Children explore conservatively when learning novel word extensions

Martin Zettersten (mzettersten@ucsd.edu)
University of California San Diego, Department of Cognitive Science

Jaime W. Chen, Casey Lew-Williams
Princeton University, Department of Psychology

Abstract

Children are active, curious learners. How might children's curiosity shape their curriculum during word learning? Past research suggests that children's tendency to explore can lead them to discover novel information during learning. This exploratory tendency could be especially useful when learning word meanings: exploring potential meanings for words broadly could help children efficiently probe a word's possible extension. To investigate this question, we tested how children (5-8 years of age) and adults sample information when presented with a novel word and tasked with uncovering the word's extension. Overall, we found that children explored novel word extensions conservatively. Children (as well as adults) favored sampling choices that confirmed a novel word meaning, as opposed to exploring broader possible meanings. Younger children's sampling choices were especially conservative, with children often sampling the narrowest possible generalization option. Older children were more exploratory, probing broader possible word extensions more frequently. Counter to proposals that children are generally more exploratory at younger ages, our results suggest that when children test the extension of novel word meanings, they are often more likely to confirm their hypotheses than to explore.

Keywords: word learning; word extension; generalization; active learning; information-seeking; exploration

Introduction

When learners first encounter a word (like "fish") together with a novel referent (e.g., a clownfish), they are faced with fundamentally ambiguous information: the possible extension of the word is underdetermined by the input (Quine, 1960; Xu & Tenenbaum, 2007). For example, "fish" might refer to a specific kind of fish, clownfish (a subordinate category level), fish in general (the basic level), all sea creatures (a superordinate category level), or a much more general category such as "things found in the ocean" (a more general hypernym). Given a limited set of exemplars associated with a given word, the possible extension of a word remains potentially vast, with any set of learning experiences consistent with multiple possible generalizations. How do children learn to successfully generalize novel word meanings despite ambiguous input?

One solution is that learners could actively explore new information to clarify word meanings in the face of ambiguous input (Coenen et al., 2019; Settles, 2012). For example, for "fish," they could test if the word also generalizes to referents at different category levels, such as other fish like blue tang, other sea creatures like octopuses, or more general ocean-related items such as seaweed. The hypothesis that exploratory tendencies could be helpful to learners is appealing given evidence that children are predisposed to explore their learning environments broadly (Gopnik, 2020). Past research suggests that children may be more flexible than adults when searching for new hypotheses and exploring for rewards (Blanco & Sloutsky, 2021; Gopnik et al., 2017; Lucas et al., 2014). Children's propensity to explore helps them discover new structure, sometimes even finding regularities that adults miss (Liquin & Gopnik, 2022; Sumner et al., 2019). How might children's tendency to explore support their ability to learn about the extension of novel words?

Computational modeling work has demonstrated that there are potentially substantial benefits to adaptive sampling during word learning (Hidaka et al., 2017; Gelderloos et al., 2020). For example, Hidaka et al. (2017) found that when word learning models sample adaptively, focusing on less frequently encountered words, they can learn a large vocabulary much more rapidly compared to when learning events are drawn randomly from Zipfian distributions. This work is mirrored by empirical findings demonstrating that both adults and children will actively sample object-label associations that are informative or associated with higher uncertainty given past experience (de Eccher et al., 2024; Kachergis et al., 2013; Zettersten & Saffran, 2021). However, much of the prior computational and empirical work shares a similar limitation: typically, the learning task is to form straightforward mappings between individual labels and single referents, which is a mismatch to the often ambiguous and complex nature of lexical-semantic knowledge (Wojcik et al., 2022). How do learners sample and generalize in ambiguous contexts with many words and many possible referents?

In the current study, we tested how children (5-8 years of age) and adults seek information when learning to generalize a novel word to multiple possible category levels. We focused on 5-8-year-olds because past work suggests that children are increasingly likely to make sampling choices driven by uncertainty within this age range (Zettersten & Saffran, 2021). Adapting an existing experimental paradigm (Xu & Tenenbaum, 2007; Zettersten et al., 2023), we systematically manipulated the initial exemplars participants were trained on when learning a novel word. For example, learners might see a novel word together with a single image of a clownfish, multiple images of a clownfish, or three different types of fish. Past research has shown that learners will adapt how they generalize a word's meaning in response to differing exemplar sets, e.g. inferring a more narrow word meaning when viewing the word with multiple clownfish as compared to a single clownfish or multiple exemplars of varying types of fish. While the precise mechanisms driving this effect are debated (Lewis & Frank, 2018; Spencer et al., 2011; Wang et al., 2022), all accounts agree that participants shift their expectations about a novel word's extension in response to differing exemplar sets. In the current paradigm, we use these training manipulations to investigate how children and adults subsequently sample new information about a word's extension. To what extent do we observe broad exploration in children's information-seeking when learning to generalize novel words? We predicted that children would show a greater tendency to explore broader hypotheses about a word's meaning than adults, and that this tendency might lead children to generalize novel word meanings differently.

