

# Photographic Identification of Individual Domestic Cats: Comparing the Accuracy of Life Science University Students Versus Cat Advocate Citizen Science Volunteers



CITIZEN SCIENCE:  
THEORY AND PRACTICE

RESEARCH PAPER

SABRINA AELURO 

JON VAN OAST

*\*Author affiliations can be found in the back matter of this article*

]u[ubiquity press

## ABSTRACT

There have been no long-term field studies of the potential effect of spay/neuter programs on free-roaming domestic cat population sizes. To address that gap via citizen science, we are developing a novel approach to photographic mark-recapture population research that engages volunteers as both smartphone-wielding data collectors and as online data processors in building capture histories from submitted photos. Here, we present a validation study testing the accuracy of cat advocate volunteers at matching smartphone photos of cats, and we compare their success to a reference group of life science university students. We also examine feline photographic identification from two additional perspectives: what makes a volunteer better at cat identification, and what makes a cat photo more identifiable? 151 cat advocates and 17 students completed 37,800 pairwise photo comparisons using our online platform. Cat advocates' matching attempts ( $n = 34,080$ ) were correct 98.1% of the time compared with students' 97.5% ( $n = 3,720$ ). Volunteers who reported a pet cat increased their accuracy. Volunteers who held less than a bachelor's degree, or those who volunteered with cats previously, had reduced accuracy. If a cat was a color other than black, its ability to be identified increased. We demonstrated that our citizen science volunteer sample was not only adequate at identifying individual cats in smartphone photos, but performed better than our sample of life science students—a labor pool commonly trusted to organize data from camera trap research. While photographs are the data foundation of many studies of free-roaming cats, we are the first to analyze by-eye visual identification in this species.

CORRESPONDING AUTHOR:

**Sabrina Aeluro**

Kitizen Science, US

[sabrinaaeluro@gmail.com](mailto:sabrinaaeluro@gmail.com)

KEYWORDS:

photo identification; citizen science; validation study; free-roaming cats; urban ecology; mark-recapture

TO CITE THIS ARTICLE:

Aeluro, S and Van Oast, J. 2022. Photographic Identification of Individual Domestic Cats: Comparing the Accuracy of Life Science University Students Versus Cat Advocate Citizen Science Volunteers. *Citizen Science: Theory and Practice*, 7(1): 7, pp. 1–14. DOI: <https://doi.org/10.5334/cstp.465>

## INTRODUCTION AND BACKGROUND

One of the most inflammatory topics in animal welfare and wildlife conservation is free-roaming domestic cat (*Felis catus*) management (Loss and Marra 2018; Lynn et al. 2019). Free-roaming domestic cats, as we use the term here, cover a behavioral spectrum from fractious feral animals to sociable outdoor pets. There is a need to control their populations to protect threatened and endangered wildlife species (Loss and Marra 2018), safeguard public health from disease and injury (Kravetz and Federman 2002), and curtail nuisance behaviors such as noisy mating activities (Levy, Isaza, and Scott 2014), urine spraying, and fighting with pets (Gramza et al. 2016).

The health and well being of free-roaming cats themselves is an important issue to cat welfare proponents, as well as to the public, which generally supports non-lethal management of cats (Wald, Jacobson, and Levy 2013). Cat advocates, focused on reducing feline deaths, champion the use of spay/neuter (often called trap-neuter-return or TNR, and community cat programs) with the aim of shrinking free-roaming cat populations over time as sterilized adults do not replace themselves with new kittens (Boone et al. 2019). While evidence exists to support the arguments that spay/neuter efforts are linked to reductions in animal shelter intake and euthanasia (Spehar and Wolf 2019), no long-term field research has demonstrated that these programs are affecting the metric of concern to conservation stakeholders: the sheer number of free-roaming cats on the landscape. Two controlled studies of spay/neuter interventions in North America ran for only a year apiece and concluded that the treatment areas did not experience a population decline (Bissonnette et al. 2018; Kilgour et al. 2016). This research is dismissed by cat advocates as having failed to collect data for a sufficient amount of time.

There is a need for long-term monitoring of the potential effects of spay/neuter programs on free-roaming cat population sizes. To be acceptable to all stakeholders, this research must combine science-based population estimation methods and multi-year time spans. Such an initiative could answer questions such as whether spay/neuter programs can reduce population sizes of free-roaming cats, what proportion of cats in an area need to be sterilized to see a population reduction, and how long it takes to see that decline. While computer simulations have resulting in widely varying answers, these ideas have never been field tested. Further, these answers are likely to vary based on location-specific context, so there is a need to replicate the same style of cat population monitoring in multiple locations so program efficacy can

be compared based on factors such as climate and level of urbanization.

The gold standard for quantifying animal populations is photographic mark-recapture modeling, which estimates a population size based on how often animals in an area are detected and photographed, often combined with spatial data. An animal is “captured” and “marked” when it is first photographed and identified, and “recaptured” when photographed again, creating sighting histories for analysis. While there are statistical methods being developed to perform this type of modeling with only partial identification of individuals (Augustine et al. 2018), the ability to accurately identify each individual is preferable.

Photo identification using an animal’s natural marks has been tested across many terrestrial and marine species (e.g., Gibbon, Bindemann, and Roberts 2015; Gilkinson et al. 2007; Gomez-Salazar, Trujillo, and Whitehead 2011; Zheng et al. 2016). These studies sometimes compare the accuracy of identifiers from two groups: those who are considered experts or professionals, and those presumed to be less experienced, such as citizen science volunteers (Chesser 2012). However, despite an abundance of free-roaming cat population research that relies on photos as data (e.g. Cove et al. 2017), and individual photo identification studies in other felids (Johansson et al. 2020; Park et al. 2019), no whole-body photo identification work has been published about domestic cats.

Researchers from the world of machine learning and computer vision have had success in testing algorithms for identifying domestic cats by their noses (Chen et al. 2016) and faces (Lin and Kuo 2018; Klein 2019), but these approaches rely on high-resolution, well-lit photos of selected areas of the body in cooperative subjects. Such images would not be obtainable by field biologists collecting free-roaming cat photos via camera traps, smartphones, or possibly even DSLR cameras with telephoto lenses. Artificial intelligence tools are not yet advanced to the point where they can be used for identifying individual domestic cats in real-world research applications.

