ABC 2026, 13(1):51-66
DOI: https://doi.org/10.26451/abc.13.01.03.2026

Animal Behavior and Cognition

Virtual Reality in Free-Flying Honey Bees: A Replication of Abramson et al. (1996)

Ian T. Jones[1], **James W. Grice**[2,*], **and Charles I. Abramson**[2]

[1] Japan Tobacco International, Raleigh, NC
[2] Oklahoma State University

*Corresponding author (Email: james.grice@okstate.edu)

Citation – Jones, I., T., Grice, J. W., & Abramson, C. I. (2026). Virtual reality in free-flying honey bees: A replication of Abramson et al. (1996). *Animal Behavior and Cognition, 13*(1), 51-66. https://doi.org/10.26451/abc.13.01.03.2026

Abstract – In two previous experiments, Abramson and colleagues (1996) used a mirage device to demonstrate that honey bees could be influenced by a virtual reality illusion. The current study sought to replicate several key findings of their second experiment by simultaneously presenting honey bees with both a real and illusory target from which to feed. Results for sixteen subjects largely replicated those reported by Abramson and colleagues, but only when individual response patterns were considered. As a group, illusion vs. real target error rates decreased linearly over study trials. However, using organism-centered (i.e., person-centered) analyses, this effect was discovered to be attributable to only two honey bees. Considered as individuals, and consistent with previous findings, the error rates for most of the bees in this sample did not decrease over time. Results therefore support the use of the mirage device in future studies of visual illusions in insects. Implications for analyzing error rates and other important outcomes in such studies are further discussed.

Keywords – Honey Bees, Illusions, Insects, Organism-Centered, Person-Centered, Replication

---

Investigations into how honey bees (*Apis mellifera*) respond to visual illusions, like the Ebbinghaus circle or Cornsweet illusions, can reveal important knowledge about their visual processing systems and provide insights into the universality of perceptual mechanisms across species (see Gatto, et al., 2022, for a review). For example, if honey bees, with their compound eyes and simple neural architecture, are susceptible to the same visual illusions as humans, it suggests certain perceptual shortcuts or biases are ancient, hardwired traits, not just quirks of mammalian brains. Moreover, honey bees serve as vital pollinators in many ecosystems through their complex foraging behavior. Studying visual illusions may help researchers understand how bees carry out this important behavior.

Abramson and colleagues (1996), for instance, demonstrated how a visual illusion could be used to induce frustration in free-flying honey bees during a foraging task. Using a mirage device (described below), they created a visual illusion of a sucrose-rich target to which the bees had previously been trained. Results revealed the honey bees could not discriminate between the illusory and real targets, a vulnerability that could affect their ability to locate authentic food sources amidst natural or artificial distractors, impacting pollination efficiency. Abramson and colleagues went on to discuss how the mirage device could further be used in studies of learning in both free-flying and constrained honeybees, opening "…a series of new experiments on frustration based on purely perceptual phenomena as frustrating stimuli, as opposed to such traditional variables as water as the non-reinforcing stimulus" (p. 42). They also discussed how the device could be used in experiments investigating the neural structure and processing mechanisms underlying honey bee behavior, in line with illusion studies of bees and other insects (e.g., Finke et al., 2023; Horridge et al., 1992; Howard & Dyer, 2024; Sakiyama & Gunji, 2016).

Given the mirage device's potential for studying honey bee behavior and the broader significance of exploring visual illusions in invertebrates, we sought to replicate the second experiment from Abramson et al. (1996). A successful replication would bolster confidence in the mirage device for modern researchers and reinforce the findings of Abramson and colleagues. Replication underpins scientific knowledge – a principle underscored by the recent replication crisis across the sciences and efforts to address the crisis to maintain public trust (e.g., see Bialek, 2018; Korbmacher et al., 2023; Munafo, 2018; Nissen et al., 2016). Our replication effort additionally permits the application of newer, advanced statistical techniques to examine the behavioral patterns of individual honey bees. These methods ensure more robust conclusions as they avoid the ecological fallacy, which assumes aggregate results accurately reflect the behavior of most individual organisms in a given study (Fisher et al., 2018; McManus et al. 2023).

In Abramson and colleagues' (1996) original study, 16 bees were individually trained to fly back and forth between the hive and a feeding station by associating a colored target with a rich sucrose solution. After training, the bees were presented with both a real and an illusory target using a mirage device. Based on their observations and analyses, Abramson and colleagues concluded that the bees could not discriminate between the two stimuli, "…honey bees, like humans, can be easily fooled by the MIRAGE machine" (p. 42). We attempted to replicate this experiment and these findings using a new sample of honey bees collected at a different site. Moreover, we utilized modern methods of statistical analysis (Grice et al., 2020) to determine if the behaviors of individual honey bees were consistent with expectation.

Methods

Ethics Statement

The honey bees in this study were treated in accord with the American Zoological Association's ethical guidelines for research with animals. Oklahoma State University does not require an Institutional Animal Care and Use Committee (IACUC) review for research conducted with invertebrate animals.

Subjects

The subjects of the current study were experimentally naive honey bees (*Apis mellifera ligustica*) from one maintained hive in central Oklahoma at a site approximately 16 km from the original study location. Replicating Abramson et al.'s (1996) sample size, 16 foraging honey bees were included. Individual foragers were taken from the feeding station in a small matchbox and brought to a separate training table. The training table was a white circular table, stationed approximately 18.29 m away from the hive and 5.49 m away from the feeding station (Figure 1). The matchbox containing the bee was opened slightly so that the bee could stick out its proboscis to feed from a 50% (by volume) sucrose solution (for more details refer to Abramson, 1990; Abramson et al., 1996; Rodriguez et al., 2023). The 50% solution was presented on a 6 cm (diameter) clear, circular petri dish with a 6 cm (diameter) laminated grey (brand: Glidden, Color & Code: Granite Grey, 00NN 37/000) target affixed to the top of the dish. A 25 ml droplet of 50% sucrose solution was placed on the center of the grey training target.

