



Visual Discrimination of Geometric and Complex 2D Shapes in Goldfish (*Carassius auratus*)

Caroline M. DeLong^{1*}, Ashlynn M. Keller^{1,3}, Kenneth Tyler Wilcox^{2,4},
Irene Fobe¹, and Susan Ann Keenan¹

¹Department of Psychology, Rochester Institute of Technology, Rochester, NY

²School of Mathematical Sciences, Rochester Institute of Technology, Rochester, NY

³AMK is now at the Department of Psychology, Tufts University

⁴KTW is now at the Department of Psychology, University of Notre Dame

*Corresponding author (Email: cmdgsh@rit.edu)

Citation – DeLong, C. M., Keller, A. M., Wilcox, K. T., Fobe, I., & Keenan, S. A. (2018). Visual discrimination of geometric and complex 2D shapes in goldfish (*Carassius auratus*). *Animal Behavior and Cognition*, 5(3), 300-319. <https://doi.org/10.26451/abc.05.03.05.2018>

Abstract – Visual object recognition allows fish to locate prey, avoid predators, recognize individuals, choose mates, and navigate. Previous studies have examined shape discrimination in goldfish, but without controlling multiple dimensions of the stimuli to closely examine which features are salient for fish. The goal of this study was to explore features that goldfish (*Carassius auratus*) may use to discriminate between 2D geometric and complex shapes (e.g., length, width, area, diameter, symmetry). We assessed shape perception in fish using a two-alternative forced choice task. Five stimulus pairs were presented to goldfish: (1) a solid circle and rectangle matched for area; (2) a solid circle and rectangle matched for length; (3) a solid circle and square matched for length and width; (4) an unfilled circle and four-blade fan matched for diameter, area, and symmetry; and (5) two concentric unfilled circles and six-blade fan matched for diameter, area, and symmetry. Fish were tested using the ‘correction method’ in Experiment 1 and the ‘non-correction method’ in Experiment 2. Goldfish visually discriminated between the geometric shapes (stimulus pairs 1-3), but failed to discriminate between complex shapes (stimulus pair 5). Only the fish in Experiment 1 learned to discriminate between the complex shapes in stimulus pair 4. This study provided preliminary evidence for shape discrimination in goldfish when stimulus features are controlled. However, further work is needed to determine whether goldfish use a single feature that changes from stimulus pair to stimulus pair, a combination of features, or holistic processing to discriminate among shape stimuli.

Keywords – Goldfish, Object recognition, Shape discrimination, Visual perception

Visual object recognition is important for fish when selecting mates, recognizing familiar and unfamiliar individuals, engaging in social learning, avoiding predators, foraging, and navigating (Brown, Laland, & Krause, 2011). Fish use information received visually and with their other senses to engage in complex cognitive tasks such as pattern recognition and categorization (Schluessel, Fricke, & Bleckmann, 2012), quantity discrimination (Agrillo, Dadda, & Bisazza, 2007; Agrillo, Dadda, Serena, & Bisazza, 2008, 2009; Agrillo, Piffer, Bisazza, & Butterworth, 2012), memory formation (Aronson, 1971; Lucon-Xiccato & Dadda, 2014), tool use (Brown et al., 2011), and navigation using cognitive maps (Lopez, Broglio, Rodriguez, Thinus-Blanc, & Salas, 1999). Fish perceptual and cognitive abilities compare very well to other vertebrates on most tasks (Brown, 2015; Brown et al., 2011). There are parallels between fish and nonhuman primates on many social cognition phenomena such as individual recognition,

cooperation, eavesdropping, social learning, cooperative hunting, cheating, punishment, and altruism (Brown et al., 2011; Bshary, Wickler, & Fricke, 2002). Sometimes fish even succeed where nonhuman primates fail on visual cognition tasks [e.g., a complex foraging task (Salwiczek et al., 2012), although changing the cues improved the performance of the nonhuman primates (Pretot, Bshary, & Brosnan, 2016)].

The aim of the current study was to investigate visual shape discrimination in goldfish. Many different terrestrial and aquatic animals have demonstrated the ability to discriminate between 2D or 3D stimuli varying in shape. Mammals that have been tested include dolphins (Harley, Roitblat, & Nachtigall, 1996; von Fersen, Schall, & Gunturken, 2000), sea lions (Mauck & Dehnhardt, 1997; Schusterman & Thomas, 1966), ferrets (Pollard, Beale, Lysons, & Preston, 1967), opossum (James & MacFarland, 1966), rats (Minini & Jeffery, 2006; Oldfield-Box, 1969), monkeys (Tanaka, 2000), and white rhinos (Daniel & Mikulka, 1998). Birds that have displayed the ability to discriminate shape include pigeons (Young, Peissig, & Wasserman, 2001), jungle crows (Bogale & Sugita, 2013), parakeets (Cogan, Inmam & Gambrel, 1971), and chickens (Zolman, 1969). Invertebrates that have been tested include octopuses (Sutherland, 1959, 1969) and honeybees (Srinivasan, 2010). Finally, fish tested for this capability include damselfish (Siebeck, Litherland, & Wallis, 2009), goldfish (Mackintosh & Sutherland, 1963), and sharks (Fuss, Bleckmann, & Schluessel, 2014b; Fuss & Schluessel, 2015). To succeed at discriminating between shape stimuli in these tasks, animals can learn about individual features (e.g., local luminance, absolute size, height) or process shape holistically (e.g., using the ratio of height to width). The features of the stimuli must be carefully controlled to determine which of those features influenced performance on the task. In addition, probe tests can be conducted after discrimination learning to further investigate what information animals use to discriminate shapes. For example, Minini and Jeffery (2006) found that rats did not use global shape processing when discriminating between triangles and squares, but instead used local luminance differences (amount of black and white space) in the lower hemifield. When they were prevented from using luminance cues, rats seemed to prefer using single features like stimulus width, and used aspect ratio only after prolonged training. Animals that generalize a shape discrimination when stimuli are altered in size, rotated, or translated may be using holistic shape processing (Mauck & Dehnhardt, 1997; Pollard et al., 1967).

Early studies of visual shape discrimination in fish included a variety of 2D geometric and complex shapes (Bowman & Sutherland, 1969, 1970; Mackintosh & Sutherland, 1963; Sutherland, 1969; Sutherland & Bowman, 1969). In these studies, fish were trained in a two-alternative forced choice task to discriminate between two shapes during training, then were presented with modified stimuli during transfer trials (modifications could include changing the position of one part of the stimulus while retaining the overall shape, changing the shape of the stimuli, rotating the stimuli, or conversion to outline forms). Sutherland and his colleagues often suggested fish were using low-level features to discriminate among the stimuli. For example, they claimed that fish appear to rely on the presence versus absence of oblique contours and the presence versus absence of vertical and horizontal lines in stimuli (Sutherland, 1969), fish use the number of points or knobs (protrusions) present in each shape (Bowman & Sutherland, 1969), and fish rely more heavily on differences in features at the tops of the shapes (Sutherland & Bowman, 1969).

Unfortunately, in these early studies of shape discrimination in fish, the features of all the stimuli (width, length, diameter, area, brightness) were uncontrolled or unreported. In most studies, some measurements of the training stimuli are reported but the test stimuli measurements are unknown, which makes it likely that the size or brightness (area of black vs. white space) were not well controlled. These uncontrolled or unreported measurements make it difficult to evaluate the features goldfish use to recognize shape. Further, Ohnishi (1991) showed that goldfish trained with compound stimuli (color and pattern) when one aspect is easily learned and the other is difficult will select only the easily learned aspect to perform the discrimination. This suggests that fish may have relied on some aspect other than holistic shape to discriminate among stimuli in previous studies if low level features were available to use as cues (e.g., area or brightness).

In more recent research, there is a body of evidence that many species of fish attend to visual stimuli as a whole, rather than attending to only certain features of objects. Redtail splitfin (*Xenotoca eiseni*), goldfish (*Carassius auratus*), and bamboo sharks (*Chiloscyllium griseum*) fall prey to visual illusions in the same manner as humans when they perceive illusory contours within Kanizsa figures (Fuss, Bleckmann, & Schluessel, 2014a; Sovrano & Bisazza, 2009; Wyzisk & Neumeyer, 2007). Redtail splitfin have shown the ability to visually complete partly occluded objects, known as amodal completion (Sovrano & Bisazza, 2008). Redtail splitfins also showed a global preference when processing visual hierarchical stimuli (Truppa, Sovrano, Spinozzi, & Bisazza, 2010). Archerfish (*Toxotes chatareus*) learned to discriminate between human faces even when experimenters controlled for brightness, color, and head shape (Newport, Wallis, Reshitnyk, & Siebeck, 2016). Cichlids (*Pseudotropheus* sp.) learned to distinguish between the categories of ‘fish’ and ‘snail’ by visually examining 2D line drawings and photos (Schluessel, Fricke, & Bleckmann, 2012). When individual features of the stimuli were selectively removed (eyes or fins of the fish, or the shell or antennae of snails) but the overall shape remained, the fish were typically still able to categorize stimuli into the two categories. Schluessel et al. (2012) concluded that fish were likely attending to the entire stimulus rather than individual features. Cichlids (*Pseudotropheus* sp.) showed the ability to discriminate between 3D models of frogs and turtles (Schluessel, Kraniotakes, & Bleckmann, 2014). The fish continued to successfully discriminate when the models were rotated or varied in size and color but retained their general shape. In these studies, the fish learned about the stimulus in its entirety instead of isolated individual stimulus elements.

The focus of the current study was to examine visual shape discrimination in goldfish. It improves upon earlier studies of shape discrimination in goldfish (e.g., Bowman & Sutherland, 1970) by controlling stimulus features such as length, width, diameter, area, and brightness. We used a series of geometric and complex shapes similar to stimuli used by Siebeck et al. (2009) to test shape learning and discrimination in the reef damselfish *Pomacentrus amboinensis*. It is uncommon to see the performance of more than one fish species compared on the same task using the same stimuli (*cf.* McAroe, Craig, & Holland, 2016; Schluessel, Kortekamp, Ortiz Cortes, Klein, & Bleckmann, 2015). Past studies have used different methods to test fish, including the ‘correction method’ (Mackintosh & Sutherland, 1963, Wyzisk & Neumeyer, 2007) and the ‘non-correction method’ (Experiment 5 in Siebeck et al., 2009; Sovrano & Bisazza, 2008). The ‘correction method’ allows subjects to immediately correct their errors and earn reinforcement on each trial but may take more time, whereas the ‘non-correction method’ can speed up the test procedure but may lead to side biases. We used the ‘correction method’ in Experiment 1 and the ‘non-correction method’ in Experiment 2 to show that goldfish can learn to discriminate between shape stimuli while various stimulus features were controlled using either method. We predicted that the fish would be able to learn the shape discriminations with either method.

