

Popping Up Balloons for Science: a Research Proposal

JÉRÉMY BARBAY*, DANIEL FREIRE-FERNÁNDEZ, and DANKO LOBOS-BUSTAMANTE,

Departamento de Ciencias de la Computación (DCC), Universidad de Chile, Chile



Fig. 1. Two Monk Parakeet parrots playing a commercial Pop the Balloons application.

Some video games were developed to entertain non human animals while measuring their abilities, logged in a file which can be analyzed later. Using such games to measure the limits of such abilities is problematic, as it requires the subjects to be exposed to instances that they cannot solve, potentially frustrating them. Could presenting the subjects with a mix of instances of various difficulties at the same time, and measuring their interaction with those, yield useful information about their abilities and inabilities, without frustrating the subjects? We propose to design, develop and validate a web game presenting several instances at once, inspired by existing ones such as 'Pop the Balloons', so that the subject can be exposed to a mix of "easy" and "difficult" instances in parallel, and to validate that this does not produce frustration by studying the correlation of the subjects' assiduity in playing the game with the rate of "difficult" instances.

CCS Concepts: • **Applied computing** → **Computer-assisted instruction**; *Interactive learning environments*; *Agriculture*; **Computer games**; • **Human-centered computing** → *User interface design*.

Additional Key Words and Phrases: Animal Computer Interaction, Comparative Psychology, Digital Life Enrichment.

ACM Reference Format:

Jérémy Barbay, Daniel Freire-Fernández, and Danko Lobos-Bustamante. 2022. Popping Up Balloons for Science: a Research Proposal. In *Ninth International Conference on Animal-Computer Interaction (ACI'22)*, December 5–8, 2022, Newcastle-upon-Tyne, United Kingdom. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3565995.3566029>

*Contact author.

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the owner/author(s).

© 2022 Copyright held by the owner/author(s).

Manuscript submitted to ACM

1 INTRODUCTION

Some human guardians of non human animals (referred to as “Other Animals Than Humans” or OATHs from now on) use the Android application **Pop the Balloons** [1], an application where the users is requested to touch balloons passing on the screen, to entertain them. Such an application was designed for humans: its visual interface was designed for human eyes, and the rhythm of the animation and game was designed for human brains.

Different species have developed different characteristics among them, where we can find both sensory and cognitive ones, so it is interesting to wonder how different species could interact with digital applications and if such interaction could provide information about their abilities. In this context, when entertaining AOTHs with applications such as **Pop the Balloons** [1], it is not clear which parts of the application can be seen by the subjects, and if the application is perceived as (too) slow or (too) fast by OATH subjects. Moreover, colors are displayed slightly differently from one device to another, and some devices might run faster than others, in ways which might be barely noticeable by human subjects but important for OATH subjects. Could a loggable version of such an application, used by subjects of various species, allow to draw a preliminary map of the interoperability of OATH visual and time sensing abilities and digital devices visual rendering abilities and velocities?

We propose to design, implement and validate 1) a simple loggable version of the **Pop the Balloons** [1], partially navigable by OATH subjects, with protected areas accessible only by humans for configuration purposes; 2) a central back-end for various instances of such games to report their usage logs, noting the colors of the objects interacted with and the speed at which they were interacted with; and 3) a website for researchers to visualize and analyze such usage logs. We propose to validate the usability of our solution with four types of users: A) some OATHs, to validate the usability of a simple version of the game by OATH subjects; B) some non-scientists guardian of OATHs, to validate the usability of the configuration part of the game by human subjects; C) some researchers in biology and animal sensory abilities, to validate the usability of the website presenting the usage logs; and D) some fellow programmers, to validate the feasibility to program similar applications contributing similar information to the same back-end and website. The validations will occur in the order they were presented, with the first ones occurring in the very early stages of the work.

After describing more formally the formal background of our work in term of sensory diversity, the use of life enrichment applications in comparative psychology and citizen science projects in Section 2, we describe the system that we propose to develop (Section 3), the experiments we propose to realize in order to validate its usability and assess its impact (Section 4), some preliminary results (Section 5) and the final results hoped for such experiments (Section 6).

