Journal of Comparative Psychology

Manuscript version of

Some Phenomena of the Cap-Pushing Response in Honey Bees (Apis mellifera spp.)

Sierra Dee Rodriguez, Riley J. Wincheski, Ian T. Jones, Michael G. De Jesus-Soto, Skylar J. Fletcher, Troy Joseph Pretends Eagle, James W. Grice, Charles I. Abramson

Funded by:

- National Science Foundation

Some Phenomena of the Cap Pushing Response in Honey Bees (*Apis mellifera spp.*)

Sierra Dee Rodriguez [1], Riley J. Wincheski [5,6], Ian T. Jones[5], Michael G De Jesus-Soto [2], Skylar J. Fletcher [3], Troy Joseph Pretends Eagle [4], James W. Grice[5], & Charles I. Abramson[5,6]

[1]Department of Biology, Texas A & M University – San Antonio

[2]Department of Biology, University of Puerto Rico – Rio Piedras

[3]Department of Biology, Southeastern Oklahoma State University

[4]Department of Biology, North Dakota State University

[5]Department of Psychology, Oklahoma State University

[6]Laboratory of Comparative Psychology and Behavioral Biology, Oklahoma State University

Authors Note

This research was primarily performed Malemi Organic Hotel located in Skala Kallonis, Lesvos, Greece 39.208896, 26.204095. We would like to acknowledge George and Effy Kapsalis for their hospitality during our stay. The Observation Oriented Modeling software can be freely downloaded from http://www.idiogrid.com/OOM. This research was supported by NSF REU grants 1560389, 1950805, and NSF PIRE grant 1545803. All data for these experiments will be made available at request.

Please send correspondence to Charles I. Abramson. Email: Charles.abramson@okstate.edu

Abstract

The cap pushing response (CPR) is a new free flying technique used to study learning and memory in honey bees. Bees fly to a target where they push a cap to reveal a hidden food source. When combined with traditional odor and color targets, the CPR technique opens the door to additional choice preference tests in honey bees. To facilitate the use of the CPR technique, three experiments were conducted. Experiment 1 investigates the impact of extended training on the CPR response and its role in extinction. Experiment 2 explores the role of CPR in overshadowing, and experiment 3 explores the effects of electric shock punishment on the CPR technique.

Keywords: cap pushing response, conditioning, honey bee, learning

Some Phenomena of the Cap Pushing Response in Honey Bees (*Apis mellifera spp)*

Introduction

The purpose of the present series of experiments is to provide additional data on the cap pushing response (CPR) of honey bees. The CPR technique is a novel free flying technique where bees are trained to push a cap to reveal a hidden food source (Abramson et al., 2016). With the advent of the CPR technique, a manipulative response can now be added to the stable of free flying techniques. When combined with odor, color, and position, we expect the CPR technique to provide more challenging experiments that test the limits of choice preference in honey bees. These findings will help further our knowledge of honey bee capabilities and behavioral importance in species comparisons at the individual level, which is explored using non-traditional statistics.

The experiments reported here provide data on aspects of the CPR not covered in our previous experiments (Abramson et al., 2016; Chicas-Mosier et al., 2019). Through a series of three experiments, we provide researchers with fundamental and practical data on various aspects of the CPR, including punishment, the effect of extended training, and overshadowing. Observation Oriented Modeling (OOM) is a type of non-traditional statistics that was utilized to analyze our data. These experiments outline why OOM is a better method for our research than traditional aggregate statistical.

Observation Oriented Modeling departs from traditional analyses based on means, standard deviations, or variances (Grice, 2011; Grice et al., 2020). OOM is a suite of nonparametric methods which permits the researcher to examine patterns within the data at the level of the individual or organism under study. Emphasis is placed on the persons or entities (i.e., organisms) in the experiment in order to answer the question, “how many people [or entities] in the study [or experiment] behaved or responded in a manner consistent with theoretical expectation" (Grice et al., 2020, p. 2; see also Grice, 2011; Grice et al., 2012). To determine how many organisms behaved as expected according to one‟s hypotheses, OOM computes a person-centered effect size, referred to as the Percent Correct Classification (PCC) index. The PCC index is the computed proportion of individuals/organisms who conformed or behaved as expected within the scope of the experimenter‟s hypotheses. Moreover, a randomization test on the PCC index can be conducted to determine whether the resulting value should be explained as having arisen by physical chance (see Grice, 2021).

The randomization test is conducted by randomly shuffling the observations a set number of times (e.g., 1,000 iterations), similar to bootstrapping, and a PCC is computed for each iteration. The software then tallies each time the randomized PCC is equal to or greater than the observed PCC, thus yielding a total proportion of occurrences (referred to as a *c*-value) that the observed PCC was achieved through randomizing the data. For example, if a PCC of 85% was found and the randomization test revealed that this PCC could only be achieved for 5 out of 1000 iterations, then the *c*-value would be 0.005, indicating that the observed PCC is unlikely to be a product of chance (Grice, 2021). Therefore, all OOM software analyses are accompanied by a randomization test using 10,000 iterations unless otherwise indicated.

All data were analyzed utilizing the OOM software (Grice, 2011). Traditional statistics were also utilized to compare the two types of statistical methods. The OOM software can be downloaded gratis following the link provided within the acknowledgments of the manuscript.

Standards for Openness and Transparency

We report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study in each of the methods sections for each of the experiments.

Experiment 1: The effect of extended training on the cap pushing response and extinction in honey bees

In Experiment 1, we sought to investigate the effects of extended training on the extinction of the CPR. We wanted to know if there was a difference in extinction rates for the responses between honey bees trained using 6 or 12-training trials. In our initial investigations (Abramson et al., 2016) we noticed a peculiar error. Bees that were trained to push the cap would continue to push the cap even though the cap was no longer directly above the feeding hole. If the cap was moved to the side of the feeding hole and was replaced with a cross now directly over the feeding hole, the bee would continue to push the cap several times before it would push the cross. We hypothesize that these errors were the result of a lack of experience with the target and the more experience with the target would result in less of these types of errors (Abramson et al., 2016). Additionally, previous research found that resistance to extinction increased from 1 to 6 trials and then decreased as the number of training trials increased to 12 (Couvillon & Bitterman, 1980). We wanted to see if we can replicate this effect with the CPR.

Materials and Methods

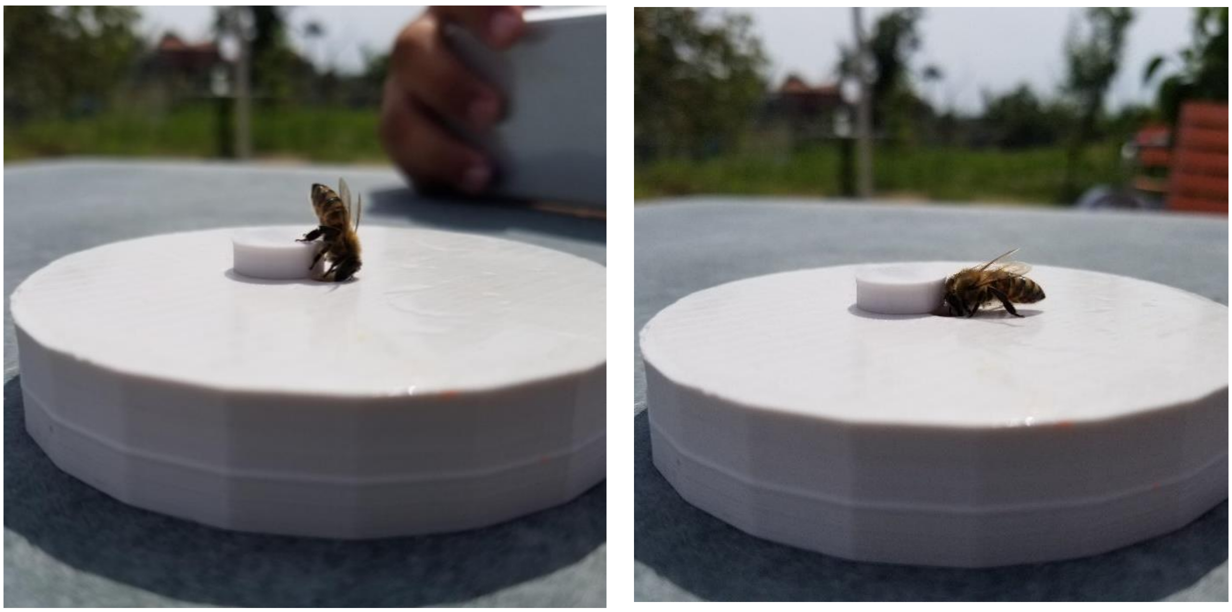
Twenty-eight bees were trained to land on a target containing a feeding well with a 50% sucrose by volume (Figure 1) to begin the steps to train the cap training process. Hives were managed in Greece and were well maintained and in healthy condition. The experiment was conducted outdoors. The bees were trained to come to a feeder with 8% sucrose by volume. The feeder was located four meters from a hive and was shared by three maintained colonies. An individual bee was picked up in a match box while on the feeder and brought to a gray target containing 50% sucrose. The match box was opened slightly to allow the proboscis to make contact with the sucrose. While the bee was feeding it was allowed to leave the match box where it was marked with nail polish (the brand “LBK™” nail lacquer). When the bee returned twice on its own it was shaped to push the cap (weight = 0.47 g, height = 0.5 cm, length = 1.5 cm).

Shaping was accomplished by placing the cap so that it covered about half the feeding hole. On subsequent visits, the cap was gradually moved so that it completely covered the food well. When the bee pushed the cap to uncover the well on two consecutive visits the experiment was begun. Details of the shaping procedure can be found in Abramson et al. (2016).

The bees were divided into two groups consisting of 14 bees each. One group received 6 training trials and the other 12 training trials. The groups then received a 10 minute extinction session divided into twenty 30 second intervals, where the sucrose was replaced with tap water. We then tracked the number of target landings, cap touches, and cap pushes. We analyzed landings for the 6 and 12 trial bees to evaluate extinction differences. We further include the analysis for the touching and pushing behaviors as additional evidence supporting the hypothesis that the bees that received 12-training trials would extinguish faster than the bees that received 6-training trials. Bees were selected from a common feeder regularly visited by three different hives, and the targets were washed and cleaned based on standard methods (Couvillion & Bitterman, 1980). Our sample size was based on previous research and personal experience with this preparation. Honey bees are not listed under any ethical codes concerning humane treatment of animals in research under USA or Greek law; therefore, no ethical review of the research protocol was conducted.

