Modality-Specific Mental Imagery Abilities are Unrelated to Modality-Specific Category Learning

Casey L. Roark (casey.roark@unh.edu)
Hooman Amiri (hooman.amiri@unh.edu)
Stephanie Dodson (stephanie.dodson@unh.edu)
Hedieh Nejatmand Malari (hedieh.nejatmandmalari@unh.edu)
Department of Psychology, 15 Academic Way
Durham, NH 03824 USA

Abstract

Category learning is an important ability that underlies complex cognitive processes such as object recognition and speech perception. Categories are ubiquitous across modalities and people differ greatly in their ability to learn novel categories. Here, we addressed a *modality-specific* cognitive individual difference that may relate to category learning – mental imagery. We examined how individual differences in self-reported auditory and visual mental imagery abilities related to individual differences in auditory and visual category learning. Overall, according to Bayesian analyses, there was anecdotal to moderate evidence for the null hypothesis that differences in self-reported modality-specific mental imagery are unrelated to differences in modality-specific category learning. These results have implications for theories of category learning and raise questions regarding the functions of mental imagery in cognitive processes such as categorization and learning.

Keywords: mental imagery; category learning; modality; audition; vision; individual differences

Introduction

Categories are ubiquitous across modalities and are important for developing a coherent and efficient view of the external world. While much of category learning in the lab has focused on the visual modality (e.g., Ashby & Maddox, 2011; Minda et al., 2024; Newell et al., 2011), recent research has indicated that the sensory *modality* of category information affects learning processes and outcomes (Newell et al., 2023; Roark et al., 2021, 2023, 2024). These results indicate that at least some processes supporting category learning may be modality specific. However, it is still unclear to what extent modality-specific mental processes support modality-specific category learning. In the current study, we examine how individual differences in the ability to mentally imagine and manipulate information within a specific sensory modality relates to modality-specific category learning.

Mental Imagery across Modalities

Mental imagery is the ability of the mind to create or represent an external sensory experience internally. Visual mental imagery specifically is the ability to create an image or scene in one's mind and in some way "see" it. While mental 'imagery' typically focuses on the visual modality, mental imagery can involve different sensory modalities. For example, auditory mental imagery involves the ability to mentally imagine linguistic (e.g., voices) or non-linguistic (e.g., music) information. Both auditory and visual mental imagery are thought to vary on similar dimensions such as vividness and controllability (Halpern, 1988, 2015; Marks, 1973). Mental imagery abilities in both modalities can be assessed with **behavioral** (Cui et al., 2007; Halpern et al., 2004; Hubbard, 2010; Kosslyn et al., 1978; Shepard & Metzler, 1971), **neural** (Cui et al., 2007; Ganis et al., 2004; Halpern et al., 2004; Hassabis et al., 2007; Pearson, 2019; Yomogida et al., 2004), and **self-report measures** (Halpern, 2015; Sulfaro et al., 2024) .

Importantly, regardless of modality, there are significant individual differences in mental imagery abilities (Halpern, 2015; Hubbard, 2010; Zeman, 2024). While the vast majority of the population is thought to have some mental imagery abilities, some individuals are thought to demonstrate extremely vivid and detailed mental images (vision: hyperphantasia; audition: hyperauralia), whereas others may not be able to voluntarily produce mental images at all (vision: aphantasia; audition: anauralia; Hubbard, 2010; Zeman, 2024; Zeman et al., 2015) .

Auditory and visual mental imagery abilities are supported by a combination of shared and modality-specific mechanisms. The modality of mental imagery relates to neural activity in modality-specific perceptual cortices (Bunzeck et al., 2005; Hubbard, 2018; Huijbers et al., 2011; Regev et al., 2021; Van Caenegem et al., 2024; Zvyagintsev et al., 2013) . There are also shared cognitive and neural mechanisms in mental imagery across modalities in the specific cognitive ability used for the particular task (e.g., attention, episodic memory, decision-making; Daselaar et al., 2010; Huijbers et al., 2011; Van Caenegem et al., 2024).

Potential modality-specific mechanisms of mental imagery may result in abilities that are generally unrelated across the senses or situations where an individual may have a modality bias in mental imagery such that they have especially vivid mental imagery in one modality, but poor or average mental imagery in the other modality (Zvyagintsev et al., 2013). Conversely, potential modality-general mechanisms of mental imagery may result in a correlation between auditory and visual mental imagery abilities (Hinwar & Lambert, 2021; Hubbard, 2018). A recent study demonstrated that individuals differ not only within the individual ability to imagine perceptual stimuli across modalities, but also in their *relative* abilities for different modalities (Sulfaro et al., 2024).

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

These individual differences in auditory and visual mental imagery and the relative abilities across modalities may have direct consequences for cognitive processes that involve processing information in a specific modality.

Mental Imagery and Cognitive Processes

While mental imagery abilities across different modalities vary across individuals, it is unclear how these individual differences may affect other aspects of cognition or behavior. Does better mental imagery across different sensory abilities allow perceivers to use and access information in useful ways during modality-specific tasks? The relationship between mental imagery and one particular cognitive process has been the topic of prior research – working memory.

