Sciknow Publications Ltd.
Animal Behavior and Cognition

ABC 2014, 1(4)446-451
DOI: 10.12966/abc.11.02.2014

Let's Talk about Citizen Science: What *Doesn't* Work

Allison B. Kaufman[1*]

[1]University of Connecticut

*Corresponding author (Email: allison.kaufman@uconn.edu)

Citation – Kaufman, A. B. (2014). Let's talk about citizen science: What doesn't work. *Animal Behavior and Cognition, 1*(4), 446-451. doi: 10.12966/abc.11.02.2014

Abstract - "Citizen Science," or the idea of using interested laypeople to assist in data collection, has the potential to be a valuable resource to those who study animal cognition. This is largely due to the potential increase in sample size it can provide. However, this technique also has the potential to introduce a significant amount of error to an experiment. As a result, before citizen science can be used as a tool, it is important to determine how it best works; specifically, what methods best motivate people to participate and to provide the most accurate data. This includes sharing situations in which data collection was not successful. Presented here is a failed attempt at collecting data on mirror self-recognition (MSR) in pet parrots, orchestrated via yahoo groups listservs. The goals in presenting this unsuccessful methodology are to encourage discussion, to prevent others from repeating the same ill-fated methodology, and to encourage others to attempt variations on said methods which might be more successful.

Keywords – Citizen science, Listserv, Data collection, Parrot, Mirror self recognition, MSR, Pet

The last year or two has seen the beginnings of the use of "Citizen Science;" or the use of laypeople to collect data for scientific investigation, in animal cognition research. Nelson and Fijn (2013) recently published an exciting overview of a technique in which YouTube videos were used to asses a variety of play behaviors in several species of animals, and Hare et al. (n.d.) have developed the Dognition website, which collects a large variety of data on dog behavior and personality traits.

The excitement, of course, lies in the ability to exponentially increase sample size – traditionally a limiting factor in studies of animal behavior (Agrillo & Miletto Petrazzini, 2012; Button et al., 2013). It is, however, tempered by the introduction of incalculable amounts of variation and potential for human error. The key question then becomes – is the power produced by the sample size large enough to compensate for the inherent error introduced by the methodology?

Much discussion is necessary. In addition to learning to balance sample size and potential variation, methods by which data can be obtained must be vetted. Currently it appears that citizen science methods which involve participants playing games or matching things (e.g., vocalizations) have been particularly successful (Deecke & Janik, 2006; Prestopnik & Crowston, 2012; Veirs, Veirs, & Wood, 2011), as have been population surveys of particular species (Hochachka et al., 2012; Howe, Geissler, & Harrington, 1989; Worthington et al., 2012). This is an excellent start. However, despite the (potentially detrimental) tradition of not reporting null findings (Ferguson & Heene, 2012; Nosek, Spies, & Motyl, 2012), it seems particularly important when developing a new methodology that techniques which *do not* work be discussed as much as those which do.

Here I report a disappointing attempt to collect data via citizen science. It is the first instance that I am aware of in which participants were asked to conduct a specific protocol. The method by which participants were recruited was through a website very commonly used for communication among large groups of people with specific interests, so it is my belief that others would instinctually attempt a similar

methodology. Thus I report my lack of results in an attempt to create discussion and perhaps increase the probability of success in future attempts to collect data.

The Study

The study was an examination of mirror self-recognition (MSR) abilities (or lack thereof) in home-raised parrots. Parrot owners were contacted and asked to complete the task with their bird, submit a video of the task, and complete a brief survey detailing basic information about the bird (species, age, gender if known, experiences as a pet, etc.). The basic task involved allowing the bird five min in front of a mirror after receiving a touch on the cheek (control condition 1), allowing the bird five min away from the mirror after receiving a dot of food coloring on the cheek (control condition 2), and then allowing the bird five min in front of the mirror with the dot of food coloring. Videos were to be examined by an appropriate number of raters for self-directed behavior toward the mark (i.e., touching or scratching); this is currently the most widely accepted evidence of MSR in non-verbal animals (Gallup, 1977).