Figure 1

Free Flying Honey Bee Feeding Station

Note. Feeding station where bees were randomly captured in a matchbox and then brought to the experimental table.

Apparatus

The neutral grey target was placed on the top of a mirage illusion device (brand: Toysmith 3-D Mirascope). The mirage illusion device is a three-dimensional real imagery apparatus that reflects light rays to project an illusion from a stimulus inserted into the device (Figure 2). The device is saucer shaped and measures approximately 15.5 cm in diameter, with a 4 cm hole in the top. The inside of the saucer contains two parabolic mirrors. Objects inserted into the bottom of the parabolic devices project upwards to the 4cm hole located in the top and the images are reversed back from a singular image so that the 3-dimensional mirage is identical to the original item placed inside the device. It should be noted the mirage apparatuses used by Abramson et al. (1996) were larger in size (23 cm in diameter) and produced by a company which no longer exists. Accordingly, materials were scaled 1.48 times less than Abramson and colleagues' (1996) materials. For instance, the black flaps originally measured 12.5 cm x 12.5 cm. In the current experiment the black flaps were 8.4 cm x 8.4 cm. This scaling was done to control for proportionality between the reduced size of the stimuli.

Abramson and colleagues (1996) observed that bees could detect the illusions when flying above the mirage machine. To prevent this from occurring, an 8.4 cm square hood was made from black poster board and was placed at a 60-degree angle over the hole. To land on the target the bees had to fly around the square and then land on the grey training targets, as shown in Figure 2. The mirage can be seen in bright, ambient light, but works best under conditions of a shaded, ambient light source. Accordingly, a solid white

3.05 m x 3.05 m tent (brand: Ozark Trail) was used to cover the table to provide 9.29 m^2 of shaded, ambient light. The mirage devices were placed on a 35.7 cm x 27.6 cm laminated piece of poster board. Finally, the mirage devices were placed 21 cm apart, as measured from the center of each (for more details, see Abramson et al., 1996).

Figure 2

Mirage Device Setup with Training Targets

Note. Neutral grey training targets as they were placed on the mirage devices.

The targets used during the visual discrimination trials were two, 1.7 cm in diameter, orange, circular laminated targets (brand: Valspar color chip, Autumn Blaze, #2002-1A; hex: #e55b2b). The first 1.7 cm orange target was affixed to the center of the top of the mirage device by placing the orange laminated target on a clear, 6 cm (in diameter) clear, petri dish. The petri dish and orange target were placed inside the hole of the mirage device with double sided tape so that it was mounted at the top, henceforth referred to as the real target. The second 1.7 cm orange circular laminated target was affixed inside the mirage device so that it projected a reflection to the top of the device, henceforth referred to as the illusion target. The rewarded target (S+) was the real target, which contained a 50% drop of sucrose solution, while the nonrewarded target (S–) was the produced mirage (i.e., illusion) of a 50% drop of sucrose solution that was reflected to the top of the mirage device (Figure 3). To prevent the bees from accessing the S– target inside the mirage machine, a clear petri dish was mounted inside the mirage illusion device (see Abramson et al., 1996). However, during pilot trials, a few bees seemed capable of detecting the illusion from the real target due to the change in illumination of the orange target. The real target was visually different from the illusion target; accordingly, the petri dish covering the illusory target was removed as it was causing a potential bias in the bees' behaviors.

Figure 3

Mirage Devices with Orange Targets Shown

Note. Mirage devices with the illusion (left) and real (right) targets.

To prevent a bee from entering the illusory mirage device and feeding on the sucrose solution within, we allowed the bee to land on and crawl up to the lip of the device before nudging it with an eye dropper or gently blowing it off the device. The bee would then lift off, fly up, and be allowed to make another choice until it landed on the real target and fed from the sucrose solution. One bee nonetheless succeeded in entering the illusory mirage device. The top portion of the mirage piece was consequently removed and the bee was gently blown away and out of the mirage device. The mirage apparatus was then reset and the bee allowed to continue making choices. This bee was not included in the data analyses below, however, because it later failed to return to the table within 20 min after the 8th trial (see the procedures below for more details).

Procedures

Following common practice (e.g., Rodriguez et al., 2023; Wincheski et al., 2023), the honey bees were randomly pulled while feeding from the 8% feeder table and captured in matchboxes. Once captured, the honey bees were brought to the experimental table with the mirage machines. At the table, the bees were slowly released from the matchbox while feeding from a 50% sucrose solution (by volume) that was placed on the grey training target on top of the mirage machine. While feeding, the bees were marked on their abdomen, thorax, or both with “OPI” brand nail polish. Four colors (*big apple red*: #NLN25, *mi casa es blue casa*: #NLM92, *exotic birds do not tweet*: #NLF91, *that’s hula-rious*: #NLH65) of OPI Nail Lacquer were used.

After feeding and being marked with the nail polish, the bees were allowed to orient and fly back to the hive. Once a bee returned to the mirage device, it was required to interact with both targets on a single visit. When the bee landed on the first target (i.e., the left or the right) it was shooed off by nudging it until it flew to, and landed on, the other target. Once the bee visited both targets, two separate times, the bee began trial 1 of the experimental phase.

During the experimental phase, the two grey targets were removed from the top of the mirage device so that the orange targets, the real target and illusion target, were shown. For exactly 50% of the trials the illusion target was placed on the right and the real target was placed on the left. The mirage devices were moved in a pseudorandom (i.e., ABBABAAB) design to ensure proper counterbalancing and to control position effects.