Figure 1

Example of Bees in the Cap Pushing Response



Results and Discussion

The individual bee‟s extinction curves were computed and analyzed in OOM for each bee‟s landing, touching, and pushing behaviors across all 20 intervals. The cumulative responses were then compared with an expected ordinal pattern of monotonic increasing landings, touches, and pushes (viz., interval 1 < interval 2 < interval 3, etc.). The PCC index was computed at each interval to determine how many total bees were performing the landing, touching, and pushing behaviors. It was expected that the bees should demonstrate extinction, therefore at some point

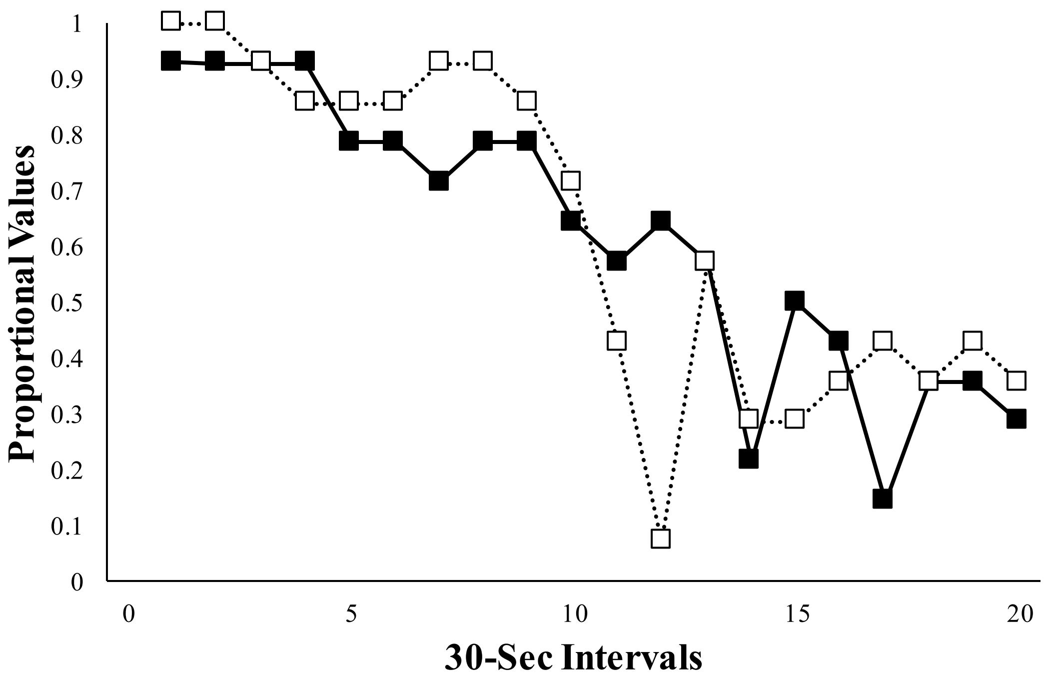
Figure 2

Six and Twelve Trial Bees' PCC Response Curves for the Landing Behavior

the monotonic pattern should stop increasing and the accompanying PCC should subsequently decrease (i.e., indicating a nonmonotonic relationship).

Based upon previous findings (Couvillon & Bitterman, 1980), it was also expected that the 12-trial bee‟s PCC response curve should decrease faster than the 6 trial bees for the landing, touching, and pushing behaviors. Therefore, following „best-practice‟ data analysis recommendations put forth by Fidler and Loftus (2009), graphical interpretations are primarily relied upon. The result will outline traditional and non-traditional statistical interpretations.

Non-traditional statistics, shown in figures 2, 3, and 4 contain the results for the PCCs for all bees at each interval across the landing, touching, and pushing behaviors, respectively. Across all three figures, a decline in the PCC response curve suggests a nonmonotonic relationship within the data, which revealed extinction occurring for most of the bees. In other words, if a decline is observed in the graph, then a majority of the bees are no longer performing the behavior and extinction was occurring. As expected, the results for the landing, touching, and pushing behaviors showed that the majority of the 12-training trial bees demonstrated an extinction response slightly more quickly than the majority of the 6-training trial bees.

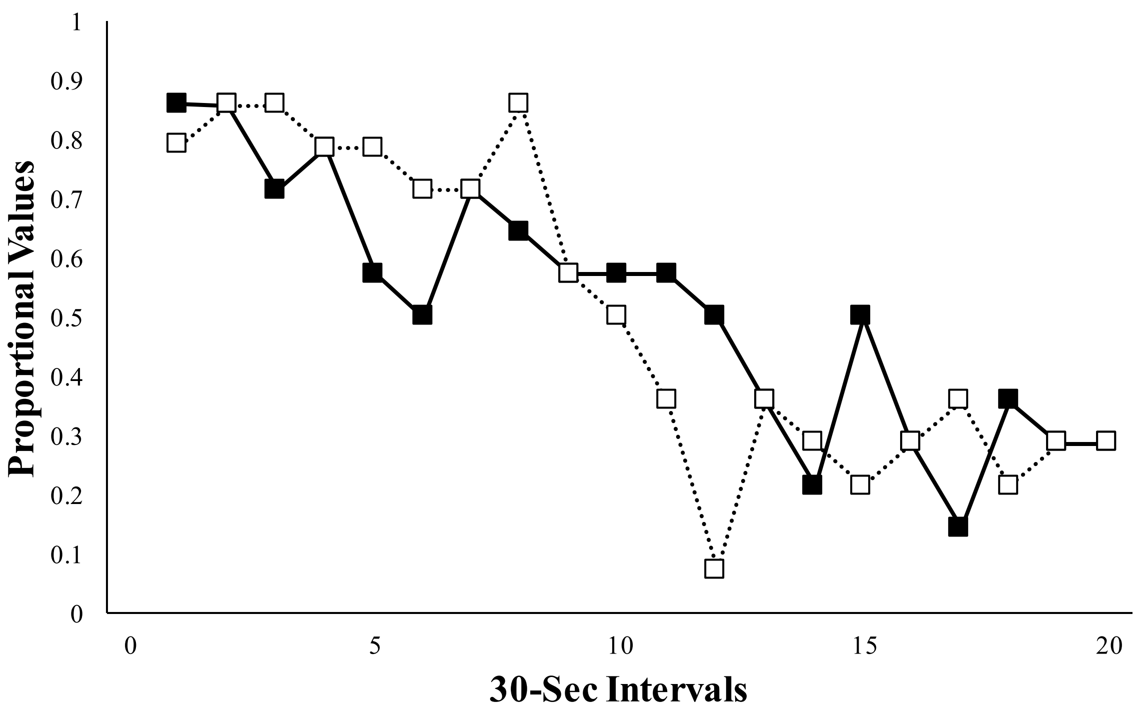


Note. 6-trial (solid black line) and 12-trial (dotted black line) PCC response curves across all bees for the landing behavior. The proportional values were obtained by converting the PCC to proportional frequencies ($\frac{PCC}{100}$). Each interval represents data produced over a 30-second time period.

Looking at Figure 3, the PCC response curve for the 6 trial bees gradually declines after the 2nd interval, then sharply decreases between the 13th and 14th intervals (360 – 390 seconds). The majority of the 6 trial bees first demonstrated extinction at the 14th interval as only 3 out of 14 bees (PCC = 21.43%) were still performing the landing behavior during this time period. The PCC response curve for the 12-training trial bees is comparable to the 6-training trial bees initially until the 10th interval, then sharply declines between the 10th and 12th intervals (300 – 360 seconds). The majority of the 12-training trial bees first demonstrated extinction at the 11th interval, as 6 out of 14 bees (PCC = 42.86%) perform the landing behavior. By the 12th interval, however, only 1 of the 12-training trial bees (PCC = 7.14%) was still performing the landing behavior.

Figure 3

Six and Twelve Trial Bees' PCC Response Curves for the Touching Behavior



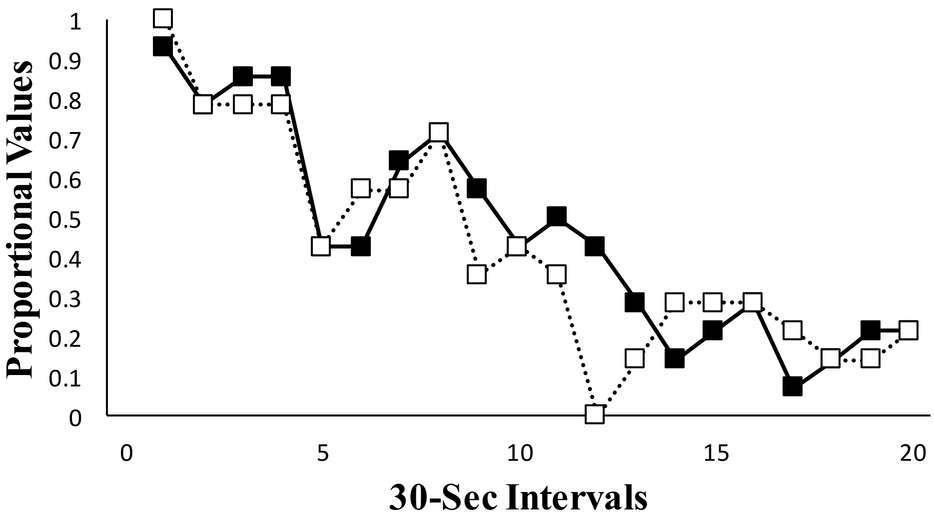
Note. 6 Trial (solid black line) and 12 Trial (dotted black line) PCC curves across all bees for the touching behavior. The proportional values were obtained by converting the PCC to proportional frequencies ($\frac{PCC}{100}$). Each interval represents data produced over a 30-second time period.

By the 12th interval, 64.5% of the 6-training trial bees were still demonstrating the landing behavior, whereas only 7% of the 12 trial bees were still demonstrating the landing behavior. The „eye-test" from Figure 3, and accompanying PCCs, supported our hypotheses as it revealed that the majority of the 12-training trial bees demonstrated extinction after 330-360 seconds, whereas it took the majority of the 6-training trial bees around 420 seconds to demonstrate extinction of the landing behavior

Looking at Figure 4, the PCC curve for the touching behavior shows the same general pattern found in Figure 3. Although at the 6^{th} interval only 50% of the 6-trial bees were performing the touching behavior, we only consider extinction to be first demonstrated at the interval where the majority of bees (i.e., < 50%) are no longer performing the behavior. Therefore, the majority of the 6-training trial bees first demonstrated extinction at the 13^{th} interval as only 5 out of 14 bees (PCC = 35.71%) were still performing the touching behavior. Comparably, the majority of the 12-training trial bees first demonstrated extinction at the 11^{th} interval, as only 5 out of 14 bees (PCC = 35.71%) were performing the landing behavior. As we saw in Figure 2, by the 12^{th} interval only one (PCC = 7.14%) of the 12 trial bees was still performing the touching behavior.

