Vision Research 191 (2022) 107971

Contents lists available at ScienceDirect

Vision Research

journal homepage: www.elsevier.com/locate/visres

Dissociations between performance and visual fixations after subordinate- and basic-level training with novel objects

Mina Elhamiasl [a], Gabriella Silva [a], Andrea M. Cataldo [b], Hillary Hadley [b], Erik Arnold [b], James W. Tanaka [c], Tim Curran [d], Lisa S. Scott [a,*]

[a] Department of Psychology, University of Florida, United States
[b] Department of Psychological and Brain Sciences, University of Massachusetts Amherst, United States
[c] Department of Psychology, University of Victoria, Canada
[d] Department of Psychology & Neuroscience, University of Colorado Boulder, United States

ARTICLE INFO

Keywords:
Perceptual expertise
Subordinate-level training
Eye-tracking
Visual fixation strategies

ABSTRACT

Previous work suggests that subordinate-level object training improves exemplar-level perceptual discrimination over basic-level training. However, the extent to which visual fixation strategies and the use of visual features, such as color and spatial frequency (SF), change with improved discrimination was not previously known. In the current study, adults (n = 24) completed 6 days of training with 2 families of computer-generated novel objects. Participants were trained to identify one object family at the subordinate level and the other object family at the basic level. Before and after training, discrimination accuracy and visual fixations were measured for trained and untrained exemplars. To examine the impact of training on visual feature use, image color and SF were manipulated and tested before and after training. Discrimination accuracy increased for the object family trained at the subordinate-level, but not for the family trained at the basic level. This increase was seen for all image manipulations (color, SF) and generalized to untrained exemplars within the trained family. Both subordinate- and basic-level training increased average fixation duration and saccadic amplitude and decreased the number of total fixations. Collectively, these results suggest a dissociation between discrimination accuracy, indicative of recognition, and the associated pattern of changes present for visual fixations.

1. Introduction

Perceptual expertise is important for a wide range of skilled tasks spanning health and safety professions. For example, visual expertise is required for radiologists scanning for evidence of disease or injury, forensic scientists analyzing fingerprints or tool marks, TSA agents matching various forms of photo identification to faces, and military specialists searching for threats in the environment (for review see: Shen, Mack, & Palmeri, 2014; Scott, Tanaka, & Curran, 2010). Visual perceptual expertise is generally characterized by increased proficiency to perceptually discriminate, identify, and recognize exemplars within visual categories (e.g., cars, birds, medical images, satellite images, fingerprints, etc.) (Tanaka & Taylor, 1991; for review see: Scott, 2011; Scott et al., 2010).

Over the last three decades, researchers have examined expert recognition and perceptual discrimination in real-world bird experts (Johnson & Mervis, 1997; Tanaka & Taylor, 1991), dog show judges (Tanaka & Taylor, 1991), fingerprint analysts (Busey & Parada, 2010; Tangen, Thompson, & McCarthy, 2011), car experts (Gauthier, Skudlarski, Gore, & Anderson, 2000), and budgerigar judges (Campbell & Tanaka, 2018). Consistent and replicable behavioral and electrophysiological findings have shown that real-world experts demonstrate faster and more accurate subordinate-level recognition and more specialized event-related brain responses than novices (e.g., Gauthier & Tarr, 1997; Gauthier, Williams, Tarr, & Tanaka, 1998; Tanaka & Taylor, 1991; Tanaka & Curran, 2001; Gauthier, Curran, Curby, & Collins, 2003). Accordingly, "family-level" characteristics facilitate "basic level" recognition and "species level" features facilitate "subordinate level" recognition. For example, categorizing birds at the family level could include classifying two bird exemplars as either a Warbler or Finch. However, one could also categorize these same birds at a more subordinate level as a Wilson's Warbler or House Finch and this type of classification often requires perceptual expertise. Several investigations have applied training protocols to promote perceptual expertise in

* Corresponding author at: Department of Psychology, University of Florida, 945 Center Drive, PO Box112250, Gainesville, FL 32611-2250, United States.
E-mail address: lscott@ufl.edu (L.S. Scott).

https://doi.org/10.1016/j.visres.2021.107971
Received 7 May 2021; Received in revised form 8 October 2021; Accepted 1 November 2021
Available online 23 November 2021
0042-6989/

natural categories (birds: Tanaka, Curran, & Sheinberg, 2005; Scott, Tanaka, Sheinberg, & Curran, 2006), human-made categories (cars: Scott, Tanaka, Sheinberg, & Curran, 2008) and artificial object categories (Gauthier & Tarr, 1997; Gauthier et al., 1998; Jones et al., 2018) and find similar improvements in perceptual discrimination accuracy and reaction time as well as increased and more specialized neural responses after training.

These previous investigations examining the abilities of experts, as well as studies that trained experts in the laboratory, have led to a better understanding of what differentiates expert and novice visual perceptual processing. However, very little is known about the factors that facilitate or constrain learning and performance early in the acquisition of perceptual expertise, including the critical visual features and strategies that foster accurate perceptual discrimination and later expertise. The current investigation examines the extent to which discrimination accuracy and visual fixations changed after adult participants were trained to identify different families of novel objects at the subordinate and the basic level.

Previous training studies highlight the importance of learning exemplars at more specific levels of abstraction in order to increase perceptual expertise (Tanaka et al., 2005; Scott, et al., 2006, 2008; Jones et al., 2018). For example, discrimination of novel exemplars of wading birds is increased after wading bird training that includes subordinate or species-level labels, but not after training at the basic level, as "other" (Tanaka et al., 2005; Scott et al., 2006). Discrimination after basic-level training also does not differ from exposure training (training that does not include labels or classification as "other") (Scott et al., 2008). Additionally, generalization of learning was reported such that improvements in performance after subordinate-level discrimination training with birds generalized to both untrained exemplars of trained species and new exemplars from new species (Tanaka et al., 2005; Scott et al., 2006). Generalization after subordinate-level training with car models was relatively limited and only generalized within car classes (i.e., within Sedans or SUVs; Scott et al., 2008). These differences in generalization between birds and cars leaves open questions about how object characteristics can moderate the effect of training on generalization of learning.

Visual surface information, like color and texture, are important details used in the recognition and discrimination of both natural and man-made objects (Tanaka & Presnell, 1999; Therriault, Yaxley, & Zwaan, 2009; Bramão, Faísca, Petersson, & Reis, 2012). Recently, real-world bird experts were found to utilize important bottom-up stimulus features including color (Hagen, Vuong, Scott, Curran, & Tanaka, 2014), shape, and the internal visual details (Hagen, Vuong, Scott, Curran, & Tanaka, 2016) when making fast and precise recognition decisions. For example, although both experts and novices relied on color to recognize birds at the basic (family) level, color also facilitated expert performance when recognizing species of birds at the subordinate level. In addition, when making species level discriminations, bird experts made faster and more accurate judgments when color information was consistent with their color knowledge compared to grayscale images or when the true colors were replaced (Hagen et al., 2014). While experts and novices both utilized color information, experts used color across all responses whereas novices used color information only at the slowest reaction times (Hagen et al., 2014). These findings suggest that for experts, color information is readily available and utilized quickly in recognition judgements, whereas novice access to color information may be more deliberate and strategic.

Spatial frequency (SF) manipulations have also been found to influence real-world expertise. When bird images were filtered over a range of SFs, bird experts categorized common birds at the family level (e.g., Robin or Sparrow) more quickly and more accurately than novices (Hagen et al., 2016). In that study, bird experts were also asked to categorize birds at the more specific, species level (e.g., Wilson's Warbler or House Finch). Expert recognition was fastest for images filtered to retain only the middle range (8–16 cycles per image (cpi)) of SFs. These results indicate that the mid-range SFs contain crucial stimulus information for subordinate-level recognition and that extensive perceptual experience increases the efficiency with which this information is utilized.

The importance of color and SF information has also been examined within the context of laboratory training of novices (Devillez et al., 2019). Participants were taught to categorize finches (or warblers) at the subordinate-species level (e.g., purple finch) and categorize warblers (or finches) at the more general family/basic level across six training sessions. Similar to Hagen et al. (2014) and Hagen et al. (2016), training images were presented in their natural (congruent) or unnatural (incongruent) colors, in greyscale, or with only low SF (LSF $<$ 8 cpi) or high SF (HSF $>$ 8 cpi) visual information. Prior to training, both congruent and incongruent colors equally improved performance relative to greyscale. However, after subordinate-level training, congruent color facilitated perceptual discrimination over and above incongruent color. This color facilitation effect was also present in scalp recorded event-related potentials (ERPs) over occipital-temporal brain regions. Although both subordinate- and basic-level training generally increased N170 and N250 amplitude, color congruency effects were only observed after subordinate level training for the N250 component. However, unlike previous findings in real world experts (Hagen et al., 2016), SF effects in this study were unrelated to training.

The findings of Devillez et al. (2019) complement results with real world bird experts (Hagen et al., 2014) showing that novices integrate color information into their object representations during the acquisition of subordinate-level concepts. However, novices likely already have some knowledge of bird species and so it is possible that they focused on color as an important feature given this prior knowledge. For the present investigation and a previous investigation that examined neural and behavioral responses in an overlapping group of adults (Jones et al., 2018), completely novel objects were created to mimic families and species of birds. This overlapping group of participants, tested on different days than the current investigation, showed enhanced discrimination of exemplars from recently trained novel object categories after subordinate-level training, regardless of whether the stimuli were presented in color, in grayscale, or with either a low (LSF, $<$8 cpi) or high (HSF, $>$8 cpi) SF filter (Jones et al., 2018). In addition, although ERPs differentiated color images from grayscale, HSF, and LSF images, this difference was present regardless of whether the stimuli were trained at the subordinate or basic level and was present both before and after training.