Method

We preregistered the study on the Open Science Framework (https://osf.io/qsfdy). All data and analysis code are openly available (https://osf.io/7bukp).

Participants

Child Participants. We recruited 83 English-speaking children (43 female, 40 male; mean age: 6.9 years, range: 5.1-9.0 years). Due to a minor experimental error, 23 participants were initially considered for exclusion and were not counted towards our preregistered target sample of N = 60. These participants completed an experiment in which an image from the Training Phase was (erroneously) shown again as an option in the Test Phase. However, given that the error did not affect the Sampling Phase (and had minimal impact on the Test Phase), we ultimately opted to include all 83 participants to maximize power. The main results reported below all follow the same basic pattern whether these 23 participants are included or not.

Adult Participants. We recruited 60 adult participants (35 female, 25 male; mean age: 33.8 years, SD = 13.1) from Prolific. Participants received \$1.00 for completing the study. The top 25% of participants with the highest test performance received a \$0.20 bonus to incentivize effort. We excluded 3 additional participants for failing an attention check on multiple rounds (n = 1) or submitting nonsensical or bot-like responses to open-ended prompts (n = 2).

Stimuli

The general task used a total of 6 category sets, one for each round. The sets each contain 15 total images. These images consisted of 5 subordinate-level images (e.g., clownfish in the sea creature category), 4 basic-level images (e.g., multiple kinds of fish), and 6 superordinate-level images (e.g., other sea creatures). The 6 sets were split into two groups, Group A (vegetables, sports items, land animals) and B (fruits, vehicles, sea creatures). Three of our category sets–vegetables, land animals, and vehicles–were adopted from Lewis and Frank (2018). We also created and normed three new sets: fruits, sea creatures, and sports items. To label training images, we used 6 nonce words: *sibu, kita, beppo, tibble, roozer, guffy*. Nonce words were randomly assigned to categories and conditions for each participant.

Design & Procedure

The task was presented using jsPsych (version 6.3; de Leeuw, 2015) and hosted online through Pavlovia (https://pavlovia.org/). Adult participants completed the task online, following on-screen instructions. Child participants visited the Princeton Baby Lab in person to participate. The task was administered through an iPad, and instructions were provided verbally by an experimenter. For the child-adapted version of this task, we used an alien cover story to facilitate engagement from child participants and motivate their best efforts in learning the novel words. In the cover task, children were introduced to an alien character who needed help finding Earth items for their Earth item collections. Children were instructed that their goal was to learn each alien's language to figure out what they were asking for and help them find Earth items.

The overall structure of the task was identical for children and adults. Participants completed six total experiment rounds. In each round, the participant was tasked with learning a different novel word. Each round consisted of a Training, Sampling, and Test Phase.

Training Phase. During the Training Phase, participants viewed either one exemplar image (1-Item condition) or three exemplar images (3-Item conditions) from a category set. Each training image was presented with a randomly assigned nonce word. Then, we presented participants with an attention check. In the child-adapted task, children were prompted to verbally repeat the novel word after all of the round's training images had been shown. In the adult-adapted task, adults were asked to type the novel word into a text box as an attention check. Of the included participants, novel word recall was highly accurate (M = 98.6%).

The six rounds were presented in a 1-Item Block and a 3-Item Block consisting of three rounds each. The order of the 1-Item Block vs. the 3-Item Block was counterbalanced. For the 1-Item condition, the single training image was always drawn from the category's subordinate-level images (e.g., a clownfish; Figure 1A). In the 3-Item Block, participants were presented with three different 3-Item training conditions, corresponding to the category level implied by the three images shown during training. In the 3-Item Narrow condition, participants viewed 3 subordinate-level examples (e.g., three clownfish); in the 3-Item Intermediate condition, participants viewed 3 basic-level examples (e.g., three different types of fish); and in the 3-Item Broad condition, participants viewed 3 superordinate-level examples (e.g., three sea creatures from different basic-level categories). Within each block, training condition order was randomized across participants. The assignment of category set groups (Group A or Group B) to each block was counterbalanced. For each participant, category sets within the 3-Item Block were randomly assigned to training conditions.