Field data collection for photographic mark-recapture research commonly uses motion-activated camera traps placed over a study area to obtain photos for building capture histories. However, these methods are not without drawbacks and limitations. Maintaining a camera trap array consumes professional labor, cameras and security boxes require purchasing and maintenance, and time is spent sifting through massive photo collections where many images feature swaying plants or the movement of humans and other non-target species. Camera trap placement permissions, theft, and vandalism are also of concern, especially if used in urban and populated areas.

We propose an alternative to traditional camera trap arrays and data management: engaging volunteers as both smartphone-wielding human camera traps and as data processors in identifying cats and building capture histories for analysis. While one free-roaming cat research program solicited members of the public to host camera traps on their property as part of the DC Cat Count (Herrera et al. 2021) and another deploys volunteers in road-based cat counts for the Hayden Island Project in Portland, OR (Cats Safe at Home, no date) no one has yet attempted a holistic, multi-role citizen science approach for studying free-roaming cat populations. To that end, a nonprofit citizen science program, Kitizen Science, is being developed to monitor free-roaming cat population sizes in urbanized areas in collaboration with spay/neuter organizations.

A flourishing body of literature appeals to practitioners to ensure that tapping into crowdsourced free labor does not cause an unacceptable decrease in the rigor of the scientific process (Kosmala et al. 2016, Parrish et al. 2018). While there is always benefit to ensuring that research methods are subject to transparency, validation, and piloting measures, this is especially true in areas with high conflict between stakeholder groups and where results could affect policy decisions.

In response to these needs, the first validation study of Kitizen Science's methods tested our intended approach's most basic assumption: whether citizen science volunteers drawn from a pool of cat advocates can identify individual free-roaming cats in smartphone photos at a reasonable level of accuracy, thereby making them capable of building credible capture histories for mark-recapture analysis. One metric for defining a reasonable level of accuracy is to compare volunteers in a citizen science program's target demographic against those in the demographic that is already trusted to perform this type of task. Here, we tested the success of cat advocate citizen science volunteers in identifying and matching cats taken with smartphone photos, and compared them with a reference group of biology and environmental science students—a labor pool commonly trusted to organize data from camera trap research. Further, our study sought to probe the question of cat identification from two perspectives: what makes a volunteer better at cat identification, and what makes a cat photo more identifiable?

## **METHODS**

### **SOFTWARE AND TECHNOLOGY**

Kitizen Science is based on the Wildbook platform: open source, cloud-based citizen science software developed by the conservation nonprofit Wild Me to support photographic mark-recapture wildlife field studies. Our configuration

features numerous additional customizations for use in free-roaming cats in populated environments. A key difference between ours and other Wildbook projects is that we engage citizen science volunteers as both photographers and animal identifiers, whereas other Wildbook projects use artificial intelligence for the latter.

### **PHOTOGRAPH COLLECTION**

We obtained photos of cats in June and July 2019 in various locations in Washington and Oregon during the late afternoon and evening when free-roaming cats are known to be active and visible. We shot most photos of cats for this study in the same manner as we would direct volunteers to do for our program, where a person walked along a street or sidewalk and photographed any observed cats. We also included some photos from a rural property where the resident gave us verbal permission to photograph their cats. We included these to supplement the urban/suburban backgrounds of the rest of the photo collection. We used an iPhone XS (Apple, Inc., California, USA) smartphone's default camera without applying the digital zoom tool or any filters. We also photographed some cats with a Nikon D5300 digital SLR camera with a 70–300 mm AF-S Nikkor zoom lens (Nikon Corporation, Tokyo, Japan) to ascertain physical features less visible with a smartphone. All photos were gathered by a single researcher (SA), allowing us to be confident in the true identities of all cat photos included in the collection, creating a validation study where we can compare volunteer responses to a known true state.

We did not crop or retouch photos to enhance brightness, contrast, lighting, or saturation to make cats more or less identifiable, but we did make alterations to some photos to blur house numbers, street signs, license plates, commercial signage and logos, human faces, cat collar tags with visible text, and other potentially identifying information to protect trademarks and the privacy of residents in areas where we collected photos.

### **VOLUNTEERS**

We recruited two types of people to form our test and reference groups. We solicited our test group of cat advocate citizen science volunteers online via cat welfare Facebook Pages and Groups, Twitter, personal contact networks, and an email list of feral cat advocates. These volunteers could be located anywhere so long as they were over 18 and had an internet-connected device. We chose these advertising methods to be similar to how we would recruit future participants for Kitizen Science. We solicited our reference group of student research volunteers at the University of Washington via flyers in environmental and life science buildings on campus, department email lists, and through making announcements in undergraduate

wildlife courses. We chose these advertising methods and single-institution scope to mimic how a traditional university-based researcher would seek assistance in processing camera trap images from a photographic mark-recapture study.

During our website signup process, volunteers agreed to an informed consent statement to participate as a research subject, signed up for a user account in agreement with a set of software terms and conditions, completed a demographics and personal background survey, and an explanation about the study and a set of instructions. Only after completing and clicking through these pages could a volunteer proceed into photo-matching tasks. We created separate enrollment pages for cat advocates and students, with slightly different survey questions, but all other aspects of the matching study were the same for both groups. We used the English language only in recruitment material and on our website. Volunteers could ask questions about study and instructions via email. Volunteers participated at their own pace within the 9-week period spanning summer and autumn of 2019 that was allotted for each group.