Combined, these results suggest that although color and SF information appear to play an important role for bird recognition in experts with years of experience (Hagen et al., 2014; Hagen et al., 2016), SF may not be a critical factor in the early learning and acquisition of expertise (Devillez et al., 2019; Jones et al., 2018). On the other hand, the extent to which color is important for the early learning and acquisition of expertise has not been clearly established as it appears to be important when learning bird species (Devillez et al., 2019) but not novel objects (Jones et al., 2018).

Measures of visual fixations have previously been used to examine visual attention and strategy use in real-world experts using eye-tracking. The extent to which experts analyze images differently than novices has been investigated across a variety of real-world expert domains, including medical professionals (e.g., radiology, cardiology), sports (e.g., soccer, baseball, chess, gymnastics), and transportation (aviation, car driving), as well as forensic science, physics, and cartography (see meta-analysis: Gegenfurtner, Lehtinen, & Saljo, 2011). Based on this meta-analysis (Gegenfurtner et al., 2011) real-world expertise is generally associated with shorter fixation durations, increased fixations to task-relevant regions/features, and longer saccadic amplitudes. Work with expert radiologists and medical professionals also suggests that, relative to novices, experts were faster and required fewer fixations to gather pertinent information and make a response (Drew, Evans, Võ, Jacobson, & Wolfe, 2012; Kundel & La Follette, 1972; Krupinski, Graham, & Weinstein, 2013). Expert radiologists also show larger saccadic amplitudes relative to novices, suggesting an increased use of global shape information when viewing medical images within their domain of expertise (Manning, Ethell, Donovan, & Crawford, 2006; Bertram, Helle, Kaakinen, & Svedstrom, 2013). These findings in radiologists are similar to findings examining visual strategy use in scene perception that shows a reduced number and increased duration of fixations when the scene is more complex or densely populated (Henderson, 2011; Henderson, 2015; Rayner, 2009). Longer fixation durations have also been found when images were of a lower quality and during memorization tasks (Loftus, 1985; Mills, Hollingworth, Van der Stingchel, Hoffman, & Dodd, 2011). Finally, when reading difficult text or a complex font, fixation duration increases and fixation number decreases (Rayner, 2009; Rayner, Reichle, Stroud, Williams, & Pollatsek, 2006).

Visual fixations and strategies can be examined using a variety of dependent measures. In the present investigation we examined visual fixations broadly and included analyses of four dependent measures including: total fixation duration during the trial (millisecond (ms)), average duration of each fixation (ms), fixation count/number, and average saccadic amplitude (the angular distance a saccade travels between fixations, measured in degrees). The duration and number of visual fixations provide important information about how visual information is processed (Gegenfurtner et al., 2020; Gegenfurtner et al., 2019; Hauser, Mottok, & Gruber, 2018; Holmqvist & Andersson, 2017; Russo, 2019). More generally, measures of average fixation duration are thought to reflect sustained attention to a location (Just & Carpenter, 1980) while fixation number reflects shifts in attention (Schlesinger, Amso, & Johnson, 2007).

Visual fixations provide a useful measure of the allocation of visual attention to important features as well as the perceptual strategies employed by experts and novices under different task conditions. However, few studies have examined whether visual fixations change as perceptual discrimination performance improves or as one acquires expertise. Therefore, the first aim of this study was to replicate previous behavioral reports (e.g., Scott et al., 2006, 2008) showing increased perceptual discrimination accuracy after subordinate-, but not basic-level training. The second aim of this study was to examine the extent to which subordinate and basic-level training impacts visual fixations measured with an eye-tracker during the early acquisition of expertise with novel objects. Finally, the third aim of this study was to determine the extent to which color and SF manipulations impacted discrimination performance and visual fixations. To address these aims, adult participants were trained to identify different families of novel objects at the subordinate and the basic level. Discrimination accuracy and visual fixations were examined prior to training and immediately after training. Following previous reports (e.g., Scott et al., 2006, 2008; Tanaka et al., 2005; Devillez et al., 2019) the current investigation trained adults at both the subordinate and basic levels (within subject design) with two different families of artificial objects. These novel objects were created to have several shapes and features that varied across individuals, species, and families. Before and after training a participants completed a within family same/different discrimination task where they were asked to determine whether two serially presented images were from the same or different species while their fixations were recorded using an eye-tracker. For some trials, color and SF were modified to examine their impact on performance and visual fixations. It was predicted that learning at more specific levels of abstraction may have a top-down impact on visual representations as indexed by differential use of visual features (e.g., color and SF) and visual fixation strategies. More specifically, subordinate-level, but not basic level, training was expected to improved discrimination accuracy ((Tanaka and Curran, 2001); Scott et al., 2006, 2008; Devillez et al., 2019). Visual fixations were expected to follow a similar pattern such that subordinate-level, but not basic-level training was expected to increase average fixation duration and saccadic amplitude and decrease the number of total fixations (Gegenfurtner et al., 2011). Generalization of learning was examined by testing two generalization conditions before and after training. Transfer and generalization effects were examined within family and included untrained exemplars of trained species, and new exemplars of untrained species. Based on previous reports (Scott et al., 2006, 2008), generalization was expected after subordinate- but not basic-level training.

2. Method

2.1. Participants

Thirty-eight adults were recruited to participate in a larger perceptual expertise training study using both electrophysiological and eye-tracking methods (see Jones et al., 2018[1] for electrophysiological results). The EEG and eye tracking tasks were completed on different days before and after a training and only the eye-tracking results are reported here.

Of the 38 trained adults, 6 participants were removed because they did not complete all of the tasks analyzed in the current study and 8 participants were removed because they were identified as having a mean 3 SDs greater/less than the group mean for accuracy ($n = 1$) and two or more variables 3 SDs greater/less than the group mean within each eye-tracking measure ($n = 7$). The final sample included 24 adults. One participant declined to provide demographic data. The remaining 23 participants (age range: 18–28; $M = 21.91$, $SD = 2.98$) included 16 females and 7 males, 78.26% identified as White and 21.73% Asian or Pacific Islander, 95.65% of participants identified as not Hispanic or Latino and 4.34% identified as Hispanic or Latino).

Additional participant demographic information (e.g., GPA, SAT Scores, Field of Study) is reported in Table 1. Participants were paid $15 per hour for pre- and post-training assessments and $10 per hour for each of the 6 training sessions. Participants received a $25 bonus for completing all sessions. All participants provided written consent and all procedures were approved by the University IRB.

2.2. Stimuli

The stimuli included 240 novel, computer-generated objects called "Sheinbugs". Stimuli were generated and edited using Modo (Luxology, LLC). All stimuli are freely available on the Open Science Framework: https://osf.io/du5ke/?view_only=ce567f0cefa543288dee46d

Table 1
Participant demographics.

Measure	Statistics
Age	$M = 21.93$, $SD = 2.98$
Race	78.26% were White and 21.73% Asian or Pacific Islander
Ethnicity	95.65% not Hispanic or Latino, 4.34% Hispanic or Latino
College GPA	$M = 3.31$, $SD = 0.48$
High School GPA	$M = 3.70$, $SD = 0.76$
SAT Score	$M = 1802.64$, $SD = 42.17$
Field of Study	45.83% Natural Sciences, 25.0% Public Health, 12.5% Social and Behavioral Sciences, 4.16% Arts and Humanities, 4.16% Management, 4.16% Linguistics, 4.16% Engineering

[1] Participants were drawn from the same sample of adults reported by Jones et al. (2018). The pre- and post-test tasks and data analyzed and reported here do not overlap with Jones et al. (2018) (i.e., no data presented here has been presented elsewhere). The main purpose of the Jones et al. (2018) study was to examine the neural responses using EEG recordings. Twenty-two participants (of our 24 analyzed data) overlapped with those of the Jones et al. study (33 participants; 66.66% overlap).

1eee1f997). Stimuli were created to form two families of objects (Family A and Family S; See Fig. 1A). Each family included 10 unique species (labeled "1" through "10"), each containing 12 exemplars (See Fig. 1B). Image exemplars were presented in different orientations. However, each individual object exemplar was always presented in the same orientation across presentations. All objects were cropped and scaled to fit within a frame of 500 by 500 pixels and presented on a gray background. Participants sat approximately 60 cm away from a 17-inch Acer (AL1717) monitor with the following features: 150° (H) 135° (V) viewing angle, 500:1 contrast ratio, 30–81 kHz horizontal refresh rate, 55–76 Hz vertical refresh rate, refresh rate 60 Hz, 12 ms response time, 1280 × 1024 resolution, 300 cd/m^2 brightness. The stimuli were presented at a visual angle of approximately 13.85–16.13 horizontally and vertically. Each stimulus was 17 cm × 17 cm on the monitor.

For training, all images were presented in full color. To examine the effect of color and SF information on discrimination accuracy and visual fixations, stimuli were presented across four different image manipulation conditions for the pre- and post-test assessments. These image manipulations included: full color images, grayscale images, high SF (HSF) images, and low SF (LSF) images. The HSF and LSF images were created by band-pass filtering the grayscale images for two ranges of spatial frequencies: >8 cpi (HSF) and <8 cpi (LSF). A cutoff of 8 cpi was used based on previous studies on object recognition showing the importance of midrange SFs, typically defined around 6–10 cpi (Collin & McMullen, 2005; (Harel & Bentin, 2009); Hagen et al., 2016). A mask of the external bird contour was applied to keep the external contour constant for the two SF ranges (Hagen et al., 2016). Fig. 1C illustrates the stimuli and the transformations used in this experiment.