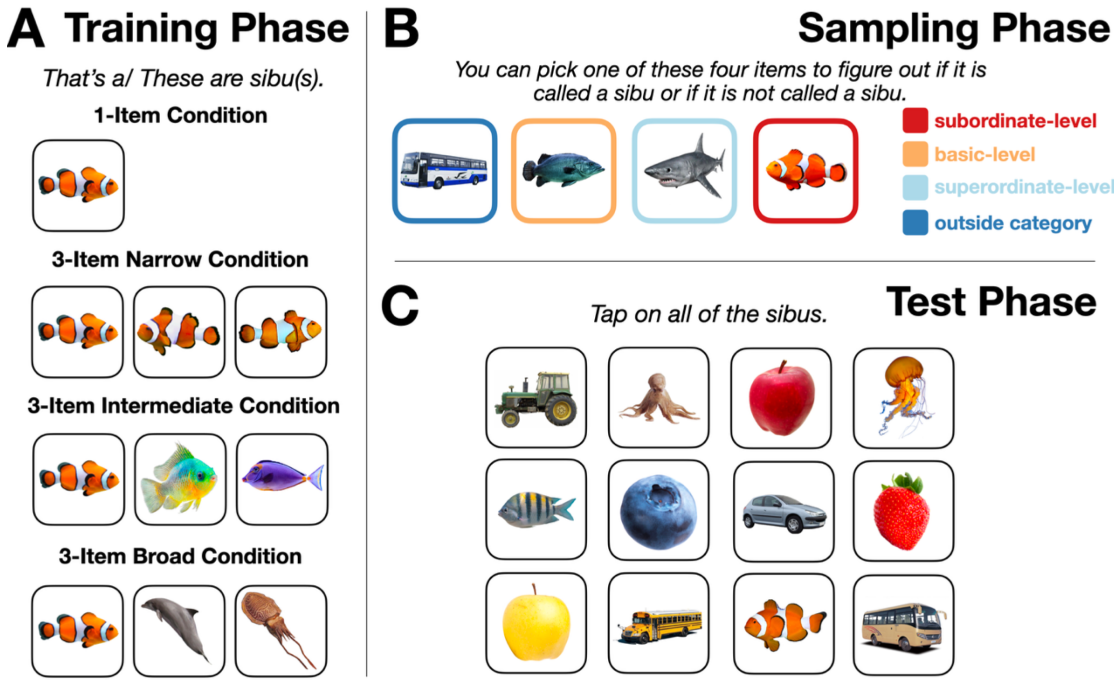

Figure 1: Overview of the task design: Training Phase (A), Sampling Phase (B), and Test Phase (C). Colors (not visible to participants) indicate different choice types (subordinate-level, basic-level, superordinate-level and outside category choices).

Sampling Phase. Next, participants were given the opportunity to sample one novel image in order to learn whether it could or could not be labeled by the round's novel word. The sampling choice array consisted of four total images (Figure 1B). Three of these images were from the same category set as was shown during training: one at the subordinate level, one at the basic level, and one at the superordinate level. The fourth image was an outside category option, an object from a different category set within the same group (A or B). Participants were instructed to select a single image to learn about. Once they made their selection, the next screen showed the chosen image labeled as either the novel word (e.g., "*beppo*") or not the novel word (e.g., "not a *beppo*") based on the "ground truth" meaning of the novel word.

A "ground truth" for the meaning of each of the six novel words was specified in order to supply participants with feedback based on their selections during the Sampling Phase (e.g., whether their selection is or is not a *beppo*). By varying the meaning of each novel word, we ensured that participants could discover new information about each word through sampling. The ground truth category level for each round was randomly selected from one of two possibilities based on the training condition. For the 3-Item training conditions, the ground truth level was either at the level implied by the training examples or at one level higher in the category hierarchy. For example, if 3 clownfish were shown for the word *beppo*, *beppo* might mean "clownfish" (same level, subordinate-level ground truth) or "fish" (one level up, basic-level ground truth). The ground truth for the three 1-Item rounds was selected randomly from the same three combinations used in the 3-Item rounds.

