, 1999).

In addition to vocabulary weaknesses, PCs exhibit poor performance in experimental tasks tapping word-level meaning throughout development. PCs are slower and less accurate than typically-developing (TD) individuals on semantic processing tasks, including synonym judgment and generation of low-frequency names during picture identification (Nation, Marshall, & Snowling, 2001; Nation & Snowling, 1998). PCs appear to build semantic networks based on surface-level features (e.g., co-occurrence, frequency) rather than deeper information that informs category membership (e.g., causal reasoning). They show weak category coordinate priming when prime and target do not frequently co-occur in spoken language (e.g., *sheep – cow*; (Nation & Snowling, 1999). They also show reduced priming for subordinate meanings (e.g., *flower – bulb)* but typical priming to more frequent meanings of words (e.g., *light – bulb*; Henderson, Snowling, & Clarke, 2012). Although the existence of lexical-semantic weaknesses in PCs is well documented, little is known about why PCs fail to develop strong knowledge of word meanings. To explore the development of this deficit, some research has examined word learning in PCs.

Studies of word learning in PCs primarily focus on the ability to learn novel words from a story context as well as through direct instruction on word definitions (e.g., “A *mup* is a small rounded ball.”). PCs produce less accurate definitions of the novel words than TD children in both tasks (Cain, Oakhill, & Lemmon, 2004). These findings suggest that PCs have difficulty learning novel words from sentential context. Other studies of word learning in PCs teach novel word forms to children that are associated with a picture and additional semantic information, either from a list of features or a story context. These studies show that while PCs show similar performance to their TD peers on word meaning learning, they show reduced retention of this information after a delay (Nation, Snowling, & Clarke, 2007; Ricketts, Bishop, & Nation, 2008) While these tasks have additional nonverbal semantic information (i.e., pictures), much of the semantic learning is still done using sentences. PCs frequently show sub-clinical deficits in sentence processing more generally (e.g., weak grammatical processing; Tong, Deacon, & Cain, 2013). Thus, it is difficult to separate word learning difficulties from broader language deficits in studies where word learning is embedded primarily in sentence processing tasks. In the current study, we use a task where participants learn novel categories organized by a set of features. In this way, we can teach participants meaningful relations among items without using a sentence context.

A key manipulation in the current study is the inclusion of categories that have both verbal (linguistic) and nonverbal features, allowing us to discover whether PCs can use both types of features to the same extent. The inclusion of this manipulation is motivated by research in infants and adults indicating an important role for language in category and word learning (Lupyan, Rakison, & McClelland, 2007; Smith, Jones, Landau, Gershkoff-Stowe, & Samuelson, 2002). Early-acquired words tend to refer to categories that are organized by shape (e.g., ball).

Once children have learned this relationship, attention during novel word learning is directed towards shape. For example, young children who are trained to learn the names for novel categories organized by shape in the laboratory also tend to extend object names based on shape even for untrained items (Smith et al., 2002). Thus, because linguistic cues (object names) are correlated with perceptual features (shape), the presence of a linguistic cue leads infants to attend to these perceptual features.

This inter-relation between linguistic and perceptual cues in feature selection during word learning extends to adults. That is, adults take advantage of language cues even when they provide redundant information. In one study, participants learned to sort aliens into one of two categories based on head shape. A second redundant category feature was also provided; some participants were given category labels, while others were provided with an additional visual cue. Participants who were given labels showed better generalization of the new categories, suggesting that language improves category learning specifically (Lupyan et al., 2007). Currently, it is unclear whether the language-primary deficits seen in PCs include impaired use of linguistic cues like labels. If they do, PCs may have difficulty internalizing statistical regularities and contingencies that involve language, leading to poorer word learning.

If PCs are not using learned regularities to help them select relevant features for word learning, explicit instruction on these features may lead to better word learning. To address this possibility, we include a second crucial manipulation in our task. Participants complete both a directed version of the task, where their attention is explicitly drawn towards relevant features, and an undirected version, where relevant features are presented incidentally. This allows us to investigate whether having explicit instruction on relevant features provides additional benefit for PCs during category learning.

Prior research on inference-making in PCs suggests that this may be the case. PC children perform poorer than controls on inferential questions. However, if the experimenter directs their attention to portions of the text relevant for answering these inferential questions, PC children show marked improvement (Cain & Oakhill, 1999). These results indicate that PC individuals have difficulty selecting important information and benefit from direction towards relevant information. If this is true, then PCs should show better word learning when their attention is directed. Recall, however, that PCs in the word learning study done by Cain et al. (2004) showed poorer performance than TD individuals even in the direct instruction task. Thus, depending on the experimental design, PCs may not always benefit from direct instruction. Importantly, the effect of direct instruction in PCs has not been studied outside of tasks that involve reading or sentence processing, and we do not know whether it will improve category learning in PCs.