The starting position was counterbalanced so that for half of the bees the mirage started on the left and the real on the right. For the other bees, the mirage started on the right and the real on the left. Thus, there were two different experimental groups to control for the starting positions of the experiment (see Table 1). The bees were allowed to fly back and forth to both targets until landing and feeding from the S+ target. Accordingly, the S+ target was always the real target and the S- target was always the mirage target. Intertrial intervals were recorded from the moment the bee left the table until the moment the bee received the next reward (i.e., made a correct choice). The bees completed 16 total trials during the experimental phase, completing two cycles of the ABBABAAB design. Upon completion of the 16th and final trial, the bees were captured and sacrificed to ensure they were not reused and that every bee was experimentally naïve. Furthermore, any bee that did not return during an experimental trial after a duration of 20 min was removed and excluded from continuing in the study (see also Chicas-Mosier et al., 2019; Rodriguez, et al., 2023). In total four bees were removed from the study and replaced with new bees because they failed to return to the table within 20 minutes. Two out of the four bees failed to return before the 4th trial. The other two bees failed to return on the 6th and 8th trials. The four dropout bees only made three correct choices out of 15 total choices (20%) across their combined trials. Their data are not included nor considered further. Finally, it should be noted that Abramson and colleagues (1996) used the same ABBABAAB design in their study but with 20 trials. In the interest saving time and balancing the design, only 16 trials were used in the current experiment.

Table 1

Target Positioning in the Pseudorandom Design

Order	Group 1		Group 2	
	Real Target (S+)	Illusory Target (S-)	Real Target (S+)	Illusory Target (S-)
A	R	L	L	R
B	L	R	R	L
B	L	R	R	L
A	R	L	L	R
B	L	R	R	L
A	R	L	L	R
A	R	L	L	R
B	L	R	R	L

Note. The ABBABAAB pseudorandom design trials were completed twice for a total of 16 trials. S+ = rewarded target; S- = non-rewarded target; R = Right; L = Left

Behavioral Coding

There were two primary dependent variables in this experiment: 1) the number of incorrect choices made by each bee, and 2) the first choice made. The former value was determined by tallying the number of times the bee landed on the illusion (i.e., incorrect) target during each trial (e.g., see Rodriguez et al., 2023). Recall if a bee landed on the illusion target, it was shooed away and permitted to land again on either the correct or incorrect target. This process was repeated until the bee landed on the correct (real) target, filled its crop, and returned to the hive. The observed number of incorrect choices ranged in value from zero (the bee landed on the real target on its first visit) to sixteen (i.e., the bee was shooed away from the illusion target sixteen times) for the bees in this sample. The first choice variable was coded when the bee flew to the table and first landed on either target. The side (e.g., left or right) was recorded in addition to whether the choice was correct (the real target, coded as 1) or incorrect (the illusion, coded as 0) for each trial. All trials for each bee were recorded on video, and these videos were consulted to validate behavioral coding conducting during the experiment.

Results

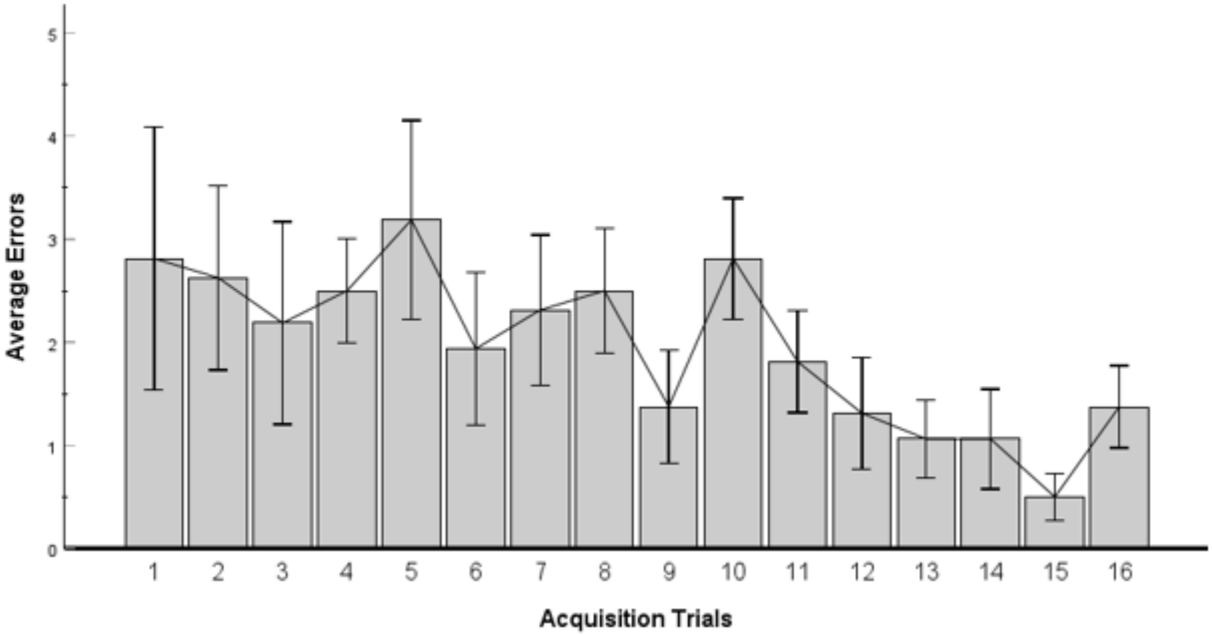
Replication Analyses

Abramson and colleagues (1996) reported results for only two analyses. First, they found no preference for either the right or left starting target position, *F*(1,19) = .511, $p = .475$. Second, while they found evidence for average differences in incorrect choices [*F*(1,300) = 6.402, $p < .001$], they did not observe a monotonic decrease in the means across their 20 trials. Specific decreases in the average number of incorrect choices were attributed to particular trials [viz., 3, 7, 11, 15, and 19] in which the S+ on the preceding trial was in that same position. Given the incorrect choices did not decrease over time, Abramson and colleagues concluded "…that bees were unable to distinguish between the real and illusion targets in this experiment" (p. 42).