The immediate implication of such a task being performed by laypeople in their own homes is the inability to control the environmental variables surrounding the test. Noise, location, conditions, and even the movements of the owner were bound to vary widely. It was my hypothesis that if a large enough sample size could be obtained it would serve as sufficient evidence to at least make a case for the ability, despite the lack of a strict experimental setting. This is an unusual outlook and I in no way mean for it to replace current experimental practices. It is, at the moment, only my own theoretical idea born of my knowledge of sample size and experimental design (Rosenthal & Rubin, 1982; 2003). However, I believe that it merits some consideration, if only as a method to collect preliminary data, which can later be confirmed under strict empirical settings (R. Rosenthal, personal communication, May 16, 2014). To illustrate my point, if 500 birds from across the country participated in the test and 100 showed evidence of MSR, it would be illogical to completely disregard the results despite the lack of complete control.

Logistics

A website for the study was created which contained task instructions (Appendix 1), a demonstration video and two links. One link was to the survey, hosted by Qualtrexx (Appendix 2), and the second was to the online file sharing/hosting service YouSendIt, which graciously donated an account for the duration of the study. Participants were instructed to complete the task – which took 15minutes and required minimal equipment – and to upload the video and complete the survey. They were told that for every video uploaded they would be entered into a drawing for one of five book prizes. There was also an optional language supplement included in which interested owners could record and upload their bird's language use during a one week period.

Participant Recruitment

Recruitment was focused on parrot themed "yahoo groups" ("Yahoo Groups," n.d.). 46 such groups were originally identified for recruitment, although the final number varied slightly as some groups had gone inactive and additional ones were found or cross posted to in the process of recruitment. An email was composed and sent to the moderators of these groups (Appendix 3). The email introduced me, the project, discussed some background of and the project's importance, provided a brief explanation of the task, and included a link to the study website (where instructions and the sample video could be found). List moderators were asked to post the email to their lists (to avoid my having to join over 40 mailing lists; however, later in the course of the study, this did turn out to be necessary).

Outcomes

After all attempts were made to recruit participants, two videos were submitted with the matching surveys completed. In addition, nine more surveys were completed, but videos were not submitted to match. Several other attempts to obtain videos – some successful – were undertaken and will be discussed further below.

Primary and Secondary Postings

Initial emails were sent to the moderators of 40 groups (some of the originally identified groups were later identified as defunct), requesting these moderators to post the recruitment information to their groups. From this, nine posts were made by moderators to groups. The two videos stemmed from these posts, as did 20 enthusiastic emails noting an individual's interest or requesting clarification.

After approximately a month and no additional videos, I joined (or requested to join, some groups are by permission only) the groups and posted either the original message (18 groups; where the moderator had not posted the original message) or a reminder (eight groups; one did not approve my request to join). Three groups where the moderator had not posted the original message also denied my membership request or to post the message (it should be noted that some of these were due to time; if there is no response from a moderator within a certain amount of time, the yahoo groups site automatically declines the membership or message). No additional videos responses stemmed from these postings, although there were a few response emails.

Home Visits

After these dismal results, I contacted a local bird club. I initially planned to attend a club meeting and test birds there, but a variety of circumstances made this impossible. Instead, the bird club president discussed the study at the meeting and, as it was a local club, announced that I would be willing to make "house calls" to test birds. A sign-up sheet produced two interested pet owners, in addition to the club president (who had multiple birds). It is probable that my absence was directly related to this small showing, as I was unavailable to answer questions or generate excitement for the study.

Home visits to the three residences recruited from the bird club yielded nine videos, although again, only two were from traditional citizen homes (the remaining seven birds belonged to the club president, who also maintained the birds for an animal education business). Testing at a local parrot store yielded an additional six. The "have test, will travel" methodology seemed promising, and as a result I opted to offer it via the yahoo groups in two geographic areas. This yielded one more home visit in the area where I was living at the time (three videos), an offer of a second local visit which could not be arranged due to other circumstances, and a list of ten potential home visits in a second geographic area (to which I was relocating). Interestingly, this post also yielded six additional, enthusiastic emails from people who did not live in either specified area, inquiring if they could run the test at home. Directions and links were sent to these people, but no videos were submitted.

Post Outcomes

In total, the recruitment email post was seen in 30 yahoo groups, whose members total 28,357 people (although this estimate is slightly inflated as some people join multiple groups). After the secondary posting, I remained on the lists and discussion occurred on five lists, however this was meager (the exchange of two to four emails) did not cause additional participation, which it had the potential to do. The majority of these discussions were anecdotes about birds and mirrors. In only one case did an owner express reluctance to involve his or her bird in the test; this was the only negative response to the request on any of the lists

The Avian Cognition Listserv

At one point I was introduced to a smaller group interested specifically in avian cognitive abilities. This group of approximately 40 people was able to brainstorm a little with regards to the lack of responses. The group suggested that a more in-depth explanation of the methodology of the study be included, in addition to more information to reassure bird owners that no harm would come to their animals. The group also pointed out a flaw in the experimental procedure which made it very hard to complete with flighted birds (i.e., a flighted bird was far less likely to stay on the perch in front of the mirror for an extended period of time). However, although they were all bird owners, and there was interest and discussion within the group, none of the members participated in the experiment. When asked, members either cited the above reasons (primarily the complication of flighted birds) or did not provide a reason at all.