Experiment 1

Method

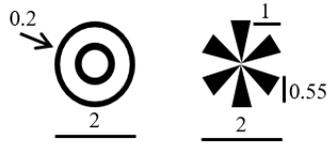
Subjects. The subjects were three experimentally-naïve goldfish (*Carassius auratus*), 5.5–6.0 cm in length, obtained commercially. Individual fish were identified by their phenotypic characteristics, including coloration pattern (but sex was not determined phenotypically). Subjects were housed in 10 gallon tanks (50 cm long x 26 cm wide x 31 cm high), and each lived with one other goldfish that was not participating in the current study. Each tank contained aquarium gravel and flat glass marbles, one artificial plant, and an Aqueon Quiet Flow 10 filter unit that provided aerated and filtered water. Aqueon water conditioner was added to each tank, which neutralizes chlorine and chloramines and detoxifies heavy metals, ammonia, and other elements released from fish waste. The top of each tank contained one Aqueon 10-watt mini-compact fluorescent light bulb that simulated a natural light/dark cycle (there were no windows in the room). The tanks were covered on three sides by blue cellophane on the exterior walls. Water temperature was between 22°C - 25°C. Water changes (40% of tank) were carried out on a weekly basis and tanks were monitored regularly for pH and waste levels with Tetra Aquarium Test Strips (Tetra

U.S., Blacksburg, VA). Training and testing sessions were conducted during daylight hours (morning and/or afternoon). On days with training or test sessions, subjects were food restricted because food was their reward during sessions. On weekdays when the fish were not trained or tested, they were fed twice; once with TetraFin flakes (Tetra GMBH, Melle, Germany) and once with API Premium pellets (Mars Fishcare Inc., Chalfont, PA). They began the study after acclimating to their home tanks for four to five days. One fish was excluded from the experiment after failing to cooperate (would not approach stimulus board) during testing for stimulus pair 1.

Stimuli. Table 1 shows the five stimulus pairs. We adapted two stimulus pairs that Siebeck et al. (2009) presented in 3D (a disc and a bar) such that they were presented in 2D (stimulus pairs 1-2: a circle and a rectangle). We used the same three stimulus pairs that Siebeck et al. (2009) presented in 2D (stimulus pairs 3-5: a circle and a square, outline of a circle and four-blade windmill, and two concentric circles and eight-blade windmill), except the outer edges of the segments in windmills in the current study were straight instead of curved. Each pair of stimuli was matched for different features so that the fish could not rely on any one spatial feature for all five stimulus pairs. Stimulus pair 1 was matched for area, stimulus pair 2 was matched for length, and stimulus pair 3 was matched for length and width. Stimulus pairs 4 and 5 were both matched for diameter, area of black vs. white zones, and symmetry. Stimuli were printed using black ink on standard white paper and then laminated (as in Experiment 5 in Siebeck et al., 2009). Each stimulus was centered on a 4 cm x 4 cm square (the stimulus card). The stimulus cards were attached to a white plastic corrugated stimulus board (25.5 cm high x 16.5 cm wide x 0.4 cm thick) with hook and loop fasteners (as in Experiment 5 in Siebeck et al., 2009).

Table 1

Summary of Stimuli

Stimulus Pair	Shapes	Features Matched
1		Area (3.14 cm ²)
2		Length (2 cm)
3		Length and width (2 cm)
4		Diameter and area of black vs. white zones (black = 1.1 cm ²)
5		Diameter and area of black vs. white zones (black = 1.7 cm ²)

Note. All stimuli were laminated printouts (2D) with the stimulus centered on a white background (total size of each stimulus card was 4 cm length by 4 cm width). All measurements are shown in cm. In stimulus pairs 2 and 3, the area of the circle was 3.14 cm². The area of the rectangle in pair 2 was 2 cm² and the area of the square in pair 3 was 4 cm². These stimuli were adapted from Siebeck et al. (2009). The S+ is on the left for each pair.

Experimental Set-up. Figure 1 shows the experimental set-up. During training and testing sessions, subjects were moved to individual 3.7 gallon test tanks (34.0 cm long x 19.5 cm wide x 21.5 cm high). Test tanks contained water and Aqueon water conditioner but contained no gravel or plants. Blue cellophane covered three of four exterior walls to block external stimuli during testing (the stimulus board was positioned on the uncovered wall). There were three identical test tanks and fish rotated among the three tanks according to a random schedule. During training and testing sessions, the stimulus board was inserted in the front of the tank and rested against the tank wall. On one side of the stimulus board, a single stimulus card could be presented during the training stage that was centered on the board (see Figure 1B). On the other side of the board, two stimulus cards could be presented during the testing stage (see Figure 1C).

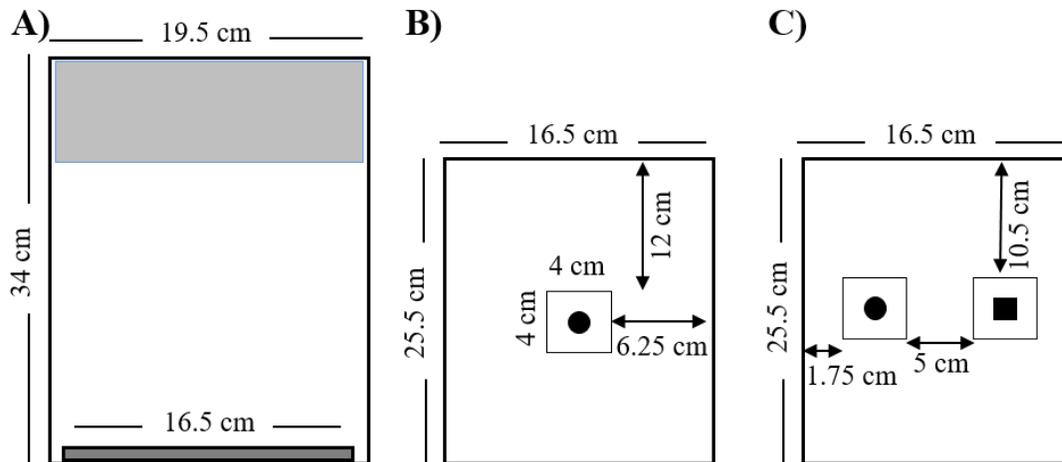


Figure 1. Schematic diagram of the experimental set up. **A)** Test tank showing the position of the stimulus board (dark gray shaded region). The fish had to be located within the light gray shaded region when the stimulus board was placed in the tank at the beginning of a session. **B)** The stimulus board side used for the training phase (S+ only). **C)** The stimulus board side used for the testing phase (S+ and S-).

Procedure. Sessions were conducted once or twice per day, always around the same time of day in the morning and afternoon (this resulted in a food deprivation period of about four to five hours between sessions on the same day). There were typically 5-10 sessions per week. Each test tank was filled with water and water conditioner, then a fish was placed into each test tank using a small net and allowed to acclimate to the tank for a minimum of 60 min. Stimulus cards were placed on the hook and loop fasteners on the stimulus board. Before the stimulus board was placed in the tank the fish was ushered into the back of the tank by gently tapping on the water in the back of the tank (see Figure 1A). This ensured the fish was swimming in the back of the tank when the stimulus board was placed in front of the tank and prompted the fish to examine the board before making a choice.

The fish were reinforced for correct choices with food paste; a mixture of TetraFin flakes and water (similar to Siebeck et al., 2009). Food paste was delivered to the fish using a 1.0 mL Luer-Lok tip syringe (BD, Franklin Lakes, NJ), and 0.01 mL of food paste was delivered for each correct choice. The fish had to tap (bite or push) the rewarded stimulus once before the food reward was given (although they would occasionally tap a stimulus multiple times before feeding from the syringe). [The choice to reinforce the fish after a single tap was based on Siebeck et al.'s (2009) results showing poor first choice performance when a reward follows 10 taps.] The experimenter held the syringe out of sight from the fish behind the test tank wall until a correct choice was made. As soon as the subject tapped once on the rewarded stimulus, the experimenter lowered the syringe into the tank positioned near the fish's head. If a fish tapped the incorrect stimulus it was not reinforced.

For each session, the total session time, the stimulus first visited by the fish for each trial, and the number of taps the fish made to each stimulus for each trial were recorded. Total session time was defined

as the time from when the stimulus board was first placed in the test tank to the time when the board was removed after the fish received five food reinforcements (rounded to the nearest minute). Individual trial times were not recorded. One experimenter (the trainer) stood in a position that allowed her to view the fish's behavior and deliver the food paste while holding the syringe out of view of the fish, while a second experimenter (the recorder) out of view of the fish recorded the fish's choices and session time. There were multiple trainers and recorders throughout the experiment.

Pre-training. Prior to training, the fish had to acclimate to the test tanks and feeding from the syringe. The first phase of pre-training was habituation to the test tanks with the white corrugated plastic board (stimulus board) inserted against the tank wall. In the first session, the fish were placed individually in test tanks and allowed to habituate to the test tank (minimum 60 min). In subsequent sessions, the syringe was inserted into the tank at varied locations and the fish were reinforced for approaching the syringe and eating from it. Figure 2A shows a fish eating from the syringe in the test tank. Each time the fish ate from the syringe was considered one trial (there were five trials per session). The fish were reliably eating from the syringe in three to five sessions. Training started four days after pre-training was complete.