These lines of research suggest two possibilities we sought to explore with our study. Impaired learning of word meanings in PCs could be due to inadequate usage of linguistic cues (e.g., labels) as well as failure to extract relevant features without explicit instruction. Neither of these possibilities has been investigated in PCs to date. We manipulate the type of information available (verbal vs. nonverbal) and the type of instruction (directed vs. undirected) to try to understand more about how PCs learn novel categories. Considering that PCs seem to have a language-primary deficit (Pimperton & Nation, 2010), we hypothesize that they will show poorer performance than TD individuals when learning verbal features as compared to nonverbal features. Additionally, since PCs show a benefit from directed instruction on relevant information for inference-making tasks (Cain & Oakhill, 1999) we hypothesize that they will show better performance in the directed version of our task as compared to the undirected version. Specifically, we predict an interaction between group and instruction. We expect both groups to benefit from directed instruction, but we predict that the PC group will show a larger benefit than the TD group. This pattern of results would suggest that PCs overall have difficulty selecting relevant features from their environment for word learning, with additional difficulty in using verbal features.

Methods

This study protocol was approved by the Yale University Human Investigation Committee, whose review processes meet the standards set forth by the US Federal Policy for the Protection of Human Subjects. The authors report no conflicts of interest.

Participants

Two groups of adolescents in middle and high school (grades 7-12) participated in this experiment. Inclusion and exclusion criteria (listed below) ensured that variability in reading comprehension was not simply due to poor word reading or general cognitive function. We conducted sensory testing with all participants to ensure that hearing and vision were within normal limits. For more details on participant age and gender, see Table 1. Parental informed consent and adolescent assent were collected for adolescents under 18 years of age, while informed consent was collected from adolescents who were 18 years old.

Assessments of Reading and General Cognitive Function. Reading comprehension was measured using the Kaufman Test of Educational Achievement Second Edition (KTEA-II) reading comprehension test (Kaufman & Kaufman, 2004). In this task, participants read short passages and answer multiple-choice as well as free response questions orally. Both literal and inferential questions are included. Decoding was measured using the Word Attack (WA) subtest of the Woodcock-Johnson III (Woodcock, McGrew, Mather, & Schrank, 2001). In this task, participants read nonwords aloud at their own pace. Scores are based on the number of correctly decoded nonwords. Finally, general cognitive function was assessed with the Wechsler Abbreviated Scale of Intelligence II (WASI; Wechsler & Hsiao-pin, 2011). This task has four subtests: block design, matrix reasoning, vocabulary, and similarities.

Group Assignment. Participants in the PC group had average decoding (≥ 95 SS on WA) and general cognitive function (≥ 80 SS on WASI performance IQ) and poor reading comprehension skill (≤ 90 SS on KTEA). Participants in the TD group had average decoding (≥ 95 SS on WA), general cognitive function (≥ 80 SS on WASI PIQ), and average or above average reading comprehension (≥ 100 SS on KTEA). These cutoffs are consistent with previous work on adolescent PCs published by our group (Breen, Kaswer, Van Dyke, Krivokapić, & Landi, 2016).

Group descriptive statistics. For reading comprehension, the PC group showed significantly lower standardized scores than the TD group. The PC group also showed significantly lower performance IQ than the TD group. Finally, the PC group showed marginally lower decoding than the TD group. For more details, see Table 1. Importantly, group differences found in decoding and performance IQ are driven by better than average performance of the TD group rather than a deficit exhibited by the PC group. Still, we include both decoding and performance IQ in subsequent analyses as additional predictors to add additional statistical control.

Category Learning Task

Our core task tested participants' ability to learn novel categories based on nonverbal and verbal features with two different types of instruction. In this task, participants learned novel "families" of items. Each family included three robot images that shared two features, a motion pattern (nonverbal) and a label (verbal). These two features were unique to each family of robots, so that motion patterns and labels were always paired together. Some families were learned in directed training, where the training task was directly related to the family features. Other families were learned in undirected training, where the training task was unrelated to the family features, which were presented incidentally. Throughout the task, participants learned a total of eight novel robot families.

Stimuli. The images used in the study were cartoon robots. Each family of robots was assigned a motion pattern and a label. The motion patterns included paths such as zig-zags and arcs. Motion patterns are ideal nonverbal features because they are visual but not always present. Names were consonant-vowel-consonant (CVC) constructions. To reduce phonological interference, each family name had a unique onset and none of the names rhymed.

Visual Norming. 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." Families were constructed to have a member rated as visually similar to the other two (mean rating = 1.74, SD = 0.55). However, these latter two robots were rated as less visually similar (mean rating = 3.6, SD = 0.35). This method ensured that visual similarity was an unreliable cue to category membership.