Test Phase. After the Sampling Phase, we tested how participants generalized the word to novel exemplars. Participants were shown an array of 12 new images from which they were instructed to select all the objects that could be labeled by the novel word and nothing else. The round's training image(s) and sampled image continued to be shown with their labels (e.g., "*beppo*" or "not a *beppo*") on the screen as a visual reminder.

The 12 images in the test array consisted of 4 images from the round's category set and 4 from each of the other two category sets within the same group (Figure 1C). These test images were not shown during the Training or Sampling Phase (i.e., these images were never labeled). Each category set's 4 test images were selected as follows: 1 subordinate-level object, 1 basic-level object, and 2 superordinate-level objects. The same 12 images were used in the test array for the rounds that used the Group A category set for training and sampling, and a different collection of 12 images were consistently used for the rounds using Group B category sets.

After participants finished selecting images from the array, they were asked what they thought the novel word meant. Children's verbal responses were transcribed by the researcher. Adults entered their answer into a text box on the screen. At the conclusion of this phase, the round ended as well. Children received a sticker to thank them for helping the alien with each round as a way of motivating their continued efforts throughout the task.

Results

Sampling Choices

Adult learners–but not children–flexibly shifted their sampling choices depending on the training condition. To test whether training condition affected sampling choices in general, we fit a multinomial logit model (separately for children and for adults) using the mlogit package (Croissant, 2020) in R (version 4.3.2; R Development Core Team, 2023). The model predicted participants' category-based sampling choice type (four options: the within-category subordinate, basic, or superordinate exemplars, or an outside-category exemplar) from training condition (3-Item Narrow, Intermediate, or Broad; dummy coded) in the 3-item block. For adults, a likelihood-ratio test indicated a significant effect of training condition on participants' sampling choices, $\chi^2(6) = 23.15$, $p < .001$. This suggests that adults shifted their pattern of sampling choices depending on the 3 exemplars presented during training. However, contrary to our preregistered prediction, children's sampling choices did not differ across training condition, $\chi^2(6) = 1.47$, $p = .96$.

Children made more superordinate choices and fewer subordinate choices as they grew older. Given our broad age range, we explored whether choice patterns were predicted by age and found a strong significant effect of child age on sampling choices, $\chi^2(3) = 28.85$, $p < .001$, though there was no age by training condition interaction ($p = .51$). This means that children's sampling choices changed markedly with age, though children did not necessarily change their sampling choices in a way that was tuned to the specific composition of the exemplars in the training condition. Follow-up analyses demonstrated that the age effect was driven by the fact that children became much more likely to choose the superordinate exemplar across all three training conditions as they grew older (logistic mixed-effects model; $b = 0.97$, $z = 3.67$, $p < .001$), while subordinate choices declined with age ($b = -0.57$, $z = -2.36$, $p = .02$).

Both children and adults tended to make (similar amounts of) confirmatory choices. To further characterize how children and adults approached the word learning task, we investigated the likelihood that children and adults made exploratory choices during the Sampling Phase. We considered all sampling choices that probed word meanings within the category level implied by the training exemplars confirmatory (1-Item and 3-Item Narrow condition: within-category subordinate choice; 3-Item Intermediate condition: within-category subordinate or basic-level choice; 3-Item Broad condition: all 3 within-category choices in the sampling grid). All sampling choices that probed word meanings outside of these confirmatory options were considered exploratory, because they probed a word's possible extension beyond the scope introduced during training. Using this binary variable of participants' sampling choice type, we compared children's and adults' sampling behavior and investigated developmental changes.