Visual mental imagery is related to visuospatial working memory abilities (Albers et al., 2013; Fiore et al., 2011; Jacobs et al., 2018; Keogh & Pearson, 2011, 2014). Individuals with better visual mental imagery abilities also tend to have better visuospatial working memory abilities. Having access to accurate and detailed visual mental images may be particularly beneficial in working memory tasks that require mentally moving or orienting spatial information (Shepard & Metzler, 1971). Auditory mental imagery abilities are also related to working memory (Baddeley & Andrade, 2000; Gelding et al., 2021). Having access to accurate and detailed auditory mental images may support rehearsal of that auditory information in phonological working memory. Together, these results suggest that both auditory and visual mental imagery relate to working memory performance.

However, prior research has demonstrated mental imagery abilities may not relate to better performance in cognitive tasks, even in tasks that actively involve mental imagery (Pounder et al., 2022). For example, individuals with aphantasia may find other strategies to support visual working memory – using a different underlying cognitive process but yielding the same overall behavior (Keogh et al., 2021). This leaves open the possibility that individual differences in mental imagery may not relate to differences in task performance.

Mental Imagery in Category Learning

One modality-specific task that involves working memory is category learning (e.g., Lewandowsky, 2011; Roark & Chandrasekaran, 2023). In the current study, we examine the potential relationship between self-reported auditory and visual mental imagery abilities and task-derived auditory and visual category learning abilities. Recent research has demonstrated important modality-specific differences in category learning even when people learn similar and comparable categories across auditory and visual modalities (Roark et al., 2021, 2022, 2023, 2024) . As such, category learning provides an excellent testbed to understand whether modality-specific mental imagery abilities may enable better modality-specific learning.

Why may category learning relate to mental imagery abilities? To determine the category identity of a stimulus, learners must compare the stimulus information to either previously encountered stimuli or an abstract representation of the category. This process may involve mental imagery through a perceptual similarity matching process. That is, a learner may actively conjure a mental image representing a category to compare with the stimulus being categorized to determine which category it belongs to.

If perceptual category learning involves this comparison of mental images of perceptual category information (i.e., previously encountered stimuli or abstract representations) with the current stimulus, then it is reasonable to expect that modality-specific mental imagery would relate to modality-specific category learning. Specifically, better auditory mental imagery may relate to better auditory category learning and better visual mental imagery may relate to better visual category learning. Just as with studies of mental imagery, there are large individual differences in perceptual category learning abilities across modalities (Lewandowsky, 2011; Lloyd et al., 2019; Roark et al., 2021), but it is not yet clear whether individuals have a bias for learning in one modality over another.

In the current study, we tested two related questions (1) Does better self-reported modality-specific mental imagery (e.g., auditory, visual) relate to better modality-specific category learning (e.g., auditory, visual)? (2) Does a relative bias in self-reported mental imagery abilities for one modality over the other relate to a relative bias for one modality over the other in learning?

Methods

We examined the relationship between self-report measures of auditory and visual mental imagery abilities (Sulfaro et al., 2024) and task-based measures of auditory and visual category learning abilities. We also tested auditory and visual category generalization abilities with post-tests in each modality using novel stimuli not encountered during training.

Participants

Participants were 90 students at the University of New Hampshire, ages 18-24 (8 M, 82 F), who participated for partial course credit. One additional participant was run but excluded due to failure to pass an *a priori* accuracy criterion on attention catch trials (see below). The study was conducted in person using the Gorilla Experiment Builder (Anwyl-Irvine et al., 2020). Procedures were approved by the Institutional Review Board at University of New Hampshire.

Stimuli

Participants separately learned two unidimensional rule-based auditory and visual categories (Figure 1A). Auditory stimuli were static spectrotemporal ripples (Figure 1B: spectrogram) varying in spectral modulation rate (cycles per octave) and temporal modulation rate (Hz). The category-relevant dimension was temporal modulation rate, with stimuli slower than 8.33 Hz belonging to category B and stimuli faster than 8.33 Hz belonging to category A. Visual stimuli were Gabor patches (Figure 1C) varying in spatial frequency (cycles per pixel) and orientation (degrees). The category-relevant dimension was spatial frequency, with stimuli with a spatial frequency less than 0.055 cycles/pixel belonging to category B and stimuli greater than 0.055 cycles/pixel belonging to category A.

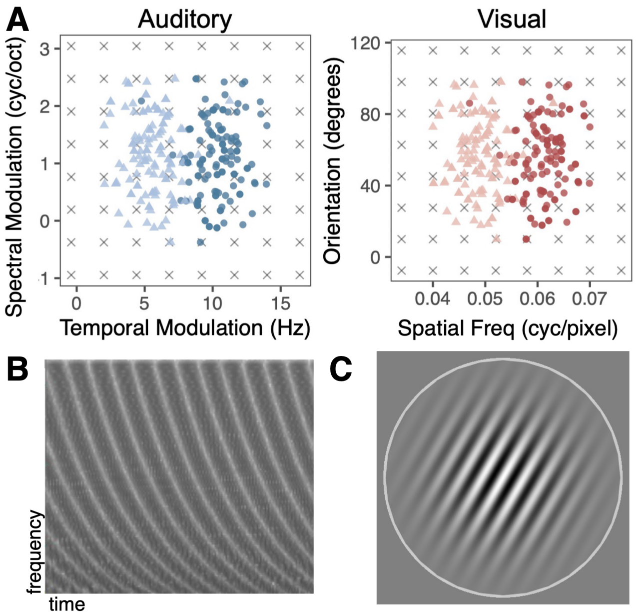

Figure 1: Category distributions and example stimuli.