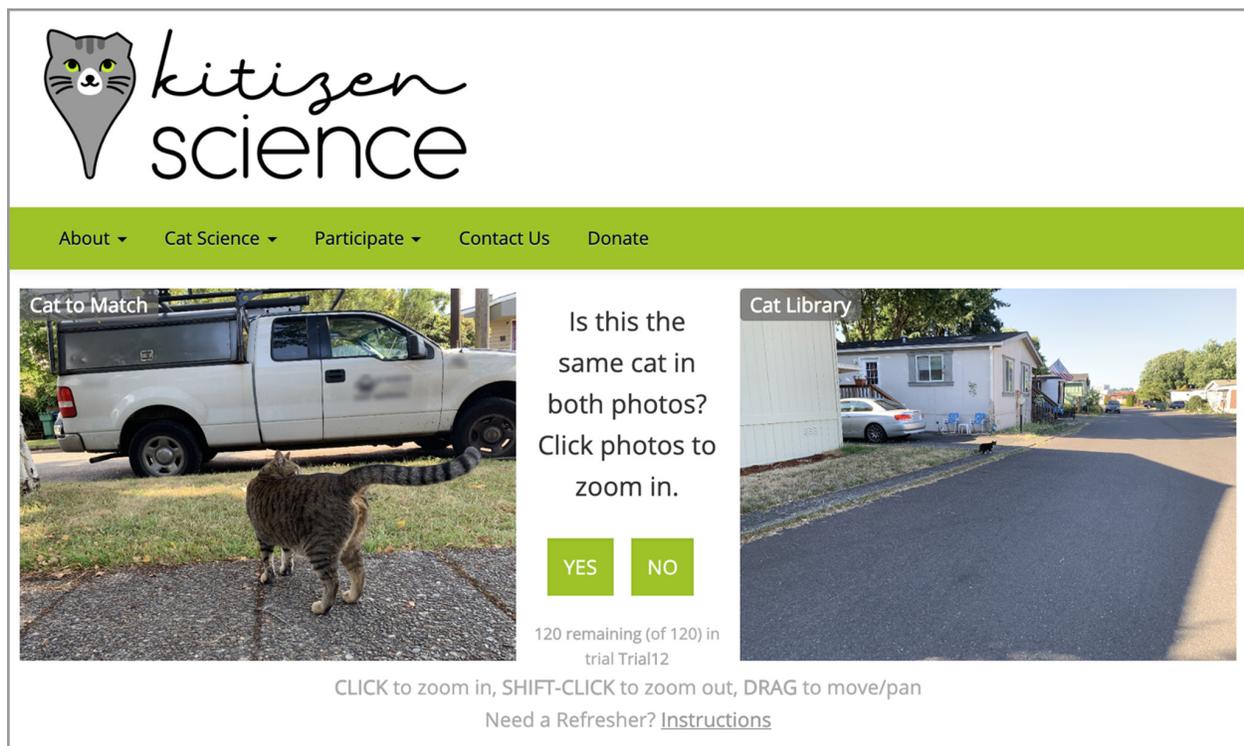
## STUDY WEBSITE

We organized photo-matching tasks into trials of 120 pairwise photo comparisons in which our website presented a volunteer with 1 of 50 randomized photos

of a different cat to match and a randomized photo of a potential match, and the volunteer was asked a simple yes/no question of whether the same cat appeared in both photos (*Figure 1*). Each trial contained 0 to 5 correct matches for each cat to match within the library of 120 images. This arrangement was chosen to be a simplified version of how our citizen science website workflow will function: volunteers will be presented with a cat to match and then look through potential matches in search of the same individual. Volunteers were required to make a binary “yes” or “no” decision on each pair of photos (Chesser 2012; Gibbon, Bindemann, and Roberts 2015), as previous research found that photo classification participants offered an uncertain option overused it (Swanson et al. 2016). We restricted volunteers to completing a maximum of 2 trials (240 photo pair comparisons) per 24-hour period to avoid observer fatigue (Chesser 2012). For analysis, we excluded incomplete matching trials (Gould, Clulow, and Clulow 2021) on the assumption a user was merely curious about the project and not putting in genuine effort (Van Horn et al. 2014).

## COVARIATES

To investigate which traits make people better at identifying cats or make a cat photo more identifiable, we summarized matching results at the level of volunteers or cat photos along with covariates of interest for each. With the human



**Figure 1** A screenshot of our study website that presented volunteers with two cat photos for comparison. Users could zoom/pan to explore detail.

models, we examined only the volunteers from our target demographic of cat advocates.

Human model covariates included the random effect of the user, and the fixed effects of gender, whether they care for feral cats, whether they have pet cats, whether they have ever volunteered with cats, their level of education collapsed to whether they hold a bachelor's degree or higher, if they have previously participated in other image-classification citizen science projects, and the mean amount of time they spent completing a trial (with outlier pauses of 120 seconds or more removed on the assumption users had stepped away from the task). We believe that the small bias created by these time deletions, which resulted in occasionally undercounting total completion time by likely seconds, was more acceptable than the bias of overcounting total completion times by up to thousands of minutes.

Cat photo model covariates included the random effect of each photo, and fixed effects of whether the cat's color/pattern was black, how much of the cat's face was visible, the proportion of the frame occupied by the cat (a proxy for distance from camera and therefore cat resolution), whether the cat had a differentiator in the form of a visible collar or visible removed ear tip (an indicator of a sterilized feral cat), and the mean amount of time users spent viewing the photo (with outlier pauses of 120 seconds or more removed). Face visibility was grouped into three categories: none (face obscured), partial (one quarter to three quarters of the face visible), or full (face toward camera with markings on both sides visible). In calculating the proportion of the photo's frame occupied by a cat, the cat's longest aspect was measured in pixels using the ruler tool in Adobe Photoshop 2020 (Adobe, Inc., California, USA) and divided by the horizontal width of the photo (4032 pixels) to create a proportion of the frame occupied by the cat. We excluded cats' tails from inclusion in width calculations as they are not always fully extended or visible. Measurements maximized cat length to account for body positions captured at different angles. We measured walking cats from a cranial point such as snout or front paws to a caudal point such as hind feet or rump, or in seated cats from rump to ears. For cats obscured by objects, our measurement only includes the visible portion of the cat. To account for minor variations in measurement, each cat was measured twice and we used the mean number of pixels.

After data collection was completed, we discovered that one cat photo in our collection had a black cat hiding in the shadows under bushes in the background. This cat was difficult to see and we decided to retain this photo's matching results data in our analysis on the assumption

that the second cat was very unlikely, although not impossible, to have been spotted by volunteers and confused for the cat in the center of the photo frame to be compared.

## ANALYSIS

We extracted user response data from our website such that all pairwise comparison responses were converted to a binary of correct/incorrect and whether they were true positives (correct match), true negatives (correct non-match), false positives (volunteers selected that two cat photos were a match when they were not a match), or false negatives (volunteers selected that two cat photos were not a match when they were a match). We compared the results of our cat advocates and student volunteers and summarized volunteer survey demographics for each group.