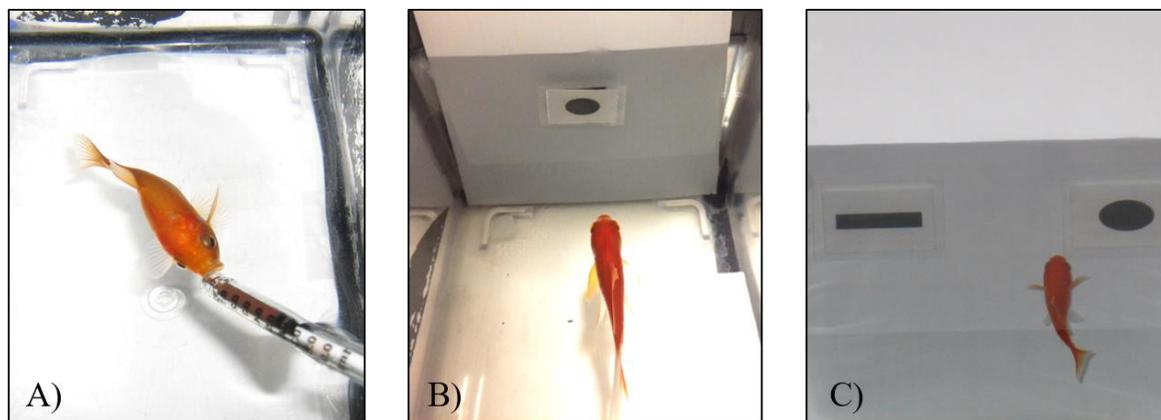


Figure 2. Photographs of fish engaged in training and testing. **A)** A fish eating flake paste from the syringe during the pre-training phase. **B)** A fish approaching the S+ during the training phase. **C)** A fish during the testing phase with stimulus pair 1 approaching the S+. After tapping the S+ the syringe was inserted from above and positioned near the fish's head.

Training. In the training phase, the fish learned to tap the S+ (positive stimulus) to receive a food reward. The stimulus used in this phase was the S+ for stimulus pair 1 (the solid circle). The experimenter used successive approximation to reinforce the fish for behaviors that brought them closer to tapping the S+. First, the experimenter placed the syringe 5-10 cm away from the stimulus board to induce the fish to approach the board. Next, the experimenter placed the syringe 1 cm from the stimulus board. At the end of the training phase, the subjects were required to approach the stimulus board and tap the S+ once (without the syringe in view) to receive the food reward. Figure 2B shows a fish approaching the S+. The experimenter inserted the syringe in the test tank five times per session (i.e., the fish ate five times in one session), just like the pre-training phase. The stimulus board remained in the same position in the test tank for the duration of each session. The fish reliably tapped on the S+ on every trial (without being cued by the placement of the syringe in front of the S+ for three sessions in a row) within 7-10 sessions. Testing started the day after the training phase was complete.

Testing. In this phase, the fish were required to make a choice between the S+ and the S- (negative stimulus). A test session was defined as the period in which a fish made five correct choices (i.e., it was fed five times from the syringe). Trial 1 was defined as the time period from the stimulus board's insertion into the tank to the fish tapping on the positive stimulus the first time (followed by the food reinforcement). The second trial began after the fish finished eating the food reward from trial 1 and swam away from the board, and ended after the fish tapped on the positive stimulus the second time and was fed. Thus, each trial began after the food reinforcement was eaten from the previous trial and ended

when the food reinforcement was consumed for the present trial. The ‘correction method’ was used, which means that incorrect choices during trials were allowed to be corrected (i.e., the fish could tap the S- as many times as necessary until it tapped the S+ and was fed to end the trial). The fish had a maximum of five minutes to correct an incorrect choice, although it was rare for fish to take longer than 30 s per trial. A session ended after five complete trials. Sessions averaged 2.2 min. (*Minimum* = 1 min, *Median* = 2 min, *Maximum* = 12 min).

The stimulus board remained in the test tank without being removed between trials for the duration of the five-trial session for stimulus pairs 1-4 (this followed Wyzisk & Neumeyer’s (2007) procedure in which fish made 25 ‘choices’ before the stimuli positions were changed). The stimulus board was removed after a correct choice and reinserted for each trial for stimulus pair 5 (the inter-trial interval was about one minute). The stimulus board was not removed after incorrect choices, because the fish were allowed to continue to make choices until they tapped on the S+. The position of the stimuli remained the same within a given session for the entire experiment (i.e., the S+ remained in the same position for all five trials within a given session). For each test session, the position of the S+ was randomly determined prior to the experiment using a pseudorandom Gellerman series (Gellerman, 1933); however, the S+ was never shown more than three sessions in a row on the same side to prevent side biases from developing. Figure 2C shows a fish approaching the S+ for stimulus pair 1.

Previous research showed that learning success for fish engaged in similar visual object discrimination tasks occurred within 350-400 trials (Gierszewski, Bleckmann, & Schluessel, 2013; Schluessel et al., 2012; Sovrano & Bisazza, 2009). Thus, we presented the fish with 77 sessions (385 total trials) for each stimulus pair to ensure there would be an adequate number of trials to learn the discrimination. We also wanted to ensure that we had enough trials to ensure adequate statistical power of 0.75 with $\alpha = 0.05$ to detect a medium effect (Cohen’s $d = 0.5$) corresponding to an odds ratio of 2.5 using a logistic regression model and a two-sided Z-test for each stimulus pair (Hasselblad & Hedges, 1995) given that we had two subjects. The required sample size was 149 for each stimulus pair or 75 sessions per fish.

Experiment 1a. Each fish completed 77 sessions per stimulus pair with a few exceptions for failed sessions in which the fish stopped interacting with the stimulus board within the session time limit (fish 2 completed only 75 sessions during the test for stimulus pair 1, and fish 1 completed only 76 sessions during the test for stimulus pair 2). The stimulus pairs were presented in order, starting with pair 1, and ending with pair 5 without any breaks between stimulus pairs. Testing began on July 27, 2011 and ended on February 20, 2013 (each pair was tested over three to four months).

Experiment 1b. To address the possibility that the performance of the fish was influenced by order effects, we conducted a second set of tests. We presented 15 additional test sessions for stimulus pairs 1 and 2 twelve days after the completion of the last test session for stimulus pair 5. Fifteen sessions for stimulus pair 1 were followed immediately with 15 sessions for stimulus pair 2. The additional test sessions for pair 1 began on March 4, 2013 and ended on March 18, 2013, and the additional test sessions for pair 2 began on March 19, 2013 and ended on March 28, 2013. The additional test sessions for stimulus pair 1 began 15 months after the fish completed the first set of test sessions for pair 1, and the additional test sessions for stimulus pair 2 occurred 11.5 months after the fish completed the first set of test sessions for stimulus pair 2.

Data Analysis. Statistical analyses were performed using R v3.4.3 (R Core Team, 2017). An experimental type I error level of $\alpha = .05$ was used. The p -values were adjusted using Holm’s (1979) method to control the experimental error level at $\alpha = .05$, except in the case of post-hoc comparisons, for which p -values were adjusted using Tukey’s HSD (1949) method to control the experimental error at $\alpha = .05$. For both models presented below, logistic regression was used to determine if treatment factors (stimulus pair: 1, 2, 3, 4, 5; S+ position: right or left), differences between blocking factors (fish, test tank, trainer), or covariates (session number or trial number) affected discriminative accuracy as measured by a choice to the S+ (correct stimulus) or to the S- (incorrect stimulus). Model goodness-of-fit was assessed using likelihood ratio tests. Tests of overall accuracy against chance performance for an experiment were

performed using the predicted marginal probability of a correct choice to the S+ from the logistic regression model at the middle of testing (i.e., the first trial of session 40).

Only the first choice made by the fish on trial one of each five-trial session was analyzed. This is because the S+ remained in the same location throughout the session, and the subjects' choice on trials two to four could be influenced by the spatial position of the S+. For example, if the S+ was on the right, after the fish was reinforced for going to the right on trial one, it could adopt a spatial strategy of continuously choosing the stimulus on the right side for the remainder of the trials in that session. Because we used the 'correction method,' the fish could tap both the S+ and S- within each trial. Instead of counting all taps and reporting the percentage of taps to the S+ (e.g., Wyzisk & Neumeier, 2007), we adopted more stringent criteria and analyzed only the first tap on the first trial of each session.

Results

Experiment 1a. The overall accuracy after controlling for all variables was significantly better than chance (i.e., 50%), $Accuracy = 69.7\%$, $SE = 8.7\%$, $Z = 2.17$, $p = .0302$, 95% CI [52.6, 86.8]. The logistic regression model was a significantly better fit to the data than a model with no variables as suggested by a likelihood ratio test of model goodness-of-fit, $\chi^2(16) = 49.84$, $p < .0001$.

A sequential analysis of deviance using likelihood ratio tests was performed to test the statistical significance of main effects of session number, test tank, trainer, fish, S+ position, and stimulus pair as well as the significance of an interaction between time and stimulus pair. Adjusting all p -values using Holm's (1979) method, there was a statistically significant main effect of session number, $\chi^2(1) = 7.15$, $p = .0449$, and no statistically significant main effects for test tank, $\chi^2(3) = 1.72$, $p = 1$, trainer, $\chi^2(2) = 0.24$, $p = 1$, fish, $\chi^2(1) = 0.83$, $p = 1$, S+ position, $\chi^2(1) = 0.50$, $p = 1$, and stimulus pair, $\chi^2(4) = 8.72$, $p = .3427$.

The interaction between session number and stimulus pair was statistically significant, $\chi^2(4) = 15.08$, $p = .0318$. For stimulus pair 1, performance accuracy increased significantly across sessions, (Rate = 0.63% per trial, 95% CI [0.15, 1.11]). For stimulus pair 2, there was no significant change in accuracy across sessions, (Rate = 0.00% per trial, 95% CI [-0.34, 0.34]). For stimulus pair 3, performance accuracy significantly increased across sessions, (Rate = 0.30% per trial, 95% CI [0.02, 0.58]). For stimulus pair 4, performance accuracy increased significantly across sessions, (Rate = 0.42% per trial, 95% CI [0.08, 0.77]). Finally, performance did not change significantly across sessions for stimulus pair 5, (Rate = -0.26% per trial, 95% CI [-0.59, 0.08]). The interaction between session number and stimulus pair is shown in Figure 3. To summarize, performance improved significantly across sessions for stimulus pairs 1, 3, and 4 and remained unchanged for stimulus pairs 2 and 5.

There were significant differences in acquisition rate among the five stimulus pairs. The acquisition rate for stimulus pair 1 was significantly greater than the acquisition rate for stimulus pair 2, $Difference\ in\ Slopes\ on\ Log-Odds\ Scale = 0.0343$, $SE = 0.0154$, $z = 2.23$, $p = .0256$, and significantly greater than the acquisition rate for stimulus pair 5, $Difference\ in\ Slopes\ on\ Log-Odds\ Scale = 0.0455$, $SE = 0.0154$, $z = 2.96$, $p = .0031$. The acquisition rate for stimulus pair 3 was significantly greater than the acquisition rate for stimulus pair 5, $Difference\ in\ Slopes\ on\ Log-Odds\ Scale = 0.0315$, $SE = 0.0123$, $z = 2.55$, $p = .0106$. Finally, the acquisition rate for stimulus pair 4 was significantly greater than the acquisition rate for stimulus pair 5, $Difference\ in\ Slopes\ on\ Log-Odds\ Scale = 0.0296$, $SE = 0.0108$, $z = 2.74$, $p = .0061$. All other comparisons of acquisition rates were not statistically significant, all $p > .05$.