Potential Flaws in Study Procedure

I do not claim that the MSR procedure given to participants was perfect. There are a myriad of potential issues with the methodology which may have gone undetected before requests for participation were made. It is possible that the use of video media is too complex or is not available in the average household. However, with the exception of the issue of flighted birds, I received no feedback on methodological issues from potential participants. Those without designated video cameras generally had smartphones with video capabilities. In addition, the actual methodology of the experimental, while relevant is not necessarily the issue at hand. My purpose in writing this is to address the lack of participation in the experiment (which in turn, would have made for a more rigorous test of methodology).

Conclusions

For an attempt at citizen science, the yahoo groups appear to be an ideal choice at first glance. Parrots are specialized pets due to their long lifespan and the large amount of cognitive stimulation they require, and as a result parrot owners in general tend to be very dedicated (Carpenter, 2013; Hergovich, Mauerer, & Riemer, 2011; Hoppes & Gray, 2010). Those who participate in email discussions on a regular basis would logically be assumed to be even more so. It seems to follow that this population would be invested in advancing the cognitive research on their species of interest.

This makes the almost complete lack of participation particularly puzzling. Could thousands of people really just "not get around to" something? Or is there some unforeseen deterrent? I suspect the type of task is a large factor – despite the fact that the survey was useless without the accompanying video for this particular project, there were still more surveys completed than videos. Despite the popularity of smartphones with video abilities, it is still much easier to remain at one's computer and click through to a survey than to actually set up the experiment. While I absolutely expected to see this lack of follow through, I was surprised by just how extreme the effect was. This may also be the determining factor for video submission from those who sent in the enthusiastic emails or participated in discussion on the avian cognition listserv. Again, while both of these situations indicate interested, thoughtful groups of people who might be ideal participants, writing an email or participating in an online discussion can be done immediately upon reading a post, with no preparation or even the necessity to get up from the computer.

The question remains, however, does the effort of setting up the camera completely preclude any type of experimental request beyond what can be completed at the computer? Does there exist a (reasonable) motivation to encourage active data collection? Ornithologists often successfully rely on laypeople for bird counts and species surveys – is this success simply a result of the natural activities of the typical birdwatcher (Cooper, Loyd, Murante, Savoca, & Dickinson, 2012)? Put differently, does citizen science in this situation work because birdwatchers are already out in the field watching birds, and

are then asked to report what they see, as opposed to being asked to go with the express purpose of data collection?

The potential audience in social media also needs to be examined. In the case of sites like Facebook or Youtube, users tend to be younger and tech savvy. This seems to be less so on discussion lists like Yahoo. There are multiple recruiting via social media might impact results – the ability/willingness of people to use technological resources such as a video camera; the time they have available (retirees vs students, for example), or their interest in research, to name just a few. Socio-economic status should also be considered; owning a pet bird, computer, and a video camera or smart phone are most likely linked to higher SES.

It is my hope that by reporting my (lack of) results here will help encourage dialog on how we might use citizen science in study of animal cognition. I suspect it can be a very powerful tool, and the work by Nelson and Fijn (2013) and Hare et al. (n.d.) provide excellent support for this idea. But, if we are to have citizen science as a tool in our toolbox, we must refine it to the point where it can be used as a reliable methodology -as opposed to the subject of the experiment in and of itself. I do not have the solutions to the problems encountered in this study, but I believe a common discussion and the reporting of both positive and negative findings will be the most expedient way to obtain them.

Acknowledgements

I would like to thank all those who participated in the study, with a particular thanks to Kim and Pete at The Perfect Parrot and Lauren at Reptacular Animals (both of Los Angeles), who allowed me to test multiple birds at their facilities. I would also like to thank YouSendIt for donating an account for the duration of the study. Additional thanks to the members of the Avian Cognition yahoo group, CB Buckley, Alan Kaufman and James Kaufman.