Category training distributions were first generated using a bivariate Gaussian distribution in a normalized 0 to 1 space. The dimensions were then transformed to auditory or visual dimensions such that the categories have the same underlying category distributions but are in different modalities. Category test distributions were generated using an 8 x 8 equal grid in the normalized space and then transforming to modality-specific spaces. These stimuli serve as a novel generalization test as participants never encountered these stimuli during training and the test stimuli fall both within and outside the range of trained stimuli. As such, the test performance allows us to examine how participants' categorization knowledge generalized beyond what they encountered during training.

Procedure

Within a single session, participants completed consent procedures, a demographics questionnaire, and then learned the auditory and visual categories in counterbalanced order, followed by the auditory, visual, and modality rankings items from mental imagery questionnaire (Sulfaro et al., 2024). The order of auditory and visual questionnaire item sets were counterbalanced across participants.

Category Learning and Generalization We operationalized auditory and visual *learning* and *generalization* abilities based on the final block and test accuracy within the specific tasks. We computed measures of modality bias during perceptual category learning and generalization by subtracting each learner's auditory accuracy from their visual accuracy. We expected that participants would demonstrate substantial variability in performance in both modalities based on prior work (Roark et al., 2021).

Participants completed four training blocks in each modality with 25 trials each for a total of 100 trials per modality. At the end of each modality task, they completed a generalization test of novel stimuli with 64 trials. Participants were not told the categorization criteria prior to learning and, instead, needed to discover the optimal rule through corrective feedback.

Training trials were similar across modalities. Participants heard the 1 sec auditory stimulus or saw the visual stimulus for 1.5 sec. We increased the duration of stimulus presentation slightly in the visual modality to account for relatively poorer temporal processing speed in the visual relative to the auditory modality (Kubovy, 1988; O'Connor & Hermelin, 1972). After stimulus presentation, participants saw the prompt "Which category? 1 or 2" and had an unlimited amount of time to respond by pressing the corresponding button on the keyboard. They then received immediate feedback in the form of "Correct" or "Incorrect" for 1 sec followed by a 1 sec inter-trial interval. Generalization test trials were very similar to training trials, but participants did not receive any feedback.

We included five attention check trials in each learning block to ensure that participants were maintaining consistent attention to the task. The attention check required participants to view a typically white fixation cross turn red and press the spacebar as soon as possible. They were required to press the spacebar within 2 sec or they received feedback for 1.5 sec – "Respond faster! Press the spacebar." We decided *a priori* to not include participants who failed to respond on more than 15% of attention check trials in either task (i.e., hit rate of >85% in *both* auditory and visual tasks). One participant was excluded due to this criterion. The remaining 90 participants responded very accurately (M = 98%) and quickly (M = 676 ms) to the attention check trials.

Mental Imagery Questionnaire We were particularly interested in understanding variation in self-reported auditory and visual mental imagery abilities across individuals and the relative bias of modality-specific imagery abilities within an individual. As such, we examined participants' responses to subscales from Sulfaro et al. (2024) assessing the quality of auditory and visual mental imagery as well as those directly comparing auditory and visual mental imagery abilities.

To assess the quality of auditory and visual mental imagery, participants were first asked to close their eyes and think about either the *appearance* (visual) or *sound* (auditory) of the letter 'O'. They then answered questions about their experiences to assess their mental imagery ability and quality of their mental images. We operationalized auditory (Q5 in Sulfaro et al., 2024) and visual (Q4) imagery abilities by creating composite scores that involved the measures of **intentionality** (Q.4.5/Q5.5; "Did your thought automatically seem to involve seeing[hearing] what you were thinking about in some way"), **consistency** (Q4.9/Q5.9; "Was your thought content visible[audible] in some way consistently during the period you were thinking about it?"), **proportion** (Q4.11/Q5.11; "What proportion of your thinking period involved seeing[hearing] your thought content in some way?"), **latency** (Q4.16/Q5.16; "Approximately how many seconds does it take before your thought content becomes visible[audible] in some way?"; slider response from 0 to 10 seconds in increments of 0.1) and **duration** (Q4.18/Q5.18; "Approximately how many seconds can you definitely and continuously keep your thought content visible[audible] in some way?"; slider response as described above).

To obtain a normalized score of each measure to compare self-reported abilities across individuals and to ensure each measure contributed equally to the composite, we converted responses to values ranging from 0 to 1, where 0 represented the lowest level of the ability or absence of mental imagery (e.g., least automatic, least consistent, lowest proportion, slowest latency, shortest duration) and 1 represented the highest level of the ability (e.g., most automatic, most consistent, highest proportion, fastest latency, longest duration). For each modality, the mean was computed across measures to obtain modality-specific ability scores.

