

METHOD

Just-in-Time Training Improves Accuracy of Citizen Scientist Wildlife Identifications from Camera Trap Photos

Roshni Katrak-Adefowora^{*†}, Jessica L. Blickley^{‡§} and Amanda J. Zellmer^{*†}

Citizen scientists can help professional scientists amass much larger datasets than would be possible without their input, but the quality of these data may impact their utility. Therefore, it is imperative to develop standard practices that maximize the accuracy of data produced by citizen scientists. One method increasingly used to improve data accuracy in citizen science-based projects is just-in-time training (JITT), in which volunteers are given on-demand resources to train them on the spot or in conjunction with the research they are performing. In this article, we examine whether JITT improves citizen scientist accuracy of subject identification, specifically wildlife identification from camera trap photos. Ninety-four participants with varying degrees of experience in biology were asked to identify photos from camera traps in Los Angeles, California set to capture photos of wildlife in an urban habitat. Without access to JITT, citizen scientists with no background in biology had lower accuracy than professional biologists (no background: mean = 51.8%, standard error [SE] = 6.0%; professional biologist: mean = 77.6%, SE = 2.1%). However, when participants with no background in biology received JITT, they were able to identify wildlife with a similar level of accuracy as professional biologists (no background: mean = 81.9%, SE = 3.6%; professional biologist: mean = 85.1%, SE = 2.5%). There was a significant interaction between biology background and training treatment (F-ratio = 7.61, $p = 0.0009$). The increase in accuracy of novice citizen scientists who received JITT was due primarily to fewer misidentifications of species overall but also to increased confidence in classification of species (participants selected the “Don’t Know” option less frequently). From these results, we conclude that the use of JITT can significantly improve subject identification accuracy for citizen scientists with no background in biology.

Keywords: citizen science; wildlife images; trail camera; urban wildlife; human computation; Zooniverse

Introduction

Citizen science is a powerful tool for garnering interest in science from the nonscientific community, as well as for allowing researchers to collect data at greater volumes and at a larger scale than would be feasible with a more limited number of professional scientists (Bhattacharjee 2005; Bonney et al. 2009; Silvertown 2009). Even without any formal scientific background, citizen scientists have contributed to ecological research by successfully identifying millions of camera trap images (Swanson et al. 2016), by quantifying species diversity (Casanovas, Lynch and Fagan 2014), and by contributing to global biodiversity datasets such as eBird (Sullivan et al. 2014) and iNaturalist (White et al. 2015). However, large-scale citizen science projects and incorporation of these datasets into research are not as common as they could be because many

researchers are skeptical of the accuracy of data produced by non-experts (Bonney et al. 2014; Kosmala et al. 2016; Swanson et al. 2016). In fact, there can be considerable variation in accuracy of data among citizen scientists and even among expert scientists (Newman et al. 2010; Gollan et al. 2012; Starr et al. 2014). Identifying the factors that contribute to the consistent collection of highly accurate data by volunteers is necessary to help researchers design better volunteer training and data collection protocols, thus making citizen science more useful for biological and other scientific research.

Of the few studies that have investigated the accuracy of citizen science data, most have focused on accuracy variation as it relates to the amount of training the volunteers received, suggesting that training improves accuracy (Prysbay and Oberhauser 2004; Sauer et al. 2013; Danielsen et al. 2014; Ratnieks et al. 2016; van der Wal et al. 2016). In addition, various studies assess accuracy as it pertains to the difficulty of the task, with greater accuracy associated with easier tasks (versus harder—e.g., identifying familiar vs. rare species) (Prysbay and Oberhauser 2004; Delaney et al. 2007; Crall et al. 2011; Casanovas, Lynch and Fagan 2014; Kelling et al. 2015; Swanson et al. 2016),