2.3. Procedure

The experiment consisted of 6 training sessions (approximately 1.5 h per session) over a 2–3-week period (for a total of 9 h), and 1 pre-training and 1 post-training assessment (approximately 2 h each). During the training, participants completed a naming task in which they were asked to determine which species the object belonged to using numbers on a keyboard. During the pre-training and post-testing assessments, participants completed a "same/different" serial matching task in which accuracy (d') and visual fixation data were measured. In this task, participants pressed a key on a keyboard to indicate whether the two presented images were from the same or different species.

2.3.1. Training

The 6 days of training included 5 different species within the two different object families (Family A, Family S) using a naming task (Scott et al., 2006, 2008). Participants were trained with a subset of 6 exemplars per species per family. Of the 24 included participants, 14 participants completed subordinate-level training with novel object Family A and basic-level training with Family S and 10 participants completed subordinate-level training with Family S and basic-level training with Family A. Participants completed a total of 25 blocks and a total of 900 trials during each of the 6 training sessions. In each block, participants were asked to label between 1 and 5 species per family using numbers on a keyboard. Exemplars were randomized across blocks between sessions, and all trained exemplars were presented during each session.

During each trial, participants viewed a single object for 1000 ms before being prompted to respond. Participants had 2000 ms to respond before the next trial began. When shown an exemplar from the subordinate-trained family, participants were asked to press the numeric key that corresponded to the correct species label (e.g., "2" for "Species 2") (See Fig. 1C). When shown an exemplar from the basic-trained family, participants were asked to always respond to exemplars by pressing the spacebar for "other" on the keyboard. Feedback was provided after each trial, and the correct label (i.e., "Species 1, 2, 3, 4, 5" or "Other") was shown if participants answered incorrectly. The remaining 5 species within each family were untrained and used during the pre- and post-test assessments to examine generalization of learning. Trained and untrained species were counterbalanced across participants.

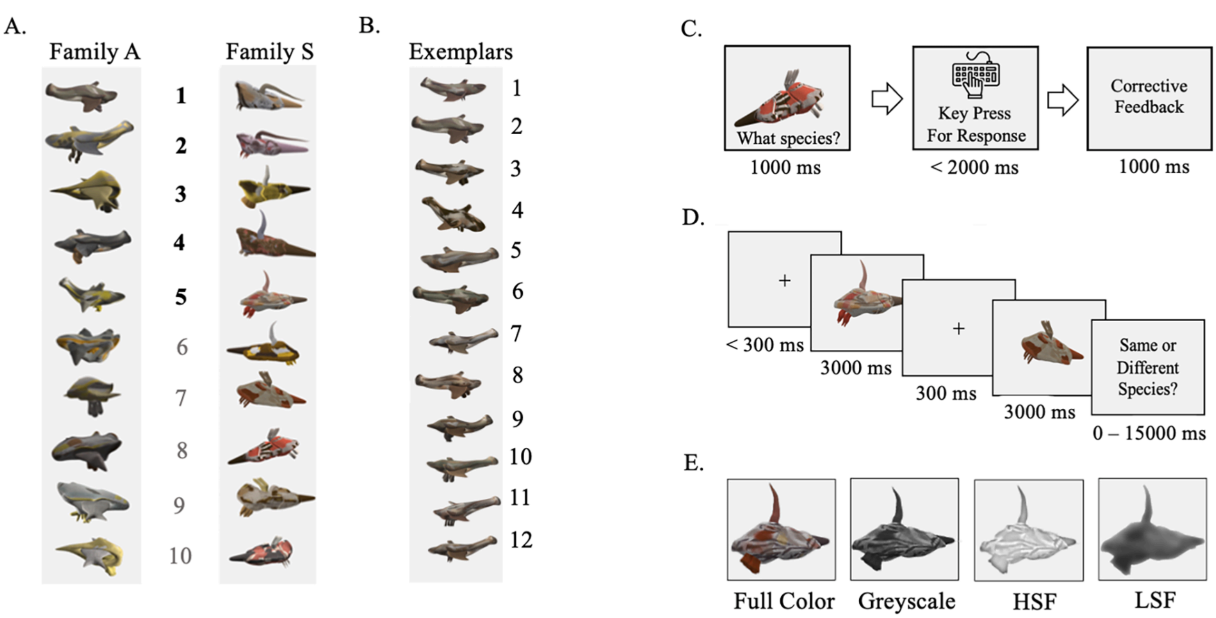

Fig. 1. Examples of novel stimuli, tasks, and image manipulations. A) Stimuli were two families (Family A, Family S) of novel, computer-generated objects called "Sheinbugs". Each family included 10 different species, and the trained species were labeled 1-5 (bolded). Species 6-10 (labeled in gray) were untrained and used to test generalization of learning. Trained and untrained exemplars were counterbalanced across participants. Each species included 12 total exemplars, but only one exemplar within each species is shown. B) Examples of 12 exemplars of Species 1 of Family A. C) Example of the training task. Each training trial consisted of the presentation of one image for 1000 ms, followed by a response screen prompting the participant to press a key corresponding to the correct species number. Corrective Feedback was provided. D) Example of the pre- and post-training serial matching task. Each trial began with a central fixation cross, followed by the presentation of one species exemplar, another fixation cross, and finally a second exemplar. Participants were then prompted to press a key to indicate whether the two objects were from the same or different species. E) During the serial matching tasks trials included images shown either in full color, greyscale, with HSF information visible, or with LSF information visible..

The difficulty of the training was manipulated between-subjects by varying the number of trained species per family. For the difficulty manipulation, participants were assigned to one of three training difficulty conditions: increasing difficulty, decreasing difficulty, or randomized difficulty. In the "increasing difficulty" condition, participants began each session with blocks consisting of one species per family and gradually increased by one species at a time for a total of five species per family. In the "decreasing difficulty" condition, participants began each session with blocks consisting of five species per family and gradually decreased by one species per family. In the "random difficulty" condition, the number trained in each block varied randomly (from 1 to 5 without replacement). However, no significant differences in performance were found across these three manipulations and so participants were collapsed across the difficulty manipulations conditions for analyses.

2.3.2. Pre- and post-test assessment: same/different serial matching task

Before and after training, participants completed a serial matching task (see Fig. 1D). During the serial image matching task, discrimination accuracy (d', calculated from an equal variance yes/no signal detection model; Macmillan & Creelman, 2005), total fixation duration, average fixation duration, fixation count, and average saccadic amplitude were measured. For each assessment (pre and post), participants completed a total of 240 trials divided equally across the two training levels (basic [120], subordinate [120]). Within each training level there were three generalization levels (trained [40], untrained exemplars of trained species [40], untrained exemplars of new species [40]), and four color (full color [30], grayscale [30]) and SF (LSF [30], HSF [30]) image manipulations. Untrained and new exemplars were examined to measure generalization of learning. The stimuli were randomly ordered and paired, with the same stimuli included in both the pre- and the post-training assessments.

During testing, participants were seated 60–70 cm away from an LCD monitor. During each test trial, a fixation cross was first presented in the center of the screen for at least 300 ms; once the fixation cross was visual fixated for 300 ms, the trial advanced. The first of the two objects was presented for 3000 ms, followed by another fixation cross for 300 ms. The second object was then presented for 3000 ms. After the second object was presented, a response screen then appeared until participants pressed a key indicating whether the two presented images were from the same or different species (See Fig. 1D). Visual fixations were measured for both the first (encoding) and the second (decision-making) object within the serial matching task (Rojas-Hortelano, Concha, & de Lafuente, 2014). Image manipulations were implemented for both the first and the second object (presented serially) within a trial. Same species trials included two different exemplars from the same species and different species trials included one exemplar from each of two separate species. Fixation data was collapsed across "same" and "different" trial types to increase power.

2.3.2.1. Eye-tracking procedure. An EyeLink 1000 remote camera eye tracker (SR Research Ltd, Mississauga, Ontario, CA) was used to record participants' visual fixations while they viewed the objects presented on a 17-inch LCD monitor mounted on an adjustable arm. The SR EyeLink 1000 eye-tracker uses a real-time and timing-sensitive operating system allowing for low variability. Fixation location and duration were recorded during both the encoding stimulus (object 1) and the test stimulus (object 2) with an average accuracy of 0.5° and a sampling rate of 500 Hz using a 35 mm lens and a 940 nm infrared illuminator. A fixation was defined by a threshold of 100 ms and the saccade velocity threshold was 30 deg/s and saccade acceleration threshold was 8000 deg/s. For each data sample, instantaneous velocity and acceleration was computed and compared with these thresholds. If either was above threshold, a saccade signal was generated.

Allowable head movement without accuracy reduction was approximately 22 × 18 × 20 cm (horizontal × vertical × depth). The gaze-tracking range was approximately 32° horizontally and 25° vertically. Head movement was minimized through the use of a head stabilizer with a chin rest sitting approximately 60 cm from the monitor. An eye track was recovered within 2 ms of losing the track.

2.3. Data analysis

All behavioral and fixation data were analyzed using MANOVAs in SPSS. Specific factors included in each analysis are described for each measure in the results section. For all analyses, paired sample t-tests were used to follow-up significant interactions and both corrected (using the Bonferroni method) and uncorrected p-values are reported.