Finally, we computed a single measure of modality bias in self-reported mental imagery by creating a composite score of the auditory-visual modality comparison questions (Q6.9 in Sulfaro et al., 2024; which is easiest, easiest to initiate, quickest to initiate, most sustainable, most consistent, least movement needed, most location control, most automatic, clearest, most vivid, and realest). For each comparison question, participants were given a score of 1 (visual), 0 (auditory), 0.5 (no bias), or N/A (do not experience mental imagery). The composite score was then an average across these scores. NA scores were ignored in the computation of the composite score. As such, the scores could range from 0 (complete auditory bias) to 1 (complete visual bias). Values near 0.5 would represent no bias to either modality.

Results

As a reminder, we tested two questions (1) Does better self-reported modality-specific mental imagery (e.g., auditory, visual) relate to better modality-specific learning? (2) Does a relative bias in self-reported mental imagery abilities for one modality over the other relate to a relative bias for one modality over the other in learning? We first summarize overall performance and variability in both learning and mental imagery measures and then pursue answers to the questions. Analyses were conducted in R using the *rstatix* (Kassambara, 2023) and *BayesFactor* (Morey & Rouder, 2024) packages with data plotted using *ggplot2* (Wickham, 2016) and *ggthemes* (Arnold, 2024) packages. Bayes factors (BFs) are reported for all tests and are interpreted based on recommendations from Lee and Wagenmakers (2014) with the interpretation of evidence in favor of the null hypothesis with BFs < 1 and evidence in favor of the alternative hypothesis with BFs > 1.

Category Learning

Participants learned the categories quite well (Figure 2A), with extreme evidence that performance was better than chance (50%) in the final block and generalization test for both auditory (final block: $M = 70\%$, $t(89) = 14.6$, $p < .0001$, $d = 1.54$, $BF_{10} = 2.01 \times 10^{22}$; test: $M = 75\%$, $t(89) = 15.2$, $p < .0001$, $d = 1.60$, $BF_{10} = 1.98 \times 10^{23}$) and visual categories (final block: $M = 75\%$, $t(89) = 14.6$, $p < .0001$, $d = 1.54$, $BF_{10} = 1.41 \times 10^{22}$, generalization: $M = 79\%$, $t(89) = 16.1$, $p < .0001$, $d = 1.69$, $BF_{10} = 6.96 \times 10^{24}$).

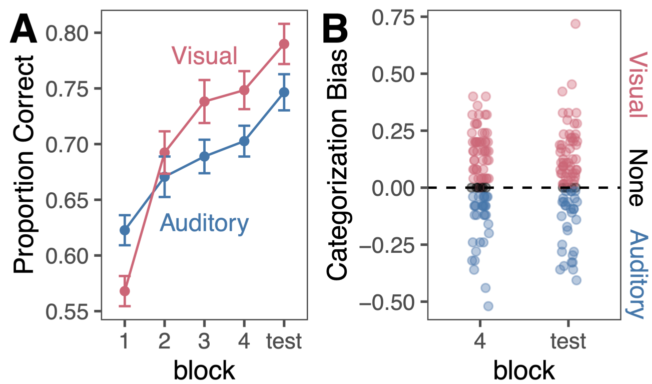

Figure 2: Category learning and generalization test performance and categorization modality bias.

We calculated each individual's **modality bias in categorization** as the difference in their performance in the visual and auditory modalities (Figure 2B). For this measure, negative numbers represent a bias towards the auditory modality, positive numbers represent a bias toward the visual modality, and 0 represents no bias. On average, there was anecdotal evidence for a very small visual bias in the final block of training ($M = 0.046$, $SD = 0.19$; one-sample t-test vs. 0: $t(89) = 2.31$, $p = .023$, $d = 0.24$, $BF_{10} = 1.43$) and anecdotal evidence for no bias in the generalization test ($M = 0.043$, $SD = 0.19$; one-sample t-test vs. 0: $t(89) = 2.12$, $p = .037$, $d = 0.22$, $BF_{10} = 0.97$). Importantly, participants demonstrated substantial variability in modality bias during categorization.

Mental Imagery

There was significant variability in self-reported **auditory and visual mental imagery abilities** as measured by the composite scores (Figure 3A). There was anecdotal evidence for no differences in self-reported mental imagery abilities across auditory ($M = 0.70$; $SD = 0.26$) and visual modalities ($M = 0.62$; $SD = 0.22$; ($t(89) = 2.06$, $p = .042$, $d = 0.22$, $BF_{10} = 0.88$). Importantly, participants ranged across the spectrum from low ability (~ 0) to high ability (~1) for both modalities.

We calculated each individual's **modality bias in mental imagery** as the composite average across all comparison (modality ranking) questions (Figure 3B). For this measure, 0 represents the strongest bias towards the auditory modality, 1 represents the strongest bias towards the visual modality, and 0.5 represents no bias. On average, there was moderate evidence for no bias towards either the auditory or visual modalities ($M = 0.49$, $SD = 0.35$; one-sample t-test vs. 0: $t(89) = -0.15$, $p = .88$, $d = -0.016$, $BF_{10} = 0.12$). Importantly, participants demonstrated substantial variability in the composite modality bias score.