* Department of Biology, Occidental College, Los Angeles, CA, US

† Arroyos and Foothills Conservancy, Pasadena, CA, US

‡ Center for Digital Liberal Arts, Occidental College, Los Angeles, CA, US

§ Natural Sciences Division, Pasadena City College, Pasadena, CA, US

Corresponding author: Amanda Zellmer (zellmer@oxy.edu)

increased experience performing a task (Jiguet 2009; Kelling et al. 2015; Swanson et al. 2016), and increased background experience in the related scientific field (Ratnieks et al. 2016). The majority of these studies have focused on longer-term training programs (e.g., 2–3 days of training before beginning a project or a lifetime of birding experience). Unfortunately, longer-term training can be a significant commitment for volunteers, thus deterring citizen scientists from participating in research, and consequently discouraging researchers from attempting to attract volunteers. As a result, some researchers have begun to utilize JITT for their projects (Sullivan et al. 2014; Kosmala et al. 2016; Swanson et al. 2016), training volunteers on the spot—that is, in conjunction with the research they are performing (Jones 2001). This approach provides the necessary resources for participants to use at their discretion. The term JITT originated within industry and manufacturing as a way to provide on-the-job training by making resources available to employees as needed (Jones 2001). This training method can be used for other purposes as well, such as subject identification. Studies that have assessed subject identification accuracies in the absence of any form of training have found the accuracies to be either inconsistent across individuals or low, both for experts and non-experts (Austen et al. 2016; Roy et al. 2016). These studies demonstrate the need for tools and training to assist citizen scientists performing identification tasks. However, it remains unclear whether JITT is a sufficient tool for training citizen scientists.

Although there are multiple applications for JITT for citizen scientists completing subject identification tasks, this training may be particularly useful in the analysis of camera trap images. Camera traps, also known as trail cameras, are motion-sensitive cameras used to take photos of wildlife. These cameras are helpful tools for researchers who are monitoring wildlife in areas where they do not want to interfere with the animals; this is especially useful when dealing with elusive creatures (Harmsen et al. 2017). However, camera traps can produce hundreds or even millions of images (Swanson et al. 2016), making it difficult to process the resulting data. Citizen scientists can assist by identifying organisms in photos through what is referred to as human computation, in which humans carry out tasks that computers are not yet able to perform, thereby allowing researchers to more rapidly process the datasets (von Ahn 2009). Online platforms, such as Zooniverse (www.zooniverse.org), provide a point of access for citizen scientists to find and participate in research. In a recent study, citizen scientists on Zooniverse contributed to identifying more than a million and a half photos of wildlife in Tanzania (Swanson et al. 2016). Equally important, camera trap datasets provide a great opportunity for volunteers to get involved in a citizen science project because they can participate in research anywhere or at any time they have access to the internet. JITTs have been used with camera trap identification projects (e.g., Snapshot Serengeti; Swanson et al. 2016), but the impact of these trainings on the accuracy of data collected has not been explored.

We developed an experiment that compared the impacts of online JITT on the data accuracy of citizen scientists

with varying levels of biology experience using a baseline of groups that received no online JITT. Participants were asked to identify wildlife photos from camera traps set on an urban college campus. We hypothesized that if JITT improves accuracy, then citizen scientists with limited to no background in biology who receive training will be able to correctly identify wildlife as accurately as participants with a more extensive background in biology. Alternatively, if training did not improve accuracy, then we expected volunteers with a background in biology to maintain a significantly higher accuracy than volunteers without a biology background, even when those volunteers received training. Further, we explored the different ways in which accuracy was impacted, comparing the frequency with which participants selected the wrong species, did not spot the organism in the photo, or could not decide which species was present. Finally, we assessed the differences in accuracy across the different species observed in our study site. Here, by quantifying accuracy in identifications of wildlife images from camera traps, we investigate the impact that JITT has on the quality of data collected by citizen scientists.

Methods

To test our hypothesis, we collected photos of wildlife from camera traps set up on the Occidental College campus in Los Angeles, California. Using the Zooniverse platform, participants viewed and identified the species appearing in each photo. We grouped participants based on their biology experience and whether they received training, and then assessed the accuracy of their identifications.