3. Results

3.1. Accuracy: "same/different" serial matching task

d' was analyzed using a 2 × 2 × 3 × 4 MANOVA, with two levels of test (pre-test, post-test), two levels of training (basic, subordinate), three levels of generalization (trained exemplars, untrained exemplars of trained species, untrained/new exemplars of new species), and four levels of image manipulation (full color/SF, greyscale, HSF, LSF). Means and SDs by training levels, generalization conditions, and image manipulation are reported in Supplementary Materials Table 1.

Results showed a significant main effect of test, $F(1, 23) = 32.21$, $p < 0.001$, $\eta_p^2 = 0.58$, such that post-test d' ($M = 0.99$, $SD = 0.16$) was significantly greater than pre-test d' ($M = 0.78$, $SD = 0.20$). There was also main effect of generalization, $F(2, 22) = 7.57$, $p = 0.003$, $\eta_p^2 = 0.41$, such that d' for trained exemplars ($M = 0.91$, $SD = 0.15$, $p = 0.002$) and untrained exemplars of trained species ($M = 0.93$, $SD = 0.20$) was significantly greater than d' for new species exemplars ($M = 0.80$, $SD = 0.19$, $p = 0.001$), corrected p's < 0.05. Finally, there was a significant main effect of image manipulation, $F(3, 21) = 7.88$, $p = 0.001$, $\eta_p^2 = 0.53$, such that d' was greater for full color images ($M = 0.96$, $SD = 0.16$) relative to LSF ($M = 0.81$, $SD = 0.19$, $p < 0.001$) and HSF images ($M = 0.84$, $SD = 0.24$, $p = 0.01$), corrected p's < 0.05. d' for grey scale images ($M = 0.91$, $SD = 0.19$) was marginally greater than LSF images, corrected $p = 0.053$. There was no significant difference between full color/SF and grayscale images. Image manipulation did not significantly interact with test, training level, or generalization.

In addition to the reported main effects, reliable interactions were found between test and training level, $F(1, 23) = 14.98$, $p = 0.001$, $\eta_p^2 = 0.39$ (Fig. 2). Follow-up paired t-tests showed that subordinate level d' increased from pre-test ($M = 0.74$, $SD = 0.25$) to post-test ($M = 1.07$, $SD = 0.23$), $t(23) = -8.62$, $p < 0.001$, corrected $p < 0.001$. However, the d' difference between pre-test ($M = 0.81$, $SD = 0.21$) and post-test ($M = 0.91$, $SD = 0.22$) for basic-level training was not significant, $t(23) = -1.68$, $p = 0.11$. Finally, d' was significantly greater for subordinate- ($M = 1.07$, $SD = 0.23$) relative to basic- ($M = 0.91$, $SD = 0.22$) level trials at post-test, $t(23) = -2.53$, $p = 0.02$, corrected $p < 0.05$.

There was also a significant interaction between test (pre, post) and generalization (trained exemplars, untrained exemplars of trained species, new species exemplars), $F(2, 22) = 8.29$, $p = 0.002$, $\eta_p^2 = 0.43$. d' significantly increased for trained exemplars from pre- ($M = 0.75$, $SD = 0.21$) to post-test ($M = 1.07$, $SD = 0.17$), $t(23) = -6.78$, $p < 0.001$, corrected $p < 0.001$. The d' increase from pre ($M = 0.85$, $SD = 0.25$) to post ($M = 1.02$, $SD = 0.22$) was also significant for untrained exemplars of trained species, $t(23) = -3.45$, $p = 0.002$, corrected $p < 0.006$. The d' increase from pre- ($M = 0.74$, $SD = 0.23$) to post-test ($M = 0.87$, $SD = 0.23$) for the new species exemplars was marginally significant after correction, $t(23) = -2.54$, $p = 0.018$, corrected $p = 0.058$.

However, interpretation of these two-way interactions are superseded by a significant three-way interaction between test (pre, post), training level (basic, subordinate), and generalization (trained exemplars, untrained exemplars of trained species, new species exemplars), F

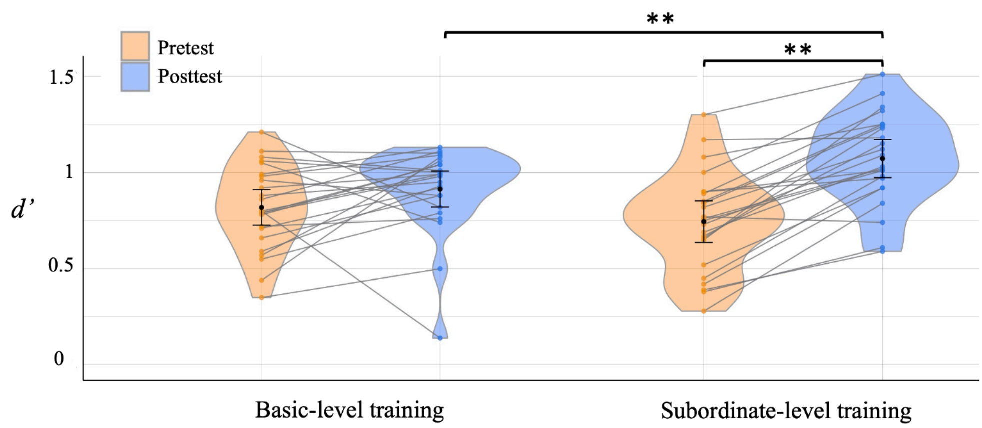

Fig. 2. Mean *d'* difference from pre-test to post-test for basic and subordinate-level trials. For each bar, the black dot represents the mean difference, the error bar represents the 95% confidence interval, and each participant's mean is marked with a colorful dot. Gray lines represent within individual *d'* changes from pre- to post-test for basic- and subordinate-level trials. Corrected ** $p < 0.001$.

(2, 22) = 5.44, $p = 0.012$, $\eta_p^2 = 0.33$ (Fig. 3, Table 1, supplementary materials). For subordinate-level trials, d' significantly increased from pre- to post-test for trained exemplars, $t(23) = -11.35$, $p < 0.001$, corrected $p < 0.001$, (pre-test: $M = 0.69$, $SD = 0.29$; post-test: $M = 1.25$, $SD = 0.23$), and untrained exemplars of trained species, $t(23) = -4.96$, $p < 0.001$, corrected $p < 0.001$, (pre-test: $M = 0.81$, $SD = 0.33$; post-test: $M = 1.09$, $SD = 0.31$). The increase from pre-test to post-test for new species from the subordinate-level trained family was not significant after correcting for multiple comparisons, $t(23) = -2.14$, $p = 0.042$, corrected $p = 0.126$, (pre-test: $M = 0.71$, $SD = 0.31$; post-test: $M = 0.86$, $SD = 0.31$). There were no significant changes from pre-test to post-test for any of the basic-level comparisons.

3.2. Eye-tracking

Four dependent measures were used to analyze visual fixations. Total fixation duration during the trial (ms), average fixation duration for each fixation (ms), fixation count, and average saccadic amplitude (degree) were analyzed separately in response to the first and second images of the serial matching task. Total fixation duration, count, and