2 BACKGROUND

Each species have developed a selection of senses across generations, which current results differ from one to another (e.g. dogs' vision is dichromatic, humans' is trichromatic and monk parakeet parrots' is cuadri-chromatic.): we summarize the current knowledge about such difference in term of vision abilities in Section 2.1. As we propose to design and use digital life enrichment applications to measure such abilities, we review in Section 2.2 the history of using life enrichment applications in general and digital life enrichment

applications in particular in the field of comparative Psychology. As we propose to scale up such study via a citizen science project, we review the principles and some examples of such projects in Section 2.3.

2.1 Sensory Differences Between Species

The colors displayed by digital displays and the sound frequencies played by devices are optimized for the majority of humans. It is not always clear how much and which colours and sound can be seen and heard by individual of each species, an issue absent from purely physical experimental set-ups such as that of Al Aïn et al. [2]. The solution proposed presents extensive parameters to vary the colours displayed and the sounds played to the subject. Even less intuitively, species can differ in their Critical Flicker Fusion Frequency (CFFF) [10], the frequency at which they perceive the world and can react to it (in some species, such frequency even vary depending on the time of the day or of the season [7, 12]). For instance, dogs have higher CFFF while cats have lower ones, and the CFFF of reptiles vary with the ambient temperature. Such variation might affect not only their ability to comprehend the visual display and sound play from devices, but might also affect how they comprehend some application designs over others. The solution proposed presents extensive parameters to vary the time between each exercise and which game, so that part of the rhythm of the application can be adjusted by the experimenter to the CFFF of the subject, but more research is required in order to automatically adapt the rhythm of such applications to the CFFF of individuals from a variety of species.

2.2 Life Enrichment Applications

The study of the abilities of OATHs and the use of life enrichment activities in general, and digital ones in particular, have been interconnected from their very beginning. In 1990, when Richardson et al. [13] describe a Computerized Test System to measure some abilities in a population of rhesus monkeys, they mention that “*the animals readily started to work even when the reward was a small pellet of chow very similar in composition to the chow just removed from the cage*”, and that “*the tasks have some motivating or rewarding of their own*”.

Furthermore, OATH subjects seem to enjoy participating in cognitive studies involving game-like digital applications. Washburn [15] describes, among various other anecdotes, how game-like application for apes were developed as early as 1984, and how the subjects “*chose to work on joystick-based tasks, even though they did not need to perform the game-like tests in order to receive food*”, and “*opted for computer task activity over other potential activities that were available to them*”. Lastly, he describes evidence that the subjects were not only motivated by food rewards, but also by the enjoyment of the tasks themselves: when given a choice between completing trials for pellets or receiving pellets for free but not being able to play the game-like tasks during the free-pellet period, the subjects chose to work for their reward, a behavior described ever since 1963 [8] and christened as “*Contrafereeloading*” [11] in 1977.

2.3 Citizen Science

The term “*Citizen Science*” refers to scientific projects conducted, in whole or in part, by amateur (or nonprofessional) scientists [6]. It is sometimes described as “public participation in scientific research”, with the dual objectives to improve the scientific community’s capacity, as well as improving the public’s understanding of science and conscience about the research’s themes [3]. Citizen Science has become a means

of encouraging curiosity and greater understanding of science whilst providing an unprecedented engagement between professional scientists and the general public.

Such methodology must be used with care, in particular about the validity of volunteer generated data. Projects using complex research methods or requiring a lot of repetitive work may not be suitable for volunteers, and the lack of proper training in research and monitoring protocols in participants might introduce bias into the data [14]. Nevertheless, in many cases the low cost per observation can compensate for the lack of accuracy of the resulting data [5], especially if using proper data processing methods [9].

In the following sections, we describe a potential citizen science project using digital life enrichment applications to draw an preliminary map of the interoperability of the visual senses of OATH species with the visual display abilities of various digital devices.

3 SYSTEM PROPOSED

We propose to design, implement and validate a system allowing

- guardians to configure a graphical application where objects (e.g. balloons) slowly drift on the screen, such that the application emits different sounds (e.g. "correct" and "incorrect" according to the category) and make them disappear when such objects are selected (e.g. popping), so that the guardian can reward the selection of a category over another one, and to measure the ability of the subjects to distinguish between the categories;
- subjects to exercise their agency by starting and stopping such application at any time, and selecting various options concerning the background and setting of the application; and
- researchers to define categories of objects for the guardians to select, and to collect a log of the users' interactions with the application in order to measure the overall discriminating ability and the evolution over time of such discriminating ability of the subjects.