Using the number of incorrect choices as the dependent variable, we conducted a 2 x (16) split-plot ANOVA on the current sample of honey bees. Consistent with the original study, the main effect for starting target position was nonsignificant, *F*(1, 14) = 0.08, $p = .79$, $\eta_p^2 = .01$ ($M_{right} = 2.02$, $SE_{right} = 0.28$, 95% $\text{CI}_{\text{right}} = 1.42, 2.62$; $M_{leftt} = 1.91$, $SE_{leftt} = 0.28$, 95% $\text{CI}_{\text{left}} = 1.31, 2.51$). Contrary to the original study the main effect for trials was not statistically significant, *F*(15, 210) = 1.21, $p = .23$, $\eta_p^2 = .08$. A test for linear trend, however, was significant, *F*(1,14) = 9.26, $p = .009$, partial $\eta_p^2 = .40$. The slight decline in average incorrect choices can be seen in Figure 4 and indicates that, contrary to Abramson and colleague' (1996) findings, the bees appeared to give up on attempting to feed from the illusion target over time. It should also be noted the data violated a number of important statistical assumptions (viz., normality and equality of variances) and that adjusted analyses yielded the same conclusions.

Figure 4

Average Incorrect Choice Rates Across All Bees for Each Trial



Note. Error bars represent standard errors of the mean.

In order to focus on individual honey bees and provide a more thorough assessment of the bees' behaviors, we supplemented the ANOVA results with organism-centered analyses. Specifically, an ordinal pattern analysis (see Beechy, 2023; Grice, Craig, & Abramson, 2015) was conducted to determine if the number of incorrect choices *for each individual bee* decreased monotonically over time. This analysis examined all pairwise comparisons of incorrect choices across the sixteen trials (e.g., Trial 1 > Trial 2, Trial 1 > Trial 3, Trial 2 > Trial 3, etc.). With sixteen trials, there were 120 ($_{16}C_2$) such comparisons for each bee from which a Percent Correct Classifications (PCC) statistic could be computed. If the number of incorrect choices decreased in a perfectly monotonic fashion across all sixteen trials, the resulting PCC would equal 100%.

In addition to the PCC, the ordinal pattern analysis conducts a randomization test to compute a plausibility value, referred to as a chance value (*c*-value; See Grice, 2021). The *c*-value is computed by randomly determining the bee's choice outcomes (correct vs. incorrect), re-computing the PCC, and comparing whether the randomly generated PCC is greater than or equal to the obtained PCC. This procedure can be repeated for a set number of iterations, and in the current paper the *c*-values were all generated using 10,000 iterations. Mathematically, the process is represented as follows:

$$c\text{-value} = \frac{\sum_{\mathrm{i}=1}^{\mathrm{k}} (PCC_i \geq PCC_{obs})}{k}$$

where k = the total number of iterations, PCC_k = the randomly generated PCC from the k^{th} iteration, and $\mathrm{PCC}_{\mathrm{obs}}$ = the observed PCC computed from the raw data.

The chance value aids in assessing whether the observed PCC index can be explained by physical chance. In other words, could the pattern of the individual bee's behaviors have been produced haphazardly (i.e., by accident)? As a proportion the *c*-value ranges from 0 to 1, and if the computed result is high, then it is concluded the data pattern is likely a product of physical chance. If the *c*-value is low (e.g., < .05), then it is concluded the data pattern is not likely to be a product of physical chance and some other explanation can be offered (e.g., the bee has learned to give up on feeding from the illusion target and fly to the real target). An added advantage of conducting these analyses is that they are free of the restrictive statistical assumptions underlying the split-plot ANOVA conducted above.

The results from the ordinal pattern analysis for each of the sixteen bees are reported in Table 2. As can be seen the vast majority of PCC values were low (< 50%) and accompanied by modest *c*-values (> .20), indicating these bees' number of incorrect choices did *not* decrease over time. The two exceptions are bees #6 (PCC = 62.50%, *c* = .02) and #16 (PCC = 57.50, *c* = .08). Figure 5 shows example individual error rates for three bees over the 16 trials. As can be seen, the error rates for bee #6 ranged from zero to fifteen and decreased monotonically, but the pattern was not visually compelling. By comparison, the error rates for bee #12 also varied considerably across the sixteen trials and did not decrease monotonically (PCC = 44.17%, *c* = .32). Most of the honey bees' patterns were similarly inconsistent across trials, revealing they did not consistently give up on attempting to feed from the illusory target over time.

Table 2

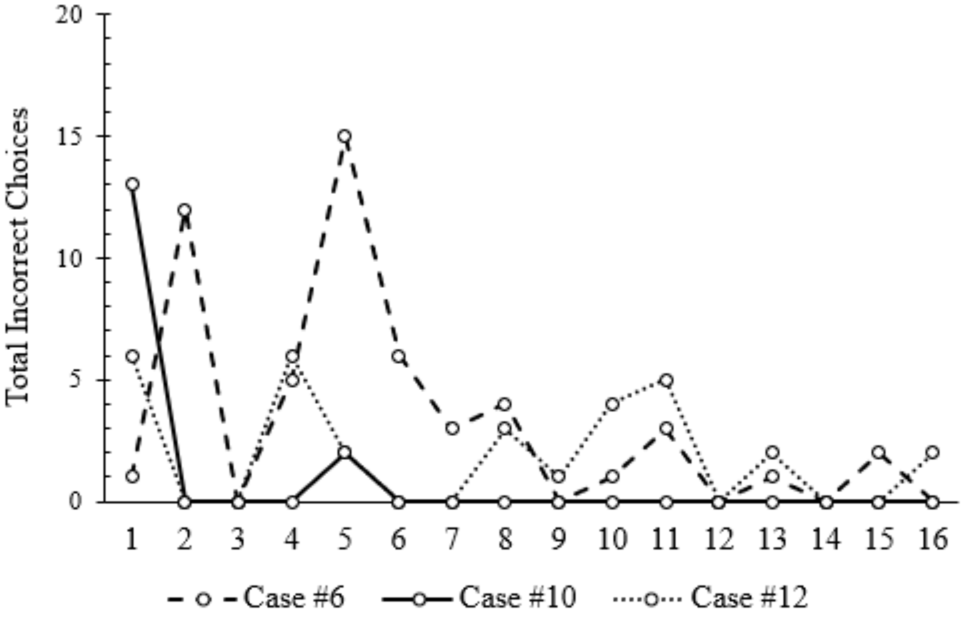
Monotonic Reduction in Incorrect Choices Over Time