Post-hoc analysis of performance accuracy was conducted for the first trial of sessions 10, 40, and 70 for each stimulus pair. These analyses revealed that performance started at chance for stimulus pair 1 but was statistically significantly better than chance from the middle of testing to the end. For stimulus pair 2, performance did not differ from chance at all time points. For stimulus pair 3, performance was statistically significantly better than chance at all time points. For stimulus pair 4, performance was not significantly different from chance at the beginning and middle of testing, but was significantly better than chance at the end of testing. Finally, for stimulus pair 5, performance did not differ from chance at all time points. These results suggest that both fish were able to learn to discriminate between stimuli successfully for stimulus pairs 1, 3, and 4, but not for stimulus pairs 2 and 5. Performance at the end of

testing (session 70) is shown in Figure 4. Average accuracies and confidence intervals are shown in Table 2 for both fish on each stimulus pair and suggested that both fish performed very similarly to each other across the five stimulus pairs.

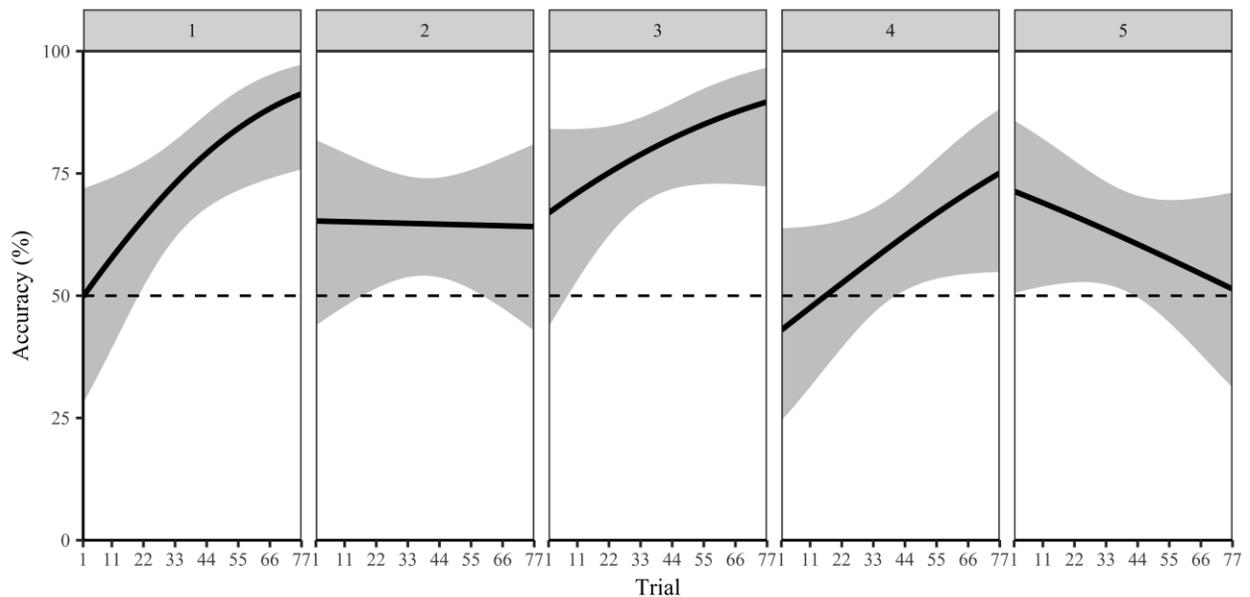


Figure 3. The performance accuracy of the fish in Experiment 1a shown for each stimulus pair 1-5. A significant interaction between session number (represented as trial number on the x-axis, since only the first trial per session was analyzed) and stimulus pair is shown. Model estimated accuracies are shown along with 95% Tukey-adjusted confidence intervals. Pair 1 = area matched, pair 2 = length matched, pair 3 = length and width matched, pairs 4 and 5 = diameter and area matched.

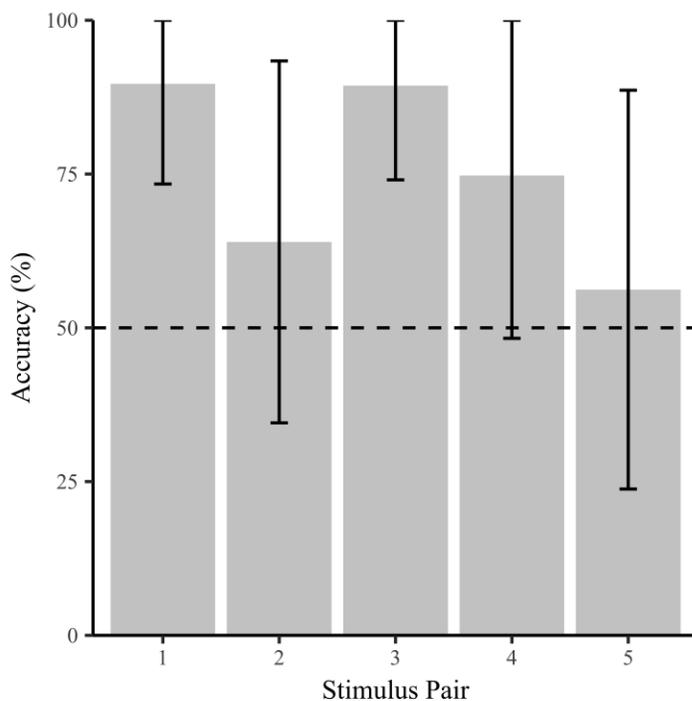


Figure 4. The performance accuracy of the fish in Experiment 1a. A main effect of stimulus pair is shown at the end of testing at trial 70. Model estimated accuracies are shown along with 95% Tukey-adjusted simultaneous confidence intervals. Pair 1 = area matched, pair 2 = length matched, pair 3 = length and width matched, pairs 4 and 5 = diameter and area matched.

Table 2

Experiment 1a Post-Hoc Contrasts for the Stimulus Pairs for each Fish

Fish	Stimulus Pair	Accuracy (%)	95% CI (%)
1	1	76.6	(56.1, 97.1)
	2	65.7	(39.9, 91.4)
	3	83.0	(64.7, 100.0)
	4	64.6	(36.6, 92.6)
	5	66.1	(38.4, 93.7)
2	1	73.9	(51.8, 96.1)
	2	62.3	(35.4, 89.3)
	3	80.9	(60.7, 100.0)
	4	61.2	(32.1, 90.4)
	5	62.8	(33.9, 91.6)

Note. Marginal discriminative accuracy is given with corresponding Tukey-adjusted 95% confidence intervals adjusting for multiple comparisons among stimulus pairs for a given fish.

Experiment 1b. The overall accuracy in Experiment 1b after controlling for all variables was significantly better than chance (i.e., 50%; Accuracy = 69.7%, $SE = 2.6\%$, $z = 6.77$, $p < .0001$, 95% CI [64.6, 74.8]). The logistic regression model was a significantly better fit to the data than a model with no variables as suggested by a likelihood ratio test of model goodness-of-fit, $\chi^2(5) = 23.05$, $p = .0003$.

As in Experiment 1a, a sequential analysis of deviance using likelihood ratio tests was performed to test the statistical significance of main effects of session number, fish, S+ position, and stimulus pair as well as the significance of an interaction between session number and stimulus pair. Adjusting all p -values using Holm's (1979) method, there were no statistically significant effects for session number, $\chi^2(1) = 2.94$, $p = .2588$, or stimulus pair, $\chi^2(1) = 0.11$, $p = .9696$. There was a statistically significant difference between the two fish, $\chi^2(1) = 7.49$, $p = .0248$; while both fish performed significantly better than chance, Fish 1 (Accuracy = 76.8%, 95% CI [70.1, 83.5]) performed significantly better than Fish 2 (Accuracy = 62.6%, 95% CI [55.0, 70.2]). There was also a statistically significant difference between when the S+ was on the left or right, $\chi^2(1) = 13.92$, $p = .0010$; while performance was significantly better than chance regardless of S+ position, performance was significantly better when the S+ was on the right (Accuracy = 79.7%, 95% CI [73.5, 86.0]) than when the S+ was on the left (Accuracy = 59.7%, 95% CI [51.6, 67.9]). Finally, the interaction between session number and stimulus pair was not significant, $\chi^2(1) = 0.49$, $p = .9696$.

Accuracies in Experiment 1a and Experiment 1b for both fish on stimulus pairs 1 and 2 are shown in Table 3. Whereas performance was not better than chance for stimulus pair 2 in Experiment 1a at the end of testing for either fish, performance was better than chance for both fish on both stimulus pairs in Experiment 1b. As shown by non-overlapping 84% confidence intervals in Table 3, the performance of Fish 2 on stimulus pair 1 was significantly better in Experiment 1a than in Experiment 1b, but did not differ significantly on stimulus pair 2 between experiments (Tryon, 2001). The performance of Fish 2 did not differ significantly between experiments on either stimulus pair.

Table 3

Comparison of Performance Accuracy in Experiment 1a and Experiment 1b

Stimulus Pair	Fish	Experiment 1a		Experiment 1b	
		Accuracy	84% CI	Accuracy	84% CI
1	1	90.3%	(81.9%, 98.8%)	78.3%	(72.7%, 83.8%)
	2	89.0%	(79.5%, 98.5%)	64.5%	(57.8%, 71.2%)
2	1	65.6%	(49.7%, 81.6%)	75.4%	(69.5%, 81.3%)
	2	62.3%	(45.7%, 79.0%)	60.8%	(53.9%, 67.6%)

Note. Accuracies and 84% confidence intervals are shown for Experiment 1a at the end of testing (session 70) and at the midpoint of testing in Experiment 1b (session 8). The 84% confidence intervals allow for direct comparison of performance for a stimulus pair and fish between experiments (no significant difference between experiments correspond to overlapping confidence intervals).