Procedure. Participants completed training and testing in four conditions. All training and testing tasks were presented on a computer. The four conditions corresponded to the two main task manipulations: instruction type (undirected versus directed) and information type (verbal versus nonverbal). Training always preceded testing, and nonverbal tasks always preceded verbal tasks. This was designed to loosely mimic development, where conceptual representations are sometimes formed prior to attaching a label. Instruction type order (undirected or directed first) was counterbalanced across subjects.

Training (see Fig. 1)**.** The training task was designed to teach participants the family features for each robot, ideally leading them to group the robots mentally into their respective families. In nonverbal training blocks, participants learned to associate robots with their motion patterns. All feedback in nonverbal training blocks was also nonverbal (i.e., visual). In verbal training blocks, participants learned to associate labels with robots. In verbal blocks, feedback was only verbal, presented aurally. The directed training blocks asked participants to make responses based on the family features (i.e., the motion patterns and labels). In this way, directed training drew participants' attention towards these features. The undirected training blocks had a main task that was unrelated to the family features. The family features were presented incidentally between trials. Thus, family feature information was present but was not the focus for the undirected training. Overall, training blocks had a trial-and-error structure, where participants' first responses were essentially random. After each response, feedback was provided. Extensive feedback including the family features was provided after a correct response. This feedback structure allowed participants to actively engage in learning the different families.

Participants began all trials by clicking on a fixation cross. In **directed nonverbal** training, participants then saw a black box make a motion. This motion pattern was identical to the motion pattern of one of the robots on the screen. 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 **directed verbal** training, participants saw two robots and heard, “Find a *[label]*.” 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 *[label]*.” In **undirected nonverbal** training, one of the robots 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. The goal was to click on the robot inside the shape that matched the empty center shape. Feedback was identical to directed nonverbal blocks. In **undirected verbal** training, participants saw the same 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. Participants completed 54 trials of training in each condition, which included every combination of robot without presenting any from the same family simultaneously. Aside from 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.

Testing. The testing block was identical for all types of training. At test, participants completed a triad task, similar to those used in other categorization studies (Gelman & Markman, 1986; Lupyan, 2009). Participant saw three robots arranged in a triangle on the screen. Two of the robots were from the same family. 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. 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.

Results

Overall Task Performance

Accuracy at test for all training conditions was significantly greater than chance for TD individuals (see Table 2 and Fig. 2) However, performance was only significantly greater than chance for the PC group in the directed-verbal condition. At an uncorrected threshold, PCs also showed learning greater than chance in the directed-nonverbal condition. Thus, significant learning was found for TD individuals, but significant learning was only found in one condition for the PC group when using a corrected threshold. In particular, the directed-verbal condition showed very high accuracy for both groups, with 12/20 TD individuals and 4/19 PCs showing accuracy of 95% or greater. The reliabilities for each individual condition were high, Cronbach's $\alpha > 0.95$. The task as a whole also had high reliability, Cronbach's $\alpha = 0.89$.

Effects of information and training type. We used a linear mixed-effects model to examine the effects of the experimental manipulations on accuracy at test. Accuracy for each subject and condition was logit transformed. First, we confirmed that the random effect of participant significantly improved the base model. The relationship between accuracy and task condition showed significant variance in intercepts across participants, $SD = 1.14$, $\chi^2(2) = 38.43$, $p < .0001$. Next, we tested fixed main effects, including group (PC vs. TD), training type (undirected vs. directed), and information type (verbal vs. nonverbal). In addition, performance IQ and decoding score were added as fixed effects to control for group differences in these scores. Performance IQ was not a significant predictors of accuracy ($b = 0.004$, $t(35) = 0.24$, $p = .81$). However, decoding did significantly predict accuracy ($b = 0.051$, $t(35) = 2.06$, $p = 0.05$). Still, adding the task-related predictors into the model after decoding significantly increased fit $\chi^2(7) = 50.65$, $p < .0001$. This model revealed a significant main effect of training type, $b = -1.11$, $t(111) = -4.54$, $p < .0001$. Participants were generally more accurate in directed blocks rather than undirected blocks. The model also revealed a significant main effect of information type, $b = 0.56$, $t(111) = 2.30$, $p = .023$. Participants overall showed greater accuracy in verbal blocks as compared to nonverbal blocks. The main effect of group was not significant, $b = -0.81$, $t(35) = -1.51$, $p = 0.14$. However, there was a significant interaction between group and training type, $b = 0.85$, $t(111) = 2.44$, p – 0.016.

To break down this interaction, we conducted separate multilevel models separated by group. These models revealed that training type did not have a significant effect on accuracy in the TD group, $b = -0.26$, $t(54) = -1.24$, $p = 0.22$. However, training type did have a significant effect on accuracy in the PC group, $b = -1.11$, $t(57) = -3.94$, $p = 0.0002$. PCs were more accurate in directed blocks than they were in undirected blocks. This suggests that PCs but not TD individuals benefit from explicit direction in a category learning task.