Case #	Classifiable Pairs	PCC	*c*-value
Group #1			
1	120	31.67	.64
2	120	30.00	.38
3	120	22.50	.52
4	120	41.67	.38
5	120	36.67	.25
6	120	62.50	.02
7	120	48.33	.24
8	120	34.17	.85
Total	960	38.44	.14
Group #2			
9	120	36.67	.22
10	120	21.67	.04
11	120	45.83	.17
12	120	44.17	.32
13	120	25.00	.94
14	120	43.33	.27
15	120	50.83	.26
16	120	57.50	.08
Total	960	40.63	.04

Note. For Group 1 the S+ starting position was on the right, and for Group 2 the S+ starting position was on the left. PCC = Percent Correct Classifications index. The c-value is a probability statistic from a randomization test in which a low value indicates a PCC is not likely due to physical chance.

Figure 5

Case #6, #10, and #12 Incorrect Choice Frequencies Across All Trials



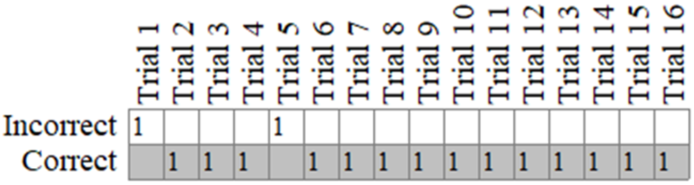
Aggregating the error responses across the bees in each group yielded overall PCCs less than 50% (Table 2). However, an important pattern of responses was noted for several bees. As can also be seen in Figure 5, for example, bee #10 had numerous ‘zero’ values, indicating this bee made few incorrect first choices and consequently few errors. While this bee’s number of incorrect choices did not decrease over time (PCC = 21.67%, *c* = .04) it nonetheless may have been able to discriminate between the two targets from the beginning of the experimental trials. It should be noted that the *c*-value for this particular case was low, despite the low PCC, because the bee made few errors altogether, resulting in numerous instances of tied quantities of zero errors made. With so many ties, PCCs from randomized data will necessarily be close to zero, so even values as low as 21.67% will occur infrequently. The patterns of error responses for cases #2 and #3 were similar to those for case #10, revealing that additional analyses are needed to determine if individual bees were not fooled by the illusion and could therefore discriminate between the two targets.

First Choice Analyses

For a bee to show a monotonic decrease in errors across the sixteen trials, it must also show a consistent inability to discriminate between the two targets when first landing. In other words, the bee’s first choice of target should be equal to approximately 50% for the real target across trials. We therefore analyzed the first choice of the two targets made by each bee using a concatenated pattern analysis. The analysis is based on the individual responses of each bee in comparison to an expected pattern of results (see Jones et al., 2024). Figure 6 shows the responses for bee #10 in a table comprised of two rows and sixteen columns. The top row represents the incorrect first choices (viz., the illusion), while the bottom row represents the correct first choices (viz., the real target). The shaded portion depicts the expectation that each bee will exhibit a perfect tendency to select the correct target (if it is able to discriminate between the two targets). The 1’s in each of the pairs of 16 boxes represent the bee’s actual choice outcomes for the trials.

Figure 6

Case #10’s Expected and Observed Pattern of First Choices for All Trials



	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	Trial 9	Trial 10	Trial 11	Trial 12	Trial 13	Trial 14	Trial 15	Trial 16
Incorrect	1				1											
Correct		1	1	1		1	1	1	1	1	1	1	1	1	1	1

Note. Expected pattern (grey cells) with the bee’s observed first choices (1’s) for bee #10. The PCC is computed by counting the number of 1’s in a grey cell (i.e., 14/16) and multiplying that proportion by 100 to convert it into a percentage.

For the individual bee in Figure 6, 14 out of 16 of its responses were in the correct row. Converting this result to a percentage yields an impressively high percent correct classifications (PCC) index equal to 87.50%. Fifty percent can here be adopted as a baseline value if the bee is not discriminating between the two targets successfully. Results from the accompanying randomization test yielded an impressively low chance-value (*c* = .002) as well. As noted above, this bee appears quite capable of discriminating between the two targets. It is therefore no surprise this bee did not show a monotonic decrease in errors across the sixteen trials (PCC = 21.67%), as reported in Table 2 above.

Table 3 reports the computed PCCs and *c*-values from the concatenated pattern analysis for the individual bees in the sample. As can be seen, only one bee had a PCC equal to 50%, indicating that it performed exactly as expected if it was not discriminating between either target. There were 6 bees who appeared to be able to discriminate between the two sets of targets at a rate above 50% (PCCs > 50%), but only two of these bees (cases # 3 and #10) had PCCs equal to 75% or higher, indicating at least 12 out of

16 correct first choices, and correspondingly low *c*-values (.04 and .002, respectively). Case #2's PCC was nearly as high, 68.75%, with a low *c*-value (.09). Another 5 bees showed slight tendencies to choose the illusory target over the real target, but the results were not extreme (PCCs < 40%, *c*'s > .89). Interestingly, bee #16 in fact chose the illusory target on 12 of its 16 trials, a rate that exceeds chance if the expected pattern in Figure 6 is reversed and tested (*c* = .04). This is also one of the two bees to show a slight monotonic decrease in number of errors across trials (Table 2).