Experiment 2

We conducted a second experiment to ensure that the fish could learn to discriminate between stimuli with controlled features using either method. Studies on visual perception in fish have employed both the ‘correction method’ (e.g., Frech, Vogtsberger, & Neumeier, 2012; Mackintosh & Sutherland, 1963; Neumeier, 1992, 2003; Wyzisk & Neumeier, 2007) and the ‘non-correction method’ (e.g., Gierszewski et al., 2013; Schluessel et al., 2012, 2014; Siebeck et al., 2009; Sovrano & Bisazza, 2008). In Experiment 1, the ‘correction method’ was used, meaning that the fish could tap the S- as many times as necessary until they tapped the S+ and received a food reward to end the trial. In addition, the S+ remained in the same location throughout the session (although only the first choice in the first trial of each session was analyzed to eliminate potential choosing based on spatial location). In Experiment 2, the ‘non-correction method’ was used, meaning that the fish were allowed to choose only one stimulus on each trial (i.e., they could not tap both the S+ and the S- in the same trial). Further, the location of the S+ changed from trial to trial within a session (all trials in each session were analyzed).

Method

Subjects. The subjects were two experimentally-naïve goldfish (*Carassius auratus*), 4.0–5.5 cm in length, obtained commercially. Individual fish were identified by their phenotypic characteristics (but sex could not be determined phenotypically). These fish were not the same subjects used in Experiment 1. Subjects were housed together in a 10-gallon tank (50 cm long x 26 cm wide x 31 cm high). Tank contents, water temperature, aquarium care, diet, and feeding schedule were the same as in Experiment 1. The fish began the study after acclimating to their home tank for two days.

Stimuli. The stimuli were the same as in Experiment 1.

Experimental set-up. The experimental set-up was the same as in Experiment 1.

Procedure. Sessions were conducted once a day in the morning at approximately the same time each day on week days. There were typically five sessions per week. The test tank set-up and procedure were the same as in Experiment 1, with a few exceptions. First, the fish were allowed to acclimate to the test tank for a minimum of 15 min instead of 60 min (no adverse effect of a shorter acclimation time was detected in a different study; DeLong, Barbato, O’Leary, & Wilcox, 2017). Second, a single experimenter delivered the food reinforcement to the fish as well as recorded the data in between trials (instead of a separate recorder). The stimulus choice, number and location of taps to the stimulus, and total session time (rounded to the nearest minute) were recorded. Total session time was defined as the time the stimulus board was placed in the test tank to begin trial one to the time the board was removed at the end of trial five. Only one person served as experimenter for this study.

Pre-training. Pre-training proceeded similarly to Experiment 1. In the first session, the fish were placed in the test tank individually and allowed to habituate to the test tank. In subsequent sessions, the syringe was inserted into the tank at varied locations and the fish were reinforced for approaching the syringe and eating from it. The fish were reliably and promptly eating from the syringe in 19–30 sessions. These fish were initially more tentative about approaching the syringe in the correct location and less likely to explore the test tank than the fish in Experiment 1, so they required more sessions (Experiment 1 fish needed three to five sessions). Training started the day after pre-training was complete.

Training. The S+ used for the training phase was the same as in Experiment 1 (solid circle from stimulus pair 1). In this phase, the fish had to tap the S+ to receive food reinforcement. The experimenter used successive approximation to reinforce the fish for behaviors that brought them closer to tapping the S+ (e.g., holding the syringe near the stimulus card). However, unlike the fish in Experiment 1, these fish did not begin to spontaneously tap the S+ on trials where the syringe was not held in front of the S+ (after about 10 sessions of training). To induce them to tap the S+, a small drop of flake paste was applied directly to the S+ stimulus card for four to five sessions. The paste adhered to the card until the fish tapped the card and ate the food. For the following two to four sessions, food was applied directly to the S+ on some trials and food was delivered via the syringe on the other trials. The fish tapped the S+ for

both types of trials. Both fish reliably tapped on the S+ (without being cued in any trials by a drop of food on the S+) for five to seven training sessions before moving to the test phase. Testing started the day after the training phase was complete.

Testing. The test procedure was different from Experiment 1 in that the ‘non-correction method’ was used and the location of the S+ changed from trial to trial within a session. Each test session still contained five trials. However, the fish were allowed to choose only one stimulus on each trial (they could not tap both the S+ and the S- in the same trial), and the stimulus board was removed between each trial. The location of the S+ was determined prior to the experiment using a pseudorandom Gellerman series (Gellerman, 1933); however, the S+ was never shown more than two trials in a row on the same side to prevent side biases from developing. The S+ appeared on the right and left an equal number of times across two five-trial sessions. These modifications ensured that the fish could not select the S+ based on a spatial position cue after the first trial. Data from all five trials in each session were analyzed (instead of only the first trial). Test sessions lasted an average of 3.9 min (*Minimum* = 1 min, *Median* = 4 min, *Maximum* = 20 min).

The number of test sessions and timeframe for Experiment 2 was different from Experiment 1. In Experiment 1, there were 75-77 sessions per stimulus pair (only the first trial was analyzed for each session, thus 75-77 total trials per pair). The entire test phase lasted 19 months (each pair was tested over a three to four month period). In Experiment 2, there were 16 sessions per stimulus pair (with all five trials analyzed for a total of 80 trials per pair). The entire test phase lasted about four months (each pair was completed in three and a half weeks). As in Experiment 1, the stimulus pairs were presented in order, starting with stimulus pair 1, and ending with stimulus pair 5 without any breaks between stimulus pairs. Testing began on August 12, 2016 and ended on December 14, 2016. Testing was continuous except for one break between November 23, 2016 and December 4, 2016 during stimulus pair 5.

Data Analysis. The same statistical analyses were performed as in Experiment 1, except all trials were analyzed.

Results

The overall accuracy in the second experiment after controlling for trial number was significantly better than chance (i.e., 50%), *Accuracy* = 62.3%, *SE* = 1.7%, *Z* = 6.85, *p* < .0001. The logistic regression model was a significantly better fit than a model without any variables as suggested by a likelihood ratio test of model goodness-of-fit $\chi^2(16) = 111.95$, *p* < .0001. A sequential analysis of deviance using likelihood ratio tests was performed to test the statistical significance of main effects of trial number, test tank, fish, S+ position, and stimulus pair as well as the significance of an interaction between trial number and stimulus pair. Adjusting all *p*-values using Holm’s (1979) method, there were no statistically significant main effects of trial number, $\chi^2(1) = 3.32$, *p* = .34, test tank, $\chi^2(5) = 4.36$, *p* = 1, fish, $\chi^2(1) = 0.08$, *p* = 1, or S+ position, $\chi^2(1) = 0.40$, *p* = 1. The interaction between trial number and stimulus pair was not statistically significant, $\chi^2(4) = 1.22$, *p* = 1.

There were statistically significant differences in accuracy among stimulus pairs, $\chi^2(4) = 19.90$, *p* = .0031, after controlling for the other effects (see Figure 5). Controlling for all other variables and time by comparing marginal accuracies at the end of testing (trial 70), performance was significantly better than chance for three of the five stimulus pairs. Post-hoc analyses of the main effect of stimulus pairs revealed that performance on the first stimulus pair (*Accuracy* = 85.4%, 95% CI [75.4, 97.6]), second stimulus pair (*Accuracy* = 76.4%, 95% CI [61.0, 91.8]), and third stimulus pair (*Accuracy* = 84.3%, 95% CI [72.8, 95.8]), were significantly better than chance while performance on the fourth (*Accuracy* = 48.7%, 95% CI [32.0, 65.4]), and fifth stimulus pairs (*Accuracy* = 50.2%, 95% CI [32.9, 67.4]) were not significantly better than chance. Performance did not differ significantly among the first three stimulus pairs and did not differ significantly between the fourth and fifth stimulus pairs. Performance was significantly better on the first three stimulus pairs than it was on the fourth and fifth stimulus pairs.

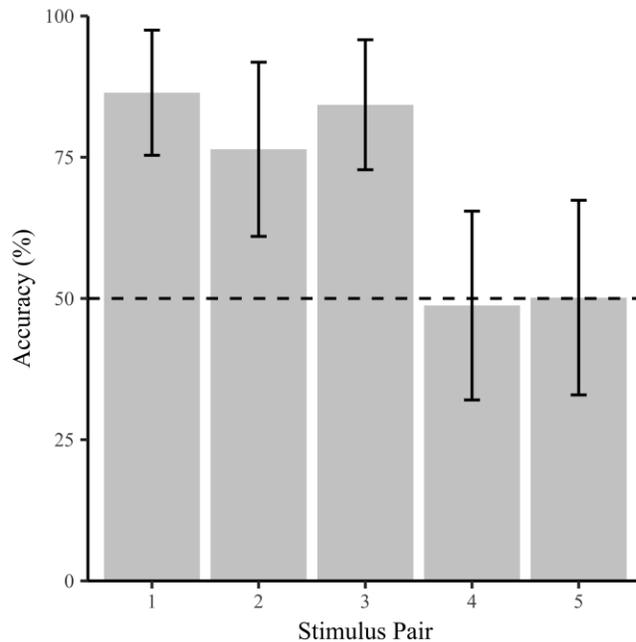


Figure 5. The performance accuracy of the fish in Experiment 2. A main effect of stimulus pair is shown at the end of testing at trial 70. Model estimated accuracies are shown along with 95% Tukey-adjusted simultaneous confidence intervals. Pair 1 = area matched, pair 2 = length matched, pair 3 = length and width matched, pairs 4 and 5 = diameter and area matched.

Although the main effect of trial number was not statistically significant, a post-hoc analysis of performance across trials (a contrast at trial 10, trial 40, and trial 70) revealed that discriminative accuracy improved across trials. The average accuracy at trial 10 for a stimulus pair was not statistically significantly different from chance, *Accuracy* = 54.3%, 95% CI (47.7, 61.0). However, average performance improved by the middle of testing (trial 40) to be significantly better than chance, *Accuracy* = 62.2%, 95% CI (58.2, 66.1). By the end of testing (trial 70), average accuracy was significantly better than chance and significantly better than accuracy at the beginning of testing, *Accuracy* = 69.2%, 95% CI (63.2, 75.2). Figure 6 shows the performance across trials for each of the five stimulus pairs. Average accuracies and confidence intervals are shown in Table 4 for both fish on each stimulus pair and suggested that both fish performed very similarly to each other across the five stimulus pairs as in Experiment 1.