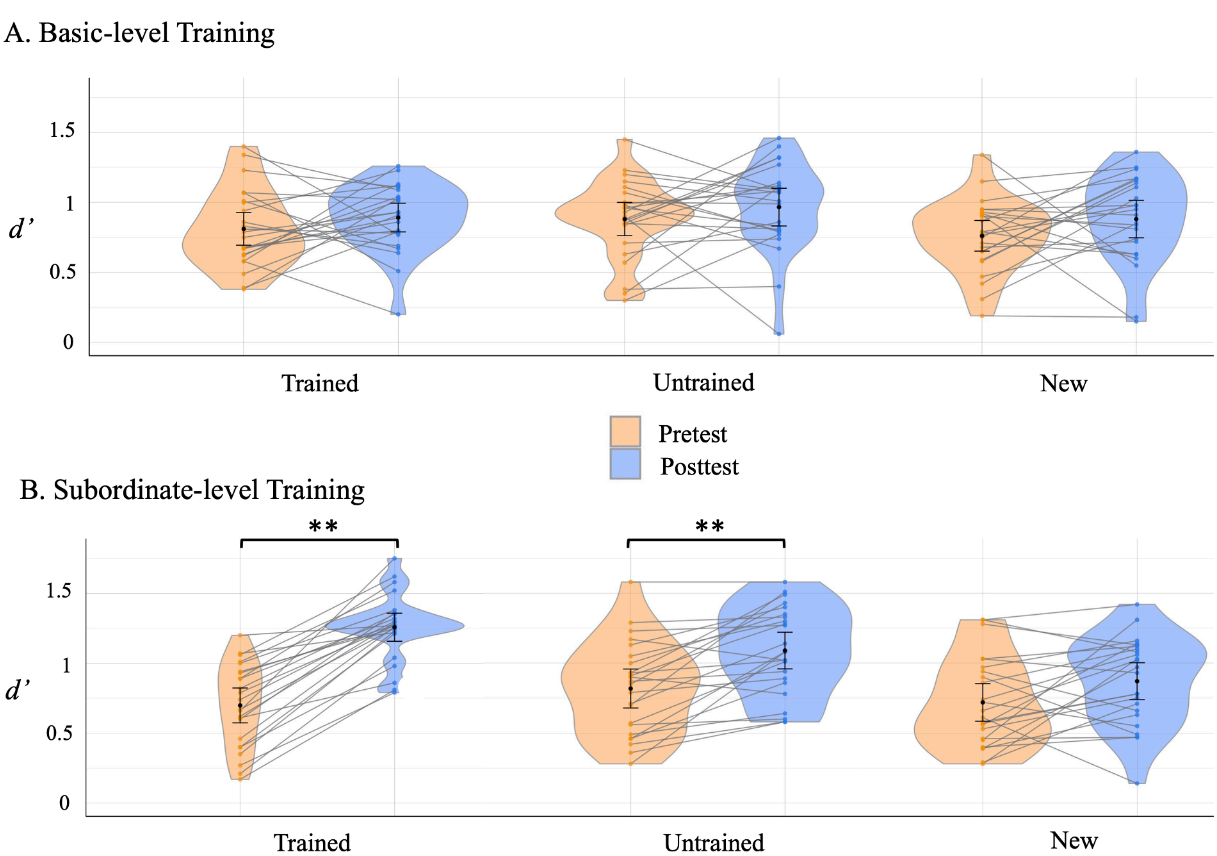

Figure 3. Accuracy (*d'*) mean differences for each generalization condition (trained exemplars, untrained exemplars of trained species, and new exemplars of untrained species) from pre-test to post-test for A. basic-level and B. subordinate-level training. Means are collapsed across image manipulation conditions. For each bar, the black dot represents the mean difference, the error bar represents the 95% confidence interval, and each participant's mean is marked with a colorful dot. Gray lines represent within individual *d'* changes from pre- to post-test for basic- and subordinate-level trials. Corrected **$p < 0.001$.

average fixation duration analyses were analyzed for an ROI that contained the surface area of the object and immediately adjacent areas (not for the entire screen). Analyses of average saccadic amplitude measure fixations for the entire screen.

ROI was measured based on size of the screen (768 × 1024 pixels) and the stimuli (500 × 500 pixels), which were shown in the center of the screen (coordinates 384, 512 pixels). To determine the interest area of each image, the difference in pixels between the height of the screen (768 pixels) and the height of the image (500 pixels) was divided by two to get the area of pixels above and below the image (134 pixels). The difference in pixels between the length of the screen (1024 pixels) and the length of each image (500 pixels) was divided by two to get the area of pixels to the right and left of the image (262 pixels). The final ROI included coordinates (262, 134 pixels) as the top left corner, coordinates (762, 134 pixels) as the top right corner, coordinates (262, 634 pixels) as the bottom left corner, and coordinates (762, 634 pixels) as the bottom right corner.

For each measure, analysis included a 2 × 2 × 2 × 3 × 4 factor MANOVA with two levels of test (pre-test, post-test), two levels of training (basic, subordinate), two levels of object (object 1, object 2), three levels of generalization (trained exemplars, untrained exemplars of trained species, new species exemplars) and four levels of image manipulation (full color, greyscale, HSF, LSF). Means and SDs by training levels, generalization conditions, and image manipulation for each object level are reported in Supplementary Materials Table 2. Visual fixation means are also shown in Figs. 4 and 5.

3.2.1. *Total fixation duration*

There were no main effects of test or training (see Fig. 4). There was a significant four-way interaction between test, training, object, and generalization, $F(2, 22) = 7.01$, $p = 0.004$, $\eta_p^2 = 0.38$. However, follow-up comparisons, corrected for multiple comparisons, revealed no significant condition differences.

3.2.2. *Average fixation duration*

There was a significant main effect of test (pre, post), $F(1, 23) = 4.59$, $p = 0.043$, $\eta_p^2 = 0.16$. Average fixation duration increased from pre-test ($M = 289.68$, $SD = 63.38$) to post-test ($M = 328.61$, $SD = 73.38$) (Fig. 4). There was also a main effect of object (object 1, object 2), $F(1, 23) = 6.38$, $p = 0.019$, $\eta_p^2 = 0.21$, such that there was a greater average fixation duration in response to object 2 ($M = 319.72$, $SD = 65.44$) than to object 1 ($M = 298.56$, $SD = 44.71$). Finally, there was a main effect of generalization condition (trained, untrained, new), $F(2, 22) = 4.28$, $p = 0.027$, $\eta_p^2 = 0.28$. Average fixation duration for trained exemplars (M = 313.69, $SD = 56.24$) was significantly greater than for new species exemplars ($M = 305.16$, $SD = 51.58$). There was no difference between new ($M = 305.16$, $SD = 51.58$) and untrained ($M = 308.59$, $SD = 50.67$) or between trained (M = 313.69, $SD = 56.24$) and untrained exemplars ($M = 308.59$, $SD = 50.67$).

These main effects were qualified by a significant five-way interaction between test, training, object, image manipulation, and generalization, $F(6, 18) = 3.28$, $p = 0.023$, $\eta_p^2 = 0.52$. However, none of the pairwise comparisons were significant after adjusting for multiple comparisons.

3.2.3. *Average fixation count*

There was a significant main effect of test, $F(1, 23) = 12.52$, $p = 0.002$, $\eta_p^2 = 0.35$, such that there was a decrease in the number of fixations from pre-test ($M = 9.06$, $SD = 1.91$) to post-test ($M = 7.72$, $SD = 0.89$) (Fig. 4). There was also a main effect of generalization condition, $F(2, 22) = 11.58$, $p < 0.001$, $\eta_p^2 = 0.51$, such that exemplars were fixated a greater number of times for new objects ($M = 8.48$, $SD = 1.18$) than for

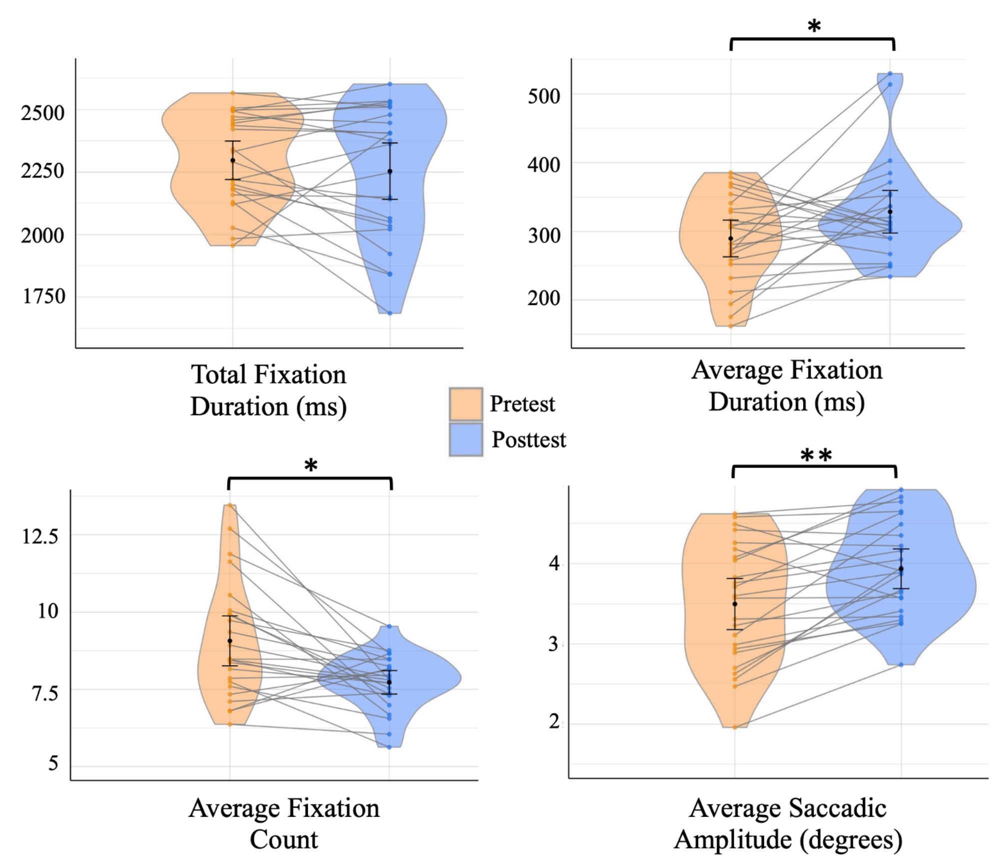

Figure 4. Mean measures of visual fixations for pre-test and post-test collapsed across training level, generalization, image manipulations, and object. The black dot represents the mean difference, the error bar represents the 95% confidence interval, and each participant's mean is marked with a colorful dot. Gray lines represent within individual changes in visual fixation measures from pre- to post-test for basic- and subordinate-level trials. Corrected $^*p < 0.05$; $^{**}p < 0.001$.