Figure 4

Six and Twelve Trial Bees' PCC Curves for the Pushing Behavior



Note. Six Trial (solid black line) and 12 Trial (dotted black line) PCC curves across all bees for the pushing behavior. The proportional values were obtained by converting the PCC to proportional frequencies ($\frac{PCC}{100}$). Each interval represents data produced over a 30-second time period.

A similar pattern of results found in the landing behavior was also found in the touching behavior. By the 12th interval, 50% of the 6-training trial bees were still demonstrating the touching behavior, whereas only 7% of the 12-training trial bees demonstrated the touching behavior. The majority of the 12-training trial bees demonstrated extinction after 330 seconds, whereas it took the majority of the 6-training trial bees 390 seconds to first demonstrate extinction of the touching behavior.

The PCC response curves for the 6-training trial bees and 12-training trial bees slightly support the expectations. The PCC response curves for both the 6 and 12-training trial bees are nearly identical from intervals 1 – 6. At the 5th interval, both the 6-training trial bees and 12-training trial bees revealed that a majority of bees demonstrated extinction of the pushing behavior as only 6 out of 14 (PCC = 42.86%) of both groups of bees were performing the pushing behavior. However, the greatest occurrence of extinction in the 6-training trial bees did not occur until the 17th interval, where one bee (PCC = 7.14%) from the 6-training trial bees was still performing the pushing behavior. Comparably, the greatest occurrence of extinction in the 12-training trial bees occurred by the 12th interval, where none (PCC = 0%) of the 12-training trial bees were still performing the pushing behavior.

By the 12th interval, 43% of the 6-training trial bees demonstrated the touching behavior, whereas none (i.e., 0%) of the 12-training trial bees demonstrated the landing behavior. The "eye-test" from Figure 4 revealed nearly identical results initially; however, towards the middle of the experiment the majority of the 12-training trial bees extinguish the pushing behavior prior to the majority of the 6-training trial bees.

Across all three behaviors using OOM and SPSS, the majority of the 12-training trial bees extinguished the behavior earlier than the 6-training trial bees, as exhibited by the sharper declining slope and low PCC indices across each interval. The "eye-test," accompanying PCC indices, and *c*-values are more than sufficient at establishing the differences in the patterns of observations across the three behaviors (landings, touches, pushes) between the 6 and 12-training trial bees.

Across all three behaviors, the majority of the 6 trial bees consistently demonstrated extinction after the 13th and 14th intervals (~390-420 seconds), with the most bees no longer performing the behaviors at the 17th interval. By comparison, the majority of the 12-training trial bees consistently demonstrated extinction around the 11th and 12th intervals (~330-360 seconds), with the most bees no longer performing the behaviors at the 12th interval. These results align with the results previously demonstrated by Couvillon and Bitterman (1980). Tables 1, 2, and 3 below summarizes the findings across all 6-training trial and 12-training trial bees across each interval. These results are further supported when considering the traditional aggregate approach and descriptive statistics (see Table 1).

Table 1

Six and Twelve Trial Bees‟ Complete Results for the Landing Behavior

Timepoints	6-Trial Bees (*N* = 14)			12-Trial Bees (*N* = 14)		
	Total Bees Responding	PCC	C-Value	Total Bees Responding	PCC	C-Value
Interval 1	13	92.86	< 0.001	14	100	< 0.001
Interval 2	13	92.86	< 0.001	14	100	< 0.001
Interval 3	13	92.86	< 0.001	13	92.86	0.001
Interval 4	13	92.86	< 0.001	12	85.71	0.001
Interval 5	11	78.57	0.012	12	85.71	0.002
Interval 6	11	78.57	0.011	12	85.71	0.002
Interval 7	10	71.43	0.039	13	92.86	< 0.001
Interval 8	11	78.57	0.011	13	92.86	< 0.001
Interval 9	11	78.57	0.012	12	85.71	0.001
Interval 10	9	64.29	0.114	10	71.43	0.039
Interval 11	8	57.14	0.258	6	42.86	0.646
Interval 12	9	64.29	0.117	1	7.14	1.000
Interval 13	8	57.14	0.256	8	57.14	0.238
Interval 14	3	21.43	0.985	4	28.57	0.927
Interval 15	7	50.00	0.444	4	28.57	0.935
Interval 16	6	42.86	0.650	5	35.71	0.816
Interval 17	2	14.29	0.997	6	42.86	0.633
Interval 18	5	35.71	0.827	5	35.71	0.821
Interval 19	5	35.71	0.839	6	42.86	0.640
Interval 20	4	28.57	0.936	5	35.71	0.824

Note. These data represent the complete data across all 6-trial and 12-trial bees and across all intervals for the landing behavior. If the PCC is above 50%, then the majority of the bees are still performing the behavior and not demonstrating extinction. If the PCC is below 50%, then the majority of the bees are not performing the behavior and are demonstrating extinction. It is worth noting that the *c*-value is inversely related to the PCC, as one increases the other will decrease and vice versa.

Table 2

Six and Twelve Trial Bees" Complete Results for the Touching Behavior

Timepoints	6-Trial Bees ($N = 14$)			12-Trial Bees ($N = 14$)		
	Total Bees Responding	PCC	C-Value	Total Bees Responding	PCC	C-Value
Interval 1	12	85.71	0.001	11	78.57	0.01
Interval 2	12	85.71	0.001	12	85.71	0.001
Interval 3	10	71.43	0.03	12	85.71	0.002
Interval 4	11	78.57	0.01	11	78.57	0.01
Interval 5	8	57.14	0.21	11	78.57	0.004
Interval 6	7	50.00	0.39	10	71.43	0.02
Interval 7	10	71.43	0.03	10	71.43	0.03
Interval 8	9	64.29	0.08	12	85.71	0.001
Interval 9	8	57.14	0.21	8	57.14	0.19
Interval 10	8	57.14	0.20	7	50.00	0.37
Interval 11	8	57.14	0.20	5	35.71	0.77
Interval 12	7	50.00	0.40	1	7.14	1.00
Interval 13	5	35.71	0.79	5	35.71	0.78
Interval 14	3	21.43	0.98	4	28.57	0.90
Interval 15	7	50.00	0.39	3	21.43	0.97
Interval 16	4	28.57	0.92	4	28.57	0.90
Interval 17	2	14.29	1.00	5	35.71	0.78
Interval 18	5	35.71	0.78	3	21.43	0.97
Interval 19	4	28.57	0.92	4	28.57	0.90
Interval 20	4	28.57	0.93	4	28.57	0.91

Note. These data represent the complete data across all 6-trial and 12-trial bees and across all intervals for the touching behavior. If the PCC is above 50%, then the majority of the bees are still performing the behavior and not demonstrating extinction. If the PCC is below 50%, then the majority of the bees are not performing the behavior and are demonstrating extinction. It is worth noting that the *c*-value is inversely related to the PCC, as one increases the other will decrease and vice versa.

Table 3

Six and Twelve Trial Bees" Complete Results for the Pushing Behavior

Timepoints	6-Trial Bees ($N = 14$)			12-Trial Bees ($N = 14$)		
	Total Bees Responding	PCC	C-Value	Total Bees Responding	PCC	C-Value
Interval 1	13	92.86	0.001	14	100	< 0.001
Interval 2	11	78.57	0.004	11	78.57	0.002
Interval 3	12	85.71	< 0.001	11	78.57	0.002
Interval 4	12	85.71	< 0.001	11	78.57	0.001
Interval 5	6	42.86	0.52	6	42.86	0.44
Interval 6	6	42.86	0.51	8	57.14	0.10
Interval 7	9	64.29	0.08	8	57.14	0.11
Interval 8	10	71.43	0.02	10	71.43	0.01
Interval 9	8	57.14	0.15	5	35.71	0.66
Interval 10	6	42.86	0.52	6	42.86	0.43
Interval 11	7	50.00	0.31	5	35.71	0.67
Interval 12	6	42.86	0.51	0	0.00	1.00
Interval 13	4	28.57	0.89	2	14.29	0.99
Interval 14	2	14.29	0.99	4	28.57	0.83
Interval 15	3	21.43	0.96	4	28.57	0.85
Interval 16	4	28.57	0.88	4	28.57	0.84
Interval 17	1	7.14	1.00	3	21.43	0.95
Interval 18	2	14.29	0.99	2	14.29	0.99
Interval 19	3	21.43	0.96	2	14.29	0.99
Interval 20	3	21.43	0.96	3	21.43	0.95

Note. These data represent the complete data across all 6-trial and 12-trial bees and across all intervals for the pushing behavior. If the PCC is above 50%, then the majority of the bees are still performing the behavior and not demonstrating extinction. If the PCC is below 50%, then the majority of the bees are not performing the behavior and are demonstrating extinction. It is worth noting that the *c*-value is inversely related to the PCC, as one increases the other will decrease and vice versa.

Experiment 1 Aggregate Results

A 3 behavior (Landing, Tou

ching, Pushing) by 20 Time (Intervals 1-20) by 2 Group (6-Trial, 12-Trial) mixed model Analysis of Variance (ANOVA) with repeated measures on the first and second factors was conducted in the Statistical Package for Social Sciences (SPSS) V25 (IBM Corporation, Chicago, IL) to compare frequencies in behaviors over time for each training group. It is worth noting that the full multivariate tests could not be conducted because of insufficient degrees of freedom, due to the small sample size ($N = 28$) and number of within-subject factors. Specifically, the behavior by time interaction as well as the behavior by time by group (6 trial vs 12 trial) interaction could not be conducted.

Nevertheless, the sphericity assumption was violated for both behavior ($\chi^2(2) = 13.82$, $p = 0.001$) and time ($\chi^2(189) = 286.95$, $p < 0.001$). This tells us that an F correction must be used, therefore, the Greenhouse-Geisser correction was chosen; $\varepsilon_{\text{behavior}} = 0.70$, $\varepsilon_{\text{time}} = 0.44$. The tests of within-subjects effects revealed that there were large differences in the behaviors performed according to Cohen‟s (1988) effect size conventions; $F(1.40, 36.50) = 10.01$, $p = 0.001$, $\eta_p^2 = 0.28$. In considering the highest-ordered contrasts, we observed a quadratic contrast for the frequencies of the performed behaviors (landings, touches, or pushes); $F(1, 26) = 21.13$, $p < 0.001$.