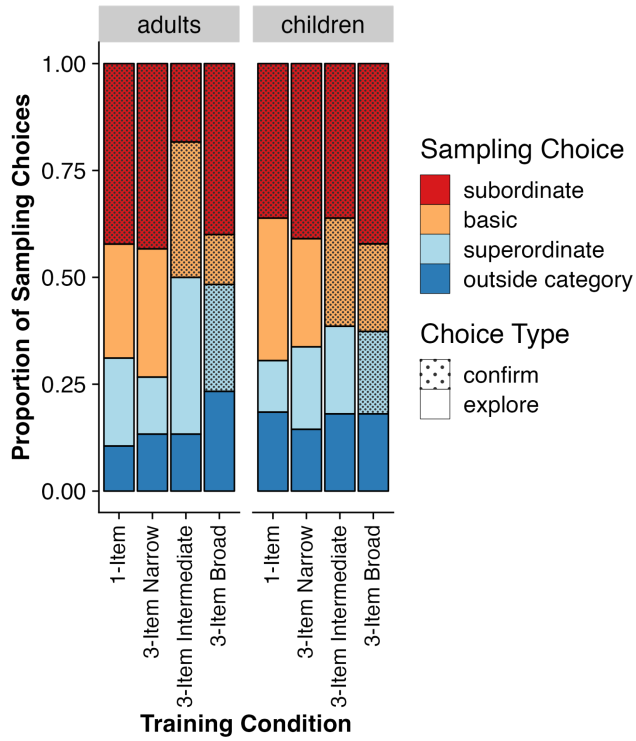

Figure 2: Adults' and children's overall sampling choices across training condition. The dotted pattern indicates confirmatory (vs. exploratory) sampling choices.

In a preregistered analysis, we fit a logistic mixed-effects model predicting the likelihood of making an exploratory choice from age group (children vs. adults; centered) while controlling for training condition. The model included a by-participant random intercept. There was no significant difference between children's and adults' likelihood of making an exploratory choice ($\chi^2(1) = 0$, $p = .99$; Figure 2). In exploratory analyses, we found no significant interaction between training condition and sampling group, though the effect was marginal ($\chi^2(3) = 6.92$, $p = .07$). Using a variable offset that adjusted for chance differently at each training condition, we found that both children ($b = -0.53$, z = -2.31, p = .02) and adults ($b = -0.54$, z = -2.03, p = .04) made more confirming (as opposed to exploratory) choices than would be expected by chance.

Children's exploration increased with age after training on single exemplars. To investigate how children's exploratory choices changed with age, we fit a logistic mixed-effects model predicting whether children made an exploratory choice from age, training condition (dummy coded, with 1-Item condition as the reference level), and their interaction. The model included a by-participant random intercept. We found a significant effect of age ($b = 0.57$, $\chi^2(1) = 4.21$, $p = .04$), indicating that older children had a higher tendency to explore in the 1-Item training trials (Figure 3). There was a significant effect of training condition ($\chi^2(3) = 65.07$, $p < .001$) and no age by training condition interaction ($\chi^2(3) = 2.47$, $p = .48$).

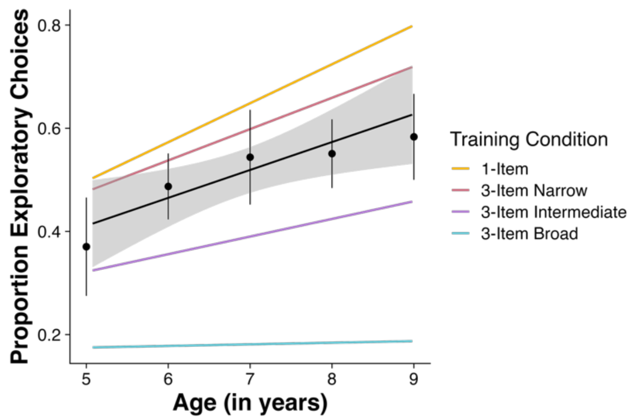

Figure 3: Children's tendency to make exploratory choices increased with age in the 1-Item condition. Points represent the average proportion of exploratory choices within a given age bin (with a range of [-0.5, +0.5] around each year) across all conditions, with +1/-1 SEs. Lines represent linear regression fits; error bands are 95% confidence intervals.