We used Generalized Linear Mixed Models (GLMMs) to analyze mixed effects data, along with Generalized Linear Models (GLMs) for those without random effects. We modeled our response variable under the binomial distribution where the number of pairwise photo comparisons were treated as trials and the number of correct matches as successes. Our GLMMs included 1) models with random effect plus one predictor variable, 2) a full model with all potential predictors, and 3) an intercept-only GLM. Based on Akaike Information Criterion corrected for small sample sizes (AICc) rankings of these models and which variables were emerging as common to best-performing models, additional combinations were created and compared. For our human accuracy question, we fit ten additional models that included the following combinations of covariates: user, whether they have pet cats, and gender; user, whether they have pet cats, and if they have participated in other image classification citizen science projects; user, whether they have pet cats, and the mean amount of time spent completing a trial; user, whether they have pet cats, and whether they care for feral cats; user, whether they have pet cats, and whether they hold a bachelor's degree or higher; user, whether they have pet cats, and whether they have ever volunteered with cats; user, whether they have pet cats, whether they have ever volunteered with cats, and whether they care for feral cats; user, whether they have pet cats, whether they have ever volunteered with cats, and whether they hold a bachelor's degree or higher; user, whether they have pet cats, whether they have ever volunteered with cats, whether they care for feral cats, and whether they hold a bachelor's degree or higher; user, whether they have pet cats, whether they hold a bachelor's degree or higher, and whether they care for feral cats. For our cat photograph

identifiability question, we fit four additional models that included the following combinations of covariates: photo and the proportion of the frame occupied by the cat; photo and whether a cat had a visible collar or removed eartip; photo and the mean time users viewed the photo; and photo, proportion of the frame occupied by the cat, and the mean time users viewed the photo. Upon determining the best model of each question, a GLM was built without random effects to determine their importance. Lastly, we re-ranked our models by AICc to determine the new best model for each question.

We performed analyses using R version 3.6.2 (R Core Team 2019) in RStudio version 1.2.5033 (RStudio Team 2015) with the packages lme4 (Bates et al. 2015), MuMIn (Barto 2018), and plyr (Wickham 2011).

## RESULTS

In total, 151 cat advocates and 17 life science university students participated in this study by completing a total of 37,800 pairwise photo comparisons using our online platform (See Supplemental Files 1–5 for data analyzed).

Our volunteer background and demographics survey revealed similarities and differences between groups. The mean age for cat advocates was 47.1 (range 19–76) and 23.9 (range 18–50) for students. In both volunteer groups, the majority were women (cat advocates 90.7%, students 82.3%) and selected white as their race/ethnicity (cat advocates 90.1%, students 70.6%). Many in both groups had a pet cat or cats (cat advocates 90.1%, students 58.8%), with 35.1% of cat advocates and no students reporting that they cared for feral/free-roaming cats. Only 7.9% of cat advocates had previously participated in an online citizen science project doing image identification or classification, compared with 23.5% of students. A majority (62.2%) of cat advocates held a bachelor's degree or higher. Regarding experience volunteering with cats, 57.6% of cat advocates were currently involved, 17.9% previously volunteered, and 24.5% reported having never volunteered with cats. Among students, 17.6% had previously volunteered to do image identification or classification as part of research that was not online citizen science, such as viewing camera trap images for University of Washington wildlife researchers. This student group was comprised largely of undergraduates (82.3%), with a few master's and doctoral students (5.9% and 11.8%, respectively). See volunteer demographics summarized in [Table 1](#).

Cat advocates' matching attempts ( $n = 34,080$ ) were correct 98.1% of the time compared with students' 97.5% ( $n = 3,720$ ). Among cat advocates, there were 33,089 true negatives, 329 true positives, 143 false negatives, and 519

false positives. The students had 3,594 true negatives, 34 true positives, 10 false negatives, and 82 false positives. While both groups performed highly, students made 27.3% more errors than cat advocates. This difference between groups was statistically significant using a chi-squared test ( $\chi^2 = 4.831$ ,  $df = 1$ ,  $p\text{-value} = 0.0280$ ) using an alpha of 0.05. In considering the two different error types, which add confusion to animal capture histories and impact population estimates, cat advocates were 3.6 times more likely to make false positive than false negative errors, compared with the students' 8.2-fold increase.

Our best model (by AICc) to determine what influences a human's ability to match cat photos accurately was the GLMM, which contained user as a random effect and whether one had a pet cat, whether one had volunteered with cats previously, and whether one held a bachelor's degree or higher as fixed effects ([Table 2](#)). Reporting a pet cat increased a volunteer's cat photo identification accuracy, and holding less than a bachelor's degree or having ever volunteered with cats reduced accuracy. Model diagnostics supported the model as fitting the data. In performing an ANOVA test between our best model and an intercept-only GLM to determine the importance of including user as a random effect, our best model was significant ( $p < 0.05$ ). A Chi-squared goodness-of-fit test indicated no lack of fit ( $p > 0.05$ ). The model was underdispersed due to a high accuracy rate by most volunteers.

Our best model (by AICc) to determine what makes a cat photo more identifiable was the GLMM which contained photo as a random effect and whether the cat was black as a fixed effect ([Table 3](#)). If a cat was a color other than black, its ability to be correctly identified increased. Model diagnostics supported the model as fitting the data. In performing an ANOVA test between our best model and an intercept-only GLM to determine the importance of including photo as a random effect, our best model was significant ( $p < 0.05$ ). A Chi-squared goodness-of-fit test indicated no lack of fit ( $p > 0.05$ ). As with above, this model was underdispersed due to a high matching accuracy rate by most volunteers.

## DISCUSSION

Here, we demonstrate that our sample of cat advocate citizen science volunteers were not only adequate at identifying and matching individual cats in smartphone photos, but performed better at the task than our sample of life science university students. These findings support our desire to engage cat advocates in matching cats and building capture histories for use in photographic mark-recapture population estimation. Further, by casting a wide

	CAT ADVOCATES	STUDENTS
Number	151	17
Are you currently involved in volunteering with cats in some way?		
Yes	57.6%	
No	24.5%	
Not now, but in the past	17.9%	
Do you have a disability or personal limitation (such as being a parent/caregiver) that prevents you from volunteering with cats in a typical offline setting like a shelter?		
Yes	8.6%	
No	78.2%	
Sometimes	13.2%	
Do you currently have a cat/cats in your care?		
Yes, a pet cat/cats	58.3%	58.8%
Yes, I care for feral/free-roaming cats	3.3%	0.0%
Yes, a pet cat/cats AND Yes, I care for feral/free-roaming cats	31.8%	0.0%
No	6.6%	41.2%
Have you ever participated in an online citizen science project doing image identification or classification?		
Yes	7.9%	23.5%
No	92.1%	76.5%
Have you ever volunteered to do image identification or classification as part of research that is NOT online citizen science, such as viewing camera trap images for UW wildlife researchers?		
Yes		17.6%
No		82.4%
Age		
Mean (range)	47.1 (19-76)	23.9 (18-50)
Retired		
Yes	21.2%	0.0%
No	78.8%	100.0%
Gender		
Man	8.0%	11.8%
Woman	90.7%	82.3%
Nonbinary/Other	1.3%	5.9%
Race/ethnicity		
White	90.1%	70.6%
All other options (including mixed race selections that included white)	9.9%	29.4%
Highest level of education		
Less than bachelor's degree	33.8%	
Bachelor's degree or higher	66.2%	
What is your current standing in school?		
Undergraduate		82.3%
Master's Student		5.9%
Doctoral Student		11.8%