Camera trap photos

Reconyx HC500 camera traps (Reconyx, Holmen WI) were used to capture the wildlife photos. The cameras were set to high-sensitivity motion activation and were adjusted to capture either 5 or 10 pictures after motion was detected. All the camera traps were secured to trees on the Occidental College campus. They were attached approximately 1 ft (30.48 cm) off the ground. There were three camera stations: Station 1 camera was set up on February 9, 2017; Station 2 camera on March 28, 2017 (this camera was removed on May 13, 2017 because arborist work was blocking the camera); and Station 3 camera on June 5, 2017. The cameras were checked once per week. We went through each photo and removed the images with people, identified the wildlife species present, then unmethodically selected 966 photos to upload to Zooniverse. Though photos were taken by camera traps in bursts of either 5 or 10, they were displayed to participants individually rather than as a consecutive series. Approximately 89% of the photos selected had an organism visible, while the remaining had no wildlife.

Participants

The experiment ran from June through July of 2017, in March and in October of 2018, and finally from April through June of 2019 to increase our sample size. To attract participants for our study, we advertised through email, on social media, and on SurveyCircle

(www.surveycircle.com), a site specifically designed for recruiting survey participants. Participants were given a chance to win an Amazon gift card for making the most identifications on Zooniverse or through a raffle. The volunteers were required to specify whether they had no background in biology, some background (e.g., some high school or college biology), or an extensive background (a degree and/or career in biology) in which they were considered professional biologists.

Accuracy experiment

To quantify accuracy of photo identification by citizen scientists with varying backgrounds in biology, we either provided or did not provide JITT during the identification process. We used the citizen science website Zooniverse to create two separate conditions under which participants would identify images: the JITT treatment that offered resources to the volunteers and the control (no JITT) that did not. Participants were randomly assigned to one of these conditions and were required to classify a minimum of 5 images. In both treatments, the participants were asked to determine if an animal was present, iden-

tify the species, and to indicate the total number of individuals visible in the photo. The species options included: bird, bobcat, cat, coyote, dog, mouse, possum, raccoon, rat, skunk, and fox squirrel. The remaining options were “Other,” “Don’t Know,” and “Nothing Here.” The identification process was repeated as many times as the participant desired.

In the “No JITT” control treatment, participants were directed to a Zooniverse interface in which a camera trap image appeared along with a multiple-choice list of the possible species (see the previous paragraph). However, participants received no images, descriptions, or other resources to help them identify the image (Figure 1*i*). In contrast, participants who received the “JITT” treatment were presented with a different Zooniverse interface that provided images, descriptions, and additional identification resources for all of the potential species they may be asked to identify (Figure 1*ii*). On this interface, participants were first presented instructions on how to use the interface and the resources that were available to them. For each photo needing identification, each of the possible species on the multiple-choice list was accompanied