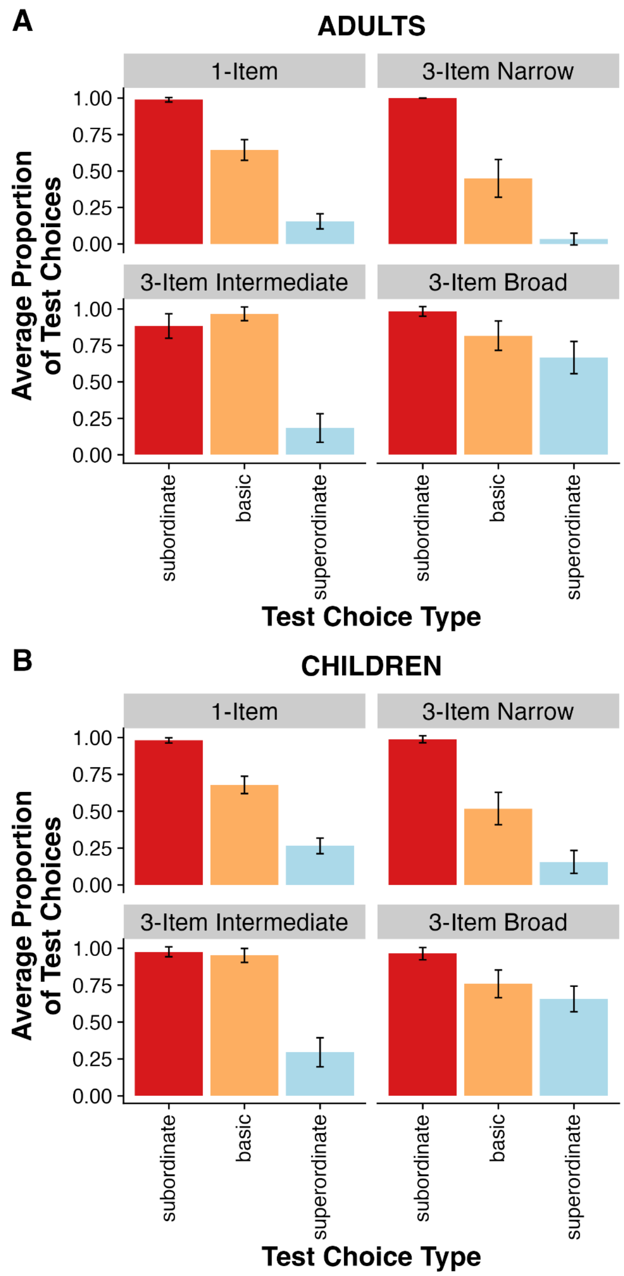

Figure 4: Proportion of test choices at different category levels during the Test Phase, depending on training condition (facetted) for (A) adults and (B) children. Error bars represent 95% CIs.

Test Performance

Training condition modulated both children's and adults' choices at test. To investigate children and adults' generalizations in the Test Phase, we computed the proportion of images selected at each category level (subordinate, basic, superordinate) for each participant (as in Xu & Tenenbaum, 2007) and investigated whether each of these proportions were predicted by training condition using a preregistered mixed-effects regression approach. For the superordinate level choices, we computed the proportion of superordinate-level options that are chosen (out of 2) and fit a linear mixed-effects model. For the subordinate and basic level models, there was only one available option, so we fit logistic mixed-effects models. We fit all models using the full preregistered random effects structure, including random intercepts for participant and category set and by-participant random slopes for training condition, and pruned random effects until models converged. Degrees of freedom were estimated using the Satterthwaite approximation. Models were fit separately for adults and children.

Children's and adults' choices at test were strikingly similar across all training conditions (Figure 4). For children, training condition was a robust predictor of the proportion of subordinate-level ($\chi^2(3) = 14.91$, $p = .002$), basic-level ($\chi^2(3) = 40.55$, $p < .001$) and superordinate-level choices ($\chi^2(3) = 135.07$, $p < .001$). Similarly for adults, training condition also predicted the likelihood of making all three choice types (subordinate-level: $\chi^2(3) = 27.61$, $p < .001$; basic-level: $\chi^2(3) = 62.55$, $p < .001$; superordinate-level: $\chi^2(3) = 164.19$, $p < .001$). In general, subordinate-level choices remained high across all training conditions, while the likelihood of basic-level and superordinate-level choices changed based on the training manipulation, suggesting that learners shifted their generalization of each word's meaning flexibly from a narrow, subordinate-level interpretation to a broad, superordinate-level interpretation depending on training.

Both children and adults generalized more to the basic-level in the 1-Item vs. the 3-Item Narrow training condition. To investigate the classic finding from Xu and Tenenbaum (2007) on differential generalization depending on the number of training exemplars, we focused an additional preregistered analysis on the data from the 1-Item and the 3-Item Narrow conditions. We fit a logistic mixed-effects predicting the likelihood of making a basic-level choice (one option in the test array, i.e. a binary variable) from condition (1-Item condition vs. three-Item Narrow condition; centered) for both the child group and the adult group. We included random intercepts for participant and category type. Both children ($b = 1.28$, $z = 3.38$, $p < .001$) and adults ($b = 1.11$, $z = 2.98$, $p = .003$) were more likely to select the basic-level item at test in the 1-item condition than in the 3-item Narrow condition, replicating the original effect (see top row in Figure 4A and 4B). We found no evidence that this effect differed between children and adults ($p = .79$).