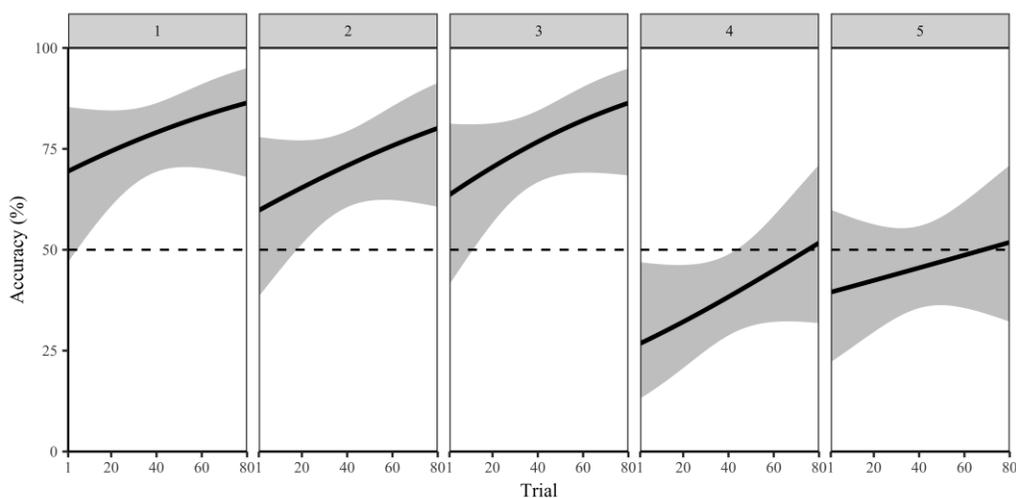


Figure 6. The performance accuracy of the fish in Experiment 2 shown for each stimulus pair 1-5. Because the interaction between trial number and stimulus pair was not significant, none of the slopes were significantly different from each other. Model estimated accuracies are shown along with 95% Tukey-adjusted confidence intervals. Pair 1 = area matched, pair 2 = length matched, pair 3 = length and width matched, pairs 4 and 5 = diameter and area matched.

Table 4

<i>Experiment 2 Post-hoc Contrasts for the Stimulus Pairs for each Fish</i>			
Fish	Stimulus Pair	Accuracy (%)	95% CI (%)
3	1	79.4	(70.4, 88.4)
	2	70.1	(58.7, 81.5)
	3	76.6	(67.1, 86.2)
	4	38.0	(26.8, 49.2)
	5	44.9	(33.3, 56.6)
4	1	80.2	(71.1, 89.2)
	2	71.1	(59.5, 82.7)
	3	77.5	(68.1, 86.8)
	4	39.1	(27.8, 50.4)
	5	46.1	(34.3, 57.9)

Note. Marginal discriminative accuracy is given with corresponding Tukey-adjusted 95% confidence intervals adjusting for multiple comparisons among stimulus pairs for a given fish.

General Discussion

The goal of this study was to assess the ability of goldfish to discriminate 2D geometric and complex stimulus pairs while various stimulus features were controlled (e.g., length, width, area, diameter, symmetry). Our results show that the goldfish in both experiments were able to successfully discriminate between the geometric shapes (circle vs. rectangle or square) in stimulus pairs 1-3. Only the fish in Experiment 1 learned to discriminate between the complex shapes in stimulus pair 4 (single circle outline vs. four-blade windmill). The fish in both experiments failed to discriminate between the complex shapes in stimulus pair 5 (double circle outline vs. six-blade windmill).

The stimuli in the current study were designed such that different stimulus features (area, length, width, diameter, symmetry) were matched between stimuli in each stimulus pair so that none of these features could be used across all five pairs (see Table 1). There were potentially other features, such as the presence or absence of straight lines – that were consistent across all five stimulus pairs. Sutherland (1969) suggested that fish rely on the presence vs. absence of oblique lines and the presence vs. absence of vertical and horizontal lines in stimuli. The S+ (solid or unfilled circles) were constructed of only curved lines whereas the S- (rectangle, square, four-blade windmill, six-blade windmill) contained only straight lines. This single cue would allow successful discrimination performance for all stimulus pairs. If the fish had attended only to that single feature throughout the study (or another single feature that differentiates all the negative stimuli from all the positive stimuli such as the presence vs. absence of corners), there should have been no significant difference in accuracy between the five stimulus pairs. Because the fish were unable to discriminate between stimuli in pair 5, one single feature either was not utilized for all discriminations or used in combination with other features but insufficient on its own.

Assuming that the fish did not find one feature that was effective for all five stimulus pairs, they could have used a number of different strategies: (A) use of a single feature that changed from stimulus pair to stimulus pair, (B) use of a combination of features that changed from stimulus pair to stimulus pair, or (C) holistic processing. If the fish used a single feature that changed for each stimulus pair, they could have used length for stimulus pair 1, then switched to using width for stimulus pair 2, then switched to using area or perimeter for stimulus pair 3 (many other features were available as well). If the fish used a combination of features that changed from stimulus pair to stimulus pair, they could have used length and width for stimulus pair 1, then switched to using surface area and width for stimulus pair 2, then switched to using area and perimeter for stimulus pair 3. The same features used for stimulus pairs 1-3 could not be used for stimulus pairs 4-5 (except perimeter), but there are other features the fish could have extracted from those stimuli (e.g., the number of individual elements contained within each stimulus, local luminance differences in a certain location within the stimulus).

Alternatively, the goldfish could have been using holistic shape processing instead of attending to local stimulus features. Minini and Jeffery (2006) defined holistic processing as using information

extracted from two dimensions simultaneously, i.e., the aspect ratio (ratio of height to width). Using aspect ratio would work to discriminate between the shapes in stimulus pairs 1 and 2, but would not work for stimulus pairs 3-5. A different way to assess holistic processing is through the lens of visual illusions. Organisms that perceive illusory contours have visual systems that promote the processing of objects as wholes and underlying neural mechanisms that represent object boundaries even when they are not present in the sensory input (Rosa Salva, Sovrano, & Vallortigara, 2014). Fish perceive a wide variety of illusions (Agrillo, Miletto Petrazzini, & Dadda, 2013; Rosa Salva et al., 2014). Goldfish, redbtail splitfins (*Xenotoca eiseni*), and bamboo sharks (*Chiloscyllium griseum*) perceive illusory contours in Kanizsa figures (Fuss, Bleckmann, & Schluessel, 2014a; Sovrano & Bisazza, 2009; Wyzisk & Neumeyer, 2007). Although no Kanizsa figures were used in the current study, the fish may have perceived an illusory figure of a circle while looking at the negative stimuli for stimulus pairs 4 and 5. An outline of a circle can be perceived if one joins together the outer edges of the fan blades. This effect is particularly strong for stimulus pair 5 with six fan blades. If the fish perceived an illusory circle contour when viewing the negative stimuli, then they would have difficulty discriminating between the negative and positive stimuli. This could explain their poor performance on stimulus pair 5. To summarize, the goldfish could have used a strategy focusing on local stimulus features or holistic processing. To gain more insight into strategies used by fish during shape perception, future studies are needed that include stimulus pairs that manipulate stimuli so that to succeed fish must use holistic processing as was done was rats in Minini and Jeffery's (2006) study. Poor performance would indicate a failure to process global shape information.

The difference in accuracy over time across the two experiments may provide some insights into their strategies for the discrimination task. In Experiment 1, performance improved significantly over time for stimulus pairs 1, 3, and 4 but did not change over time for stimulus pairs 2 and 5 (Figure 3). Whatever the fish learned during stimulus pair 1 did not appear to help them during stimulus pair 2, which could mean they were focusing on the length of the stimuli (which was equivalent between stimuli in pair 2). Their accuracy during stimulus pair 3 improved across trials and ended very high, which means they found effective cues (perhaps surface area). What they used as discriminatory cues for stimulus pair 3 did not appear to work for when they first encountered stimulus pair 4, but, by the end of the 77 sessions, they had devised an effective strategy. However, the cues they used for stimulus pair 4 apparently did not work for stimulus pair 5. The pattern of performance shown in Figure 3 may indicate that the Experiment 1 fish solved each new discrimination problem in context of the previous stimulus pair. In contrast, the Experiment 2 fish always began each stimulus pair with performance at chance and improved their performance over time (Figure 6). Whatever the fish in Experiment 2 learned to use for one stimulus pair did not seem to carry over to the following pair. An alternative explanation is that the fish in Experiment 2 started with an ineffective strategy from the previous stimulus pair but learned more or better cues to use across the 80 trials.

The better performance of the fish in the current study on stimulus pairs 1-3 compared to stimulus pairs 4 and 5 may be a product of the training method we employed, which may have affected the strategy they used during the test. During the training phase, the fish were exposed to the S+ for stimulus pairs 1-3 (solid black circle). Once they learned to tap the solid black circle, they began the test phase for pair 1. Upon completing the test phase for pair 1, they were immediately presented with the test phases for pairs 2-5 without any further training. Thus, their superior performance on pairs 1-3 may be attributed to their exposure to the S+ for pairs 1-3 during training. Had they been previously exposed to the S+ for each stimulus pair prior to the test phase for each pair, they may have performed differently on pairs 4-5.

We explored the ability of goldfish to discriminate between carefully controlled shape stimuli using two methods. We used the 'correction method' in Experiment 1 and the 'non-correction method' in Experiment 2. In the 'correction method,' incorrect choices during trials were allowed to be corrected (although only the first choice on the first trial of a session was analyzed) and the S+ remained in the same location within a session. In the 'non-correction method,' the fish were not allowed to correct their first choice (all trials analyzed) and location of the S+ changed from trial-to-trial. In both methods, the fish were reinforced only after a correct choice, and across a set of two sessions the S+ was viewed equally often on the right and left. The 'non-correction' method is more efficient in that data can be

collected quicker, but the disadvantage is that sometimes fish will develop a side bias (Gierszewski et al., 2013; Schluessel et al., 2014).

