Response to Comments on “Ducklings imprint on the relational concept of ‘same or different’”

Antone Martinho III and Alex Kacelnik

Two Comments by Hupé and by Langbein and Puppe address our choice of statistical analysis in assigning preference between sets of stimuli to individual ducklings in our paper. We believe that our analysis remains the most appropriate approach for our data and experimental design.

Both Comments (1, 2) on our Report (3) question the assignment of individuals to a “preference” class. We defined the sign of each subject’s preference according to which of two revolving stimuli elicited more approaches. To control for the effect of exposure to novel stimulus pairs on the effect we were measuring, ducklings were tested in a single, short session. We defined “an approach” operationally, taking into account the large interindividual qualitative variation in behavior. Because preference is not expressed in any natural, countable, unit of behavior, we used an arbitrary but unbiased construct. A duckling that showed maximum bias by faithfully following a single stimulus was assigned one “approach” each time the stimulus completed a quarter of a circle, thus scoring four approaches per revolution. Another that showed a series of brief approaches and halted within the quarter revolution also scored a single count per quarter. The significance of population bias was computed by a sign test on the number of ducklings that scored either way. The null hypothesis (i.e., lack of sensitivity to the relational properties of the stimulus pairs) was consistent with a nonsignificant difference in such a test. We used 4 pairs of shapes and 10 pairs of colors, and our results led to a convincing rejection of the null hypothesis. Pooling shape-same, shape-different, color-same, and color-different conditions, we tested 152 individuals, 39 of which did not move at all, leaving 113 who expressed a preference (we return to this issue below). Seventy-seven of these preferred the training relation, a result for which the expected frequency under the null hypothesis is 8 in 10,000, a reliable outcome by most biological standards. We explored the sensitivity of our preference criterion by requiring a difference of at least 5 or 10 approaches to be counted. As the number of ties increased, significance fell to 0.05 for the five criterion and became nonsignificant for the ten criterion.

There is no information on why some ducklings did not move, but this is not surprising. Following a suggestion from our critics to favor the null hypothesis, the immobile ducks could be assumed to have precisely no preference. (This is unjustified, because imprinting is likely to have occurred, even if we could not observe it; but we can explore it.) If we assign 19 ducks to the training preference and 20 to the other, the number of observations raises to 152, of which 96 preferred the training relation. The probability of this outcome under the null hypothesis is 15 out of 10,000, which, while lowering confidence, is well above currently accepted criteria. Consequently, we are confident that the result claimed was properly supported by the data.

Both Comments suggest that individuals should only have been assigned a preference if they had exceeded some within-subject fixed statistical criterion (such as $P = 0.05$), using the number of approaches as independent observations in within-subjects sign tests, but this is not an assumption of the sign test, is overconservative and arbitrary, and has no statistical justification. Our use of short, single sessions per duckling reflects the biology of the problem. Ducklings that have formed an imprint still approach novel stimuli (4), acting under an evolved behavior by which young fowl in their critical period for imprinting investigate potential substrates for their imprint—i.e., any appropriately sized moving object (5). Ducklings imprint in as little as 15 minutes (6) but remain plastic for some time, called the “sensitive period” (7). Thus, as a duck is tested, its preference changes, diluting the training treatment. To prevent this, we restricted testing to 10 min and tested each animal once.

Rejecting the null hypothesis is not the same as establishing the likelihood of one particular explanation. Further transformations of the stimuli and protocol should be explored to identify what mechanism underlies the observed phenomenon. It is possible, for instance, that even though we used a variety of shapes and colors, the particular set of objects and colors used allows for some nonrelational property to generate our results. Experimental replication rather than statistical wizardry is, as always, the only safe way forward, and we are engaged in that.

Because the full results are available in the original paper, we have referred here only to pooled data. Both Comments examined the results separating different tests. Langbein and Puppe found that bias was weaker in color than in shape tests. This is not surprising: As one partitions the results in a larger number of categories, the strength in each treatment must decline. However, it is possible that different dimensions are different in their effects, and this should be investigated. They suggest that shape is likely to be more relevant than color. This is reasonable, although it is hugely improbable that ducklings do not use color at all, given that they discriminate colors and are visually active. Hupé notes an interesting interaction between subgroup and imprinted concept in experiment 1 that is also worth exploring in future research.

Hupé further delves into the disadvantages of overreliance on P values and favors greater reliance on confidence intervals. We do not agree that this is necessary for our study but find his analysis of the effect of the criterion on P value enlightening. At present, however, the P value is the prevalent metric in our field, and it fully supports our conclusions.

REFERENCES


30 August 2016; accepted 5 December 2016 10.1126/science.aaf6997
Response to Comments on "Ducklings imprint on the relational concept of 'same or different'"

Antone Martinho III and Alex Kacelnik

*Science* 355 (6327), 806.
DOI: 10.1126/science.aai8397