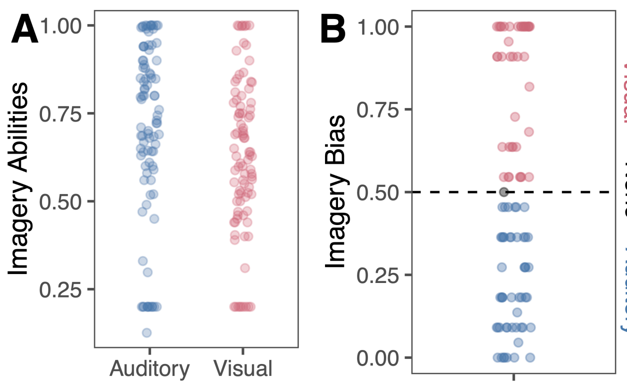

Figure 3: Mental imagery abilities and modality biases.

There was a strong positive correlation between the difference between separately computed auditory and visual modality abilities and the modality bias ranking score ($r(89) = 0.58$, $p < .0001$, $BF_{10} = 7.01\text{x}10^6$), indicating that participants self-report of their abilities strongly relates to their self-reported ranking of the abilities across modalities.

Relation between Modality-Specific Learning and Self-Reported Mental Imagery Abilities

We examined the correlation between self-reported *auditory* imagery and *auditory* category learning and generalization (Figure 4A). Overall, there was moderate evidence that auditory imagery abilities were unrelated to auditory category learning abilities during training ($r(89) = 0.011$, $p = .92$, $BF_{10} = 0.24$). However, there was anecdotal evidence that better auditory imagery abilities related to better performance in the auditory generalization test ($r(89) = 0.22$, $p = .033$, $BF_{10} = 2.05$). The relationship between auditory mental imagery abilities and category generalization was stronger than the relationship with category learning ($p = .032$, 95% CI: -0.41, -0.018; compared using *cocor* package in R; Diedenhofen & Musch, 2015).

We examined the correlation between self-reported *visual* imagery and *visual* category learning and generalization (Figure 4B). Overall, there was anecdotal evidence that visual mental imagery abilities were unrelated to visual category learning abilities ($r(89) = 0.07$, $p = .66$, $BF_{10} = 0.30$) and moderate evidence that visual mental imagery abilities were unrelated to performance in the visual generalization test ($r(89) = -0.0078$, $p = .94$, $BF_{10} = 0.24$). There were no significant differences in the strengths of these correlations ($p = .35$, 95% CI: -0.085, 0.24).

Modality Biases in Learning and Mental Imagery

Finally, we examined whether there was a correlation between modality biases in during category learning and modality biases in self-reported mental imagery abilities (Figure 4C). There was moderate evidence that modality biases in mental imagery were unrelated to modality biases in categorization during either training ($r(89) = -0.092$, $p = .39$, $BF_{10} = 0.34$) or generalization ($r(89) = -0.0094$, $p = .93$, $BF_{10} = 0.24$). Reporting relatively better mental imagery in one modality did not relate to better category learning or generalization in that modality.

Discussion

In this study, we examined whether there was a relationship between self-reported auditory and visual mental imagery abilities and task-based auditory and visual category learning performance. Overall, we found that while there was substantial variability in both self-reported modality-specific mental imagery abilities and modality-specific category learning abilities, there was *anecdotal to moderate* evidence that these abilities were *unrelated.* These results have important implications for understanding how individual differences in self-report measures of modality-specific mental imagery abilities may relate to more complex modality-specific behaviors such as categorization.

Modality Biases in Mental Imagery

Overall, as in Sulfaro et al. (2024), we found substantial individual differences in self-reported mental imagery abilities across auditory and visual modalities. We developed

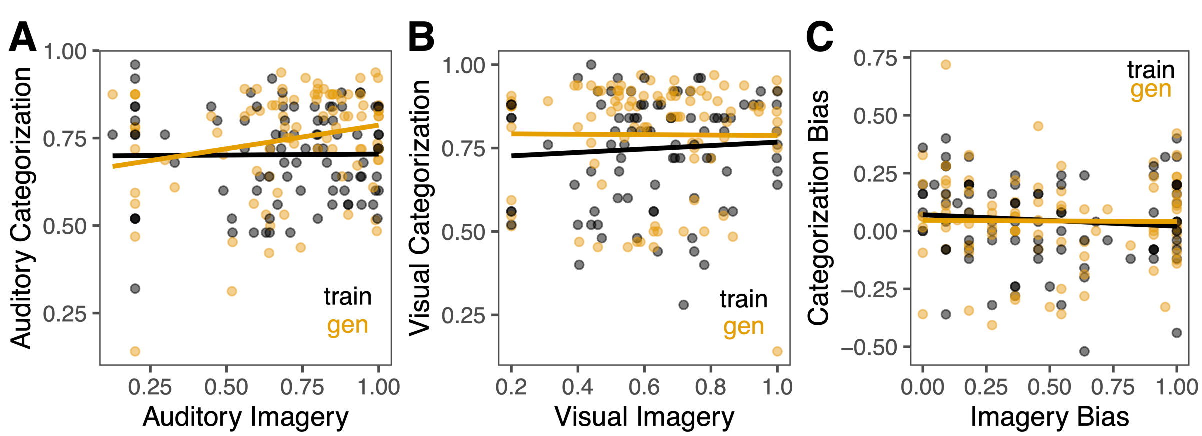

Figure 4: Correlation between mental imagery and categorization abilities.

novel composite measures of auditory and visual mental imagery abilities and modality rankings to relate to category learning performance. Because we focused on single composite measure, it is not yet clear whether specific features or components of mental imagery (e.g., vividness or precision) may relate to category learning. Future research should clarify the potential role of specific features of mental imagery that may relate to performance by examining individual differences in specific measures. Determining these features would require a larger sample but would reveal more precise possible relationships between features of mental imagery and learning.