As predicted, we found that the goldfish were able to discriminate between the shapes using either method. The goldfish tested with both methods performed similarly in that they successfully discriminated stimulus pairs 1-3 but did not succeed with stimulus pair 5. Because of the low sample size in this study, it is not possible to directly compare the performance of the fish in Experiment 1 and 2. Our implementation of the ‘correction method’ did take more time overall than when we used the ‘non-correction method’ (Experiment 1 lasted approximately a year and a half whereas Experiment 2 lasted only four months), although the individual test sessions were not substantially longer using the ‘correction method.’ While we did not observe any side biases in using the ‘correction method’ in Experiment 1a or the ‘non-correction method’ in Experiment 2, we did observe a side bias using the ‘correction method’ in Experiment 1b. Side biases have been found in other visual discrimination studies with fish (e.g., Colwill, Raymond, Ferreira, & Escudero, 2005; Gierszewski et al., 2013; Graeber & Ebbesson, 1972; Schluessel & Bleckmann, 2005; Schluessel et al., 2014). It would be useful to replicate the current study with more subjects in each experiment so that the two methods can be directly compared, as we are not aware of any studies that have done so using the same species and the same stimuli (Siebeck et al. (2009) used both methods, but with different pairs of stimuli).

A limitation of the current study is that we always presented the stimulus pairs in the same order (1-5), so that we could replicate the order used in Siebeck et al.’s (2009) study. The failure of the goldfish on pair 5 could be due to order effects. To address this possibility, we conducted Experiment 1b, in which the fish completed an additional 15 test sessions with stimulus pairs 1 and 2 after the completion of stimulus pair 5 in Experiment 1a. There was no difference in performance between Experiment 1a and Experiment 1b for Fish 1, but Fish 2 performed significantly worse on stimulus pair 1 in Experiment 1b. Thus, order of presentation may have had an effect on performance, at least for one fish. Interestingly, both fish failed to discriminate the stimuli in pair 2 during Experiment 1a, but succeeded with stimulus pair 2 during Experiment 1b. The order in which stimulus pairs are presented to the fish could have an impact on how they process shape information, utilize features, and develop strategies to discriminate shapes. If the first stimulus pair they encounter is equated for surface area, they may learn not to rely on that feature, whereas if surface area is available as a cue in the first pair, they may default to using it on subsequent stimulus pairs. A future study could present stimulus pairs in different orders to different groups of fish to investigate this possibility.

Another limitation of the current study is that only two to three goldfish participated in each experiment. We should use caution in interpreting results from few subjects because within-species differences between fish can occur (Gierszewski et al., 2013; Lucon-Xiccato, Dadda, & Bisazza, 2016; Miletto Petrazzini & Agrillo, 2016; Schluessel et al., 2012, 2015; Siebeck et al., 2009; Wyzisk & Neumeyer, 2007). However, other studies have utilized two to five goldfish per experiment or test (Douglas, Eva, & Guttridge, 1988; Freche et al., 2012; Neumeyer, 1992, 2003; Wyzisk & Neumeyer, 2007). Some studies utilize larger groups of goldfish (six or more per experiment, as in McAroe et al., 2016 or Onishi, 1991). The strategy in the current study was to test fewer subjects but over an extended period of time with more test trials than are typically included in fish research (see DeLong et al., 2017). Because it takes a larger investment of time to train and test subjects in this manner, a low sample size is conventional. This strategy of lower sample size combined with extensive training and testing has been used more often with nonhuman primates, marine mammals, and birds (e.g., Beran, 2007; Jaakkola, Fellner, Erb, Rodriguez, & Guarino, 2005; Pepperberg, 2006; Tomonaga, 2008; Yaman, Kilian, von Fersen, & Gunturkun, 2012) than with fish. Even though we compensated for fewer subjects with more trials, it is still important to replicate these results with additional goldfish subjects in future research.

Our goldfish performed differently in this task than Siebeck et al.’s (2009) coral reef damselfish *Pomacentrus amboinensis*. Whereas the damselfish succeeded on all stimulus pairs, our fish successfully discriminated between the geometric shapes (stimulus pairs 1-3), but did not successfully discriminate between the complex shapes in stimulus pair 5 (and only the goldfish in Experiment 1 succeeded with the complex shapes in stimulus pair 4). We were not aiming to directly replicate Siebeck et al.’s (2009) study

and we had differences in subjects, stimuli, and methods, so differences in performance are not surprising. Our stimulus pairs 3-5 were the same as theirs (except our windmill blades had straight edges and theirs were curved; see Table 1). All our stimuli were 2D, whereas their stimulus pairs 1 and 2 were 3D (see Table 1 in Siebeck et al., 2009). Goldfish performed significantly better with 3D stimuli vs. 2D stimuli when discriminating among rotated objects, so that aspect of the stimuli may have affected their performance in the current study as well (DeLong, Fobe, Wilcox, & O'Leary, 2017; DeLong, Gardner, Rachfal, Morrison, & Wilcox, 2018). All of our fish viewed all stimulus pairs, whereas Siebeck et al. (2009) tested a separate group of fish on the complex stimulus pairs. Perhaps learning to discriminate between the geometric shapes first influenced the way our goldfish performed with the complex shapes. It is very likely that goldfish and reef fish may have used different strategies on this task since they are distantly related and live in completely different habitats. Reef fish live in a colorful and densely-populated ocean environment characterized by high biodiversity whereas goldfish live in shallow slow-moving freshwater ponds and streams (Froese & Pauly, 2018). In addition, the individual subjects used in Siebeck et al.'s (2009) study were freshly-caught reef fish who had recently experienced their natural environment whereas the subjects in the current study were bred in captivity, purchased at age one month from a commercial aquarium, and kept in standard tanks with gravel and a single plant.

A number of fish species from a wide range of families appear to be able to discriminate between shape stimuli (e.g., Gierszewski et al., 2013; Herter, 1953; Northmore, Volkmann, & Yeager, 1978; Schluessel et al., 2012). These studies suggest that visual shape recognition may be a fundamental ability present in a variety of fish species who live in diverse habitats. In order to determine the salient cues fish use to discriminate among shapes, carefully controlled stimuli should be utilized. The current study provided preliminary evidence for discrimination of shape stimuli with controlled features in goldfish, but a further study with more subjects is required to examine whether goldfish use global shape or individual stimulus features.

Compliance with Ethical Standards

The experiments described here were reviewed by the Rochester Institute of Technology Institutional Animal Care and Use Committee (IACUC) and comply with the current laws of the country (USA) in which they were performed. All applicable international, national, and institutional guidelines for the care and use of animals were followed. The authors declare they have no conflict of interest.

Acknowledgments

This project was supported by the RIT Paul A. and Francena L. Miller Research Fellowship and a grant from the RIT College of Liberal Arts Faculty Research Fund to CMD. We thank Amanda Heberle, Brigid Colgan, Kayla Mata, Amy Lonczak, and Brittany Priddy for their assistance with the experimental sessions and maintenance of the animals and aquaria.

References

- Agrillo, C., Dadda, M., & Bisazza, A. (2007). Quantity discrimination in female mosquitofish. *Animal Cognition*, *10*, 63–70.
- Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2008). Do fish count? Spontaneous discrimination of quantity in female mosquitofish. *Animal Cognition*, *11*, 495–503.
- Agrillo, C., Dadda, M., Serena, G., & Bisazza, A. (2009). Use of number by fish. *PloS ONE*, *4*, e4786.
- Agrillo, C., Miletto Petrazzini, M. E., & Dadda, M. (2013). Illusory patterns are fishy for fish, too. *Frontiers in Neural Circuits*, *7*, 1–3.
- Agrillo, C., Piffer, L., Bisazza, A., & Butterworth, B. (2012). Evidence for two numerical systems that are similar in humans and guppies. *PLoS One*, *7*, e31923.
- Aronson, L. R. (1971). Further studies on orientation and jumping behavior in the goby fish, *Bathygobius soporator*. *Annals of the New York Academy of Science*, *188*, 378–392.

- Beran, M. J. (2007). Rhesus monkeys (*Macaca mulatta*) enumerate large and small sequentially presented sets of items using analog numerical representations. *Journal of Experimental Psychology: Animal Learning and Cognition*, 33, 42–54.
- Bogale, B. A., & Sugita, S. (2013). Shape discrimination and concept formation in the jungle crow (*Corvus macrorhynchos*). *Animal Cognition*, 17, 105–111.
- Bowman, R., & Sutherland, N. S. (1969). The discrimination of ‘W’ and ‘V’ shapes by goldfish. *Quarterly Journal of Experimental Psychology*, 21, 69–76.
- Bowman, R. S., & Sutherland, N. S. (1970). Shape discrimination by goldfish: Coding of irregularities. *Journal of Comparative and Physiological Psychology*, 72, 90–97.
- Brown, C. (2015). Fish intelligence, sentience, and ethics. *Animal Cognition*, 18, 1–17.
- Brown, C., Laland, K., & Krause, J. (2011). *Fish cognition and behavior*. Oxford, UK: Wiley-Blackwell.
- Bshary, R., Wickler, W., & Fricke, H. (2002). Fish cognition: A primate's eye view. *Animal Cognition*, 5, 1–13.
- Cogan, D., Inmam, S., & Gambrel, M. (1971). Discrimination in parakeets (*Melopsittacus undulatus*) as a function of age. *Psychological Record*, 21, 117–120.
- Cogan, D., Inmam, S., & Gambrel, M. (1971). Discrimination in parakeets (*Melopsittacus undulatus*) as a function of age. *Psychological Record*, 21, 117–120.
- Colwill, R. M., Raymond, M. P., Ferreira, L., & Escudero, H. (2005). Visual discrimination learning in zebrafish (*Danio rerio*). *Behavioural Processes*, 70, 19–31.
- Daniel, J. C., & Mikulka, P. J. (1998). Discrimination learning in the white rhinoceros. *Applied Animal Behavior Science*, 58, 197–202.
- DeLong, C. M., Barbato, S., O’Leary, T., & Wilcox, K. T. (2017). Small and large number discrimination in goldfish (*Carassius auratus*) with extensive training. *Behavioural Processes*, 141, 172–183.
- DeLong, C. M., Fobe, I., Wilcox, K. T., & O’Leary, T. (2017, April). *Visual discrimination of rotated 2D objects in goldfish (Carassius auratus)*. Paper presented at the 24th Annual International Conference on Comparative Cognition, Melbourne, FL.
- DeLong, C. M., Gardner, K., Rachfal, H., Morrison, E., & Wilcox, K. T. (2018, April). *Visual discrimination of rotated 3D objects in goldfish (Carassius auratus)*. Paper presented at the 25th Annual International Conference on Comparative Cognition, Melbourne, FL.
- Douglas, R. H., Eva, J., & Guttridge, N. (1988). Size constancy in goldfish (*Carassius auratus*). *Behavioral Brain Research*, 30, 37–42.
- Frech, B., Vogtberger, M., & Neumeyer, C. (2012). Visual discrimination of objects differing in spatial depth by goldfish. *Journal of Comparative Physiology A*, 198, 53–60.
- Froese, R. & D. Pauly. (Ed). (2018). FishBase. World Wide Web electronic publication. Retrieved from <http://www.fishbase.org>
- Fuss, T., Bleckmann, H., & Schluessel, V. (2014a). The brain creates illusions not just for us: Sharks (*Chiloscyllium griseum*) can “see the magic” as well. *Frontiers in Neural Circuits*, 8, 24.
- Fuss, T., Bleckmann, H., & Schluessel, V. (2014b). Visual discrimination abilities in the gray bamboo shark (*Chiloscyllium griseum*). *Zoology* 17, 104–111.
- Fuss, T., & Schluessel, V. (2015). Something worth remembering: Visual discrimination in sharks. *Animal Cognition*, 18, 463–471.
- Gellerman, L. W. (1933). Chance orders of alternating stimuli in visual discrimination experiments. *Journal of Genetic Psychology*, 42, 206–208.
- Gierszewski, S., Bleckmann, H., & Schluessel, V. (2013). Cognitive abilities in Malawi cichlids (*Pseudotropheus* sp.): Matching-to-sample and image/mirror-image discriminations. *PLoS ONE*, 8(e57363).
- Graeber, R. C., & Ebbesson, S. O. (1972). Visual discrimination learning in normal and tectal-ablated nurse sharks (*Ginglymostoma cirratum*). *Comparative Biochemistry and Physiology*, 42, 131–139.
- Harley, H. E., Roitblat, H. L., & Nachtigall, P. E. (1996). Object representation in the bottlenose dolphin (*Tursiops truncatus*): Integration of visual and echoic information. *Journal of Experimental Psychology: Animal Behavior Processes*, 22, 164–174.
- Hasselblad, V., & Hedges, L. V. (1995). Meta-analysis of screening and diagnostic tests. *Psychological Bulletin*, 117(1), 167–178.
- Herter, K. (1953). *Die Fischdressuren und ihre sinnesphysiologischen Grundlagen*. Berlin: Akademie Verlag.
- Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 6, 65–70.