Order effects. Results from the first analysis indicate that PCs show greater accuracy for directed blocks as compared to undirected blocks, suggesting that the directed blocks may be easier for this group. However, training type (directed vs. undirected) was counterbalanced across subjects. Thus, participants who receive the directed blocks first may be clued in to the task by the time they reach undirected blocks. This could lead to systematic variance within training type based on order. To investigate this possibility, we added order as a fixed effect to the final model (see Table 3 and Fig. 3).

Adding order to the model significantly improved fit, $\chi^2(8) = 41.83$, $p < .0001$. With this model, some significant effects emerged. In particular, a significant three-way interaction between group, training type, and order was found. This interaction was first broken down by conducting separate multilevel models on subsets of the data for each of the two orders (directed first and undirected first). The first model used only data from participants who received the directed blocks first. This model indicated no significant effect of group, $b = 0.28$, $t(17) = 0.38$, $p = 0.71$. The second model used only data from participants who received the undirected blocks first. This model revealed a main effect of group ($b = -2.07$, $t(14) = -3.24$, $p = 0.006$) and a main effect of training type ($b = -2.18$, $t(48) = -6.03$, $p < 0.001$). Overall, PCs who completed the undirected tasks first performed worse than TD individuals who completed the undirected tasks first. In addition, all participants who received the undirected blocks first performed worse on undirected blocks. Further, there was an interaction between group and training type, $b = 1.80$, $t(48) = 3.53$, $p = 0.0009$.

To further investigate this interaction, we conducted two multilevel models split by group (TD vs. PC) only for those participants who received undirected training first. The first model, using only PCs who received undirected training first, indicated these individuals showed no main effect of training type ($b = -0.40$, $t(24) = -1.08$, $p = 0.29$). However, the second model that included only TD individuals who received undirected training first, indicated that these individuals showed a main effect of training type ($b = -2.18$, $t(24) = -5.54$, $p < 0.0001$). TD adolescents who received undirected training first had better performance in directed blocks as compared to undirected blocks. Thus, the three-way interaction reflects important group differences. When directed training is received first, the TD and PC groups do not differ in categorization performance at test. However, when undirected training is received first, the effect of training type (i.e., directed versus undirected) depends on group. TD individuals who receive undirected training first show better performance in directed as compared to undirected blocks, whereas PCs who receive undirected training first show similar performance across training types.

Discussion

The current study investigated the processes underlying word and category learning in adolescent PCs, with two main manipulations. First, we presented both verbal and nonverbal category-relevant features, manipulating information type. Second, we had two types of instruction, varying whether or not attention was directed towards these relevant features. Because PCs exhibit sub-clinical oral language weaknesses, we hypothesized that they would show less learning than their TD peers on verbal blocks and similar performance to TD peers on nonverbal blocks. We did not find any evidence for this interaction. We also hypothesized that PCs would show a greater benefit from directed blocks than TD peers. We found an interaction between group and training type, such that TD individuals showed no difference in accuracy between directed and undirected blocks while PCs showed significantly better accuracy in directed blocks. However, because the four conditions were relatively similar, there may have been some crossover or learning effects between blocks. Thus, getting the directed or undirected blocks first may have led some participants towards the goals of the task earlier. To account for this possibility, we added order as an additional predictor of category learning.

Adding order to the model produced additional group-level effects. We found that there were no group differences in accuracy when directed blocks were completed first. In directed blocks, all responses during training pertain to the family features. Both groups were able to learn the families of robots to a similar extent when they were directed towards family features from the beginning of the task. Despite the fact that responses during training in undirected blocks require attention towards irrelevant features (i.e., the shape surrounding the robot), movement patterns and labels are still the family features. Thus, participants receiving directed blocks first may be clued in to the task so that by the time they get to undirected blocks they already know which features are most important.

In contrast, group differences did emerge when undirected blocks were completed first. Both groups showed relative difficulty in undirected blocks when they were encountered first. However, the TD group as a whole seemed to recover when given directed blocks afterwards, showing much greater accuracy for the directed blocks relative to PCs. There was also substantial variability within the PC group when completing the directed blocks second. This indicates that while most of the TD individuals were able to switch from a suboptimal learning strategy in the undirected block to an optimal strategy in the directed block, many of the PC individuals persisted in using a suboptimal strategy.

This inability to recover from a suboptimal strategy, if replicated further, may indicate difficulty with executive functioning more broadly in PCs. Indeed, some deficits in executive functioning have been found in PC populations. PCs have trouble inhibiting irrelevant information, even when specifically led to do so (Henderson et al., 2012). In our task, it is possible that having directed blocks first provided the PCs with the guidance they needed to overcome executive function deficits and perform similarly to their TD peers. The inability to recover may also indicate a lack of flexibility in learning strategies. Indeed, PCs exhibit cognitive flexibility deficits during card-sorting tasks (Cartwright, Bock, Coppage, Hodgkiss, & Nelson, 2017). Thus, PCs may have an underlying impairment in considering multiple potential dimensions both at a moment-to-moment timescale as well as across longer blocks. PCs may be selecting and perseverating on a single irrelevant feature.