The front-end of such a system should be navigable by the subjects in order to maintain their agency, with the access to special settings (e.g. such as the configuration of the categories and frequencies with which they appear) reserved to their guardians by tests such as an extra long press or a query usually considered unlikely to be performed by the subject (e.g. reading and typing a sequence of characters). The first version of the back-end of such a system should allow at the strict minimum to download all of the logs of the interactions of all the subjects into a large file. Given time, a separate front-end will be implemented, to allow researchers to select and download logs of the interactions of subjects according to their unique ID or species or guardian, and according to various parameters of the logs such as the dates and visualisation parameters.

We describe in the next section a subset of the experiments that such a system would allow to perform, along with the training and experimental protocol that we plan to follow.

4 EXPERIMENTS PROPOSED

After the initial phase of development of both software and devices, and extensive testing with human subjects, we plan to apply to the local Institutional Animal Care and Use Committee (IACUC) for the authorisation to run the following experiments, in order to validate the most basic hypothesis behind the design of the software developed, and to guide the ongoing design of more advanced versions. Each iterated

phase of validation by OATHs will be preceded by an extensive validation phase by human subjects, in order to avoid any frustration of the OATHs subject, both for ethical reasons and for the success of the experiment, as such frustration would anyway taint the results of the experiments

Given the set-up described in the previous section, an experiment consists in regularly presenting a given subject with the application (e.g. three daily sessions) for voluntary interactions, with rewarding schemes of increasing intricacy:

- (1) presenting the subject with various objects (e.g. balloons) slowly drifting across the screen, exploding with a soft noise when selected, each and everyone of them prompting a reward; and
- (2) presenting the subject with two categories of objects (e.g. balloons with alphabet letters and balloons without any drawing), similarly slowly drifting across the screen and exploding when selected, but such that only the selection of one category of objects (e.g. balloons with a letter of the alphabet) prompts for a reward.
- (3) repeating the previous phase with more than two categories, using criteria both from concepts (e.g. recognizing alphabet letters) and from senses (e.g. distinguishing between distinct colors)

5 PRELIMINARY EXPERIMENTS

A preliminary experiment with two monk parakeets playing the (non loggable) application **Pop the Balloons** [1] on a touch interface, well within the norms of life entertainment applications and not requiring the validation of any IACUC, showed the sequence of tasks described in the previous section to be adequate. The subjects were presented with the application **Pop the Balloons** [1] on a touch interface, and rewarded first whenever they touched a balloon, and then only when they touched a special one, such as one with an alphabet letter or a special object. It was informally observed that both subjects quickly learned to give priority to such special balloons, without having formal logs to justify such observation more formally.

Another preliminary experiment validated how varying the colors and contrasts of the visual displays affected the performance of a subject at simple discriminating tasks. The experimental protocol described by Barbay et al. [4] was replicated in several separate experiments using respectively high and low contrast (and various colors) in the visual display: see Figures 2 to 5 for some pictures of the subject with the displays in various levels of contrast and colors. As expected, the accuracy of the subject is much lower with low contrast (see Figure 6 for some statistics). Not surprisingly, the subject requested to change game much earlier in sessions with low contrast than in sessions with normal contrast: the subject was less motivated to participate in the study in a display mode with contrast values which impaired their perception. The subject was free to switch to another activity at any activity and did not suffer in the experiment. But repeating such experiments in order to gather enough data to draw a map of what a subject can or cannot see, and can or cannot count, would definitely cross the boundaries of what is ethically acceptable, while we hope that an application inspired by **Pop the Balloons** [1], offering a mix of challenges in parallel, would allow to perform such testing without frustrating the subject by mixing both challenges that the subject is assumed to be able to solve with challenges of less known difficulty to the subject.

We describe in the next section the final results which we are hoping to achieve with such an application, either directly through the experiments performed, or at a higher level by demonstrating the possibilities of our experimental protocol to be applied to prove and disprove other hypothesis.