Table 3

Correct First Choices

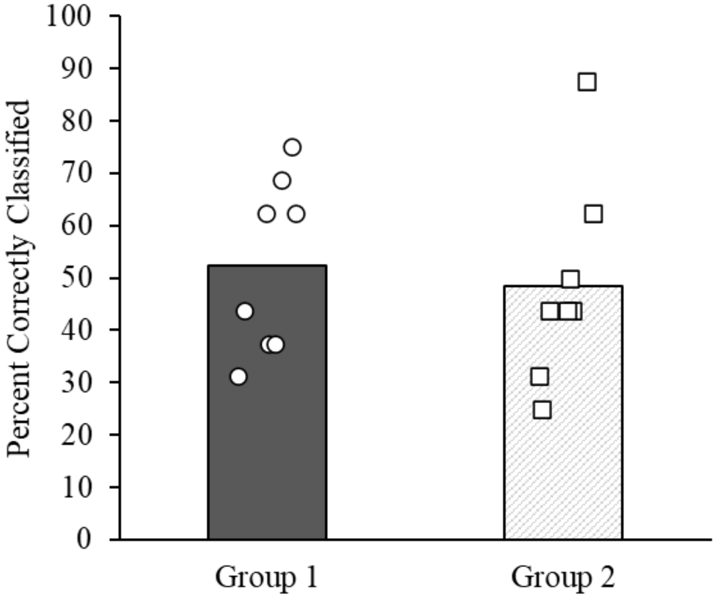
Case	# of Trials	PCC	*c*-value
Group #1			
1	16	62.50	.22
2	16	68.75	.09
3	16	75.00	.04
4	16	43.75	.77
5	16	62.50	.20
6	16	37.50	.90
7	16	37.50	.90
8	16	31.25	.96
Total	128	52.34	0.32
Group #2			
9	16	62.50	.22
10	16	87.50	.002
11	16	50.00	.58
12	16	43.75	.75
13	16	43.75	.74
14	16	43.75	.77
15	16	31.25	.96
16	16	25.00	.99
Total	128	48.44	.65

Note. For Group 1 the S+ starting position was on the right, and for Group 2 the S+ starting position was on the left. PCC = Percent Correct Classifications index. The c-value is a probability statistic from a randomization test in which a low value indicates a PCC is not likely due to physical chance.

In addition to the individual level results for each bee, the pattern analysis computes group level results by tallying the total number of correct first choices across all 16 trials for the bees assigned to group one and the bees assigned to group two (Figure 7). The results for the first group of bees revealed that only 67 out of 128 total choices (PCC = 52.34%) matched the expected pattern across all 16 trials. The *c*-value additionally showed that this result could be explained as a product of physical chance as 32% of the randomized PCCs were equal to or greater than the observed value (*c* = .32). The results for the second group of bees revealed a similar outcome as only 62 of the 128 total choices matched the expected pattern; PCC = 48.44%, *c*-value = .65.

Figure 7

Percentage of Correct First Choices Between Starting Position Groups

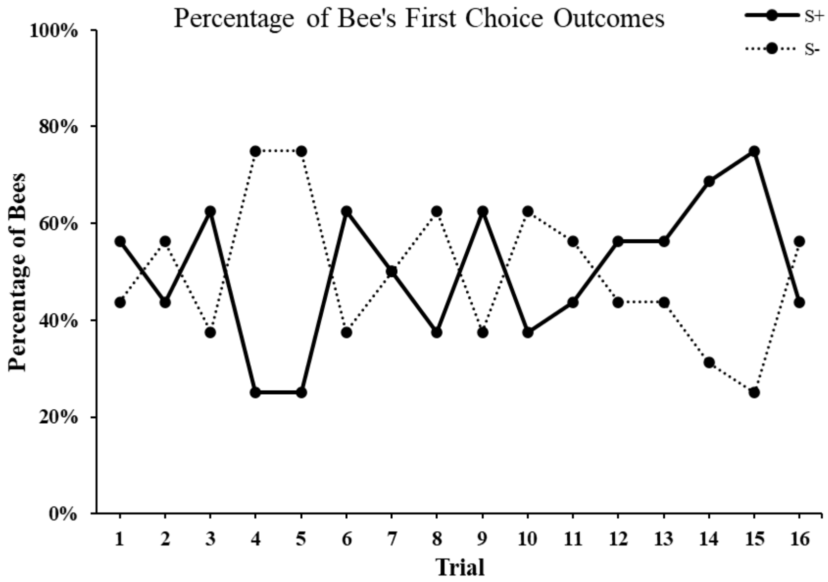


Note. Circles represent individual bee's percentage of correct first choices in group 1 and the squares are the individual bee's percentage of correct first choices in group 2. For Group 1 the S+ starting position was on the right, and for Group 2 the S+ starting position was on the left.

Finally, classical learning curves were examined in a graph of group-level first-choice outcomes (S+ or S-), presented in Figure 8. These outcomes were plotted to enable comparison with other research findings (e.g., Geng et al., 2022; Lafon et al., 2023). The plot reveals significant variability in group-level S+ choices with no clear increasing or decreasing trends over time. At trials 3, 6, 9, 14, and 15, over 60% of the bees correctly chose the S+ target. For trials 12 through 15 the correct choices were over 50%, but for trial 16 less than 50% of the bees chose the correct target. These results therefore indicate the honey bees' first-choice behaviors did not, *in the aggregate*, exhibit a typical learning curve that would support successful discrimination between the real and illusory targets. This conclusion must of course be considered in conjunction with the individual-level results reported above.

Figure 8

Percentage of the Bees' S+ and S- Choices Over Time



Note. The solid line depicts the S+, correct choices (i.e., the real target). The dashed line depicts the S-, incorrect choices (i.e., the illusion target).