The tests of within-subjects effects further revealed that there were large differences in the frequencies across time; $F(8.37, 217.69) = 12.96$, $p < 0.001$, $\eta_p^2 = 0.33$. Once again considering the highest-ordered contrasts, we observed a quadratic contrast of behavioral frequencies across time; $F(1, 26) = 10.43$, $p = 0.003$. Finally, tests of within-subjects effects further revealed that there were medium differences in the frequencies of the performed behaviors (landings, touches, pushes) across time; $F(8.43, 219.17) = 2.43$, $p = 0.01$, $\eta_p^2 = 0.09$. We observed a marginal linear contrast for the frequencies of landings, touches, or pushes across time; $F(1, 26) = 15.02$, $p = 0.052$. There were no differences observed between the two groups, behaviors, across time, or behaviors performed across time; p‟s > 0.05 and are therefore not considered or reported.

Interpreting the traditional statistics may lead one to conclude that there were no differences between the two groups of bees and that these bees extinguished the three behaviors at nearly equal rates across time. The overall interaction model was not observed to be „statistically significant,‟ thus follow-up pairwise comparisons were not considered. However, when the descriptive statistics are considered, we see the same pattern of observations as found in our OOM analysis.

Table 4 summarizes the mean frequencies of the 3 behaviors across time and between the two groups. As observed in the table and found in our OOM analyses, across time all three behaviors gradually declined and there was a difference between the performed behaviors such that the bees preferred to land more than push and push more than touch. However, by only considering the overall model from our traditional analyses, we miss the subtle differences being demonstrated between the two groups. Specifically, looking at interval 12 in table 4 you can see that there is a difference between the two groups of bees. The 12-trial bees had far fewer frequencies of landings, touches, and pushes, compared to the 6-trial bees. In fact, at the 12^{th} interval the cap was not pushed at all. The pattern observed between the two groups of bees aligns with the results of OOM, which would have likely been missed had we only considered the results of the overall model or aggregate findings.

Table 4

Average Frequencies of Landings, Touches, and Pushes for 6 and 12-Trial Bees

Timepoints	Landings		Touches		Pushes	
	6-Trial	12-Trial	6-Trial	12-Trial	6-Trial	12-Trial
	M (*CI*)	*M* (*CI*)	*M* (*CI*)	*M* (*CI*)	*M* (*CI*)	*M* (*CI*)
Interval 1	4.79 (3.75, 5.83)	4.71 (3.89, 5.53)	2.21 (1.45, 2.97)	2.21 (1.41, 3.01)	4.43 2.63, 6.23)	3.43 2.35, 4.51)
Interval 2	3.43 (2.39, 4.47)	3.86 (3.00, 4.72)	2.07 (1.27, 2.87)	2.21 (1.41, 2.87)	2.43 0.86, 4.00)	2.71 1.57, 3.85)
Interval 3	3.21 (2.13, 4.29)	3.14 (2.22, 4.06)	1.64 (0.78, 2.50)	2.14 (1.36, 2.42)	2.64 1.52, 3.76)	1.86 0.96, 2.76)
Interval 4	2.71 (1.97, 3.45)	2.57 (1.63, 3.51)	1.86 (1.00, 2.72)	1.79 (1.10, 2.55)	2.36 (1.40, 3.32)	1.79 0.91, 2.67)
Interval 5	1.93 (1.20, 2.66)	2.93 (1.85, 4.00)	0.86 (0.37, 1.35)	1.64 (1.01, 1.49)	0.93 0.24, 1.61)	1.50 (0.40, 2.60)
Interval 6	2.57 (1.33, 3.80)	2.86 (1.96, 3.76)	1.57 (0.26, 2.88)	1.50 (0.89, 2.18)	2.71 0.57, 4.85)	2.00 (0.73, 3.27)
Interval 7	1.64 (0.91, 2.37)	2.57 (1.84, 3.30)	1.14 (0.65, 1.63)	1.29 (0.56, 1.87)	1.71 (0.57, 2.85)	2.00 (0.63, 3.37)
Interval 8	2.57 (1.53, 3.61)	2.86 (1.96, 3.76)	0.93 (0.44, 1.42)	1.64 (1.11, 1.46)	2.36 (1.13, 3.59)	1.50 (0.52, 2.48)
Interval 9	1.57 (1.00, 2.14)	2.36 (1.56, 3.16)	1.14 (0.40, 1.88)	0.71 0.34, 1.51)	0.86 (0.37, 1.35)	2.00 (0.26, 3.74)
Interval 10	2.36 (1.20, 3.52)	1.79 (0.88, 2.69)	1.29 (0.51, 2.07)	1.00 (0.39, 1.90)	1.21 (0.27, 2.15)	1.36 (-0.29, 3.00)
Interval 11	1.43 (0.65, 2.21)	1.00 0.16, 1.84)	1.21 (0.52, 1.90)	0.64 (0.11, 1.74)	1.64 (0.41, 2.87)	0.36 0.11, 0.61)
Interval 12	1.86 (0.90, 2.82)	0.07 (-0.07, 0.21)	1.07 (0.38, 1.76)	0.07 (-0.07, 1.21)	1.64 (0.52, 2.76)	0.00 (0,0)
Interval 13	1.00 (0.35, 1.65)	1.07 (0.40, 1.74)	0.71 (-0.11, 1.53)	0.64 (0.15, 1.2)	1.57 (-0.92, 4.06)	0.71 (-0.43, 1.85)
Interval 14	0.71 (-0.03, 1.45)	0.93 (0.09, 1.77)	0.64 (-0.144, 1.42)	0.57 (-0.04, 1.25)	0.64 (-0.22, 1.50)	0.71 -0.15, 1.57)
Interval 15	1.07 (0.38,	0.79 (-0.09,	1.07 (0.19,	0.36 -0.09,	0.79 (-0.17,	1.50 (-0.85,

	1.76)	1.67)	1.95)	1.52)	1.75)	3.85)
Interval 16	1.00 (0.10,1.90)	0.57 (0.08, 1.06)	0.71 (-0.03, 1.45)	0.64 (0.01, 1.34)	0.93 0.05, 1.81)	0.79 (-0.01, 1.59)
Interval 17	0.21 (-0.08, 0.50)	1.43 (0.37, 2.49)	0.29 (-0.14, 0.72)	0.79 (0.16, 0.92)	0.07 (-0.07, 0.21)	0.36 (-0.09, 0.81)
Interval 18	0.86 (0.15, 1.57)	0.93 (0.07, 1.79)	0.79 (0.05, 1.53)	0.36 (-0.03, 1.18)	0.29 (-0.14, 0.72)	0.50 (-0.34, 1.34)
Interval 19	0.71 (0.12, 1.30)	0.50 (0.17, 0.83)	0.57 (0.00, 1.14)	0.36 (0.03, 0.90)	0.64 (-0.12, 1.40)	0.29 (-0.14, 0.72)
Interval 20	0.64 (-0.03, 1.31)	1.43 (0.16 , 2.70)	0.50 (0.05, 0.95)	0.79 (0.05, 1.24)	0.64 (-0.07, 1.35)	0.36 (-0.09, 0.81)

Note. The average frequencies for 6 and 12 trial bees separated out into each individual interval. The confidence intervals for each trial are reported in parentheses.
M = Mean
CI = Confidence Interval

Experiment 2: Overshadowing in the Cap Pushing Response

In our second experiment, we explored overshadowing (a decrease in response to one conditioned stimulus because of the presence of another stimulus) using three different stimuli consisting of scent, color, and odor (VandenBos, 2007). These stimuli were used to further our understanding of the CPR by incorporating it into a choice preference paradigm. Previous research had revealed that the scent of jasmine overshadowed the color orange when bees were given a choice between the two stimuli (Couvillon & Bitterman, 1980).

The rationale for this experiment is to determine whether including a cap would alter the bee‟s preference for odor and color in an overshadowing experiment. The cap acts as an obstacle that must be pushed in order for a honey bee to access the food in the well. We performed an acquisition phase which paired the three stimuli (jasmine, orange, and cap) with a target that had a well filled with sugar solution. After acquisition, an extinction phase was initiated where the three stimuli were separated, and each well was filled with tap water. This allowed researchers to analyze the CPR technique in an overshadowing experiment.

Initially, we hypothesized that the honey bee‟s original choice would be the cap and the highest frequency of choice throughout extinction. This is because the cap covered the well and therefore would have to be moved to find the omission of sucrose solution.

Materials and Methods

Foragers from three honey bee colonies were trained to an artificial feeder approximately four meters from the respective colonies. The bees were captured, marked, and shaped to push the cap as in the previous experiment. The target was located on a table approximately two meters from the artificial sucrose syrup feeder and four meters from the hive. A black poster board 27.94 cm x 35.56 cm was placed on top of the table to differentiate the target from the white tabletop on the desired platform. In addition, the brand "LBK™" nail lacquer was used to mark the bees on the abdomen and thorax to differentiate bees during the experiment. Marking took place while the bee was feeding on the desired platform approximately two meters from the artificial sucrose syrup feeder.

Honey bees (N = 96) were trained to fly to a target containing a 50% sucrose solution by volume. The resistance method of shaping was used to train the bees to push the cap (see experiment 1). Once the bee completed pushing the cap twice, acquisition trials commenced. Acquisition instruction included training the bees to push a cap on a platform which consisted of the color orange (Orange Ochre, Valspar, 2010-1, Minneapolis, MN), the scent of jasmine measured as 1 drop approximately 0.1 mL (Mary Tylor Naturals, B08JF9WZC1, Fort Myers, FL), and the physical stimulus of cap pushing (weight = 0.47 g, height = 0.5 cm, length = 1.5 cm). The platforms each had their own wells filled with a 50% sucrose solution. After 6 acquisition trials, a 10-minute extinction phase-separated into twenty 30 second intervals began.

The three stimuli were then separated approximately 15cm apart from center to center, and each of the three wells was filled with water. Initial choice preferences and the overall order of choice throughout the 10-minute extinction trials were recorded. All data for this experiment can be accessed by request.

As in our previous experiment, bees were selected from a common feeder regularly visited by three different hives, and the targets were washed and cleaned based on standard methods (Couvillon & Bitterman, 1980). Our sample size was based on previous research and personal experience with this preparation. Further, honey bees are not listed under any ethical codes concerning humane treatment of animals in research under USA or Greek law; therefore, no ethical review of the research protocol was conducted.