Children and adults incorporated sampling feedback when generalizing word meanings. Because the ground truth of word meanings was randomized across trials, learners who made exploratory sampling choices could receive negative or positive feedback. We used this design feature to investigate how learners integrate sampling feedback into their choices at test. If learners incorporate feedback from their sampling choices, they should be more likely to select items from the category level at which they sampled following positive feedback compared to negative feedback. To test this prediction, we focused our analysis on the subset of trials in which participants made exploratory sampling choices. We fit a logistic mixed-effects model predicting the likelihood of choosing test items at the category level sampled (e.g., selecting superordinate-level test items after sampling the superordinate-level choice) from the sampling feedback received (positive vs. negative; centered), age group (children vs. adults; centered), and their interaction. We included random intercepts for participant and category set, as well as a by-participant random slope for sampling feedback. Learners were much more likely to choose test items at the sampled level after positive feedback compared to negative feedback, $b = 13.89$, $z = 7.51$, $p < .001$. There was no interaction with age group ($p = .52$) and the effect of sampling feedback held separately for both children ($b = 5.22$, $z = 9.39$, $p < .001$) and adults ($b = 11.60$, $z = 4.23$, $p < .001$).

General Discussion

Exploring a word's meaning–seeking more information about the possible extension of a novel word–could be an effective way to understand how word meanings generalize. Strikingly, we do not find evidence that children are more exploratory during a novel word generalization task than adults. Both children and adults tended to prefer confirmatory choices, and only adults were more likely to shift their sampling preferences depending on training condition. Children's sampling choices tended to be conservative, sticking close to the narrowest possible generalization option (the subordinate choice), especially at younger ages. Children grew more exploratory with age, making subordinate choices more rarely and becoming more likely to probe broader possible word extensions.

These findings counter proposals that children, especially at younger ages, are generally more exploratory than adults (Gopnik et al., 2017). Even when presented with an opportunity to explore broadly, in a context when exploration could be useful for uncovering new information about a word's meaning, children often opted for more narrow, confirmatory choices. Our findings show that children's tendency to explore depends on the specific task at hand, consistent with other findings demonstrating contextual and task-based variation in children's exploratory behavior (e.g., Ruggeri et al., 2021).

One question raised by the current findings is why children did not vary their information-seeking approach depending on the training condition. Children clearly showed differential patterns of generalization at test, demonstrating that they were highly sensitive to the training manipulation in how they were considering words' extensions. Why, then, were children not similarly sensitive to differences in the training composition during the Sampling Phase? One possible explanation is that children had a bias to avoid negative feedback and hence were averse to selecting items that had a higher likelihood of generating negative evidence. In general, children's most common selection in every training condition was the most conservative choice: choosing the exemplar that matched the subordinate item the closest. A related possibility is that children were drawn to the item that was most similar to the training exemplars. Regardless of the explanation, our findings suggest that, when given the option to explore novel word meanings, children sometimes remain quite conservative in their information-seeking strategies.

One limitation of the current work is that children's opportunity to explore word meanings was highly constrained, with a single sampling trial for each word and a limited set of options. Future work could probe children's sampling strategies in more depth by allowing participants to sample multiple items from a larger array of options. This not only would allow for a more realistic experiment design (as children often have the opportunity to repeat questions or investigative tests), but would also allow us to ask whether the sampling horizon affects children's information-seeking during word learning (Wilson et al., 2014). Another limitation of the present study is that it maps novel words onto pre-existing, familiar categories. This design allows learners to quickly infer the underlying hierarchical category structure, but it may also artificially make learners explore more conservatively because they have strong expectations about how words will map to categories. Moreover, children's knowledge about these familiar categories is likely increasing with age, complicating interpretations of developmental changes in information-seeking. Future work could tackle this problem by studying sampling strategies when learners must infer novel word extensions without strong priors about how to carve up the category space.

In sum, we find that both children and adults make more confirmatory than exploratory choices when seeking information in support of generalizing novel word meanings, though children's tendency to explore broader word extensions increases gradually with age. Studying how children seek new information across a variety of learning contexts can help us understand how and when active learning promotes language learning.