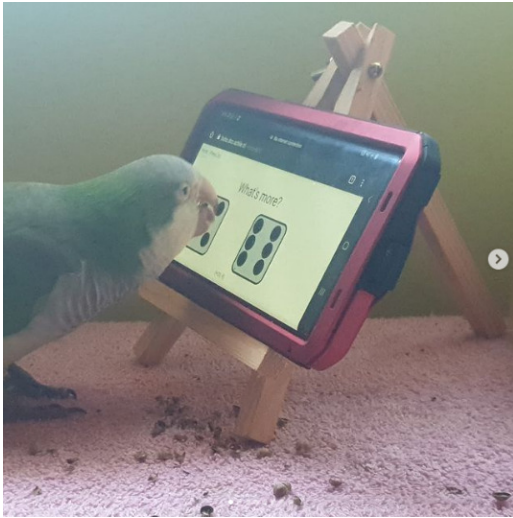


Fig. 2. A Monk Parakeet parrot selecting the largest of two quantities in a high contrast display mode.

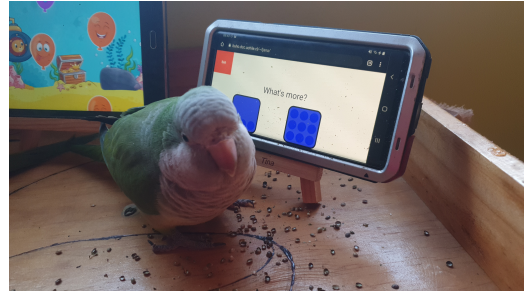


Fig. 3. A Monk Parakeet parrot selecting the largest of two quantities in a low contrast display mode, with dark blue on lighter blue.

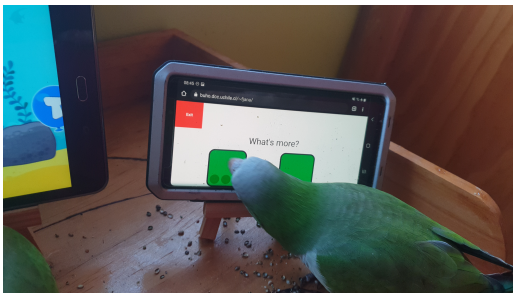


Fig. 4. A Monk Parakeet parrot selecting the largest of two quantities in a low contrast display mode, with dark green on lighter green.



Fig. 5. A Monk Parakeet parrot using the application in a low contrast display mode, with dark red on lighter red.

6 RESULTS HOPED FOR

At a first basic level, we hope that such experiments will allow to measure and compare the ability of individuals from various species, among those able to use touch interface, to distinguish between some colors, in function of the display technology used to display such colors: as such devices' parameters were tailored to human senses, it is likely that they will vary much more in their adequacy among OATHs than among humans. The same set-up will allow to measure and compare the ability of individuals from various species to distinguish between categories of images (e.g pictures of individuals from the same gender, species, or family of species versus from other ones).

At a more advanced level, we hope that such system and experiments will confirm the possibility to measure not only the abilities of subject but also their inabilities, in a way which does not frustrate them (an issue both from the point of view of ethics and of the accuracy of experimental measurements).

Value Set (x, y)	Accuracy	
	High Contrast (895 data points)	Low Contrast (115 data points)
{1,2}	80	50
{1,3}	100	50
{1,4}	100	50
{1,5}	96	100
{2,3}	66	100
{2,4}	88	0
{2,5}	91	25
{3,4}	71	33
{3,5}	100	33
{4,5}	61	ND

Fig. 6. Comparison of the performances of a Monk Parakeet subject on simple binary discrimination tasks, with high and low contrast display during a single testing session.

7 CONCLUSION

We have described our preliminary work concerning the development of a solution to test the abilities, and more particularly the inabilities, of subjects, human or not, while avoiding the ethical issue of most basic protocols to measure inabilities, frustrating the subject. This work is still its early stage, and we welcome any kind of criticism or feedback which could allow us to be more successful in this project.

ACKNOWLEDGMENTS

We wish to thank Joachim Barbay for his suggestion to use `Svelte` and `Svelte Kit` to develop the early prototypes, and his support in learning such programming environments; Jennifer Cunha for indirectly (via her instagram account) show the ability of some parrots to play the game `Pop the Balloons` [1], which inspired our solution to the problem of testing inabilities in ethical ways; and the referees for their numerous suggestions.