Results and Discussion

The data for all 96 bees were analyzed to determine whether or not the cap would potentially overshadow the other two stimuli (odor & color). As noted in the procedure, the control bees (n = 24) were not presented with all three stimuli. Instead, they were only presented with the color and odor stimuli to assess if the bees had an odor preference as found in previous literature (Couvillon & Bitterman, 1982, 1987, 1989). However, the experimental bees (n = 72) were presented with all three stimuli (color, odor, and the cap) to see if the cap stimuli influenced choice preference. The first stimulus touched (cap, jasmine, or orange) by the bee was recorded for both groups. Overall, the honey bees preferred the scent of jasmine over the color orange and the physical act of pushing the cap was their initial choice preference and overall choice preference. This pattern follows previous literature that honey bees rely highly on olfactory cues (Couvillon & Birtteman, 1982, 1987, 1988, 1989; Funayama et al., 1995).

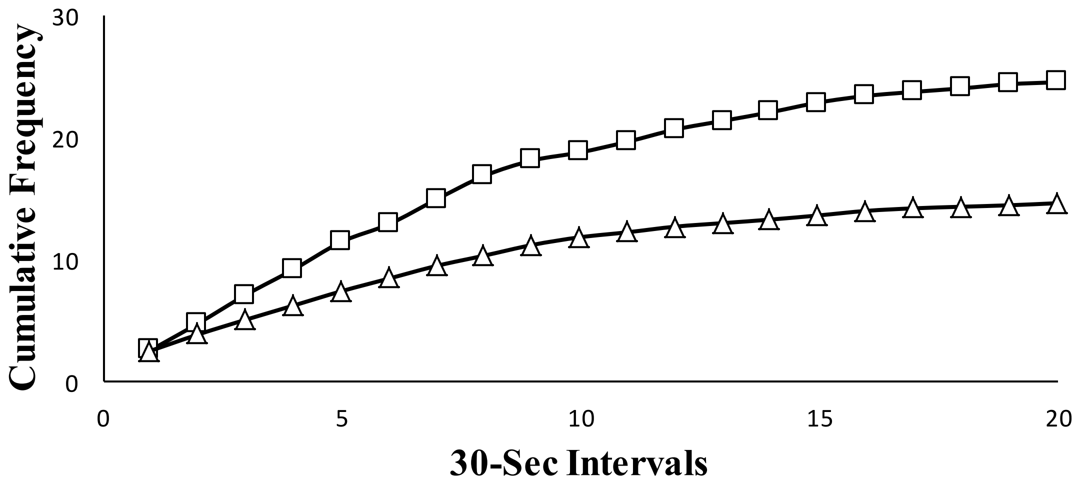
In regard to initial first touch, most control bees made contact with the orange platform first (15/24 bees, 62.50 %), followed by jasmine (9/24 bees, 37.50%). A larger minority of experimental bees also made contact with the orange platform first (34/72 bees, 47.22%), followed by jasmine (22/72 bees, 30.56%), and finally the cap (16/72 bees, 22.22%).

It was initially expected that the number of behavioral responses within the control bees would match a pattern such that the odor (jasmine, J) would be greater than color (orange, O), J > O. This pattern was compared across all control bees, such that bee *N*'s J touches would be greater than bee *N*'s O touches. For the experimental bees, it was hypothesized that if the cap (C) overshadows the other two stimuli, then the expected pattern produced should follow such that, C > J > O. This pattern was compared across all experimental bees, such that bee *N's* C touches should be greater than Bee *N's* J touches, which should be greater than bee *N's* O touches.

Results for the control bees revealed that 20 out of 24 bees were correctly classified, meaning they fit the entire ordinal pattern (J > O) as expected, yielding a PCC = 83.33%, $c < 0.001$, see Figure 5. These results indicate that the overwhelming majority of control bees preferred the odor (jasmine) stimuli more than the color (orange) stimuli. Out of the four bees that were not correctly classified, two bees went to J and O at equal rates, whereas the other two bees were exactly opposite of the expected pattern (i.e., J > O).

Figure 5

Cumulative Response Curves for the Control Bees' Choices

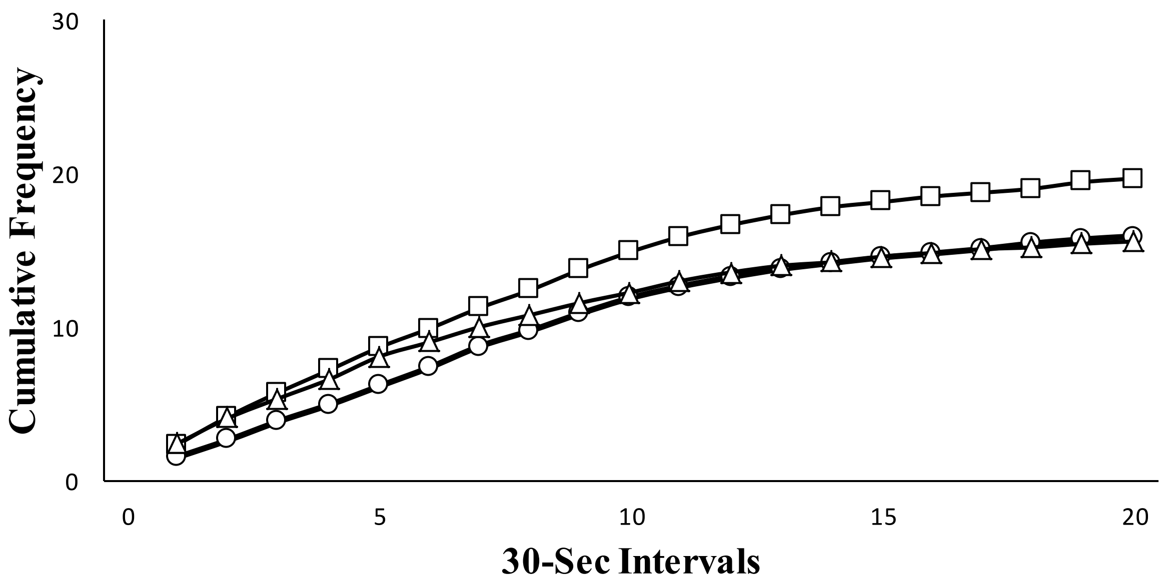


Note. This figure represents the cumulative frequencies of the jasmine (black line with squares) and orange (black line with triangles) choices for the control bees.

Results for the experimental bees did not support our initial hypothesis that the act of pushing the cap would alter the odor/color relationship. It was expected that if the cap overshadowed the odor (jasmine) and the color stimuli, then most cases should result such that C > J > O. The analysis revealed that only 46.76% (101/216) observations matched the pattern, see Figure 6. Only 10/72 bees were completely correctly classified for the entire pattern (C > J > O) yielding a meager PCC = 13.89%, *c* = 0.66. All possible combinations of pairs were compared further to break down the 101 correctly classified observations. It is worth noting that only unique pairwise comparisons are reported here; pairwise comparisons that have already been reported in our analyses (e.g., J > O) are omitted in the additional analyses to follow to avoid presenting repetitive information.

Figure 6

Cumulative Response Curves for the Experimental Bees' Choices



Note. This figure represents the cumulative frequencies of the jasmine (black line with squares), orange (black line with triangles), and cap (black line with circles) choices for the experimental bees.

The C > J pairwise comparison was unimpressive as only 20/72 cases matched, yielding a PCC = 27.78%, *c* = 1.00. The C > O pairwise comparison was similarly unimpressive as only 37/72 cases matched, yielding a PCC = 51.39%, *c* = .32. Finally, the J > O pairwise was the best pairwise comparison as 44/72 cases matched, yielding a PCC = 61.11%, *c* = .01. Due to the underwhelming results of the C > J > O, we re-conducted the analyses with two new patterns, 1) J > O > C and 2) J > C > O.

We were initially interested in exploring the possibility that the „cap" was the more salient stimuli for the bees, given that they must first push the cap in order to get to the food source. However, previous research has suggested that odors (e.g., jasmine) may be more salient

(Couvillon & Bitterman, 1982, 1987, 1989); as such, we re-conducted our analyses to determine if the odor (e.g., jasmine) was overshadowing the other two stimuli.

Results for the J > O > C analysis revealed that 56.48% (122/216) of the total observations were correctly classified and 9/72 total bees were completely correctly classified, yielding a PCC = 12.50%, $c = 0.78$. All possible combinations of pairs were compared to break down the 122 correctly classified observations. The J > C pairwise comparison was impressive as 49/72 bees matched expectations, PCC = 68.06%, $c < 0.001$. The O > C pairwise comparison revealed that 29/72 bees matched expectations, PCC = 40.28%, $c = 0.89$. It is worth noting that in the previous analysis, C > J > O, the bees preferred C to O at only a slight majority (51.39%), whereas, in the J > O > C analyses, O was selected more than C for the minority of bees (40.28%). This discrepancy is found because ties are treated as incorrect classifications (see Grice, 2011). If the bee preferred cap and orange equally, then that bee would be incorrectly classified, as was the case for eight total bees in this comparison.

Finally, the results for the J > C > O analysis were slightly better as 60.19% (130/216) of the observations were correctly classified, and 20/72 total bees were completely correctly classified, yielding a PCC = 27.78, $c < 0.001$. The pairwise comparisons are not considered, as the pairwise classifications yielded identical results to the pairwise results from the J > O > C analysis. A summary of all results and possible pairwise comparisons can be found in Table 5. For ease of comparison with previous literature (Couvillon & Bitterman, 1982, 1987, 1989), we present the aggregated cumulative response curves for our control bees and experimental bees within Figures 5 and 6, respectively.

Table 5

Complete Over-Shadowing Analyses

Pairwise Comparison	Correctly Classified Bees	PCC	*c*-value
Control Bees (N = 24)			
J > O	20	83.33	< 0.001
O > J	2	8.33	1.00
J = O	2	8.33	1.00
Experimental Bees (N = 72)			
C > J > O	10	13.89	0.66
J > O > C	9	12.50	0.78
J > C > O	20	27.78	< 0.001
Possible Pairwise Comparisons			
C > J	20	27.78	1.00
C > O	37	51.39	0.32
J > O	44	61.11	0.01
J > C	49	68.06	< 0.001
O > C	29	40.28	0.89
O > J	23	31.94	0.99
C = J	3	4.17	1.00
C = O	6	8.33	1.00
J = O	5	6.94	1.00

Note. J = jasmine, O = orange, C = cap
This table provides all primary analyses and subsequent pairwise comparisons, including cases in which the stimuli were equivalently preferred by the bees.