Perseveration on an inappropriate strategy or feature during learning could contribute to the word-level meaning weaknesses seen in PCs. Specifically, PCs may be routinely attending to irrelevant features when learning new categories or concepts. If so, this could be one potential factor leading PCs to acquire word meanings in an atypical way, perhaps leading to inefficient lexical access. Similar ideas about atypical attention to features have been proposed to explain semantic networks with unusual structure in late talkers (Beckage, Smith, & Hills, 2011). Late talkers (i.e., children with small vocabularies for their age) have been shown to exhibit some of the same early language deficits as young children who go on to become PCs (Catts et al., 2006; Manhardt & Rescorla, 2002; Rescorla, 2002). Thus, a tendency to focus on irrelevant features may contribute to these early deficits in both groups, and some late talkers may go on to have PC profiles at school age.

While the above interpretations of the results from this study seem compatible with extant research, alternative explanations must be considered. The undirected task may have been too difficult for the PCs, leading them to largely disengage from the task even when the directed blocks started. PCs in particular showed greater-than-chance accuracy in only one condition (directed verbal), which may suggest that this task may not promote significant learning in this group overall. In addition, we did not observe any effects of information type (verbal vs. non-verbal training). However, the verbal and nonverbal blocks are not equivalent. Nonverbal blocks always preceded verbal blocks, so that participants already had exposure to the families before they started verbal blocks. Thus, it is possible that differences would emerge due to this manipulation if some verbal blocks had preceded nonverbal blocks.

Findings from this study must be viewed in light of some limitations. First, it is important to note that our order analysis was exploratory on a small sample size (~10 subjects per order per group). As such, these results should be interpreted cautiously and would benefit from replication in a larger sample size. . A larger sample would also allow for more investigation of individual differences in performance on this task and how they might relate to many reading comprehension sub-processes. It is also important to note that our TD individuals are somewhat above-average in their reading comprehension, decoding, and performance IQ, while our PCs had average decoding, average to low-average performance IQ, and below average reading comprehension. Thus, generalization of our results to larger populations should be done with care.

This study investigated some of the learning processes that underlie category and word learning in PCs and their TD peers. PCs were able to perform like TD individuals when their attention was initially guided towards relevant features. PCs who were not guided toward relevant features at the start of the task performed worse than their TD peers. These PCs seemed to perseverate on suboptimal strategies for learning even when relevant features were made salient. This is consistent with prior research indicating executive function and cognitive flexibility deficits in PCs. Difficulty orienting towards relevant features also has implications for the development of strong semantic networks that play a role in later text and speech comprehension.

While our findings are limited by factors discussed above, they motivate new avenues for research. Semantic category learning is foundational for vocabulary development, yet there is almost no research on how individual variability in semantic categorization may be related to language and reading skills in older children. One limitation of our work is that the learning paradigm is completely rule-based: each category has two features for inclusion that define the category completely. Although this type of categorization is important for learning some categories, and for explicit learning of labels, many categories (i.e. most natural kinds) are similarity-based and defined by probabilistic rules for inclusion (Minda & Miles, 2010). Thus, future research that includes probabilistic, feature-based semantic category learning will be important for understanding the relationship between category learning and comprehension.

Further, our group by order findings suggest that perseveration may be a fruitful line of inquiry for better understanding PCs . Future studies could test for perseveration errors in multiple domains, including classic executive function tasks as well as during reading to see if these errors generalize to reading-related processes like inference-making. Finally, more research on learning in general is needed in PCs to help support our current knowledge about word and category learning in this population. This could include topics such as statistical learning and learning in an academic setting. With research in these topics, we will be able to better understand how PCs construct meaning from their environment and during reading.

References

Beckage, N., Smith, L. B., & Hills, T. (2011). Small worlds and semantic network growth in typical and late talkers. *PLoS ONE*, *6*(5). http://doi.org/10.1371/journal.pone.0019348

Breen, M., Kaswer, L., Van Dyke, J. A., Krivokapić, J., & Landi, N. (2016). Imitated Prosodic Fluency Predicts Reading Comprehension Ability in Good and Poor High School Readers. *Frontiers in Psychology*, *7*, 1–17. http://doi.org/10.3389/fpsyg.2016.01026

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. http://doi.org/10.1023/A:1008084120205

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. http://doi.org/10.1037/0022-0663.96.4.671

Cartwright, K. B., Bock, A. M., Coppage, E. A., Hodgkiss, M. D., & Nelson, M. I. (2017). A comparison of cognitive flexibility and metalinguistic skills in adult good and poor comprehenders. *Journal of Research in Reading*, *40*(2), 139–152. http://doi.org/10.1111/1467-9817.12101

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. http://doi.org/10.1044/1092-4388(2006/023)

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition*, *23*(3), 183–209. http://doi.org/10.1016/0010-0277(86)90034-X

Henderson, L., Snowling, M., & Clarke, P. (2012). Accessing, Integrating, and Inhibiting Word

Meaning in Poor Comprehenders. *Scientific Studies of Reading*, (June 2015), 1–22. http://doi.org/10.1080/10888438.2011.652721

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman test of educational achievement - comprehensive form.* American Guidance Service.