**Table 1** A summary of survey responses from volunteers.

MODELS	DF	AICC	WEIGHT
<b>GLMMS</b>			
user + pet_cats + volunteer_ever + degree	5	691.5	0.289
user + pet_cats + volunteer_ever + degree + feral_cats	6	692.6	0.170
user + pet_cats + degree	4	693.2	0.127
user + pet_cats + degree + feral_cats	5	693.3	0.119
user + pet_cats + volunteer_ever	4	694.4	0.071
user + pet_cats	3	695.1	0.050
user + pet_cats + feral_cats	4	695.1	0.048
user + pet_cats + volunteer_ever + feral_cats	5	695.3	0.045
user + pet_cats + time_viewing	4	695.3	0.044
user + pet_cats + citsci	4	697.0	0.018
user + pet_cats + gender	5	698.7	0.008
user + pet_cats + volunteer_ever + degree + feral_cats + gender + citsci + time_viewing	10	699.9	0.004
user + degree	3	700.7	0.003
user + time_viewing	3	702.1	0.001
user + feral_cats	3	702.4	0.001
user + volunteer_ever	3	703.0	0.001
user + citsci	3	704.1	0.001
user + gender	4	705.9	0.000
<b>GLMs</b>			
intercept-only null model	1	880.4	0.000

**Table 2** Models predicting which personal traits affect cat advocacy citizen science volunteer’s accuracy in matching cat photos, ranked by AICc values. AICc values are used in model comparison and selection. The lowest score represents the most plausible model of those considered. The weight values are the relative likelihood of a model.

MODELS	DF	AICC	WEIGHT
<b>GLMMS</b>			
photo + black	3	337.6	0.366
photo + black + prop_frame	4	338.2	0.263
photo + black + time_viewed	4	339.5	0.139
photo + black + prop_frame + time_viewed	5	339.9	0.113
photo + black + tip_or_collar	4	339.9	0.113
photo + black + face + prop_frame + tip_or_collar + time_viewed	8	345.8	0.006
photo + tip_or_collar	3	351.7	0.000
photo + prop_frame	3	351.9	0.000
photo + time_viewed	3	352.2	0.000
photo + face	4	352.9	0.000
<b>GLMs</b>			
intercept-only null model	1	558.0	0.000

**Table 3** Models predicting which cat photo traits were linked to accurate matching by cat advocacy citizen science volunteers, ranked by AICc values. AICc values are used in model comparison and selection. The lowest score represents the most plausible model of those considered. The weight values are the relative likelihood.

net through online citizen science, we attracted nearly nine times as many participants as responded to a well-advertised student volunteer position at a large university.

Our results reveal that in our program's target demographic of cat advocates, participants who were best at identifying and matching cats in smartphone photos are those who reported having a pet cat, holding a bachelor's degree or higher, and having not previously volunteered with cats. The finding that a user having volunteered with cats reduced their level of photo identification accuracy was unexpected. One potential explanation for this seemingly paradoxical result may be that participants with cat volunteering experience (one metric of expertise in cats) were over-confident in their abilities and less careful in determining matches. However, a different metric of expertise, having a pet cat, increased users' cat identification accuracy. From an applied perspective, this suggests that volunteer recruitment from an audience of cat advocates with pet cats is important, but recruiting those who already have experience volunteering with cats is unnecessary or even counter-productive. This allows us to solicit future volunteers from a broader audience of people who like cats while not incentivizing us to poach volunteers from cat welfare organizations.

Defining (and self-defining) expertise in animal photo identification is a complex phenomenon that has been explored by previous researchers. In a study which asked attendants at a herpetofauna conference to sort images of newts, Austen et al. (2018) noted that their volunteer with the highest accuracy score self-rated their newt identification skills as worse than their peers, but the volunteer with the lowest accuracy score self-rated as possessing the same level of expertise as their peers. Overall, however, no statistically significant difference in accuracy was detected in comparing groups of volunteers based on how they self-assessed their abilities.

Other research has also asked the question of which human traits lead to better success in photo identification. Similarly to our finding that possessing a bachelor's degree improved cat identification accuracy, Delaney et al. (2008) discovered that education was an important predictor of accuracy in field-based species-level identification among participants in a marine invasive species program, where volunteers with two years of university education exceeded 95% accuracy. Testing the effect of an online training program on species-level identification, researchers found that volunteers with a background in biology had a higher percentage correct than those without, although the latter's accuracy rose to a similar level with the addition of training (Katrak-Adefowora, Blickley, and Zellmer 2020). In a validation study of the potential of using citizen science volunteers for identifying individual Andean bears, Van

Horn et al. (2014) found no personal characteristics that had a meaningful effect on participant accuracy.

Both identification error types introduce bias into capture histories for photographic mark-recapture population estimation, with a false positive trend resulting in an underestimation of the true population size (Urian et al. 2015) through an increase in the capture probability (Johansson et al. 2020). In our study, cat advocates were 3.6 times more likely, and students 8.2 times more likely, to make false positive than false negative errors in cat identification. Looking at identification errors in camera trap images of giant pandas, Zheng et al. (2016) discovered a mean bias of 1.58 toward false positives. Conversely, using camera trap photos of snow leopards, researchers found that capture events were 3 times more likely to be false negatives than false positives (Johansson et al. 2020). Testing images of salamanders, Chesser (2012) found that observers were 185 times more likely to make false negative identification errors than false positives. An Andean bear photo identification study found a bias toward false negatives (Van Horn et al. 2014). Studying the feasibility of using natural markings to identify giant anteaters, 87.5% of errors were false negatives (Möcklinghoff, Schuchmann, and Marques 2018). In seeking to identify polar bears by whisker spots, researchers found that both experienced and inexperienced participants were more likely to make false negative than false positive errors (Prop, Staverløkk, and Moe 2020). These varied findings underscore the need for species-specific validation of photo identification that includes reporting not just general accuracy but also error-type trends. Chesser (2012) noted that false positives, as was the trend in our cat study, could be viewed as the preferable problem because they are less time-consuming than false negatives to identify and correct with intervention.