Discussion

Based on the results of a split-plot ANOVA, Abramson and colleagues (1996) found no preference for either the right or left starting target position. We replicated this outcome with our sample of bees as the mean number of incorrect choices for the two target positions were observed to be nearly equal. Abramson and colleagues also failed to find evidence for a monotonic decrease in the average incorrect choices across the experimental trials, which they reported as supporting the hypothesis that honey bees cannot discriminate between the real and illusory targets. Contrarily, the means for the incorrect choices in the current sample visibly decreased monotonically across the 16 trials ($\eta_p^2 = .40$), suggesting the bees gave up on attempting to feed from the illusory target over time. These ANOVA results were based on means and variances computed for the two groups of bees, and as recent empirical research has shown (Fisher, et al., 2018; McManus et al., 2023), group-level results do not always match those found at the level of the individual organisms in a given study.

Organism-centered analyses for the current sample revealed that the monotonic, decreasing trend in the group means was determined primarily by two bees (cases #6 and #16), although the evidence for their trends was not compelling statistically (PCCs of 62.5 and 57.50, respectively) or visually. The other fourteen individual bees did not show even weak evidence for decreases in their numbers of incorrect choices. While focusing on individual bees we also discovered that the original analysis strategy needed to be supplemented. Specifically, the number of incorrect choices depended upon the first choices being made by the bees on the sixteen trials. If a bee consistently chose the real target, then that bee could not show a monotonic decrease in errors. We therefore conducted additional organism-centered analyses on the first-choice responses and found three bees that were apparently able to discriminate between the real and illusory targets. These bees chose the correct, real target on at least eleven of their sixteen trials. Two of these bees (cases #2 and #3) were in Group 1 (S+ on the right), and one bee (case #10) was in Group 2 (S+ on the left).

Considering the number of incorrect choices and first choice results in tandem, five of the sixteen bees (cases #2, #3, #6, #10, and #16) appeared to be capable of discriminating between the real and illusory targets. Results from the incorrect choice analyses were less compelling than those for the first-choice analyses with regard to the magnitudes of the PCCs. It should also be noted the response patterns for case #16 were unique in that this bee showed a first-choice preference for the illusory target but then demonstrated a slight monotonic reduction in the number of incorrect errors across the sixteen trials. Remarkably, there were several trials for this bee in which an invader bee was present, possibly influencing its preference to the illusory target. Invaders are bees which are in the wild and not currently part of the study. These bees are foragers from the hive (or local area) which can follow the experimental bees back to the research table. In some cases, they could also be bees which were previously captured during the pre-training trials. Invader bees were captured immediately when they landed, and targets were cleaned and reset when possible. Rodriguez et al., (2023) recently noted that bees are capable of observational learning and others (e.g., Dawson et al., 2013) have noted such capabilities in similar species (e.g. bumble bees). In the current study, it could be the case that the experimental bee was observing another bee attempt to land on the illusory target and influenced its choices. Case #16 was the only case in which there were multiple invader bees present throughout several trials, which may have interfered with this experimental bee.

As for the other four honey bees who were apparently not completely “fooled” by the illusion, there were no notable, individual factors (e.g., invaders or atypical environmental conditions) to account for their behaviors. In sum, the majority of bees in our sample behaved in ways that were consistent with Abramson and colleagues’ conclusion that honey bees cannot reliably detect the difference between real and illusion targets presented in mirror mirage devices and that they will not immediately give up on attempting to feed from the illusory target. A minority of bees, however, could apparently discriminate between the two targets by choosing the real target across the sixteen trials in subtle ways that nonetheless exceeded chance.

With regard to the mirror device, these results and those of Abramson and colleagues (1996) suggest it may prove to be a fruitful apparatus for researchers interested in studying visual illusions in flying insects (e.g., Agrochao, et al., 2020; Srinivasan, 1993). The devices used in the current study were smaller than those used by Abramson and colleagues, but the illusion still appeared to be effective in consistently fooling a majority of bees in the sample. While we could not identify the factors contributing to the failure of the device for the minority of bees in our sample, future research should seek to identify such potential confounds.

As demonstrated above, critical to future research employing the mirror device is the use of organism-centered analyses. It is only through these analyses we were able to determine which individual bees were consistently fooled by the illusion. Moreover, it was only by focusing on individual honey bees that we recognized the limitations of analyzing only the number of incorrect choices made. Organism-centered, or person-centered, methods like those used above are simple, relatively free of statistical assumptions, intuitive, and easy to understand by researchers and laypersons alike (see Grice et al., 2015; Grice et al., 2020). They are also critical in preventing the ecological fallacy, which involves interpreting an average result as being applicable to every or most persons (or organisms) in a study. The comparisons of group-level and individual-level results above demonstrate this fact.

With these types of analyses in mind, if resources and time permit, the twenty trials used by Abramson and colleagues can therefore be regarded as superior to the sixteen trials used in the current experiment. While the current results largely supported the inference that honey bees can be fooled by the optical illusion presented in the mirror mirage device, drawing inferences about individual bees would be strengthened by a larger number of trials. In short, the more observations collected from each organism, the better. Future research can explore the benefits and costs of study design features for those wishing to use organism-centered methods.