References

Agrillo, C., & Miletto Petrazzini, M. E. (2012). The importance of replication in comparative psychology: The lesson of elephant quantity judgments. *Frontiers in Psychology*, *3*, 181. doi:10.3389/fpsyg.2012.00181

Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews. Neuroscience*, *14*, 365–76. doi:10.1038/nrn3475

Carpenter, J. W. (2013). Keeping parrots: Understanding their care and breeding. *Journal of Avian Medicine and Surgery*, *27*, 65–65. doi:10.1647/1082-6742-27.1.65

Cooper, C. B., Loyd, K. A. T., Murante, T., Savoca, M., & Dickinson, J. (2012). Natural history traits associated with detecting mortality within residential bird communities: Can citizen science provide insights? *Environmental Management*, *50*, 11–20. doi:10.1007/s00267-012-9866-x

Deecke, V. B., & Janik, V. M. (2006). Automated categorization of bioacoustic signals: Avoiding perceptual pitfalls. *Journal of the Acoustical Society of America*, *119*, 645–653.

Ferguson, C. J., & Heene, M. (2012). A vast graveyard of undead theories: Publication bias and psychological science's aversion to the null. *Perspectives on Psychological Science*, *7*, 555–561. doi: 10.1177/1745691612459059

Gallup, G. G. (1977). Absence of self-recognition in a monkey (*Macaca fascicularis*) following prolonged exposure to a mirror. *Developmental Psychobiology*, *10*, 281–284.

Hare, B. A., Kaminski, J., Miklósi, Á., Santos, L. R., Call, J., & Wrangham, R. W. (n.d.). Dognition. *Retrieved from* https://www.dognition.com/

Hergovich, A., Mauerer, I., & Riemer, V. (2011). Exotic animal companions and the personality of their owners. *Anthrozoos: A multidisciplinary journal of the interactions of people & Animals*, *24*, 11. doi: <ahref="http://dx.doi.org/10.2752/175303711X13045914865349">http://dx.doi.org/10.2752/175303711X13045914865349

Hochachka, W. M., Fink, D., Hutchinson, R. A., Sheldon, D., Wong, W.-K., & Kelling, S. (2012). Data-intensive science applied to broad-scale citizen science. *Trends in Ecology & Evolution*, *27*, 130–137. doi: 10.1016/j.tree.2011.11.006

Hoppes, S., & Gray, P. (2010). Parrot rescue organizations and sanctuaries: A growing presence in 2010. *Journal of Exotic Pet Medicine*, *19*, 133–139. doi: 10.1053/j.jepm.2010.05.003

Howe, M. A., Geissler, P. H., & Harrington, B. A. (1989). Population trends of North American shorebirds based on the international shorebird survey. *Biological Conservation*, *49*, 185–199. doi: 10.1016/0006-3207(89)90035-9

Nelson, X. J., & Fijn, N. (2013). The use of visual media as a tool for investigating animal behaviour. *Animal Behaviour*, *85*, 525–536. doi:10.1016/j.anbehav.2012.12.009

Nosek, B. A., Spies, J. R., & Motyl, M. (2012). Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. *Perspectives on Psychological Science*, *7*, 615–631. doi: 10.1177/1745691612459058

Prestopnik, N., & Crowston, K. (2012). Purposeful gaming & socio-computational systems. In *Proceedings of the 17th ACM International Conference on Supporting Group Work - GROUP '12* (p. 75). New York: ACM Press. doi: 10.1145/2389176.2389188 .

Rosenthal, R., & Rubin, D. B. (1982). A simple, general purpose display of magnitude of experimental effect. *Journal of Educational Psychology*, *74*, 166–169.

Rosenthal, R., & Rubin, D. B. (2003). r-equivalent: A simple effect size indicator. *Psychological Methods*, *8*, 492–496.

Veirs, S. R., Veirs, V. R., & Wood, J. D. (2011). Killer whale scouting: Listen live for troops J, K, and L. *The Journal of the Acoustical Society of America*, *129*, 2538. doi: 10.1121/1.3588433

Worthington, J. P., Silvertown, J., Cook, L., Cameron, R., Dodd, M., Greenwood, R. M., … Skelton, P. (2012). Evolution megalab: A case study in citizen science methods. *Methods in Ecology and Evolution*, *3*, 303–309. doi: 10.1111/j.2041-210X.2011.00164.x

Yahoo Groups. (n.d.). *Retrieved from* www.groups.yahoo.com