It is important to emphasize that the current study focused on self-report measures of mental imagery abilities, and it is not yet clear whether these subjective measures relate to task-based measures of mental imagery. Future work should examine the potential relationship between subjective and objective measures of mental imagery abilities across modalities. Finally, the questionnaire from Sulfaro et al. (2024) that we adopted here required participants to mentally imagine a relatively simple stimulus – the letter 'O'. To fully characterize the complexity of mental imagery, it will be important for studies of subjective report of one's mental imagery experiences to examine potential differences between relatively simple stimuli and more complex stimuli. It will also be important to clarify whether there are any differences between mental imagery of linguistic information (i.e., the letter 'O') and non-linguistic information.

Modality Biases in Category Learning

Regarding overall learning performance in similar auditory and visual category learning tasks, we observed that even in these highly comparable contexts, some learners performed much better in one modality relative to the other. In the most extreme cases, one visual-biased participant performed 40% better in the visual modality than the auditory modality whereas one auditory-biased participant performed 44% better in the auditory modality than the visual modality. These results indicate the potential importance of considering factors that would lead to these modality-specific biases.

One potential factor determining modality bias could be perceptual expertise such as musicianship or experience in visual arts. There is some indirect evidence to support this possibility. For example, prior work has demonstrated that auditory (but not visual) mental imagery is stronger in musicians compared to non-musicians (Talamini et al., 2023) and musicians have at least some advantages over non-musicians in auditory (but not visual) category learning problems (Roark et al., 2022). Visual artists have superior visual mental imagery abilities compared to non-artists (Bhattacharya & Petsche, 2005; Palmiero et al., 2015) but these groups have not yet been compared in auditory or visual category learning. In the current study, we assessed music experience as a part of the demographics questionnaire, but not visual art experience. An immediate next step will be examining whether modality biases in categorization relate to self-reported music abilities in this sample. Identifying why an individual learner may have superior auditory or visual learning abilities is an important avenue for future research. The current results suggest that self-reported mental imagery may not be a key factor in determining whether one learns better in a particular sensory modality.

Another potential factor to consider is whether differences in performance across modalities in the current study could have been due to the order in which participants performed category learning tasks across modalities. We ruled out the possibility that participants performed better in whichever modality they compared first (or second) with a supplementary analysis that found that modality bias in categorization did not depend on the order of modalities in category learning ($t(85.5) = 0.84$, $p = .40$, $d = 0.18$).

Mental Imagery in Category Learning

Together, these results provide an initial test of the potential relationship between individual differences in modality-specific mental imagery and category learning. These results are in line with prior work in other domains that found no relationship between *measured* mental imagery abilities and overt task performance (Keogh et al., 2021; Pounder et al., 2022). This may suggest that modality-specific mental imagery is not involved in perceptual category learning or that one's belief in their mental imagery abilities (from self-report measures) is unrelated to their ability to actually use mental imagery in a specific context like category learning.

Alternatively, it is also possible that the similarities in overall performance may be supported by distinct cognitive processes in individuals with high or low mental imagery abilities (Keogh et al., 2021). This would suggest that even without a clear relationship between mental imagery abilities and some overt behavior, mental imagery could still play a role in the task. In this experiment, we did not overtly encourage participants to use mental imagery or ask whether they used any mental imagery specific strategies to support their performance in either the training or the test. In future work, the involvement of mental imagery should be examined by directly probing participants about their use of mental imagery or otherwise encouraging learners to use mental imagery during learning.

Overall, these results demonstrate that modality-specific mental imagery abilities, assessed through self-report measures, are generally unrelated to the ability to use modality-specific information to learn novel perceptual categories. This initial investigation leaves many open questions for future research that should clarify the involvement of mental imagery in the *processes* of category learning, especially with regard to auditory generalization performance without feedback. These results suggest that self-reported measures of mental imagery may not reflect how people may actually use these processes during a complex task like category learning. The presence of modality biases in both category learning and mental imagery demonstrate the necessity of considering the role of modality in cognition more broadly.

References

Albers, A. M., Kok, P., Toni, I., Dijkerman, H. C., & de Lange, F. P. (2013). Shared representations for working memory and mental imagery in early visual cortex. *Current Biology*, *23*(15), 1427–1431.

Anwyl-Irvine, A. L., Massonnié, J., Flitton, A., Kirkham, N., & Evershed, J. K. (2020). Gorilla in our midst: An online behavioral experiment builder. *Behavior Research Methods*, *52*(1), 388–407.