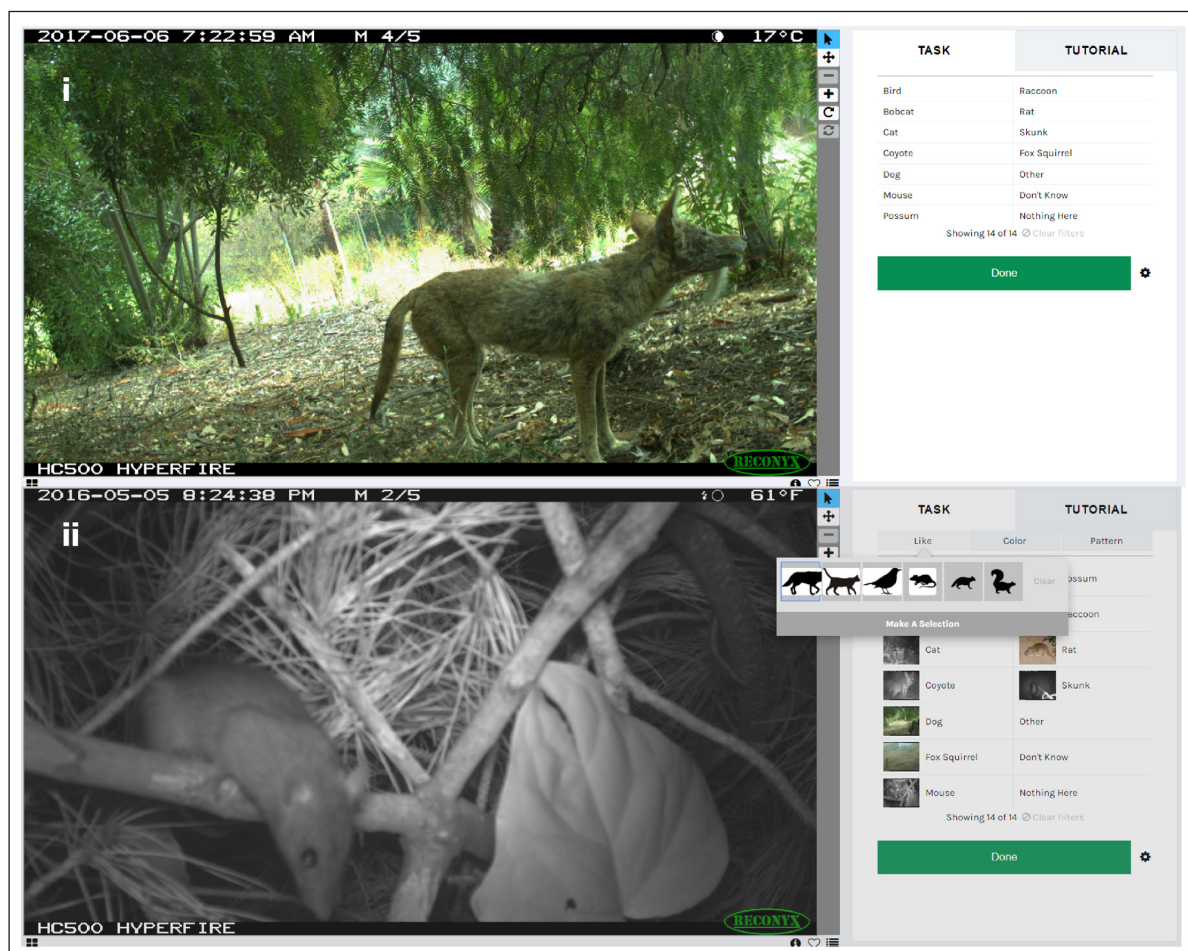


Figure 1: Zooniverse treatments for identifying wildlife images from camera traps. (i) The “No JITT” treatment includes the choices available on the right for identifying the animal, but no further assistance is provided. (ii) The “JITT” treatment includes tutorials to assist the user in identifications. Shown is what the participant would see if they selected the “Like” button, which displays the morphology choices. The “Color” and “Pattern” filters are also available to the participant with the “JITT” treatment, displaying the animals’ possible colors and coat patterns, respectively. In addition to these three categories, each animal choice has a photo associated with it, as well as a short description once that animal is selected.

by a small thumbnail image. After selecting an animal, participants were shown, via pop-ups, 2–3 additional example images and a short description before being asked to verify their choice. The pop-up images were from the same camera, location, and time period as the photo being identified to ensure that examples were similar but not identical. In addition, there was a filter available to narrow potential options based on shape, color, and pattern. If a participant was unsure about an animal, they would have the option to utilize this on-demand resource. For instance, the “Like” category displayed multiple silhouettes of wildlife, all with varying morphologies. The participant could select the morphology that they believed most accurately represented the animal in the image, and the choices would be narrowed to all the animals that fit into that morphological category. The same system was offered for the animal’s coat pattern and color, though the “Color” tab was relevant only for photos that were taken in the daylight. In addition, multiple tabs could be used at once, thereby allowing the participants to narrow their choices based on multiple factors. While the filter was not required, the example images and descriptions were presented for each photo being identified.

Analyses

A JSON parsing R script (provided by Alexandra Swanson) was used to compile raw data from Zooniverse and extract the participants’ identifications. The participants’ responses to the survey and their classifications from Zooniverse were combined. Only participants that completed both the survey and 5 or more identifications were included in the analysis. For each identification for each participant, we calculated accuracy by comparing the participant’s identification to our official identification. To increase confidence in the accuracy of the official identification, identification was determined, using photos in bursts of 5 to verify the observation, and corroborated by each of the three authors prior to image upload onto Zooniverse (Gooliaff and Hodges 2018). When calculating accuracy of participants, “Don’t Know” and “Other” responses were categorized as incorrect.