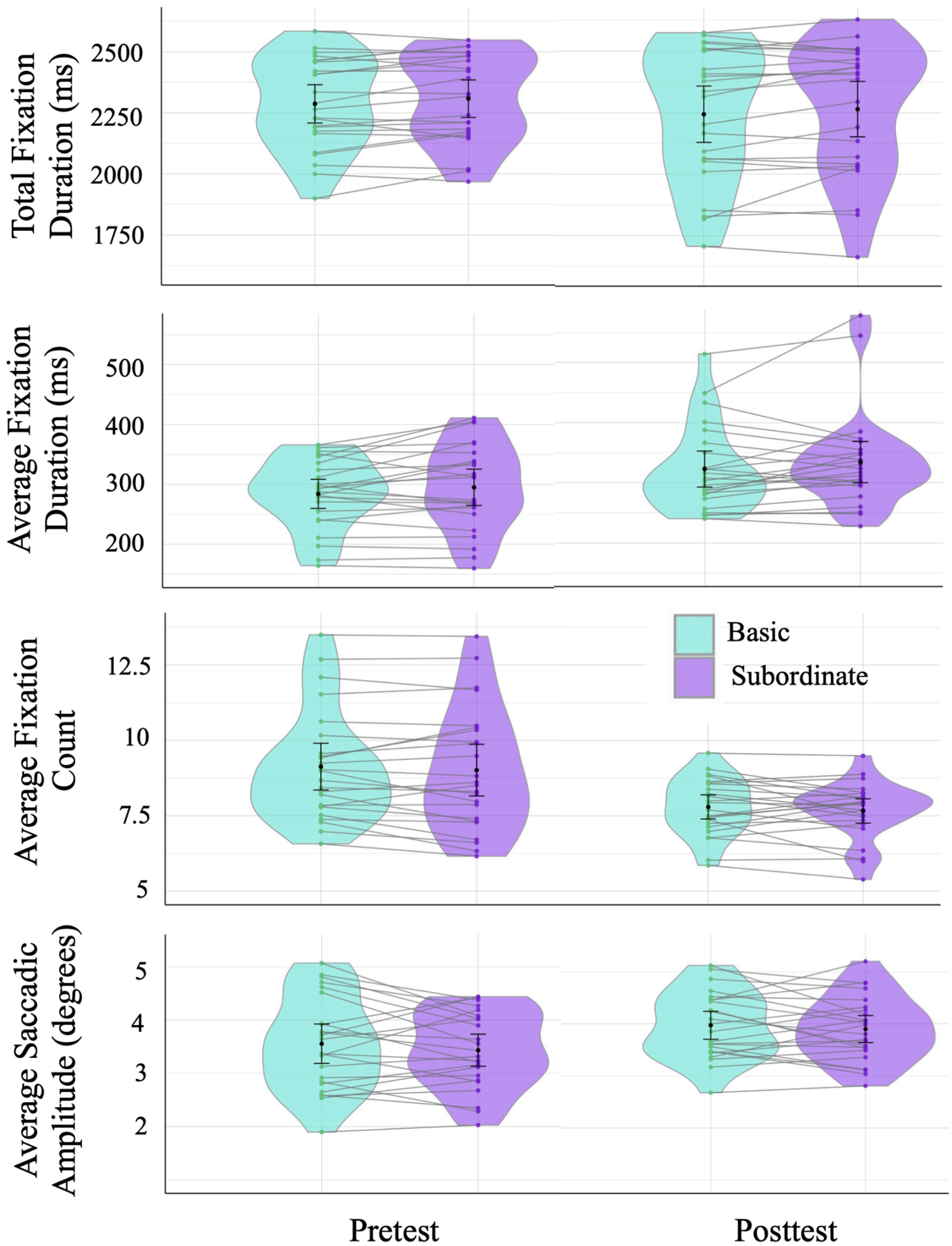

Figure 5. Mean measures of visual fixations for basic-level and subordinate-level training at pre- and post-test, collapsed across generalization, image manipulations (full color/SF) and object (first, second). From the top to bottom: Total Fixation Duration, Average Fixation Duration, Average Fixation Count, and Average Saccadic Amplitude. The black dot represents the mean difference, the error bar represents the 95% confidence interval, and each participant's mean is marked with a colorful dot. Gray lines represent within individual differences in visual fixation measures in basic- and subordinate-level trainings during pre- and post-test. Differences between basic- and subordinate-level were not significant across visual fixations measures.

trained objects (M = 8.33, SD = 1.14) and untrained objects from trained categories (M = 8.38, SD = 1.20). This effect of generalization did not interact with test and so was not considered further. There was an additional main effect of object, $F(1, 23) = 13.18$, $p = 0.001$, $\eta_p^2 = 0.36$, such that, within a trial, participants fixated object 1 ($M = 8.61$, $SD = 1.07$) more times than object 2 ($M = 8.18$, $SD = 1.33$).

The interaction between object and image manipulation was significant, $F(3, 21) = 3.28$, $p = 0.041$, $\eta_p^2 = 0.31$. Follow-up pairwise comparisons showed fixation number for grayscale object 1 ($M = 8.74$, $SD = 1.28$) was significantly greater than grayscale object 2 ($M = 8.28$, $SD = 1.43$), $t(23) = 4.20$, $p < 0.001$, corrected $p < 0.001$. For HSF exemplars, fixation number for object 1 ($M = 8.70$, $SD = 0.99$) was greater than object 2 ($M = 8.20$, $SD = 1.30$), $t(23) = 3.16$, $p = 0.004$, corrected $p = 0.016$. In addition, for LSF exemplars, fixation number was significantly higher for object 1 ($M = 8.42$, $SD = 0.84$) than object 2 ($M = 7.96$, $SD = 1.22$), $t(23) = 3.41$, $p = 0.002$, corrected $p = 0.008$. The difference between full color object 1 and full color object 2 was not significant after correction.

3.2.4. *Average saccadic amplitude*

There was a significant main effect of test, $F(1, 23) = 15.87$, $p = 0.001$, $\eta_p^2 = 0.408$, such that saccadic amplitude increased, generally, from pre-test ($M = 3.49$, $SD = 0.75$) to post-test ($M = 3.93$, $SD = 0.58$) (Fig. 4).

The main effect of object was also significant, $F(1, 23) = 9.22$, $p = 0.006$, $\eta_p^2 = 0.28$. Saccadic amplitude for object 2 ($M = 3.78$, $SD = 0.67$) was greater than for object 1 ($M = 3.64$, $SD = 0.57$). There were no significant interactions.

4. Discussion

The overarching goal of the current study was to use measures of discrimination accuracy (d') and visual fixations to examine the extent to which the use of visual features and strategies are impacted by training at the basic and subordinate levels of abstraction. The current findings demonstrated that discrimination accuracy improved after subordinate-level but not basic-level training. The improvement was evident for all image manipulations, but it was greatest for full color/SF images compared to LSF and HSF images. Furthermore, both subordinate- and basic-level training increased average fixation duration and average saccadic amplitude, and decreased fixation number from pre- to post-test for all image manipulations. Accordingly, the results suggest that whereas subordinate-level training differentially impacted perceptual discrimination performance, we found that both basic- and subordinate-level training modified visual fixation strategies.

The results suggest that discrimination accuracy, as measured by a serial matching task, improved from pre- to post-test after subordinate-level, but not basic-level training. This improvement after subordinate-level training was seen for all image manipulations (color and SF). Results from the present investigation are consistent with several other reports showing that subordinate-level, but not basic-level, training improves exemplar discrimination accuracy (Devillez et al., 2019; Scott et al., 2006; Scott et al., 2008; Tanaka et al., 2005). Increased discrimination accuracy after subordinate-level training also generalized to untrained exemplars within the subordinate trained species. These generalization findings are in line with previous findings showing generalization after subordinate-level training with either wading birds or owls (Scott et al., 2006; (Tanaka et al., 2005)). The current findings suggest that improvements in subordinate-level training discrimination generalize to untrained exemplars within trained species of novel objects. Although studies with species of birds consistently show generalization to new species of birds within the trained family, Scott et al. (2008) also report limited generalization to new models of cars within the same car class. In the future, generalization of learning should be examined after training with stimuli designed after a natural compared to an artificial kind to better understand potential generalization of learning limits. In addition, adding a pre- and post-test task that requires categorization (in addition to discrimination) may reveal increased categorization performance for the basic-training species.

In the current study, overall performance on the serial matching task was greatest for full color/SF images relative to LSF and HSF images but discrimination did not differ between full color/SF and greyscale images. The present results are consistent with results from an overlapping group of adults (tested on a different day, see Jones et al., 2018) which showed color information did not facilitate discrimination above and beyond greyscale images. Although this does not speak to the replicability of the effect, it does suggest stable performance within participants across testing sessions. Here, image manipulation did not significantly interact with test, training level, or generalization, suggesting that image manipulation effects were unrelated to the experimental manipulations. Specifically, image manipulation discrimination differences did not differ for objects from either the basic-level or subordinate-level trained families. The current findings stand in contrast to the results showing that color facilitated basic level discrimination of birds over greyscale images for both bird experts and novices (Hagen et al., 2014). That is, both experts and novices categorized the birds at the basic level (e.g., robin, sparrow) more quickly and accurately when birds were presented in their expected color than when shown in an incongruent color or in greyscale. Bird experts were also faster and more accurate when categorizing birds at the subordinate level when shown in their congruent colors than when shown in their incongruent colors or greyscale. Similarly, after two weeks of subordinate-level bird training in the laboratory, discrimination performance improved for birds presented with the congruent color relative to the incongruent color and greyscale bird stimuli (Devillez et al., 2019). The previous studies with real world and laboratory experts (Devillez et al., 2019; Hagen et al., 2014) indicate that color is an important diagnostic cue for expert-level bird discrimination. The current investigation did not show behavioral differences between the full color objects and greyscale objects for the "same/different" serial matching task. These inconsistent results may be due to the fact that the current investigation used computer-generated artificial stimuli that were novel to the participants and no information about the relevance of color was provided to them. It is possible that participants used other features (e.g., shape) to learn the distinctions during training resulting in no performance impacts when the color information was removed. In addition, the representational properties of these novel objects are largely unknown and so it is also possible that color may play a larger role for naturalistic but not artificial object categories.