Experiment 2 Aggregate Results

A 3 stimuli (Jasmine, Orange, Cap) by 20 Time (Intervals 1-20) Analysis of Variance (ANOVA) with repeated measures was conducted in SPSS V25 to compare frequencies in stimuli choice over time for the experimental group. The sphericity assumption was violated for both stimuli ($\chi^2(2) = 7.72, p = 0.02$), time ($\chi^2(189) = 522.58, p < 0.001$), and the interaction stimuli by time ($\chi^2(740) = 1478.36, p < 0.001$). This tells us that a F correction must be used, therefore, the Greenhouse-Geisser correction was chosen for the main effect of time and the stimuli by time interaction ($\varepsilon_{\text{time}} = 0.49$, $\varepsilon_{\text{stimuli*time}} = 0.49$) and a Huyn-Feldt correction was chosen for the main effect Stimuli since epsilon was greater than 0.75 ($\varepsilon_{\text{time}} = 0.93$).

The tests of within-subjects effects revealed that there was a medium difference in stimuli choice according to Cohen‟s (1988) effect size conventions; $F(1.86, 131.74) = 6.67, p = 0.002$, $\eta_p^2 = 0.09$. In considering the highest-ordered contrasts, we observed a quadratic contrast for the frequencies of the Stimuli (jasmine, orange, cap); $F(1, 71) = 18.18, p < 0.001$. The tests of within-subjects effects revealed that there were large difference in the frequencies of choices across time, according to Cohen‟s (1988) effect size conventions; $F(9.22, 654.80) = 51.77, p < 0.001$, $\eta_p^2 = 0.42$. In considering the highest-ordered contrasts, we observed a quadratic contrast for the frequencies of choices over time; $F(1, 71) = 8.64, p = 0.004$. The tests of within-subjects effects revealed that there was a small difference in the frequencies of stimuli choice across time according to Cohen‟s (1988) effect size conventions; $F(18.73, 1330.01) = 2.71, p < 0.001$ $\eta_p^2 = 0.04$. In considering the highest-ordered contrasts, we observed a linear contrast for the frequencies of the interaction stimuli across time; $F(1, 71) = 7.75, p = 0.007$.

Stimuli pairwise comparisons showed that the bees preferred jasmine ($M = 0.98$; $SE = 0.05$) more than the cap ($M = 0.80$; $SE = 0.05$) and orange; $M = 0.78$; $SE = 0.06$; p‟s < 0.01.

Table 6 presents the descriptive results for the frequencies of stimuli choice across each interval. As seen in the table, in the first interval, the bees chose the orange stimuli more than the jasmine or cap stimuli, but after the first interval the bees typically preferred the jasmine more than either the cap or orange stimuli. Moreover, the frequencies of each stimuli choice generally declined over time. Although, as observed by our quadratic contrast, there was a difference found between some of the isolated intervals (e.g., intervals 6 and 7 for jasmine, see table 6). The mean frequencies for the isolated intervals slightly increased from one interval to the next. The results produced from our aggregate findings align with the results produced from our primary OOM findings; the bees typically preferred jasmine over orange and the cap, and initially preferred the orange over the cap, but over time the bees chose the cap more frequently than the orange stimuli.

Table 6

Average Frequencies of Jasmine, Orange, and Cap Choices Across Time for Experimental Bees

Timepoints	Jasmine		Orange		Cap	
	Mean	*SE*	*Mean*	*SE*	*Mean*	*SE*
Interval 1	2.36	0.17	2.43	0.19	1.53	0.17
Interval 2	1.82	0.16	1.64	0.18	1.10	0.19
Interval 3	1.56	0.15	1.28	0.15	1.22	0.12
Interval 4	1.46	0.14	1.24	0.12	1.07	0.13
Interval 5	1.47	0.14	1.49	0.16	1.22	0.15
Interval 6	1.24	0.13	0.94	0.13	1.22	0.12
Interval 7	1.35	0.11	0.99	0.12	1.38	0.16
Interval 8	1.15	0.13	0.78	0.13	1.01	0.15
Interval 9	1.33	0.13	0.79	0.12	1.13	0.14
Interval 10	1.17	0.14	0.68	0.13	1.01	0.12
Interval 11	0.99	0.15	0.75	0.12	0.75	0.11
Interval 12	0.76	0.13	0.56	0.12	0.61	0.11
Interval 13	0.64	0.13	0.43	0.12	0.53	0.11
Interval 14	0.51	0.11	0.24	0.08	0.39	0.08
Interval 15	0.33	0.09	0.36	0.11	0.36	0.10
Interval 16	0.36	0.09	0.24	0.07	0.26	0.07
Interval 17	0.25	0.07	0.25	0.08	0.29	0.07
Interval 18	0.26	0.07	0.13	0.05	0.36	0.11
Interval 19	0.43	0.11	0.22	0.07	0.26	0.07
Interval 20	0.22	0.08	0.15	0.05	0.19	0.07
Total	**0.98**	**0.05**	**0.78**	**0.06**	**0.80**	**0.05**

Note. The total mean frequencies of each stimuli have been bolded for emphasis.
SE = Standard Error

A 2 stimuli (Jasmine and Orange) by 20 Time (Intervals 1-20) Analysis of Variance (ANOVA) with repeated measures was conducted in SPSS V25 to compare frequencies in stimuli choice over time for the control. The sphericity assumption was violated for both times ($\chi^2(189) = 287.20$, $p < 0.001$) and the interaction stimuli by times ($\chi^2(189) = 340.53$, $p < 0.001$). This tells us that a F correction must be used, therefore, the Greenhouse-Geisser correction was chosen; $\varepsilon_{time} = 0.41$, $\varepsilon_{stimuli*time} = 0.36$. The tests of within-subjects effects revealed that there was a large difference in the frequencies of stimuli choice according to

Cohen‟s (1988) effect size conventions; $F(1, 23) = 27.98$, $p < 0.001$, $\eta_p^2 = 0.55$. In considering the highest-ordered contrasts, we observed a linear contrast for the frequencies of the preferred stimuli; $F(1, 23) = 27.98$, $p < 0.001$. The tests of within-subjects effects revealed that there was a large difference in the frequencies of choices across time according to Cohen‟s (1988) effect size conventions; $F(7.80, 179.30) = 18.57$, $p < 0.001$, $\eta_p^2 = 0.45$. In considering the highest-ordered contrasts, we observed a quadratic contrast for the frequencies of choices over time; $F(1, 23) = 6.06$, $p = 0.02$. Specifically, the bees‟ frequencies of stimuli choices declined as time went on (see Table 7).

Finally, the tests of within-subjects effects revealed that there was a small to medium difference in the frequencies of stimuli choice across time according to Cohen‟s (1988) effect size conventions; $F(6.91, 159.00) = 1.62$, $p = 0.13$, $\eta_p^2 = 0.07$; although the small to medium differences were not „statistically‟ significant according to $p < 0.05$ conventions. In considering the highest-ordered contrasts, we observed a negative linear contrast for the frequencies of stimuli choice over time; $F(1, 23) = 12.85$, $p = 0.002$. Once again, we observed a similar pattern as found with the experimental bees, and our primary OOM analyses, the control bees typically preferred the jasmine stimuli over the orange stimuli (see Table 7).

Table 7

Timepoints	Jasmine		Orange	
	Mean	*SE*	*Mean*	*SE*
Interval 1	2.63	0.35	2.58	0.36
Interval 2	2.21	0.23	1.33	0.26
Interval 3	2.33	0.34	1.21	0.23
Interval 4	2.08	0.32	1.17	0.20
Interval 5	2.25	0.30	1.17	0.28
Interval 6	1.46	0.25	1.04	0.24
Interval 7	2.04	0.29	1.04	0.24
Interval 8	1.92	0.35	0.83	0.23
Interval 9	1.25	0.27	0.88	0.27
Interval 10	0.67	0.19	0.63	0.17
Interval 11	0.88	0.24	0.42	0.15
Interval 12	1.00	0.28	0.46	0.19
Interval 13	0.71	0.22	0.29	0.13
Interval 14	0.71	0.27	0.29	0.11
Interval 15	0.79	0.26	0.33	0.17
Interval 16	0.54	0.15	0.38	0.15
Interval 17	0.33	0.17	0.21	0.10
Interval 18	0.29	0.15	0.13	0.07
Interval 19	0.33	0.14	0.13	0.09
Interval 20	0.17	0.08	0.17	0.12
Totals	**1.23**	**0.09**	**0.73**	**0.07**

Note. The total mean frequencies of each stimulus have been bolded for emphasis.
SE = Standard Error

Experiment 3: Discriminate punishment of the cap pushing response in honey bees

The purpose of this experiment is to determine whether electric shock can be incorporated into the CPR paradigm. The rationale behind this experiment is to provide a full picture of how both positive and negative stimuli can influence learning the CPR. Moreover, previous free flight and harnessed honey bee proboscis extension response (PER) studies found that aversive conditioning is effective in modifying the behavior of honey bees (e.g., Abramson, 1986; Giurfa & Sandoz, 2012; Smith et al., 1991).

Materials and Methods

Each of the 16 bees received 12 training trials in a simultaneous punishment situation in which both targets were presented. For half of the bees, the punished target was the cross (weight = 0.31 g, height = 0.5 cm, length = 1.5 cm); for the remaining 8 bees, the punished target was the cap (weight = 0.47 g, height = 0.5 cm, length = 1.5 cm). The bees were shocked with 9 V 1.3 mA if they picked the wrong target (cap or cross); there seemed to be no initial bias with the targets. To account for directional biases, the cross and round caps were counterbalanced in each of the conditions. As in our previous experiment bees were chosen from a common feeder regularly visited by three different hives, shaped to push the cap or the cross, and the targets were washed and cleaned based on standard methods. Our sample size was based on previous research and personal experience with this preparation. Further, Honey bees are not listed under any ethical codes concerning humane treatment of animals in research under USA or Greek law; therefore, no ethical review of the research protocol was conducted.