To evaluate whether there were significant differences in mean accuracy among the treatment groups, we used an ANOVA. We used the proportion of correctly identified images for each participant as our response variable and biology background (none, some, and professional biologist), training treatment, and the interaction between biology background and training treatment as the explanatory variables. A Levene’s test was used to assess equality of variances among treatment groups prior to the ANOVA (Test Statistic = 2.63, $p = 0.03$). As the Levene’s test indicated unequal variances, we conducted an arcsine square root transformation on the proportion of correctly identified images per participant, which resulted in equal variances among treatment groups (Test Statistic = 1.22, $p = 0.31$). Because the ANOVA results remained consistent regardless of the transformation, we used the original data to make interpretation of the results easier. Finally, a Tukey-Kramer test was conducted to determine which treatment

groups significantly differed in mean accuracy while accounting for multiple comparisons.

Since there are multiple ways in which an identification could be incorrect, we also assessed differences in incorrect answers across the different treatment groups. The three types of incorrect responses are 1) “Don’t Know,” selected when participants were not confident in the animal’s identity or in whether an animal was present, 2) “Nothing Here,” selected when an animal was present, and 3) the wrong animal, selected by choosing either the incorrect species or “Other.” To determine which of these options was responsible for a difference in accuracy (e.g., if the participants were selecting “Don’t Know” less frequently or if they were identifying the correct species more often), we compared the percentages of incorrect identifications in each of these categories out of all identifications for participants from each background, with and without training. Finally, we calculated the proportion of correct and incorrect identifications for each image category to assess which species and which photo types were most frequently identified incorrectly. All data analyses were conducted in the R programming language (R Core Team, 2017).

Results

Participants

A total of 94 participants volunteered for the study (23 had no biology background, 37 had some background, and 34 had at least a degree and/or profession in biology). Three participants were excluded (one person with some biology background and two professional biologists) from the analysis because they did not meet the minimum requirement of five image identifications, resulting in 91 participants. A total of 3,164 classifications were made; the number of identifications made by each participant ranged from 5 to 451, with an average of 35 identifications per participant.

Accuracy experiment

Accuracy of identifications was associated with both the background of the participants and whether they received the training treatment, with a significant interaction between background and training treatment (Background: F-ratio = 5.76, $p = 0.0045$, $df = 2$; Treatment: F-ratio = 16.87, $p = 9.00e-5$, $df = 1$; Interaction: F-ratio = 7.61, $p = 0.00091$, $df = 2$) (**Table 1; Figure 2**). When the participants did not receive any training, the volunteers with biology backgrounds identified with higher accuracy than the volunteers with no background in biology (no background: mean = 51.8%, SE = 6.0%; some background: mean = 74.7%, SE = 2.6%; professional biologist: mean = 77.6%, SE = 2.1%). However, when training was provided, the disparity between volunteers with biology backgrounds and volunteers without biology backgrounds dissipated and they had similar levels of accuracy (no background: mean = 81.9%, SE = 3.6%; some background: mean = 76.3%, SE = 3.2%; professional biologist: mean = 85.1%, SE = 2.5%). As such, only the group of participants with no biology background and no training had significantly lower

Table 1: ANOVA results comparing mean accuracy of photo identifications across participants based on biology background of participants and amount of training received.

Term	DF	SS	F-ratio	P-value	η^2	95%CI Lwr	95%CI Upr
Background	2	0.29	5.76	<i>0.0045</i>	0.09	-0.03	0.25
Training	1	0.43	16.87	<i>9.00E-05</i>	0.13	0.02	0.24
Background*Training	2	0.39	7.61	<i>0.00091</i>	0.12	0.01	0.24
Residuals	85	2.15	-	-	-	-	-

Participants self-identified their background in biology as either “No Background,” “Some Background,” or “Professional Biologist.” Participants received either the treatment with no training or were provided with just-in-time training (JITT). Significant values are italicized. For each term in the model, the following are reported: degrees of freedom (DF), sum of squares (SS), F-ratio, P-value, eta-squared (η^2), and the lower and upper 95% confidence intervals (CI).