Arnold, J. B. (2024). *ggthemes: Extra Themes, Scales and Geoms for "ggplot2"* (Version 5.1.0) [R package].

Ashby, F. G., & Maddox, W. T. (2011). Human category learning 2.0. *Annals of the New York Academy of Sciences*, *1224*, 147–161.

Baddeley, A. D., & Andrade, J. (2000). Working memory and the vividness of imagery. *Journal of Experimental Psychology: General*, *129*(1), 126–145.

Bhattacharya, J., & Petsche, H. (2005). Drawing on mind's canvas: Differences in cortical integration patterns between artists and non-artists. *Human Brain Mapping*, *26*(1), 1–14.

Bunzeck, N., Wuestenberg, T., Lutz, K., Heinze, H.-J., & Jancke, L. (2005). Scanning silence: Mental imagery of complex sounds. *NeuroImage*, *26*(4), 1119–1127.

Cui, X., Jeter, C. B., Yang, D., Montague, P. R., & Eagleman, D. M. (2007). Vividness of mental imagery: Individual variability can be measured objectively. *Vision Research*, *47*(4), 474–478.

Daselaar, S. M., Porat, Y., Huijbers, W., & Pennartz, C. M. A. (2010). Modality-specific and modality-independent components of the human imagery system. *NeuroImage*, *52*(2), 677–685.

Diedenhofen, B., & Musch, J. (2015). cocor: A Comprehensive solution for the statistical comparison of correlations. *PLOS ONE*, *10*(4), e0121945.

Fiore, F., Borella, E., Mammarella, I. C., & Cornoldi, C. (2011). Mental imagery in a visuospatial working memory task and modulation of activation. *Journal of Cognitive Psychology*, *23*(1), 52–59.

Ganis, G., Thompson, W. L., & Kosslyn, S. M. (2004). Brain areas underlying visual mental imagery and visual perception: An fMRI study. *Cognitive Brain Research*, *20*(2), 226–241.

Gelding, R. W., Harrison, P. M. C., Silas, S., Johnson, B. W., Thompson, W. F., & Müllensiefen, D. (2021). An efficient and adaptive test of auditory mental imagery. *Psychological Research*, *85*(3), 1201–1220.

Halpern, A. R. (1988). Mental scanning in auditory imagery for songs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *14*(3), 434–443.

Halpern, A. R. (2015). Differences in auditory imagery self-report predict neural and behavioral outcomes. *Psychomusicology: Music, Mind, and Brain*, *25*(1), 37–47.

Halpern, A. R., Zatorre, R. J., Bouffard, M., & Johnson, J. A. (2004). Behavioral and neural correlates of perceived and imagined musical timbre. *Neuropsychologia*, *42*(9), 1281–1292.

Hassabis, D., Kumaran, D., & Maguire, E. A. (2007). Using imagination to understand the neural basis of episodic memory. *The Journal of Neuroscience*, *27*(52), 14365–14374.

Hinwar, R. P., & Lambert, A. J. (2021). Anauralia: the silent mind and its association with aphantasia. *Frontiers in Psychology*, *12*, 744213.

Hubbard, T. L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin*, *136*(2), 302–329.

Hubbard, T. L. (2018). Some methodological and conceptual considerations in studies of auditory imagery. *Auditory Perception & Cognition*, *1*(1–2), 6–41.

Huijbers, W., Pennartz, C. M. A., Rubin, D. C., & Daselaar, S. M. (2011). Imagery and retrieval of auditory and visual information: Neural correlates of successful and unsuccessful performance. *Neuropsychologia*, *49*(7), 1730–1740.

Jacobs, C., Schwarzkopf, D. S., & Silvanto, J. (2018). Visual working memory performance in aphantasia. *Cortex*, *105*, 61–73.

Kassambara, A. (2023). *rstatix: Pipe-Friendly Framework for Basic Statistical Tests* (Version 0.7.2) [R package].

Keogh, R., & Pearson, J. (2011). Mental imagery and visual working memory. *PLoS ONE*, *6*(12), e29221.

Keogh, R., & Pearson, J. (2014). The sensory strength of voluntary visual imagery predicts visual working memory capacity. *Journal of Vision*, *14*(12), 7–7.

Keogh, R., Wicken, M., & Pearson, J. (2021). Visual working memory in aphantasia: Retained accuracy and capacity with a different strategy. *Cortex*, *143*, 237–253.

Kosslyn, S. M., Ball, T. M., & Reiser, B. J. (1978). Visual images preserve metric spatial information: Evidence from studies of image scanning. *Journal of Experimental Psychology: Human Perception and Performance*, *4*(1), 47–60.

Kubovy, M. (1988). Should we resist the seductiveness of the space:time::vision:audition analogy? *Journal of Experimental Psychology: Human Perception and Performance*, *14*(2), 318–320.

Lee, M. D., & Wagenmakers, E. J. (2014). *Bayesian Cognitive Modeling: A practical course*. Cambridge University Press.

Lewandowsky, S. (2011). Working memory capacity and categorization: individual differences and modeling. *Journal of Experimental Psychology: Learning Memory and Cognition*, *37*(3), 720–738.