Kim, Y.-S. (2015). Language and Cognitive Predictors of Text Comprehension: Evidence From Multivariate Analysis. *Child Development*, *86*(1), 128–144. http://doi.org/10.1111/cdev.12293

Landi, N., & Ryherd, K. (2017). Understanding specific reading comprehension deficit: A review. *Language and Linguistics Compass*, *11*(2), 1–24. http://doi.org/10.1111/lnc3.12234

Lupyan, G. (2009). Extracommunicative functions of language: verbal interference causes selective categorization impairments. *Psychonomic Bulletin & Review*, *16*(4), 711–718. http://doi.org/10.3758/PBR.16.5.986

Lupyan, G., Rakison, D. H., & McClelland, J. L. (2007). Language is not Just for Talking: Redundant Labels Facilitate Learning of Novel Categories. *Psychological Science*, *18*(12), 1077–1083. http://doi.org/10.1111/j.1467-9280.2007.02028.x

Manhardt, J., & Rescorla, L. (2002). Oral narrative skills of late talkers at ages 8 and 9. *Applied Psycholinguistics*, *23*(01), 1–21. http://doi.org/10.1017/S0142716402000012

Minda, J. P., & Miles, S. J. (2010). The influence of verbal and nonverbal processing on category learning. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 52, pp. 117–162). Burlington: Academic Press. http://doi.org/10.1016/S0079-7421(10)52003-6

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.

http://doi.org/10.1080/01690960042000003

Nation, K., & Snowling, M. J. (1998). Semantic Processing and the Development of Word-Recognition Skills: Evidence from Children with Reading Comprehension Difficulties. *Journal of Memory and Language*, *39*(1), 85–101. http://doi.org/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. http://doi.org/10.1016/S0010-0277(99)00004-9

Nation, K., Snowling, M. J., & Clarke, P. (2007). Dissecting the relationship between language skills and learning to read: Semantic and phonological contributions to new vocabulary learning in children with poor reading comprehension. *Advances in Speech Language Pathology*, *9*(2), 131–139. http://doi.org/10.1080/14417040601145166

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. http://doi.org/10.1016/j.jml.2010.02.005

Rescorla, L. (2002). Language and reading outcomes to age 9 in late-talking toddlers. *Journal of Speech, Language, and Hearing Research : JSLHR*, *45*(2), 360–71. http://doi.org/10.1044/1092-4388(2002/028)

Ricketts, J., Bishop, D. V. M., & Nation, K. (2008). Investigating orthographic and semantic aspects of word learning in poor comprehenders. *Journal of Research in Reading*, *31*(1), 117–135. http://doi.org/10.1111/j.1467-9817.2007.00365.x

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. http://doi.org/10.1111/1467-9280.00403

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. http://doi.org/10.1177/0022219413509971

Wechsler, D., & Hsiao-pin, C. (2011). *WASI-II: Wechsler abbreviated scale of intelligence*. Pearson.

Woodcock, R. W., McGrew, K. S., Mather, N., & Schrank, F. (2001). *Woodcock-Johnson III NU tests of achievement*. Rolling Meadows, IL: Riverside Publishing.

Table 1

Descriptive and group difference statistics

Measure	Mean (SD)		*t*	*p*
	TD (n = 20)	PC (n = 19)		
Gender	11 F, 9 M	14 F, 5 M		
Age (years)	16.3 (1.6)	16.9 (1.5)	1.077	0.29
Reading comprehension	112.6 (8.82)	84. 2 (5.9)	11.891	<.001*
Decoding	107.8 (8.6)	101.8 (5.8)	2.578	0.01*
Performance IQ	114.7 (12.4)	94.5 (8.6)	5.95	<.001*

* indicates that the groups are different at Bonferroni-corrected threshold

Table 2

Summary of performance on category learning task by group and condition

Block	TD			PC		
	Mean (SD)	*t*	*p*	Mean (SD)	*t*	*p*
Directed-Nonverbal	0.82 (0.22)	6.393	<.001*	0.64 (0.22)	2.849	0.01
Undirected-Nonverbal	0.65 (0.21)	3.154	0.005*	0.58 (0.23)	1.436	0.17
Directed-Verbal	0.91 (0.11)	16.608	<.001*	0.71 (0.24)	3.776	0.001*
Undirected-Verbal	0.82 (0.18)	7.982	<.001*	0.60 (0.23)	1.822	0.085