- Jaakkola, K., Fellner, W., Erb, L., Rodriguez, M., & Guarino, E. (2005). Understanding of the concept of numerically “less” by bottlenose dolphins (*Tursiops truncatus*). *Journal of Comparative Psychology*, *119*, 296–303.
- James, W. T., & McFarland, J. (1966). A study of form discrimination in the opossum. *Journal of Psychology*, *64*, 193–198.
- Lopez, J. C., Broglio, C., Rodriguez, F., Thinus-Blanc, C., & Salas, C. (1999). Multiple spatial learning strategies in goldfish (*Carassius auratus*). *Animal Cognition*, *2*, 109–120.
- Lucon-Xiccato, T., & Dadda, M. (2014). Assessing memory in zebrafish using the one-trial test. *Behavioural Processes*, *106*, 1–4.
- Lucon-Xiccato, T., Dadda, M., & Bisazza, A. (2016). Sex differences in discrimination of shoal size in the guppy (*Poecilia reticulata*). *Ethology*, *122*, 481–491.
- Mackintosh, J., & Sutherland, N. S. (1963). Visual discrimination by the goldfish: The orientation of rectangles. *Animal Behavior*, *11*, 135–141.
- Mauck, B., & Dehnhardt, G. (1997). Mental rotation in a California sea lion (*Zalophus californianus*). *Journal of Experimental Biology*, *200*, 1309–1316.
- McAroe, C. L., Craig, C. M., & Holland, R. A. (2016). Place vs. response learning in fish: A comparison between species. *Animal Cognition*, *19*, 153–161.
- Miletto Petrazzini, M. E., & Agrillo, C. (2016). Turning to the larger shoal: Are there individual differences in small-and large-quantity discrimination of guppies? *Ethology Ecology & Evolution*, *28*, 211–220.
- Minini, L., & Jeffery, K. J. (2006). Do rats use shape to solve “shape discriminations?” *Learning & Memory*, *13*, 287–297.
- Neumeyer, C. (1992). Tetrachromatic colour vision in goldfish: Evidence from colour mixture experiments. *Journal of Comparative Physiology A*, *171*, 639–649.
- Neumeyer, C. (2003). Wavelength dependence of visual acuity in goldfish. *Journal of Comparative Physiology A*, *189*, 811–821.
- Newport, C., Wallis, G., Reshitnyk, Y., & Siebeck, U. (2016). Discrimination of human faces by archerfish (*Toxotes chatareus*). *Nature Scientific Reports*, *6* (27523), 1–7.
- Northmore, D., Volkman, F. C. & Yager, D. (1978). Vision in fishes: Color and pattern. In D. I. Mostofsky (Ed.), *Behaviour of fish and other aquatic animals*. New York: Academic Press.
- Oldfield-Box, H. (1969). On analyzing the formation of learning-sets in young and old rats. *Gerontologia*, *15*, 302–307.
- Ohnishi, K. (1991). Goldfish visual information-processing patterns in food-reinforced discrimination-learning between compound visual stimuli. *Journal of Comparative Physiology A*, *168*, 581–589.
- Pepperberg, I. (2006). Gray parrot numerical competence: A review. *Animal Cognition*, *9*, 377–391.
- Pollard, J. S., Beale, I. L., Lysons, A. M., & Preston, A. C. (1967). Visual discrimination in the ferret. *Perceptual and Motor Skills*, *24*, 279–282.
- Pretot, L., Bshary, R., & Brosnan, S. F. (2016). Factors improving the different performance of fish and primates on a dichotomous choice task. *Animal Behaviour*, *119*, 189–199.
- R Core Team (2017). *R: A language and environment for statistical computing* (3.4.3) [programming language]. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Rosa Salva, O., Sovrano, V. A., & Vallortigara, G. (2014). What can fish brains tell us about visual perception? *Frontiers in Neural Circuits*, *8*, 119.
- Salwiczek, L. H., Pretot, L., Demarta, L., Proctor, D., Essler, J., Pinto, A. I.,...& Bshary, R. (2012). Adult cleaner wrasse outperform capuchin monkeys, chimpanzees, and orangutans in a complex foraging task derived from cleaner-client reef fish cooperation. *PLoS ONE*, *7*, e49068.
- Schluessel, V., & Bleckmann, H. (2005). Spatial memory and orientation strategies in the elasmobranch *Potamotrygon motoro*. *Journal of Comparative Physiology A*, *191*, 695–706
- Schluessel, V., Fricke, G., & Bleckmann, H. (2012). Visual discrimination and object categorization in the cichlid *Pseudotropheus* sp. *Animal Cognition*, *15*, 525–537.
- Schluessel, V., Kortekamp, N., Ortiz Cortes, J. A., Klein, A., & Bleckmann, H. (2015) Perception and discrimination of movement and biological motion patterns in fish. *Animal Cognition*, *18*, 1077–1091.
- Schluessel, V., Kraniotakes, H., & Bleckmann, H. (2014). Visual discrimination of rotated 3D objects in Malawi cichlids (*Pseudotropheus* sp.): A first indication for form constancy in fishes. *Animal Cognition*, *17*, 359–371.
- Schusterman, R. J., & Thomas, T. (1966). Shape discrimination and transfer in the California sea lion. *Psychonomic Science*, *5*, 21–22.

- Siebeck, U. E., Litherland, L., & Wallis, G. M. (2009). Shape learning and discrimination in reef fish. *Journal of Experimental Biology*, *212*, 2113–2119.
- Sovrano, V. A., & Bisazza, A. (2008). Recognition of partly occluded objects by fish. *Animal Cognition*, *11*, 161–166.
- Sovrano, V. A., & Bisazza, A. (2009). Perception of subjective contours in fish. *Perception*, *38*, 579–590.
- Srinivasan, M. V. (2010). Honeybees as a model for vision, perception, and cognition. *Annual Review of Entomology*, *55*, 267–284.
- Sutherland, N. S. (1959). Visual discrimination of shape by octopus. Circles and squares, and circles and triangles. *Quarterly Journal of Experimental Psychology*, *11*, 24–32.
- Sutherland, N. S. (1969). Shape discrimination in rat, octopus, and goldfish: A comparative study. *Journal of Comparative and Physiological Psychology*, *67*, 160–176.
- Sutherland, N. S., & Bowman, R. (1969). Discrimination of circles and squares with and without knobs by goldfish. *Quarterly Journal of Experimental Psychology*, *21*, 330–338.
- Tanaka, K. (2000). Mechanisms of object recognition studied in monkeys. *Spatial Vision*, *13*, 147–163.
- Tomonaga, M. (2008). Relative numerosity discrimination by chimpanzees (*Pan troglodytes*): Evidence for approximate numerical representations. *Animal Cognition*, *11*, 43–57.
- Truppa, V., Sovrano, V. A., Spinozzi, G., & Bisazza, A. (2010). Processing of visual hierarchical stimuli by fish (*Xenotoca eiseni*). *Behavioral Brain Research*, *207*, 51–60.
- Tryon, W. W. (2001). Evaluating statistical difference, equivalence, and indeterminacy using inferential confidence intervals: An integrated alternative method of conducting null hypothesis statistical tests. *Psychological Methods*, *6*, 371–386. doi:10.1037//1082-989X.6.4.371
- Tukey, J. (1949). Comparing individual means in the analysis of variance. *Biometrics*, *5*, 99–114.
- von Fersen, L., Schall, U., & Gunturkun, O. (2000) Visual lateralization of pattern discrimination in the bottlenose dolphin (*Tursiops truncatus*). *Behavioral Brain Research*, *107*, 177–181.
- Wyzisk, K., & Neumeyer, C. (2007). Perception of illusory surfaces and contours in goldfish. *Visual Neuroscience* *24*, 291–298.
- Yaman, S., Kilian, A., von Fersen, L., & Gunturkun, O. (2012). Evidence for a numerosity category that is based on abstract qualities of "few" vs. "many" in the bottlenose dolphin (*Tursiops truncatus*). *Frontiers in Psychology*, *3*(00473).
- Young, M. E., Peissig, J. J., & Wasserman, E. A. (2001). Discrimination of geons by pigeons: The effects of variations in surface descriptions. *Animal Learning and Behavior*, *29*, 97–106.
- Zolman, J. F. (1969). Stimulus preferences and form discrimination in young chicks. *Psychological Record* *19*, 407–416.