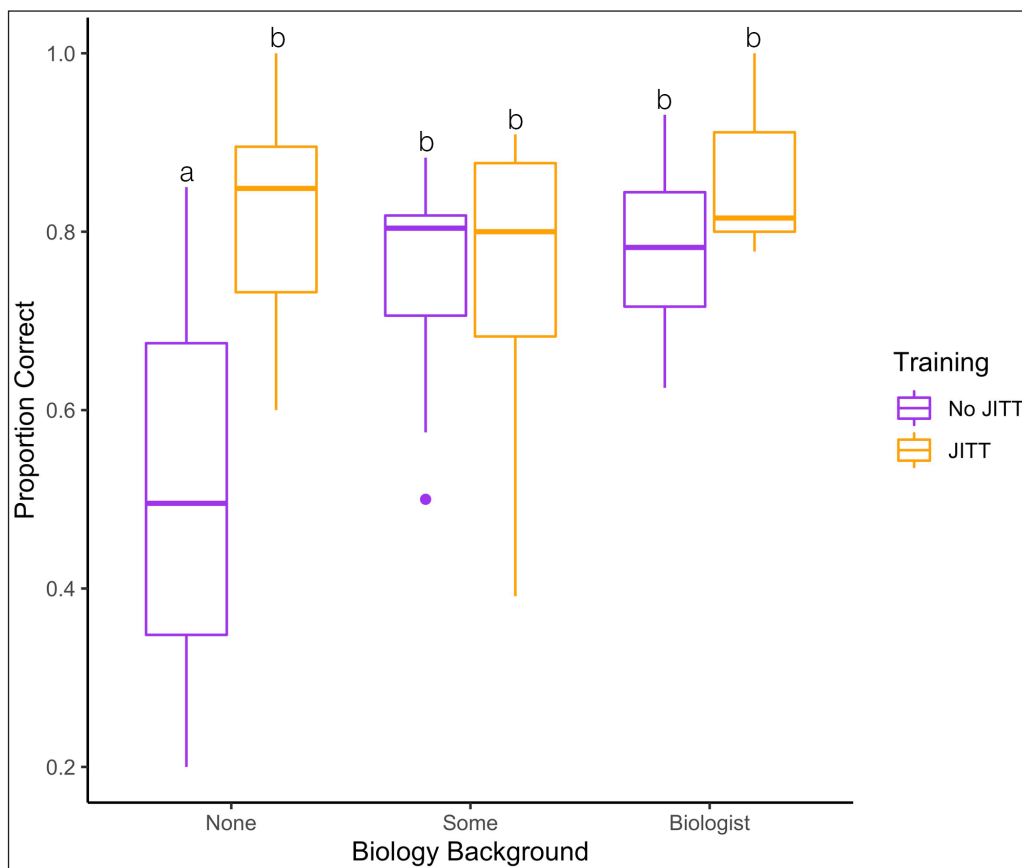


Figure 2: Accuracy of identifications (proportion of photos correctly identified) based on biology background of participants and training received. This boxplot displays the median and interquartile range for each category of biology background and treatment type (n = 91). Volunteers with no biology background were able to provide identifications that were as accurate as volunteers with biology backgrounds when training was provided but were less accurate when no training was provided (ANOVA Background by Treatment Interaction: F-ratio = 7.61, df = 2, $p = 0.00091$). Letters denote significance (Tukey-Kramer: $p \leq 0.01$).

mean accuracy than remaining groups (Tukey: $p \leq 0.01$) (Table 2). These remaining five groups did not have significantly different mean accuracies from one another (Tukey: $p > 0.05$) (Table 2). We are therefore able to reject our null hypothesis that JITTs are not associated with increased accuracy.

Incorrect identification responses

For participants with no background in biology, the proportion of incorrect observations with wrong species and “Don’t Know” was significantly lower for those with JITT

than for those without ($X^2 = 173.42$, df = 3, $p < 2.2e-16$). In contrast, the proportion of observations with “Nothing Here” that were incorrect remained constant ($X^2 = 0.21$, df = 1, $p = 0.64$). This latter result was also consistent regardless of background or training, with the number of incorrect observations marked as “Nothing Here” showing no significant difference across any of the treatment groups ($X^2 = 6.81$, df = 5, $p = 0.24$; Figure 3). When considering individual species, possums were misidentified most frequently (39.6% accuracy overall; Figure 4), whereas dogs were misidentified least frequently (97.1% accuracy; Figure 4).