* indicates that performance was different than chance (.5) at Bonferroni-corrected threshold

Table 3

Results from linear mixed-effects model including order

Predictor	*b*	SE *b*	*t*	df	*p*
Intercept	-4.72	3.00	-1.57	105	0.12
Performance IQ	0.01	0.02	0.34	33	0.74
Decoding	0.05	0.02	2.20	33	0.03*
Order	0.41	0.56	0.74	33	0.46
Group	-0.22	0.61	-0.36	33	0.72
Train Type	-0.24	0.29	-0.82	105	0.42
Info Type	0.64	0.29	2.23	105	0.03*
Order x Group	-1.17	0.79	-1.47	33	0.15
Order x Train Type	-1.94	0.43	-4.51	105	0.00*
Group x Train Type	0.09	0.42	0.21	105	0.84
Order x Info Type	-0.18	0.43	-0.42	105	0.67
Group x Info Type	-0.39	0.42	-0.94	105	0.35
Train Type x Info Type	0.34	0.41	0.84	105	0.41
Order x Group x Train Type	1.72	0.62	2.79	105	0.01*
Order x Group x Info Type	0.55	0.62	0.90	105	0.37
Order x Train Type x Info Type	0.24	0.61	0.40	105	0.69
Group x Train Type x Info Type	-0.39	0.59	-0.66	105	0.51
Order x Group x Train Type x Info Type	-0.72	0.87	-0.83	105	0.41

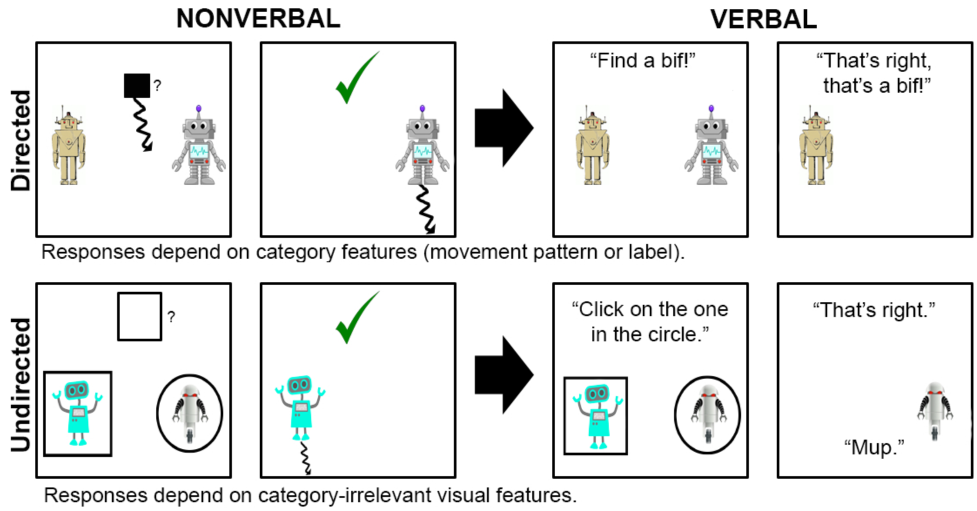
* indicates significance at a p<.05 threshold.

Figure Captions

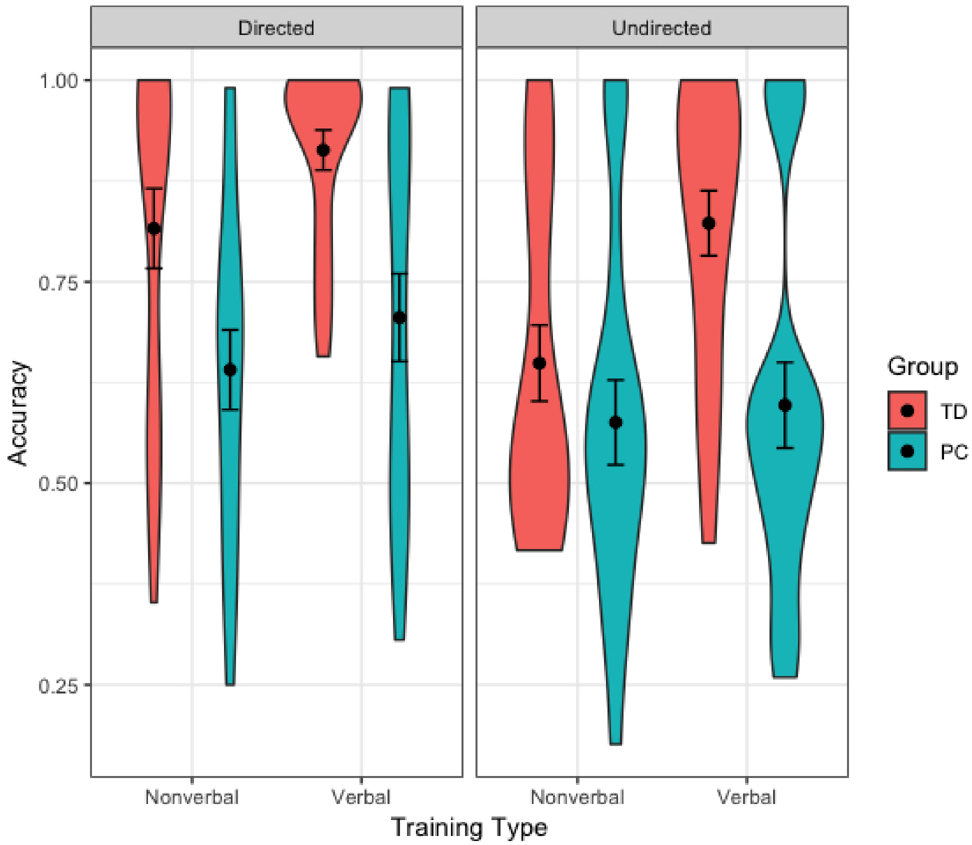
Fig. 1. Example trials for all four training conditions.