Lloyd, K., Sanborn, A., Leslie, D., & Lewandowsky, S. (2019). Why higher working memory capacity may help you learn: Sampling, search, and degrees of approximation. *Cognitive Science*, *43*(12), e12805.

Marks, D. F. (1973). Visual imagery differences in the recall of pictures. *British Journal of Psychology*, *64*(1), 17–24.

Minda, J. P., Roark, C. L., Kalra, P., & Cruz, A. (2024). Single and multiple systems in categorization and category learning. *Nature Reviews Psychology*, 1–16.

Morey, R., & Rouder, J. (2024). *BayesFactor: Computation of Bayes Factors for Common Designs* (Version 0.9.12-4.7) [R package].
Newell, B. R., Dunn, J. C., & Kalish, M. (2011). Systems of category learning fact or fantasy? *Psychology of Learning and Motivation - Advances in Research and Theory*, *54*, 167–215.
Newell, F. N., McKenna, E., Seveso, M. A., Devine, I., Alahmad, F., Hirst, R. J., & O'Dowd, A. (2023). Multisensory perception constrains the formation of object categories: A review of evidence from sensory-driven and predictive processes on categorical decisions. *Philosophical Transactions of the Royal Society B*, *378*(1886), 20220342.
O'Connor, N., & Hermelin, B. (1972). Seeing and hearing and space and space and time. *Perception & Psychophysics*, *11*(1), 46–48.
Palmiero, M., Nori, R., Aloisi, V., Ferrara, M., & Piccardi, L. (2015). Domain-specificity of creativity: a study on the relationship between visual creativity and visual mental imagery. *Frontiers in Psychology*, *6*, 1870.
Pearson, J. (2019). The human imagination: The cognitive neuroscience of visual mental imagery. *Nature Reviews Neuroscience*, *20*(10), 624–634.
Pounder, Z., Jacob, J., Evans, S., Loveday, C., Eardley, A. F., & Silvanto, J. (2022). Only minimal differences between individuals with congenital aphantasia and those with typical imagery on neuropsychological tasks that involve imagery. *Cortex*, *148*, 180–192.
Regev, M., Halpern, A. R., Owen, A. M., Patel, A. D., & Zatorre, R. J. (2021). Mapping specific mental content during musical imagery. *Cerebral Cortex*, *31*(8), 3622–3640.
Roark, C. L., Ben-Anat, Y., & Gabay, Y. (2024). Auditory and visual perceptual category learning in adults with ADHD: Implications for learning systems and domain-general factors. *PsyArXiv*.
Roark, C. L., & Chandrasekaran, B. (2023). Stable, flexible, common, and distinct behaviors support rule-based and information-integration category learning. *Npj Science of Learning*, *8*(1), 14.
Roark, C. L., Lescht, E., Wray, A. H., & Chandrasekaran, B. (2023). Auditory and visual category learning in children and adults. *Developmental Psychology*, *59*(5), 963–975.
Roark, C. L., Paulon, G., Sarkar, A., & Chandrasekaran, B. (2021). Comparing perceptual category learning across modalities in the same individuals. *Psychonomic Bulletin & Review*, *28*(3), 898–909.
Roark, C. L., Smayda, K. E., & Chandrasekaran, B. (2022). Auditory and visual category learning in musicians and nonmusicians. *Journal of Experimental Psychology: General*, *151*(3), 739–748.
Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science*, *171*(3972), 701–703.
Sulfaro, A. A., Robinson, A. K., & Carlson, T. A. (2024). Properties of imagined experience across visual, auditory, and other sensory modalities. *Consciousness and Cognition*, *117*(103598).
Talamini, F., Vigl, J., Doerr, E., Grassi, M., & Carretti, B. (2023). Auditory and visual mental imagery in musicians and non-musicians. *Musicae Scientiae*, *27*(2), 428–441.
Van Caenegem, E. E., Moreno-Verdú, M., Waltzing, B. M., Hamoline, G., McAteer, S. M., Frahm, L., & Hardwick, R. M. (2024). Multisensory approach in mental imagery: ALE meta-analyses comparing motor, visual and auditory imagery. *Neuroscience & Biobehavioral Reviews*, *167*, 105902.
Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis* (Version 3.3.5) [R package].
Yomogida, Y., Sugiura, M., Watanabe, J., Akitsuki, Y., Sassa, Y., Sato, T., Matsue, Y., & Kawashima, R. (2004). Mental Visual Synthesis is Originated in the Fronto-temporal Network of the Left Hemisphere. *Cerebral Cortex*, *14*(12), 1376–1383.
Zeman, A. (2024). Aphantasia and hyperphantasia: exploring imagery vividness extremes. *Trends in Cognitive Sciences*.
Zeman, A., Dewar, M., & Sala, S. D. (2015). Lives without imagery – Congenital aphantasia. *Cortex*, *73*, 378–380.
Zvyagintsev, M., Clemens, B., Chechko, N., Mathiak, K. A., Sack, A. T., & Mathiak, K. (2013). Brain networks underlying mental imagery of auditory and visual information. *European Journal of Neuroscience*, *37*(9), 1421–1434.