Author Contributions: I.T.J. conducted the experiment, analyzed the data, and wrote most of the manuscript which is based on his dissertation. J.W.G. wrote portions of the manuscript and assisted with data analysis. C.I.A. helped to design and conduct the experiment. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflict of Interest: Ian T. Jones is employed by Japan Tobacco International. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data Availability Statement: Data are available via the link: https://osf.io/2ynm5/overview

References

Abramson, C. I. (1990). *Invertebrate Learning: A Laboratory Manual and Source Book.* American Psychological Association. https://doi.org/10.1037/10078-000

Abramson, C. I., Buckbee, D. A., Edwards, S., & Bowe, K. (1996). A demonstration of virtual reality in free-flying honey bees: Apis mellifera. *Physiology & behavior*, *59*(1), 39-43. https://doi.org/10.1016/0031-9384(95)02023-3

Agrochao, M., Tanaka, R, Salazar-Gatzimas, E. and Clark, D. A. (2020). Mechanism for analogous illusory motion perception in flies and humans, *Proceedings of the National Academy of Sciences*. https://doi.org/10.1073/pnas.2002937117

Beechey, T. (2023). Ordinal Pattern Analysis: A Tutorial on Assessing the Fit of Hypotheses to Individual Repeated Measures Data. *Journal of Speech, Language, and Hearing Research, 66*, 347–364. https://doi.org/10.1044/2022_JSLHR-22-00133

Chicas-Mosier, A. M., Dinges, C.W., Agosto-Rivera, J. L., Giray, T., Oskay, D., & Abramson, C. I. (2019). Honey bees (*Apis mellifera spp.*) respond to increased aluminum exposure in their foraging choice, motility, and circadian rhythmicity. *PLoS ONE, 14*(6), e0218365. https://doi.org/10.1371/journal.pone.0218365

Dawson, E. H., Avargues-Weber, A., Chittka, L., & Leadbeater, E. (2013). Learning by observation emerges from simple associations in an insect model. *Current Biology*, *23*(8), 727–730. https://doi.org/10.1016/j.cub.2013.03.035

Finke, V., Scheiner, R., Giurfa, M., & Avarguès-Weber, A. (2023). Individual consistency in the learning abilities of honey bees: cognitive specialization within sensory and reinforcement modalities. *Animal Cognition, 26*(3), 909-928. https://doi.org/10.1007/s10071-022-01741-2

Fisher, A. J., Medaglia, J. D., & Jeronimus, B. F. (2018). Lack of group-to-individual generalizability is a threat to human subjects research. *Proceedings of the National Academy of Sciences, 115*(27), E6106-E6115. https://doi.org/10.1073/pnas.1711978115

Gatto, E., Loukola, O. J., Petrazzini, M. E. M., Agrillo, C., & Cutini, S. (2022). Illusional perspective across humans and bees. *Vision*, *6*(2), 28. https://doi.org/10.3390/vision6020028

Geng, H., Lafon, G., Avarguès-Weber, A., Buatois, A., Massou, I., & Giurfa, M. (2022). Visual learning in a virtual reality environment upregulates immediate early gene expression in the mushroom bodies of honey bees. *Communications Biology*, *5*(1), 130. https://doi.org/10.1038/s42003-022-03075-8

Grice, J. W. (2021). Drawing inferences from randomization tests. *Personality and Individual Differences, 179*, https://doi.org/10.1016/j.paid.2021.110931

Grice, J. W., Craig, D. P. A., & Abramson, C. I. (2015). A Simple and Transparent Alternative to Repeated Measures ANOVA. *Sage Open, 5(3)*. https://doi.org/10.1177/2158244015604192

Grice, J. W., Medellin, E., Jones, I., Horvath, S., McDaniel, H., O'lansen, C., & Baker, M. (2020). Persons as effect sizes. *Advances in Methods and Practices in Psychological Science*, *3*(4), 443-455. https://doi.org/10.1177/25152459209229

Horridge, G. A., Zhang S., & Carroll, D. (1992). Insect perception of illusory contours. *Philosophical Transactions of the Royal Society London*. B33759–64. https://doi.org/10.1098/rstb.1992.0083

Howard, S. R., & Dyer, A. G. (2024). Quantity misperception by hymenopteran insects observing the solitaire illusion. *iScience, 27*(2). https://doi.org/10.1016/j.isci.2023.108697

Jones, I. T, Kucker, S. C., Perry, L. K., & Grice, J. W. (2024). Capturing the heterogeneity of word learners by analyzing persons. *Behavioral Sciences, 14*(8), 708. https://doi.org/10.3390/bs14080708

Korbmacher, M., Azevedo, F., Pennington, C. R., Hartmann, H., Pownall, M., Schmidt, K., et al. (2023). The replication crisis has led to positive structural, procedural, and community changes. *Communications Psychology, 1*(1), 3. https://doi.org/10.1038/s44271-023-00003-2

Lafon, G., Paoli, M., Paffhausen, B. H., Sanchez, G. D. B., Lihoreau, M., Avarguès-Weber, A., & Giurfa, M. (2023). Efficient visual learning by bumble bees in virtual-reality conditions: Size does not matter. *Insect Science*, *30*(6), 1734-1748. https://doi.org/10.1111/1744-7917.13181

McManus, R. M., Young, L., & Sweetman, J. (2023). Psychology is a property of persons, not averages or distributions: Confronting the group-to-person generalizability problem in experimental psychology. *Advances in Methods and Practices in Psychological Science, 6(3)*, 1-23. https://doi.org/10.1177/25152459231186615

Rodriguez, S. D., Wincheski, R. J., Jones, I. T., Jesus-Soto, D., Michael, G., Fletcher, S. J., ... & Abramson, C. I. (2023). Some phenomena of the cap-pushing response in honey bees (*Apis mellifera* spp.). *Journal of Comparative Psychology, 137*(3), 178–190. https://doi.org/10.1037/com0000346

Sakiyama, T., & Gunji, Y. P. (2016). The Kanizsa triangle illusion in foraging ants. *Biosystems*, 142–143, 9–14. https://doi.org/10.1016/j.biosystems.2016.02.003

Srinivasan, M. V. (1993). Even insects experience visual illusions. *Current Science, 64*(9), 649-655. https://www.jstor.org/stable/24095143

Wincheski, R., Jones, I. T., Rodriguez, S., Soto, M., Fletcher, S., Eagle, T., Grice, J., & Abramson, C. (2023). Training honey bees (*Apis mellifera*) to push a cap: Shaping, observational learning, and memory. *International Journal of Comparative Psychology*, *36*(1), 1-10. https://doi.org/10.46867/IJCP.2023.36.5634