Overall, there was increased discrimination accuracy for full color/SF images compared to LSF and HSF images and for greyscale images compared to LSF images. These results are consistent with training studies showing no difference from pre- to post-training between the HSF and LSF conditions after training in the laboratory in an overlapping group of participants tested with the same objects (Sheinbugs: Jones et al., 2018) and in a different group of adults trained with birds (Devillez et al., 2019). However, SF manipulations did not differ for the subordinate- and basic-trained families. Previously, SF was shown to impact discrimination of bird exemplars in both bird experts and novices when discriminating birds at the basic level, and in real-world bird experts when discriminating birds at the subordinate level (Hagen et al., 2016). In addition, past research exploring real-world object recognition has found that high SF information (>16 cpi) is important for subordinate-level discrimination (Collin & McMullen, 2005). In the present investigation, based on these previous findings (Hagen et al., 2016; Collin & McMullen, 2005), subordinate-level training was expected to impact processing of HSF images relative to LSF images. However, no differences in performance were found between HSF and LSF conditions. The current findings suggest that while SF manipulations decreased discrimination performance generally, early learning during subordinate-level versus basic-level laboratory training was not differentially impacted by SF manipulations.

This investigation aimed to examine the extent to which visual fixation strategies for objects were impacted by training at the subordinate and basic level and across image manipulations. For eye-tracking measures, there was an overall increase in average fixation duration and average saccadic amplitude from pre- to post-test, as well as a decrease in fixation number from pre- to post-test for all manipulations. Changes in visual fixation strategies did not differ between the family trained at the subordinate-level and the family trained at the basic-level. These pre- to post-test fixation changes therefore appear to be unrelated to the level of training (subordinate versus basic). More advanced analyses of visual scan paths or fixation clusters may reveal differences between these two levels of training; however, the current set of fixation analyses do not support differential fixation strategies.

In the present investigation, no differences in total fixation duration were found. However, analyses of average fixation duration and average fixation count suggest that is not due to a lack of training effects, but that average fixation duration increases and average fixation count decreases resulting in similar total fixation durations at pre-test and post-test. These findings are consistent with reports of expert radiologists and medical professionals, who show fewer fixations to gather pertinent information and make a response relative to novices (Drew et al., 2012; Kundel & La Follette, 1972; Krupinski et al., 2013). The present increase in average fixation duration from pre-test to post-test may reflect increased sustained attention (Just & Carpenter, 1980) while the decrease in number of fixations may be indicative of a decreased number of attentional shifts (Schlesinger et al., 2007).

The extent to which experts visually analyze images differently than novices was reported in a meta-analysis (Gegenfurtner et al., 2011). Gegenfurtner et al. (2011) reported that, unlike the present results, real-world experts exhibit decreased average fixation duration relative to novices. Visual fixations across a variety of cognitive tasks also show that increases in average fixation durations are associated with more complex or densely populated visual scenes, viewing low quality images, reading difficult text or a complex font, and during memorization tasks (Henderson, 2011; Henderson, 2015; Rayner, 2009; Loftus, 1985; Mills et al., 2011; Rayner, 2009; Rayner et al., 2006). These findings suggest that as the available visual information make the task more difficult, average fixation duration increases, similar to what is seen for novices relative to experts (Gegenfurtner et al., 2011). Based on the results reported here and the results reported by Gegenfurtner et al. (2011) we predict that average fixation duration may follow a non-linear developmental trajectory during the course of learning. For example, fixation duration may increase and decrease during the course of expert-level learning, knowledge acquisition, or based on task demands. If this is the case one might expect average fixation duration to increase during early learning, indicative of sustained attention, and decrease as expertise increases over months and years.

Real-world experts visually scan images within their domain of expertise more globally than novices resulting in greater saccadic amplitudes (Manning et al., 2006; Bertram et al., 2013; Gegenfurtner et al., 2011). Consistent with these findings we show an increase in saccadic amplitudes after just 6 days of training. These results suggest that training, in general, led to more global scanning strategies and a decrease in attentional shifts as measured by the decrease in fixation number. Examining the time course of increases in saccadic amplitude across the acquisition of real-world expertise is of interest for the future. Since we find increased saccadic amplitude after both basic- and subordinate-level training, it would be interesting to see if there is a point in learning at which the impact of these two different types of training diverge.

Subordinate-, but not basic-level, training was expected to increase average fixation duration and saccadic amplitude and decrease the number of total fixations (Gegenfurtner et al., 2011). However, this expected pattern was present for both subordinate- and basic-level training. An absence of differences between basic- and subordinate-level training suggests that the eye-tracking measures reported here may not be contributing to same/different performance differences after subordinate- and basic-level learning. However, our experimental manipulation was a within-subjects design and so it is conceivable that the two families of novel objects were similar enough to lead to generalization of visual fixation strategies employed for the subordinate-trained family to the basic-trained family. In looking at the d' means for basic-level training, although not significantly different, there are numeric increases from pre- to post-test. If there was generalization from the subordinate trained family to the basic trained family, then it is possible that the fixation results reported here precede behavioral changes in time. That is, it is possible that the two trained families used in the present experiment were more similar to each other than the birds or cars used in other experiments (e.g., Scott et al., 2006, 2008), leading to generalization of learning from the species trained at the subordinate-level to the species trained at the basic level. Future work could further distinguish the two trained families of novel objects perceptually or train with the same stimuli for a longer period of time to determine the extent to which generalization of learning occurred across the subordinate- and basic-level trained families. Computational models using these stimuli may also be helpful for uncovering the factors that facilitate or constrain generalization.

There were differences between the findings reported here and studies investigating real world experts (Hagen et al., 2014, 2016). One possible explanation for differences between real-world experts with years of experience and those trained briefly in the laboratory is that real-world experts also have conceptual and contextual knowledge built up over long-term experience and training (Starbuck, 1992; Sveiby, 1997). Real-world experts have a depth of knowledge and extensive experience that adults in the present investigation did not have. This extensive knowledge may have top-down impacts on fixation strategies and may require >6 days of perceptual discrimination training to arise (e.g., Gegenfurtner et al., 2020). Adults in the current study did not have any additional information about the novel objects beyond the species or family label and the visual perceptual details. The lack of contextual details and background knowledge for these novel objects may lead to differences in processing, such that real-world experts utilize color and SF information differently than those in the early stages of learning.

In summary, the results of the present investigation contribute to our understanding of perceptual discrimination during the early acquisition of perceptual expertise. First, surface features, including color and SF information, did not impact discrimination performance for novel images; improvements in performance occurred across all image manipulations. Although color and SF are important for experts with years of experience, they may not be as important for processes contributing to the early acquisition of perceptual expertise or expertise for non-naturalistic artificial objects. Second, the current study examined the extent to which visual fixations during a discrimination task were impacted by training. Results suggest there was an overall decrease in fixation count, and an increase in average fixation duration and saccadic amplitude from pre- to post-training that did not differ for objects trained at the basic- or subordinate-level. Finally, the current study replicated performance improvements (indexed by d') and generalization of learning results reported by past expertise training experiments ((Tanaka and Curran, 2001); Scott et al., 2006, 2008) suggesting that subordinate-level training may be important for professional and educational domains that require a high level of visual perceptual expertise.

5. Author note

Portions of the data reported here were presented at the Vision Sciences Society conference in St. Pete Beach, FL in 2017 and 2018. Data will be publicly accessible on the Open Science Framework (https://osf.io/du5ke/?view_only=ce567f0cefa543288dee46d1eee1f997).

Acknowledgements

This work was supported by Army Research Institute (ARI) grant W5J9CQ-11-C-0047 to L. Scott, T. Curran, and J. Tanaka. The views, opinions, and/or findings contained in this report are those of the authors and should not be construed as an official Department of the Army position, policy, or decision. The authors report no current conflicts of interest. We thank David Sheinberg for helping with the initial stages of stimulus creation, Melody Buyukozer Dawkins and Charisse Pickron with finalizing the stimulus set, and James Calabro and Matthew Mollison for technical and programming assistance. We thank Simen Hagen, Allison Carr, Ryan Barry-Anwar, Charisse Pickron, Eswen Fava, and members of the Brain, Cognition and Development Lab for relevant discussion, additional stimulus development, and research assistance.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.visres.2021.107971.

References

Bertram, R., Helle, L., Kaakinen, J. K., & Svedstrom, E. (2013). The effect of expertise on eye movement behaviour in medical image perception. ee66169 *Plos ONE, 8*. https://doi.org/10.1371/journal.pone.0066169.