Unsurprisingly, many previous studies of photo identification in wildlife qualitatively report that "distinctive" animals in "high quality" photos are most readily identified. However, the effect of some metric of visual distinctiveness or photo quality are rarely explored at an analytical level (Theimer, Ray, and Bergman 2017). In our study, we found that whether a cat was black was the key covariate in determining its identifiability, outweighing other factors such as face visibility or the presence of a collar or ear tip. However, this result should not be construed as demonstrating that black cats as a whole are unidentifiable and that all photos containing black cats must be removed from photographic mark-recapture photo data collections. We would argue that thresholds for whether to include photos in building capture histories should not be made on coat color alone, but take into consideration that some black cats and some black cat photos will be sufficiently

distinctive for identification, just as some nonblack cats will fall below a threshold for inclusion. (Figures 2 and 3 show our four least and four most identifiable cat photos.)

Species-level identification accuracy in citizen science has been well explored in the literature. In analyzing Snapshot Serengeti, researchers tested untrained citizen



**Figure 2** The four least identifiable cat photos in our study, all of which had solid black fur.



**Figure 3** The four most identifiable cat photos in our study, none of which had solid black fur. We share these sets of images to illustrate that by-eye identifiability among domestic cats does not require a cat to be close to the camera, uniform in its pose, or even easy to initially notice within the frame.

science volunteers in species identifications and compared their success with expert opinion (Swanson et al. 2016). There was a 97.9% overall agreement between these two groups of classifiers, although accuracy was lower in less-common species. Similarly, Clare et al. (2019) reported 93.4% accuracy in crowdsourced species identifications, also with higher error rates in less-common species. In comparing the skills of volunteers self-reporting no biology background, some biology background, and professional biologists, both with and without training, Ktrak-Adefowora, Blickley, and Zellmer (2020) found that there was no statistically significant difference among five of the six groups, and only those without a biology background and without training displayed a low proportion of correct answers.

Apart from supporting our program's intended approach and establishing that individual domestic cats are generally highly identifiable from smartphone photos, our results add nuance to the way that we delineate experts versus non-experts in animal identification. People with professional experience or studying for life science degrees may not necessarily be experts at the task of identifying individuals of a given species, and enthusiastic non-professionals may possess higher levels of competency in some instances. Researchers should explore whether a hobby or subculture exists around a species or study system of interest and consider working with individuals from such groups not just to save costs and engage the public in research, but to potentially improve the quality of their data. Regardless of the entity responsible for animal identification—citizen science volunteers, student assistants, professional researchers, or artificial intelligence algorithms—validation work and pre-testing of accuracy rates and error trends should accompany all research that relies on photo identification.

## CONCLUSIONS

Discussions about the validity of citizen science approaches are often framed in terms of untrained members of the public versus professional scientists. There is a notable lack of acknowledgment that data collection and interpretation in traditional, university-based research may be performed by another type of nonprofessional: students. Student research assistants, with varying and unknown levels of skill, training, and dedication, complete a sizable amount of grunt work in some fields, which can include sifting through photos in camera trap-based population ecology. We argue that validation studies of citizen science approaches should

compare the performance of their participants not necessarily against that of professional scientists, but with those who otherwise might be tasked with a specific responsibility.

Our study sought to demonstrate that a group of people from our program's target demographic could identify individual cats at a similar level of success as students from a major research university. Our citizen science volunteers outperformed the students both in terms of quality and quantity of work. We do not argue that this would be the case in every discipline or situation, but we cannot make data quality comparisons without university-based researchers subjecting their processes to the same level of skepticism and scrutiny.

## DATA ACCESSIBILITY STATEMENTS

Data analyzed during this study are included in this published article and its supplemental data files.

## SUPPLEMENTARY FILES

The supplementary files for this article can be found as follows:

- **Supplemental File 1.** aeluro\_photos.csv. Cat photo matching results. DOI: <https://doi.org/10.5334/cstp.465.s1>
- **Supplemental File 2.** aeluro\_demographicscat.csv. Cat advocate demographics. DOI: <https://doi.org/10.5334/cstp.465.s2>
- **Supplemental File 3.** aeluro\_demographicstudents.csv. Student demographics. DOI: <https://doi.org/10.5334/cstp.465.s3>
- **Supplemental File 4.** aeluro\_matchingcat.csv. Matching results from cat advocates. DOI: <https://doi.org/10.5334/cstp.465.s4>
- **Supplemental File 5.** aeluro\_matchingstudent.csv. Matching results from students. DOI: <https://doi.org/10.5334/cstp.465.s5>

## ETHICS AND CONSENT

Photographing outdoor cats in their normal environment, without offering attractants or harassment, did not require an animal use protocol at the University of Washington, where this study was developed. Our human volunteers' participation was approved by the University of Washington's Institutional Review Board as STUDY00008004.

## ACKNOWLEDGEMENTS

Thank you to our many wonderful volunteers, without whom this study could not exist. Thank you to Mark Scheuerell, who provided help with the statistical modeling. Thank you to University of Washington professors and staff who allowed us to solicit student volunteers via class announcements, email lists, and notice boards. Thank you to cat advocates who helped us find outdoor cats for our photo collection. Lastly, thank you to the cats themselves for unknowingly contributing to this study.

## FUNDING INFORMATION

This study was funded by the corresponding author during its initial execution. We were provided with complimentary Azure Cloud Computing resources by the Microsoft for Nonprofits program. Our publication fee and manuscript development were supported by a grant from the Innovation Showdown, a collaborative project from Petco Love and The Jackson Galaxy Project that aims to invest in lifesaving animal welfare innovations. The 2020 Innovation Showdown was supported by the Annenberg Foundation & Wallis Annenberg PetSpace, ASPCA, Best Friends Animal Society, Dave & Cheryl Duffield Foundation, Maddie's Fund, Mary Jo & Hank Greenberg Animal Welfare Foundation, Michelson Found Animals Foundation, Petco Love, PetSmart Charities, and The Jackson Galaxy Project.

## COMPETING INTERESTS

Sabrina Aeluro is the Executive Director of Kitizen Science. Jon Van Oast is the Senior Engineer at Wild Me, makers of Wildbook, and the Technical Lead at Kitizen Science.

## AUTHOR CONTRIBUTIONS

Sabrina Aeluro designed the study, created the photo collections, recruited and managed volunteers, performed data analysis, and wrote the manuscript.