REFERENCES

- [1] 2022. Pop the balloons. <https://play.google.com/store/apps/details?id=ge.bebi.balloon.pop.game>. [Retrieved 2022-09-26].
- [2] Syrina Al Aïn, Nicolas Giret, Marion Grand, Michel Kreutzer, and Dalila Bovet. 2008. The discrimination of discrete and continuous amounts in African grey parrots (*Psittacus erithacus*). *Animal cognition* 12 (09 2008), 145–54. <https://doi.org/10.1007/s10071-008-0178-8>
- [3] Citizen Science Association. 2021. CitizenScience.org. Website <https://citizenscience.org/>. Last accessed on [2022-05-27 Fri].
- [4] Jérémy Barbay, Fabián Jaña-Ubal, and Cristóbal Sepulveda-Álvarez. 2022. Measuring Discrimination Abilities of Monk Parakeets Between Discreet and Continuous Quantities Through a Digital Life Enrichment Application. <https://doi.org/10.48550/ARXIV.2206.00734>
- [5] Mary M Gardiner, Leslie L Allee, Peter MJ Brown, John E Losey, Helen E Roy, and Rebecca Rice Smyth. 2012. Lessons from lady beetles: accuracy of monitoring data from US and UK citizen-science programs. *Frontiers in Ecology and the Environment* 10, 9 (2012), 471–476. <https://doi.org/10.1890/110185> arXiv:<https://esajournals.onlinelibrary.wiley.com/doi/pdf/10.1890/110185>
- [6] Trisha Gura. 2013. Citizen science: Amateur experts. *Nature volume* 496 (2013), 259–261.
- [7] Kevin Healy, Luke McNally, Graeme D. Ruxton, Natalie Cooper, and Andrew L. Jackson. 2013. Metabolic rate and body size are linked with perception of temporal information. *Animal Behaviour* 86, 4 (2013), 685–696. DOI:

- 10.1016/j.anbehav.2013.06.018.
- [8] Glen D. Jensen. 1963. Preference for bar pressing over "freeloading" as a function of number of rewarded presses. *Journal of Experimental Psychology* 65, 5 (1963), 451–454. <https://doi.org/10.1037/h0049174>
 - [9] Eva C. McClure, Michael Sievers, Christopher J. Brown, Christina A. Buelow, Ellen M. Ditria, Matthew A. Hayes, Ryan M. Pearson, Vivitskaia J.D. Tulloch, Richard K.F. Unsworth, and Rod M. Connolly. 2020. Artificial Intelligence Meets Citizen Science to Supercharge Ecological Monitoring. *Patterns* 1, 7 (09 Oct 2020). <https://doi.org/10.1016/j.patter.2020.100109>
 - [10] Mankowska ND, Marcinkowska AB, Waskow M, Sharma RI, Kot J, and Winklewski P.J. 2021. Critical Flicker Fusion Frequency: A Narrative Review. *Medicina* 57, 10 (Oct 2021), 1096.
 - [11] Steve R. Osborne. 1977. The free food (contrafreeloading) phenomenon: A review and analysis. *Animal Learning & Behavior* 5, 3 (Sept. 1977), 221–235. <https://doi.org/10.3758/bf03209232>
 - [12] E. Reas. 2014. Small Animals Live in a Slow-Motion World. *Scientific American Mind* 25, 4 (2014).
 - [13] W. K. Richardson, D. A. Washburn, W. D. Hopkins, E. S. Savage-Rumbaugh, and D. M. Rumbaugh. 1990. The NASA/LRC computerized test system. *Behavior Research Methods, Instruments, and Computers* 22 (1990), 127–131. <https://doi.org/10.3758/BF03203132>.
 - [14] Brett Thelen and Rachel Thiet. 2008. Cultivating connection: Incorporating meaningful citizen science into Cape Cod National Seashore's estuarine research and monitoring programs. *Park Science* 25 (06 2008).
 - [15] David Washburn. 2015. The Four Cs of Psychological Wellbeing: Lessons from Three Decades of Computer-based Environmental Enrichment. *Animal Behavior and Cognition* 2 (08 2015), 218–232. <https://doi.org/10.12966/abc.08.02.2015>