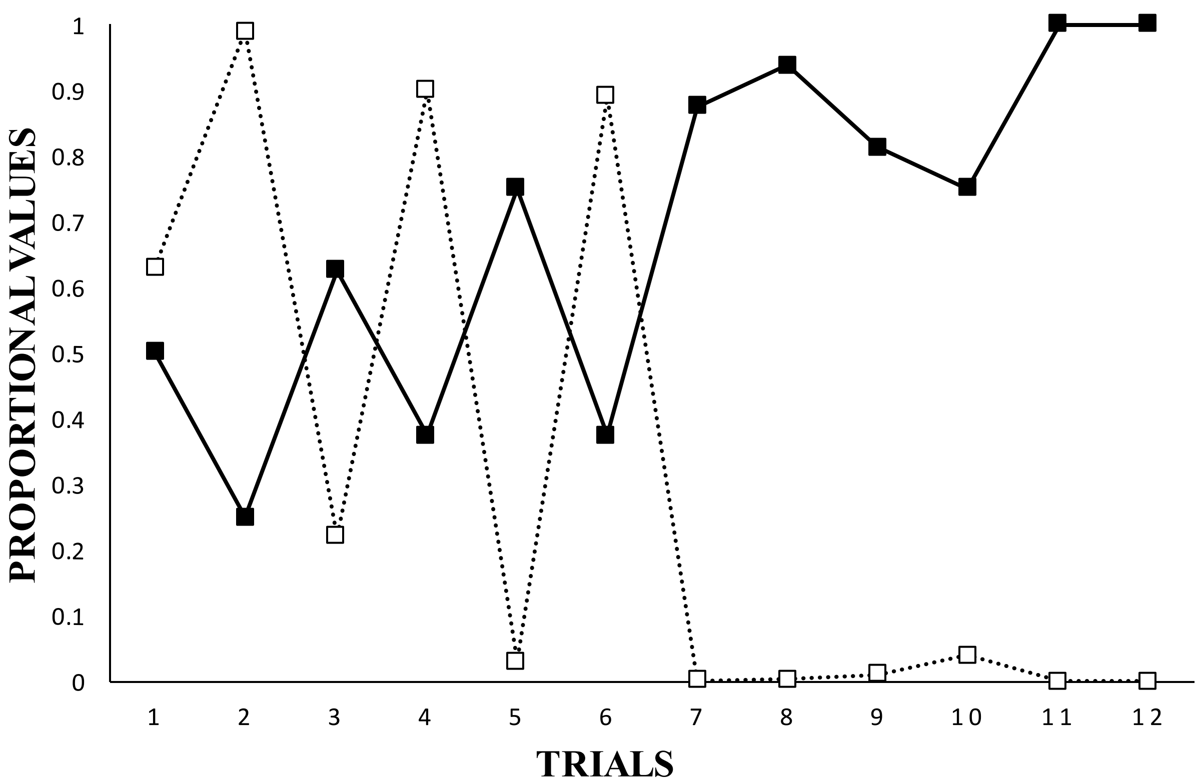
Results and Discussion

The data for all 16 bees were first analyzed utilizing an *a priori* pattern matching procedure within OOM. For this analysis, the pattern was defined such that a bee choosing the non-punished cap was classified as „correct‟ and scored as „1.‟ The bee that chose the punished cap was classified as „incorrect‟ and was scored as „0.‟ It was expected that bees could learn to differentiate between the punished and non-punished caps. Consistent with expectation, results revealed that the percentage of correctly classified (PCC) choices increased across the 12 trials. At trial 1, the PCC was only 50%, and the randomization test revealed this result could be interpreted as a product of physical chance, $c = 0.60$. Figure 8 shows the combined PCCs for all bees across the 12 trials, and as can be seen in the figure, by trial 7, the overwhelming majority of bees had learned the correct choice, PCC = 87.50%, *c* = 0.001. Moreover, by trials 11 and 12, all bees had learned the correct choice, PCC‟s = 100%, c‟s < 0.001.

Differences in the learning patterns were further examined by analyzing the cumulative frequencies of each individual bee‟s correct choices across the 12 trials. Within the OOM software, an analysis titled the *Concatenated Ordinal Analysis,* was relied upon to compare the individual‟s cumulative frequencies to an expected ordinal pattern of monotonic increasing correct responses (viz., Trial 1 < Trial 2 < Trial 3, etc.), similar to the analysis conducted in Experiment 1. If the bees were continuously choosing the correct target, then a monotonic increasing relationship should be observed for the individual bees and high PCC index should be computed for each bee. The resulting PCC indices for the individual bees varied from 63.64% to

Figure 7

Response PCC Curve for All Individual Bees



Note. The solid black line represents the proportions for the PCC's computed for the total number of correct choices across all 16 bees and across the 12 trials. The dashed black line indicates the corresponding *c*-values across all trials. The proportional values were obtained by converting the PCC across all bees correct choices to proportional frequencies $(\frac{PCC}{100})$ at each trial.

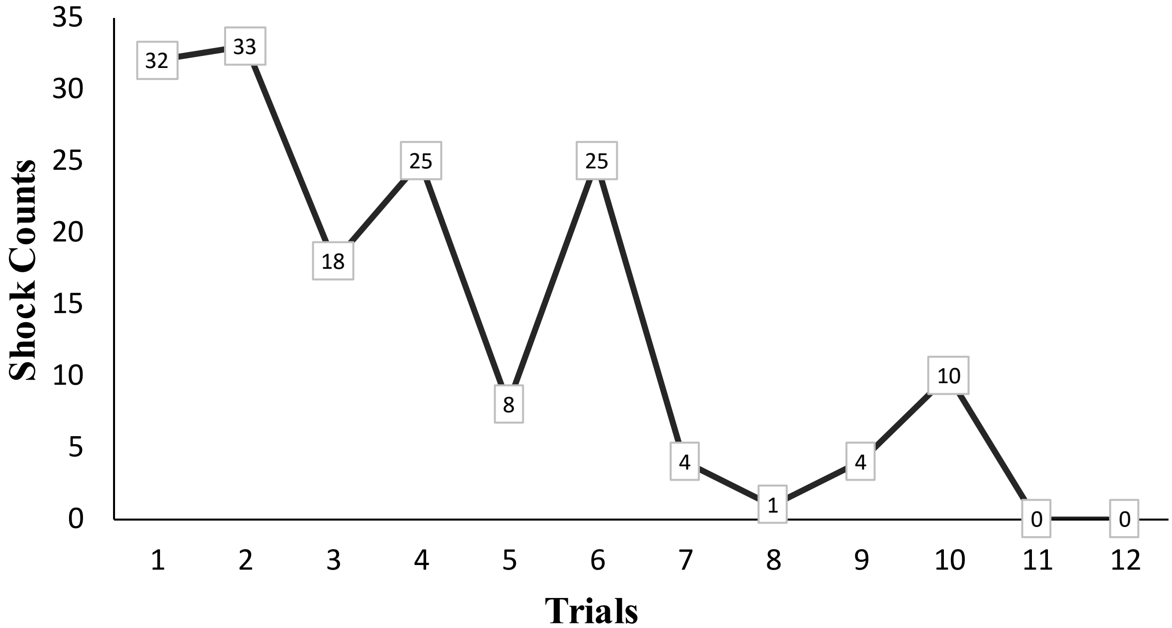
90.91%, with the randomization *c*-values all less than 0.09.

Additional ordinal analyses showed that one bee consistently chose the correct cap across all trials after an incorrect choice in the first and fourth trials. Contrarily, the other fifteen bees did not consistently choose the correct cap across the first six trials. The cumulative frequencies for these 15 bees did not show evidence of monotonic increases during the first six trials. However, analyses for the last six trials showed that the bees chose the correct cap with greater consistency (see Figure 7). Half of the bees from trials 7 through 12 had PCCs of 80% (randomization *c's* < 0.05), and the PCCs computed for the other half were all equal to 100% (randomization *c's* < 0.01).

We found that our bees readily learned to avoid shock by selecting the correct responses (i.e., shock-free target) as shown by a decrease of shock stimuli presented to the bees throughout the proceeding trials. It is important to note that as in the Smith et al (1991) proboscis conditioning study, both targets contained sucrose. The results indicated that across the 12 trials, the proportion of bees selecting the unpunished target typically increased, and the number of shocks received by the bees generally decreased. By trials 11 and 12, no bees received punishment (see Figure 8), indicating they avoided the shock (see also Abramson, 1986). Moreover, on average, the bees received fewer shocks (i.e., made more correct choices) as the trials progressed (see Table 8). Our results are similar to results previously found using a proboscis conditioning situation in harnessed forager bees and free flying situations (Abramson, 1986; Smith et al., 1991).

Figure 8

Total Shocks Across Each Trial



Note. These data represent the total shocks received at each trial for all 16 bees.

Experiment 3 Aggregate Results

A 12 Time (Trials 1-12) repeated measures ANOVA was conducted in SPSS V25 to compare the frequencies of shock over time. The tests of within-subjects effects revealed that there were large differences in the amount of shocks the bees received over time, according to Cohen‟s (1988) conventions; $F(11, 165) = 6.21$, $p < 0.001$, $\eta_p^2 = 0.29$. Specifically, we observed a negative linear contrast over time; $F(1, 11.10) = 107.22$, $p < 0.001$. These results align with those found in our primary OOM analyses, on average, the bees received fewer shocks as time progressed and by the 11th and 12th trials, no bees were shocked. See Table 8 to see shocks received over the 12-trials.

Table 8

Average Frequencies of Shock Over Each Trial

Time Point	Mean	SD
Trial 1	2.00	2.37
Trial 2	2.06	1.44
Trial 3	1.13	1.75
Trial 4	1.53	1.50
Trial 5	0.50	0.97
Trial 6	1.56	1.46
Trial 7	0.25	0.77
Trial 8	0.06	0.25
Trial 9	0.25	0.58
Trial 10	0.63	1.20
Trial 11	0	0
Trial 12	0	0

General Discussion

Summary of Findings

The cap pushing response was developed in 2016 (Abramson et al., 2016). In the current series of experiments, we looked at a progression of the cap pushing response (CPR) utilizing advanced discrimination tasks. Throughout our experiments, naïve control bees and trained experimental bees were used to show different frequencies of learning.

Experiment 1 looked at the role of extinction within the CPR. We found that the majority of the 12 trial bees stopped responding earlier than the 6 trial bees. Our results were similar to previous research where the 6 trial bees produced a longer extinction duration than the 12 trial bees (Couvillon & Bitterman, 1980). Furthermore, while the aggregate results do not offer support for the hypotheses at $p < 0.05$, the patterns of the descriptive statistics still reveal that on average, the 12-trial bees quit performing the behaviors prior to the 6-trial bees. Specifically, as observed in our OOM findings, the average frequency of behaviors performed by the 12-trial bees was lower than the 6-trial bees during the 12th interval. This experiment revealed that the 6-trial bees extinguished at a lower rate, suggesting that this number of trials is more effective for behavioral retention.

Experiment 2 sought to determine if prior learning (cap pushing) could be used to overshadow preferences for the color orange and odor jasmine. During the acquisition phase, all the stimuli were paired together while training the CPR technique. We then used extinction to see the bee‟s choice preference throughout a 10-minute extinction session split into twenty 30 second intervals. Our findings followed previous research (Couvillon & Bitterman, 1982) and suggests that honey bees rely highly on olfactory cues: the scent of jasmine overshadowed both the cap and color orange.