Jon Van Oast helped design the study, developed Wildbook software customizations, performed data formatting and extraction, and contributed to the manuscript.

## AUTHOR AFFILIATIONS

**Sabrina Aeluro**  [orcid.org/0000-0003-1101-6588](https://orcid.org/0000-0003-1101-6588)  
Kitizen Science, US

**Jon Van Oast**

Wild Me, US

## REFERENCES

- Augustine, BC, Royle, JA, Kelly, MJ, Satter, CB, Alonso, RS, Boydston, EE and Crooks, KR.** 2018. Spatial capture–recapture with partial identity: An application to camera traps. *The Annals of Applied Statistics*, 12(1): 67–95. DOI: <https://doi.org/10.1214/17-AOAS1091>
- Austen, GE, Bindemann, M, Griffiths, RA and Roberts, DL.** 2018. Species identification by conservation practitioners using online images: Accuracy and agreement between experts. *PeerJ*, 2018(1): 1–19. DOI: <https://doi.org/10.7717/peerj.4157>
- Bartoń, K.** 2018. *MuMIn: Multi-Model Inference*. Available at: <https://cran.r-project.org/package=MuMIn>.
- Bates, D, Mächler, M, Bolker, B and Walker, S.** 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1). DOI: <https://doi.org/10.18637/jss.v067.i01>
- Bissonnette, V, Lussier, B, Doizé, B and Arsenault, J.** 2018. Impact of trap-neuter-return event on the size of free-roaming cat colonies around barns and stables in Quebec: A randomized controlled trial. *The Canadian Journal of Veterinary Research*, 82(3): 192–197.
- Boone, JD, Miller, PS, Briggs, JR, Benka, VAW, Lawler, DF, Slater, M, Levy, JK and Zawistowski, S.** 2019. A long-term lens: Cumulative impacts of free-roaming cat management strategy and intensity on preventable cat mortalities. *Frontiers in Veterinary Science*, 6: 1–9. July. DOI: <https://doi.org/10.3389/fvets.2019.00238>
- Cats Safe at Home.** no date. *Hayden Island Project*, no date. Available at <https://www.catssafeathome.org/hayden-island-project> [Last accessed: 6 August 2021].
- Chen, Y-C, Hidayati, SC, Cheng, WH, Hu, MC and Hua, KL.** 2016. Locality constrained sparse representation for cat recognition. In: *MultiMedia Modeling. MMM 2016. Lecture Notes in Computer Science*. Miami, FL on January 4–6, 2016, 140–151. DOI: [https://doi.org/10.1007/978-3-319-27674-8\\_13](https://doi.org/10.1007/978-3-319-27674-8_13)
- Chesser, ME.** 2012. *An Investigation of Human-Error Rates in Wildlife Photographic Identification; Implications for the Use of Citizen Scientists*. Unpublished thesis (MS), University of Massachusetts Amherst. Available at <http://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1857&context=theses>.
- Clare, JDJ, Townsend, PA, Anhalt-Depies, C, Locke, C, Stenglein, JL, Frett, S, Martin, KJ, Singh, A, Van Deelen, TR and Zuckerberg, B.** 2019. Making inference with messy (citizen science) data: when are data accurate enough and how can

- they be improved? *Ecological Applications*, 29(2): 1–15. DOI: <https://doi.org/10.1002/eap.1849>
- Cove, MV, Gardner, B, Simons, TR, Kays, R and O'Connell, AF.** 2017. Free-ranging domestic cats (*Felis catus*) on public lands: estimating density, activity, and diet in the Florida Keys. *Biological Invasions*, 20(2): 1–12. Springer International Publishing. DOI: <https://doi.org/10.1007/s10530-017-1534-x>
- Delaney, DG, Sperling, CD, Adams, CS and Leung, B.** 2008. Marine invasive species: Validation of citizen science and implications for national monitoring networks. *Biological Invasions*, 10(1): 117–128. DOI: <https://doi.org/10.1007/s10530-007-9114-0>
- Gibbon, GEM, Bindemann, M and Roberts, DL.** 2015. Factors affecting the identification of individual mountain bongo antelope. *PeerJ*, 3. DOI: <https://doi.org/10.7717/peerj.1303>
- Gilkinson, AK, Pearson, HC, Weltz, F and Davis, RW.** 2007. Photo-Identification of Sea Otters Using Nose Scars. *Journal of Wildlife Management*, 71(6): 2045–2051. DOI: <https://doi.org/10.2193/2006-410>
- Gomez-Salazar, C, Trujillo, F and Whitehead, H.** 2011. Photo-identification: A reliable and noninvasive tool for studying pink river dolphins (*Inia geoffrensis*). *Aquatic Mammals*, 37(4): 472–485. DOI: <https://doi.org/10.1578/AM.37.4.2011.472>
- Gould, J, Clulow, J and Clulow, S.** 2021. Using citizen science in the photo-identification of adult individuals of an amphibian based on two facial skin features. *PeerJ*, 9(e11190): 1–16. DOI: <https://doi.org/10.7717/peerj.11190>
- Gramza, A, Teel, T, Vandewoude, S and Crooks, K.** 2016. Understanding public perceptions of risk regarding outdoor pet cats to inform conservation action. *Conservation Biology*, 1–11. DOI: <https://doi.org/10.1111/cobi.12631>
- Herrera, DJ, Moore, SM, Flockhart, DTT, McShea, WJ and Cove, MV.** 2021. Thinking outside the park: Recommendations for camera trapping mammal communities in the urban matrix. *Journal of Urban Ecology*, 7(1): 1–13. DOI: <https://doi.org/10.1093/jue/juaa036>
- Johansson, Ö, Samelius, G, Wikberg, E, Chapron, G, Mishra, C and Low, M.** 2020. Identification errors in camera-trap studies result in systematic population overestimation. *Scientific Reports*, 10(1): 1–10. DOI: <https://doi.org/10.1038/s41598-020-63367-z>
- Katrak-Adefowora, R, Blickley, JL and Zellmer, AJ.** 2020. Just-in-time training improves accuracy of citizen scientist wildlife identifications from camera trap photos. *Citizen Science: Theory and Practice*, 5(1): 8. DOI: <https://doi.org/10.5334/cstp.219>
- Kilgour, RJ, Magle, SB, Slater, M, Christian, A, Weiss, E and DiTullio, M.** 2016. Estimating free-roaming cat populations and the effects of one year Trap-Neuter-Return management effort in a highly urban area. *Urban Ecosystems*, 1–10. DOI: <https://doi.org/10.1007/s11252-016-0583-8>
- Klein, A.** 2019. *Pet Cat Face Verification and Identification*, Stanford University. Available at: [http://cs230.stanford.edu/projects\\_fall\\_2019/reports/26251543.pdf](http://cs230.stanford.edu/projects_fall_2019/reports/26251543.pdf).
- Kosmala, M, Wiggins, A, Swanson, A and Simmons, B.** 2016. Assessing data quality in citizen science. *Frontiers in Ecology and the Environment*, 14(10): 551–560. DOI: <https://doi.org/10.1002/fee.1436>
- Kravetz, JD and Federman, DG.** 2002. Cat-associated zoonoses. *Archives of Internal Medicine*, 162(17): 1945–1952. DOI: <https://doi.org/10.1001/archinte.162.17.1945>
- Levy, JK, Isaza, NM and Scott, KC.** 2014. Effect of high-impact targeted trap-neuter-return and adoption of community cats on cat intake to a shelter. *The Veterinary Journal*, 201(3): 269–274. DOI: <https://doi.org/10.1016/j.tvjl.2014.05.001>
- Lin, T-Y and Kuo, Y-F.** 2018. Cat face recognition using deep learning. In: *ASABE 2018 Annual International Meeting*. Detroit, MI on July 29 to August 1, 2018, 1–5. DOI: <https://doi.org/10.13031/aim.201800316>
- Loss, SR and Marra, PP.** 2018. Merchants of doubt in the free-ranging cat conflict. *Conservation Biology*, 32(2): 265–266. DOI: <https://doi.org/10.1111/cobi.13085>
- Lynn, WS, Santiago-Ávila, F, Lindenmayer, J, Hadidian, J, Wallach, A and King, BJ.** 2019. A moral panic over cats. *Conservation Biology*, 33(4): 769–776. DOI: <https://doi.org/10.1111/cobi.13346>
- Möcklinghoff, L, Schuchmann, KL and Marques, MI.** 2018. New non-invasive photo-identification technique for free-ranging giant anteaters (*Myrmecophaga tridactyla*) facilitates urgently needed field studies. *Journal of Natural History*, 52(37–38): 2397–2411. DOI: <https://doi.org/10.1080/00222933.2018.1537407>
- Park, H, Lim, A, Choi, TY, Baek, SY, Song, EG and Park, YC.** 2019. Where to spot: Individual identification of leopard cats (*Prionailurus bengalensis euptilurus*) in South Korea. *Journal of Ecology and Environment*. *Journal of Ecology and Environment*, 43(1): 1–5. DOI: <https://doi.org/10.1186/s41610-019-0138-z>
- Parrish, JK, Burgess, H, Weltzin, JF, Fortson, L, Wiggins, A and Simmons, B.** 2018. Exposing the science in citizen science: Fitness to purpose and intentional design. *Integrative and Comparative Biology*, 58(1): 150–160. DOI: <https://doi.org/10.1093/icb/icy032>
- Prop, J, Staverløkk, A and Moe, B.** 2020. Identifying individual polar bears at safe distances: A test with captive animals. *PLoS ONE*, 15(2): 1–16. DOI: <https://doi.org/10.1371/journal.pone.0228991>
- R Core Team.** 2019. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. Available at: <https://cran.r-project.org>.