Bramão, I., Faísca, L., Petersson, K. M., & Reis, A. (2012). The contribution of color to object recognition. In *Advances in object recognition systems* (pp. 73–88). https://doi.org/10.5772/34821

Busey, T. A., & Parada, F. J. (2010). The nature of expertise in fingerprint examiners. *Psychonomic Bulletin & Review, 17*(2), 155–160. https://doi.org/10.3758/pbr.17.2.155

Campbell, A., & Tanaka, J. W. (2018). Inversion impairs expert budgerigar identity recognition: A face-like effect for a non-face object of expertise. *Perception, 47*(6), 647–659. https://doi.org/10.1177/0301006618771806

Collin, C. A., & McMullen, P. A. (2005). Subordinate-level categorization relies on high spatial frequencies to a greater degree than basic-level. *Perception & Psychophysics, 67*, 354–364. https://doi.org/10.3758/bf03206498

Devillez, H., Mollison, M. V., Hagen, S., Tanaka, J., Scott, L. S., & Curran, T. (2019). Color and spatial frequency differentially impact early stage of perceptual expertise training. *Neuropsychologia*. https://doi.org/10.1016/j.neuropsychologia.2018.11.011

Drew, T., Evans, K., Võ, M.-H., Jacobson, F. L., & Wolfe, J. M. (2012). Informatics in radiology: What can you see in a single glance and how might this guide visual search in medical images? *Radiographics, 33*(1), 263–274. https://doi.org/10.1148/rg.331125023

Gauthier, I., Curran, T., Curby, K. M., & Collins, D. (2003). Perceptual interference supports a non-modular account of face processing. *Nature Neuroscience, 6*(4), 428–432. https://doi.org/10.1038/nn1029

Gauthier, I., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). Expertise for cars and birds recruits brain areas involved in face recognition. *Nature Neuroscience, 3*(2), 191–197. https://doi.org/10.1038/72140

Gauthier, I., & Tarr, M. J. (1997). Becoming a "greeble" expert: Exploring mechanisms for face recognition. *Vision Research, 37*(12), 1673–1682. https://doi.org/10.1016/S0042-6989(96)00286-6

Gauthier, I., Williams, P., Tarr, M. J., & Tanaka, J. (1998). Training "greeble" experts: A framework for studying expert object recognition processes. *Vision Research, 38*(15-16), 2401–2428. https://doi.org/10.1016/S0042-6989(97)00442-2

Gegenfurtner, A., Boucheix, J. M., Gruber, H., Hauser, F., Lehtinen, E., & Lowe, R. K. (2020). The gaze relational index as a measure of visual expertise. *Journal of Expertise, 3*(March), 32–40.

Gegenfurtner, A., Lehtinen, E., Helle, L., Nivala, M., Svedström, E., & Säljö, R. (2019). Learning to see like an expert: On the practices of professional vision and visual expertise. *International Journal of Educational Research, 98*, 280–291.

Gegenfurtner, A., Lehtinen, E., & Saljo, R. (2011). Expertise differences in comprehension of visualizations: A meta-analysis of eye-tracking research in professional domains. *Educational Psychology Review, 23*, 523–552.

Hagen, S., Vuong, Q. C., Scott, L. S., Curran, T., & Tanaka, J. W. (2014). The role of color in expert object recognition. *Journal of Vision, 14*, 1–13. https://doi.org/10.1167/14.9.9

Hagen, S., Vuong, Q. C., Scott, L. S., Curran, T., & Tanaka, J. W. (2016). The role of spatial frequency in expert object recognition. *Journal of Experimental Psychology: Human Perception and Performance, 42*, 413–422. https://doi.org/10.1167/14.10.1287

Harel, A., & Bentin, S. (2009). Stimulus type, level of categorization, and spatial-frequencies utilization: implications for perceptual categorization hierarchies. *Journal of Experimental Psychology: Human Perception and Performance, 35*(4), 1264.

Hauser, F., Mottok, J., & Gruber, H. (2018, June). Eye tracking metrics in software engineering. In *Proceedings of the 3rd European Conference of Software Engineering Education* (pp. 39-44).

Henderson, J. M. (2011). Eye movements and scene perception. In S. P. Liversedge, I. D. Gilchrist, & S. Everling (Eds.), *The Oxford handbook of eye movements* (pp. 593–606). New York, NY: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199539789.013.0033.

Holmqvist, K., & Andersson, R. (2017). Eye tracking: A comprehensive guide to methods. *paradigms and measures*.

Johnson, K. E., & Mervis, C. B. (1997). Effects of varying levels of expertise on the basic level of categorization. *Journal of Experimental Psychology: General, 126*(3), 248–277. https://doi.org/10.1037//0096-3445.126.3.248

Jones, T., Hadley, H., Cataldo, A. M., Arnold, E., Curran, T., Tanaka, J. W., & Scott, L. S. (2020). Neural and behavioral effects of subordinate-level training of novel objects across manipulations of color and spatial frequency. *European Journal of Neuroscience, 52*(11), 4468–4479. https://doi.org/10.1111/ejn.13889

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87*, 329–354. https://doi.org/10.1037/0033-295X.87.4.329

Krupinski, E. A., Graham, A. R., & Weinstein, R. S. (2013). Characterizing the development of visual search expertise in pathology residents viewing whole slide images. *Human Pathology, 44*(3), 357–364. https://doi.org/10.1016/j.humpath.2012.05.024

Kundel, H. L., & La Follette, P. S. (1972). Visual search patterns and experience with radiological images. *Radiology, 103*(3), 523–528. https://doi.org/10.1148/103.3.523

Loftus, G. R. (1985). Picture perception: Effects of luminance on available information and information-extraction rate. *Journal of Experimental Psychology General, 114*(3), 342–356. https://doi.org/10.1037//0096-3445.114.3.342

Macmillan, N. A., & Creelman, C. D. (2005). *Detection Theory: A User's Guide*. Lawrence Erlbaum Associates.

Manning, D., Ethell, S., Donovan, T., & Crawford, T. (2006). How do rediologists do it? The influence of experience and training on searching for chest nodules. *Radiography, 12*, 134–142.

Mills, M., Hollingworth, A., Van der Stingchel, S., Hoffman, L., & Dodd, M. D. (2011). Examining the influence of task set on eye movements and fixations. *Journal of Vision, 11*, 1–15. https://doi.org/10.1167/11.8.17

Rayner, K. (2009). Eye movements and attention in reading, scene perception, and visual search. *The Quarterly Journal of Experimental Psychology, 62*, 1457–1506. https://doi.org/10.1080/17470210902816461

Rayner, K., Reichle, E. D., Stroud, M. J., Williams, C. C., & Pollatsek, A. (2006). The effect of word frequency, word predictability, and font difficulty on the eye movements of young and older readers. *Psychology and Aging, 21*, 448–465. https://doi.org/10.1037/0882-7974.21.3.448

Rojas-Hortelano, E., Concha, L., & de Lafuente, V. (2014). The parietal cortices participate in encoding, short-term memory, and decision-making related to tactile shape. *Journal of Neurophysiology, 112*(8), 1894–1902. https://doi.org/10.1152/jn.00177.2014

Russo, J. E. (2019). Eye fixations as a process trace. In *A handbook of process tracing methods* (pp. 4–26). Routledge.

Schlesinger, M., Amso, D., & Johnson, S. P. (2007). The neural basis for selective attention in young infants: A computational account. *International Society for Adaptive Behaviors, 15*, 135–148. https://doi.org/10.1177/1059712307078661

Scott, L. S. (2011). Face perception and perceptual expertise in adult and developmental populations. In G. R. Rhodes, J. Haxby, M. Johnson, & A. Calder (Eds.), *Handbook of Face Perception*. Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199559053.013.0011.

Scott, L. S., Tanaka, J. W., & Curran, T. (2010). Degrees of expertise. In I. Gauthier, M. Tarr, & D. Bub (Eds.), *Perceptual expertise: Bridging brain and behavior* (pp. 107–138). Oxford: Oxford Scholarship Online. https://doi.org/10.1093/acprof:oso/9780195309607.001.0001.

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2006). A reevaluation of the electrophysiological correlates of expert object processing. *Journal of Cognitive Neuroscience, 18*, 1453–1465. https://doi.org/10.1162/jocn.2006.18.9.1453

Scott, L. S., Tanaka, J. W., Sheinberg, D. L., & Curran, T. (2008). The role of category learning in the acquisition and retention of perceptual expertise: A behavioral and neurophysiological study. *Brain Research, 1210*, 204–215. https://doi.org/10.1016/j.brainres.2008.02.054

Shen, J., Mack, M. L., & Palmeri, T. J. (2014). Studying real-world perceptual expertise. *Frontiers in Psychology, 5*, 857. https://doi.org/10.3389/fpsyg.2014.00857

Starbuck, W. H. (1992). Learning by knowledge-intensive firms. *Journal of Management Studies, 29*(6), 713–740. https://doi.org/10.1111/j.1467-6486.1992.tb00686.x

Sveiby, K. E. (1997). In *The new organizational wealth: Managing and measuring knowledge-based assets*. San Francisco, CA: Berrett-Koehler Publishers. https://doi.org/10.1016/s0956-5221(00)00026-9.

Tanaka, J. W., & Curran, T. (2001). The neural basis for expert object recognition. *Psychological Science, 16*, 145–151. https://doi.org/10.1111/1467-9280.00308

Tanaka, J. W., Curran, T., & Sheinberg, D. L. (2005). The training and transfer of real-world perceptual expertise. *Psychological Science, 16*(2), 145–151. https://doi.org/10.1111/j.0956-7976.2005.00795.x

Tanaka, J. W., & Presnell, L. M. (1999). Color diagnosticity in object recognition. *Perception & Psychophysics, 61*(6), 1140–1153. https://doi.org/10.3758/BF03207619

Tanaka, J. W., & Taylor, M. (1991). Object categories and expertise: Is the basic level in the eye of the beholder? *Cognitive Psychology, 23*(3), 457–482. https://doi.org/10.1016/0010-0285(91)90016-H

Tangen, J. M., Thompson, M. B., & McCarthy, D. J. (2011). Identifying fingerprint expertise. *Psychological Science, 22*(8), 995–997. https://doi.org/10.1177/0956797611414729

Therriault, D. J., Yaxley, R. H., & Zwaan, R. A. (2009). The role of color diagnosticity in object recognition and representation. *Cognitive Process, 10*(4), 335–342. https://doi.org/10.1007/s10339-009-0260-4