Experiment 3 helped determine whether electric shock can be incorporated into the CPR paradigm. We found that punishment is effective in modifying the choice behavior of honey bees. OOM and traditional statistics found that as bees progressed through the trials, they quickly learned which cap was shock free and which provided the punishment.

Combined, these experiments help further our knowledge of honey bees‟ learning abilities. With these results researchers will now know the ideal number of training trials needed to produce the best results for behavioral retention. They will also be able to consider the roles of stimuli preference and aversive conditioning effects on honey bees‟ learning.

Observational Orientated Modeling (OOM)

Of note, using OOM required no computations of means, standard deviations, variances, or *p*-values. Instead, the utilization of OOM provides the researcher with the exact quantity of bees that responded according to one‟s hypotheses (see Grice et al., 2020).

More importantly, as summarized in Table 9, in addition to providing the same information as the traditional aggregate statistics, the OOM analyses enabled us to observe patterns within our data that may have been missed if the traditional aggregate statistics were employed. Specifically, in Experiment 1, the primary interaction effect was not „statistically significant‟ which may lead one to believe that the bees extinguished all the behaviors at near equal rates over time. However, aggregate descriptive statistic show the same general trend as the OOM findings. Around the 12th interval the 12-trial bees are performing far fewer behaviors than the 6-trial bees. OOM, therefore, enabled us to observe potentially important and meaningful relationships, which may have otherwise been overlooked or „washed‟ within the average bee.

Table 9

Comparison Between OOM Findings and Supplementary SPSS Findings for the Experiments

Experiment	Hypotheses	OOM Findings	Aggregate Findings	Compare/Contrast
#1 Extended Training	12-Trial bees would extinguish the landing, touching, and pushing behavior earlier than 6-Trial bees.	Most 12-Trial bees extinguished the behaviors after the 11^{th} interval. Most 6-Trial bees extinguished after the 13^{th} interval.	Frequencies of bee‟s behaviors declined at „equal rates‟ as the overall interaction effect was not statistically significant.	OOM found meaningful patterns readily. Aggregate results would not have been considered further due to the non-significant interaction, though the descriptive statistics reveal the same pattern found by OOM.
#2 Overshadowing	Jasmine would overshadow the orange and cap stimuli.	The majority of experimental and control bees preferred the jasmine over both the orange and cap stimuli, but preferred the cap stimuli slightly more than orange stimuli.	On average, experimental bees preferred jasmine more than orange (except for the first interval), and the cap. The cap was preferred on average more than the orange, although only marginally. The Control bees preferred jasmine far more than the orange stimuli.	The OOM and aggregate findings offer similar conclusions. Jasmine was preferred more than the cap/orange stimuli. There is a marginal preference to the cap over orange for experimental bees.
#3 Punishment	Honey bees would be able to discriminate cap types with a shock form of negative reinforcement.	After the 6^{th} trial, the majority of bees no longer choose the punished cap and have successfully learned which cap is not punished.	On average, the bee‟s frequencies of shocks (i.e., incorrect choices) decreases over time. By the 8^{th} trial the bees receive 0.06 shocks.	The OOM and aggregate findings offer similar conclusions. As the bees progressed through the trials, they quickly learned which cap delivered punishment and which cap did not.

This paper is not the first to demonstrate the effectiveness of organism-centered data analyses. OOM has been used in a number of non-human research publications, including social reinforcement delays in honey bees (Craig et al., 2012), a comparative analysis of drone vs. worker honey bees (Dinges et al., 2013), timing in fixed-interval schedules of reinforcement in honey bees and horses (Craig, et al., 2014, 2015), and taste aversion learning to ethanol in honey bees (Varnon, et al., 2018). OOM has also been used in a number of human research studies, including terror management (Grice et al., 2012), the Stroop effect (Grice et al., 2017), memory (Grice et al., 2017), vengefulness in males (Grice et al., 2017), racial bias (Grice et al., 2017), evolutionary theory (Grice et al., 2012), epidemiology (Grice et al., 2020), and rejection in social situations (Grice, 2015).

OOM offers an alternative form of data analysis that readily enables the researcher to determine how many organisms are behaving as expected. In addition, OOM requires the researcher to produce expected patterns that the individual observations should follow. The accompanying analyses then confirm or disconfirm the proposed patterns by producing high (i.e., confirmed) or low (i.e., disconfirmed) PCC indices. Suppose a pattern does not produce a desirable PCC, such as in the case of the first pattern utilized in experiment 2. In that case, the pattern can be modified (when supported by theory or previous research) and tested again to determine if the PCC produced from one pattern is stronger than another (see also Grice, 2011; Grice et al., 2017; 2020).

Specifically, in the field of animal research, OOM provides a unique ability to explore patterns at the level of the organism, which could otherwise be missed at the level of the aggregate as shown in our SPSS analysis. Further, because the PCC is an assumption free effect size, we can use smaller sample sizes without risk of biasing our results (Grice, 2011; Grice et al., 2020), which is particularly advantageous for a field that relies upon primarily small sample sizes (Craig & Abramson, 2018). In summary, these advantages provide animal researchers more flexibility and a greater tool set to explore potentially meaning paradigms like the CPR.

References

Abramson, C.I. (1986). Aversive conditioning in honeybees (*Apis mellifera*). *Journal of Comparative Psychology*, *100*(2), 108-116. 10.1037/0735-7036.100.2.108

Abramson, C.I., Dinges, C.W., & Wells, H. (2016) Operant conditioning in honey bees (*Apis mellifera* L.): The cap pushing response. *PLoS ONE 11*(9): e0162347. https://doi.org/10.1371/journal.pone.0162347

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

Cohen, J. (1988). Chapter 8. The analysis of variance and covariance. *Statistical Power Analysis for the Behavioral Sciences; Routledge Academic: New York, NY, USA*, 273-406.

Couvillon, P. A. & Bitterman, M. E. (1980). Some phenomena of associative learning in honeybees. Journal of Comparative and Physiological Psychology, 94(5) 878-885. doi: https://psycnet.apa.org/doi/10.1037/h0077808

Couvillon, P. A., & Bitterman, M. E. (1982). Compound conditioning in honeybees. *Journal of Comparative and Physiological Psychology, 96*(2), 192–199. htts://doi.org/10.1037/h0077869

Couvillon, P. A., & Bitterman, M. E. (1987). Discrimination of color-odor compounds by honeybees: Tests of a continuity model. *Animal Learning & Behavior, 15*(2), 218–227. https://doi.org/10.3758/BF03204965

Couvillon, P. A., & Bitterman, M. E. (1989). Reciprocal overshadowing in the discrimination of color-odor compounds by honeybees: Further tests of a continuity model. *Animal Learning & Behavior, 17*(2), 213–222. https://doi.org/10.3758/BF03207637

Craig, D. P. A., & Abramson, C. I. (2018). Ordinal pattern analysis in comparative psychology – A flexible alternative to null hypothesis significance testing using an observation oriented modeling paradigm. *International Journal of Comparative Psychology*, 31. Retrieved from https://escholarship.org/uc/item/08w0c08s.

Craig, D. P. A., Varnon, C. A., Sokolowski, M. B. C., Wells, H. & Abramson, C. I. (2014). An assessment of fixed interval timing in free-flying honey bees (*Apis mellifera ligustica*): An analysis of individual performance. *PLoS One*. *9*(7): e101262. https://doi.org/10.1371/journal.pone.0101262

Craig, D. P. A., & Abramson, C. I. (2015). A need for individual data analysis for assessment of temporal control: Invertebrate fixed-interval performance. *International Journal of Comparative Psychology*, *28*, 1-39. http://escholarship.org/uc/item/847557dt.

Dinges, C. W., Avalos, A., Abramson, C. I., Craig, D. P. A., Austin, Z. M., Varnon, C. A., Dal, F. N., Giray, T., & Wells, H. (2013). Aversive conditioning in honey bees (*Apis mellifera anatolica*): A comparison of drones and workers. *Journal of Experimental Biology*, *216*(21), 4124-4134. https://doi.org/10.1242/jeb.090100

Funayama, E. S., Couvillon, P. A., & Bitterman, M. E. (1995). Compound conditioning in honeybees: Blocking tests of the independence assumption. *Animal Learning & Behavior*, *23*(4), 429-437. https://doi.org/10.3758/BF03198942

Giurfa, M., & Sandoz, J. C. (2012). Invertebrate learning and memory: fifty years of olfactory conditioning of the proboscis extension response in honeybees. *Learning & memory*, *19*(2), 54-66. doi:10.1101/lm.024711.111

Grice, J. W. (2011). Observation oriented modeling: Analysis of cause in the behavioral sciences. New York, NY: *Academic Press*.

Grice, J. W. (2015). From means and variances to persons and patterns. *Frontiers in Psychology*, *6*, 1007. https://doi.org/10.3389/fpsyg.2015.01007

Grice, J. W. (2021). Drawing inferences from randomization tests. *Personality and Individual Differences*, *179*, 110931. https://doi.org/10.1016/j.paid.2021.110931

Grice, J.W., Barrett, P., Cota, L., Felix, C., Taylor, Z., Garner, S., Medellin, E., & Vest, A. (2017). Four bad habits of modern psychologists. *Behavioral sciences*, *7*(3), 1-21. https://doi.org/10.3390/bs7030053

Grice, J. W., Barrett, P., Schlimgen, L., & Abramson, C. I. (2012). Toward a brighter future for psychology as an observation oriented science. *Behavioral Sciences*, *2*(1), 1-22, https://doi.org/10.3390/bs2010001

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/10.1177/2515245920922982

Smith, B.H., Abramson, C.I., & Tobin, T.R. (1991). Conditional withholding of proboscis extension in honeybees (*Apis mellifera*) during discriminative punishment. *Journal of Comparative Psychology*. *105*(4), 345-56. https://doi.org/10.1037/0735-7036.105.4.345

Staddon, J. E., & Cerutti, D. T. (2003). Operant conditioning. *Annual review of psychology*, *54*(1), 115-144. https://doi.org/10.1146/annurev.psych.54.101601.145124

VandenBos, G. R. (2007). *APA dictionary of psychology*. American Psychological Association.

Varnon, C., Dinges, C. W., Black, T. E., Wells, H., & Abramson, C. I. (2018). Failure to find ethanol- induced find taste aversion learning in honey bees (*Apis mellifera* L.). *Alcoholism: Clinical and Experimental Research*, *42*, 1260-1270. https://doi.org/10:1111/acer.13761.