- RStudio Team.** 2015. RStudio: Integrated Development Environment for R. Boston, MA. Available at: <https://rstudio.com/>.
- Spehar, DD and Wolf, PJ.** 2019. Integrated return-to-field and targeted trap-neuter-vaccinate-return programs result in reductions of feline intake and euthanasia at six municipal animal shelters. *Frontiers in Veterinary Science*, 6. March. DOI: <https://doi.org/10.3389/fvets.2019.00077>
- Swanson, A, Kosmala, M, Lintott, C and Packer, C.** 2016. A generalized approach for producing, quantifying, and validating citizen science data from wildlife images. *Conservation Biology*, 30(3): 520–531. DOI: <https://doi.org/10.1111/cobi.12695>
- Theimer, TC, Ray, DT and Bergman, DL.** 2017. Camera angle and photographic identification of individual striped skunks. *Wildlife Society Bulletin*, 41(1): 146–150. DOI: <https://doi.org/10.1002/wsb.743>
- Urian, K, Gorgone, A, Read, A, Balmer, B, Wells, RS, Berggren, P, Durban, J, Eguchi, T, Rayment, W and Hammond, PS.** 2015. Recommendations for photo-identification methods used in capture-recapture models with cetaceans. *Marine Mammal Science*, 31(1): 298–321. DOI: <https://doi.org/10.1111/mms.12141>
- Van Horn, RC, Zug, B, LaCombe, C, Velez-Liendo, X and Paisley, S.** 2014. Human visual identification of individual Andean bears *Tremarctos ornatus*. *Wildlife Biology*, 20(5): 291–299. DOI: <https://doi.org/10.2981/wlb.00023>
- Wald, DM, Jacobson, SK and Levy, JK.** 2013. Outdoor cats: Identifying differences between stakeholder beliefs, perceived impacts, risk and management. *Biological Conservation*, 167: 414–424. DOI: <https://doi.org/10.1016/j.biocon.2013.07.034>
- Wickham, H.** 2011. The split-apply-combine strategy for data analysis. *Journal of Statistical Software*, 40(1): 1–29. Available at: <http://www.jstatsoft.org/v40/i01/>. DOI: <https://doi.org/10.18637/jss.v040.i01>
- Zheng, X, Owen, MA, Nie, Y, Hu, Y, Swaisgood, RR, Yan, L and Wei, F.** 2016. Individual identification of wild giant pandas from camera trap photos – a systematic and hierarchical approach. *Journal of Zoology*, 300(4): 247–256. DOI: <https://doi.org/10.1111/jzo.12377>

---

**TO CITE THIS ARTICLE:**

Aeluro, S and Van Oast, J. 2022. Photographic Identification of Individual Domestic Cats: Comparing the Accuracy of Life Science University Students Versus Cat Advocate Citizen Science Volunteers. *Citizen Science: Theory and Practice*, 7(1): 7, pp. 1–14. DOI: <https://doi.org/10.5334/cstp.465>

Submitted: 19 August 2021

Accepted: 10 January 2022

Published: 24 February 2022

**COPYRIGHT:**

© 2022 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

*Citizen Science: Theory and Practice* is a peer-reviewed open access journal published by Ubiquity Press.