Fig. 2. General task performance. Violin plots show the distribution of the data. The points indicate the mean. Error bars indicate one standard error. TD = typically developing. PC = poor comprehender

Fig. 3. Task performance including order effects. Violin plots show the distribution of the data. The points indicate the mean. Error bars indicate one standard error. TD = typically developing. PC = poor comprehender.

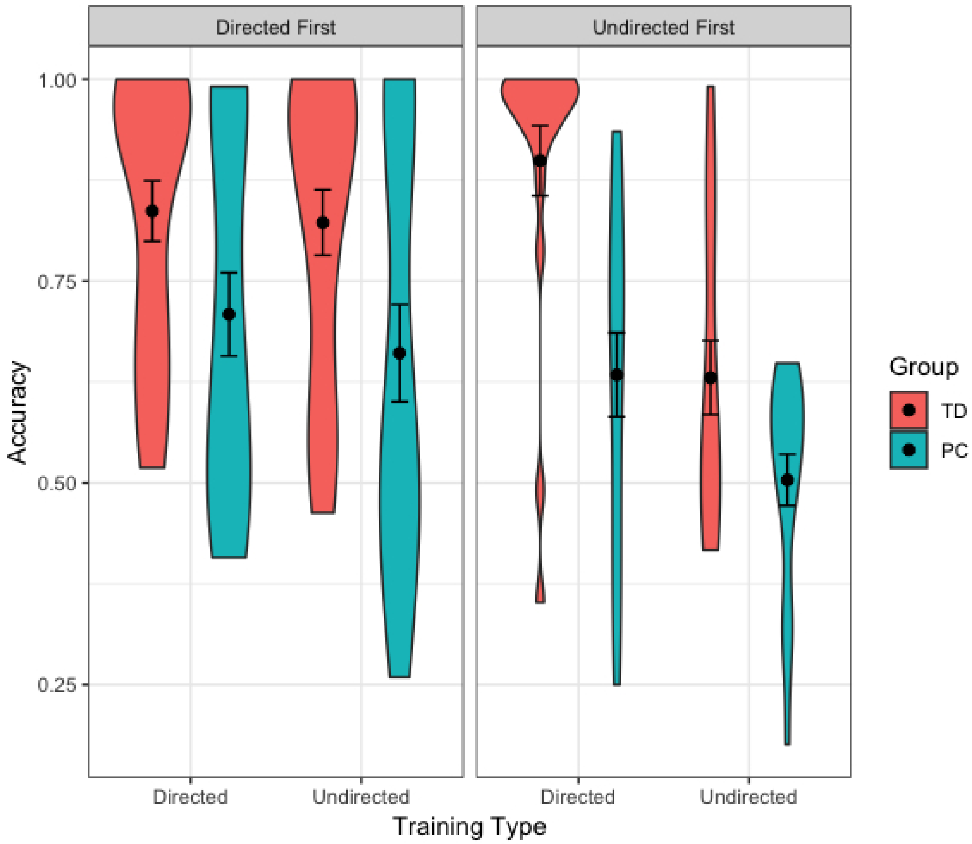


Example trials for all four training conditions.



General task performance. Violin plots show the distribution of the data. The points indicate the mean. Error bars indicate one standard error. TD = typically developing. PC = poor comprehender

200x172mm (72 x 72 DPI)



Task performance including order effects. Violin plots show the distribution of the data. The points indicate the mean. Error bars indicate one standard error. TD = typically developing. PC = poor comprehender.

204x176mm (72 x 72 DPI)