Acknowledgements

M.Z. was supported by an individual postdoctoral fellowship grant from the Eunice Kennedy Shriver National Institute of Child Health & Human Development of the National Institutes of Health under Award Number F32HD110174. We would like to thank Molly Lewis and Michael C. Frank for making a portion of the stimuli used in this study publicly available.

References

Blanco, N. J., & Sloutsky, V. M. (2021). Systematic exploration and uncertainty dominate young children's choices. *Developmental Science*, *24*(2), e13026.

Coenen, A., Nelson, J.D., & Gureckis, T. M. (2019). Asking the right questions about the psychology of human inquiry: Nine open challenges. *Psychonomic Bulletin & Review, 26*(5), 1548-1587.

Croissant, Y. (2020). Estimation of random utility models in R: The mlogit Package. *Journal of Statistical Software, 95*(11), 1–41.

de Eccher, M., Mundry, R, & Mani, N. (2024) Children's subjective uncertainty-driven sampling behaviour. *Royal Society Open Science, 11*, 231283.

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12.

Gelderloos, L., Mahmoudi Kamelabad, A., & Alishahi, A. (2020). Active word learning through self-supervision. In S. Denison, M. Mack, Y. Xu & B. C. Armstrong (Eds.), *Proceedings for the 42nd Annual Meeting of the Cognitive Science Society* (pp. 1050–1056). Cognitive Science Society.

Gopnik, A. (2020). Childhood as a solution to explore-exploit tensions. *Philosophical Transactions of the Royal Society B, 375*, 20190502.

Gopnik, A., O'Grady, S., Lucas, C. G., Griffiths, T. L., Wente, A., Bridgers, S., Aboody, R., Fung, H., & Dahl, R. E. (2017). Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. *Proceedings of the National Academy of Sciences of the United States of America, 114*(30), 7892–7899.

Hidaka, S., Torii, T., & Kachergis, G. (2017). Quantifying the impact of active choice in word learning. In G. Gunzelmann, A. Howes, T. Tenbrink & E. Davelaar (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (pp. 519–525). Cognitive Science Society.

Kachergis, G., Yu, C., & Shiffrin, R. M. (2013). Actively learning object names across ambiguous situations. *Topics in Cognitive Science*, *5*(1), 200–213.

Lewis, M. L., & Frank, M. C. (2018). Still suspicious: The suspicious-coincidence effect revisited. *Psychological Science*, *29*(12), 2039–2047.

Liquin, E. G., & Gopnik, A. (2022). Children are more exploratory and learn more than adults in an approach-avoid task. *Cognition*, *218*, 104940.

Lucas, C. G., Bridgers, S., Griffiths, T. L., Gopnik, A. (2014) When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition, 131*, 284–299.

R Development Core Team. (2023). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing.

Ruggeri, A., Walker, C. M., Lombrozo, T., & Gopnik, A. (2021). How to help young children ask better questions? *Frontiers in Psychology*, *11,* 586819.

Settles, B. (2012). Active learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning, 6*(1), 1-114.

Sumner, E., Li, A. X., Perfors, A., Hayes, B., Navarro, D., & Sarnecka, B. W. (2019). The exploration advantage: Children's instinct to explore allows them to find information that adults miss. *PsyArXiv*.

Spencer, J. P., Perone, S., Smith, L. B., & Samuelson, L. K. (2011). Learning words in space and time: Probing the mechanisms behind the suspicious-coincidence effect. *Psychological Science*, *22*(8), 1049–1057.

Wang, F. H., & Trueswell, J. (2022). Being suspicious of suspicious coincidences: The case of learning subordinate word meanings. *Cognition, 224*, 105028.

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General, 143*(6), 2074–2081.

Wojcik, E., Zettersten, M., & Benitez, V. (2022). The map trap: Why and how word learning research should move beyond mapping. *WIREs Cognitive Science, 13*(4), e1596.

Xu, F., & Tenenbaum, J. B. (2007). Word learning as Bayesian inference. *Psychological Review*, *114*(2), 245–272.

Zettersten, M., Cutler, M., & Lew-Williams, C. (2023). Active information-seeking in support of learning extensions of novel words. *Proceedings of the 45th Annual Meeting of the Cognitive Science Society*.

Zettersten, M., & Saffran, J. R. (2021). Sampling to learn words: Adults and children sample words that reduce referential ambiguity. *Developmental Science*, *24*(3), e13064.