Table 2: Tukey-Kramer Honestly Significant Difference results comparing differences in mean accuracy between treatment groups.

Comparison	Difference	95%CI lwr	95%CI upr	Adj P-value
Some–None	0.09	0.01	0.18	<i>0.021519</i>
Biologist–None	0.14	0.05	0.22	<i>0.000558</i>
Biologist–Some	0.04	–0.03	0.12	0.344381
JITT–No JITT	0.11	0.05	0.16	<i>0.00019</i>
Some*No JITT–None*No JITT	0.23	0.09	0.37	<i>0.000144</i>
Biologist*No JITT–None*No JITT	0.26	0.12	0.39	<i>4.00E-06</i>
None*JITT–None*No JITT	0.30	0.14	0.46	<i>4.00E-06</i>
Some*JITT–None*No JITT	0.25	0.11	0.38	<i>2.40E-05</i>
Biologist*JITT–None*No JITT	0.33	0.17	0.49	<i>1.00E-06</i>
Biologist*No JITT–Some*No JITT	0.03	–0.09	0.15	0.980451
None*JITT–Some*No JITT	0.07	–0.07	0.22	0.702743
Some*JITT–Some*No JITT	0.02	–0.11	0.14	0.998839
Biologist*JITT–Some*No JITT	0.10	–0.05	0.25	0.339966
None*JITT–Biologist*No JITT	0.04	–0.10	0.18	0.948098
Some*JITT–Biologist*No JITT	–0.01	–0.13	0.11	0.999552
Biologist*JITT–Biologist*No JITT	0.07	–0.07	0.22	0.658372
Some*JITT–None*JITT	–0.06	–0.20	0.09	0.867683
Biologist*JITT–None*JITT	0.03	–0.13	0.20	0.992882
Biologist*JITT–Some*JITT	0.09	–0.06	0.23	0.516319

Participants self-identified their background in biology as either “No Background,” “Some Background,” or “Professional Biologist.” For each comparison, mean difference is shown with the lower and upper 95% confidence intervals (CI) and the adjusted P-value. Significant values are italicized.

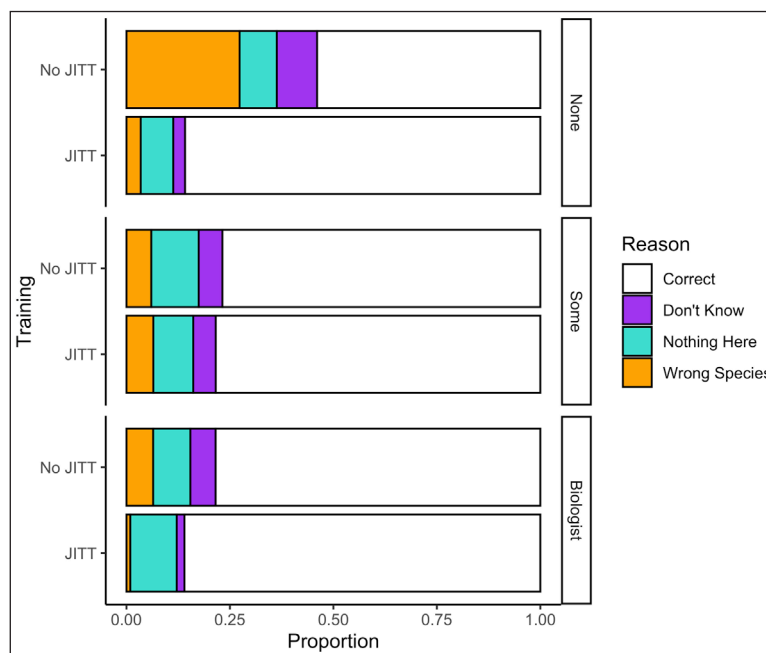


Figure 3: Proportion of correct and incorrect identifications for participants with and without training. Incorrect identifications were split into three categories: 1) “Don’t Know” was assigned to pictures identified as having an organism but the participant was unsure of the species, 2) “Nothing Here” was assigned to pictures identified as having no organisms in them when in fact there were organisms present, and 3) “Wrong Species” was assigned to pictures identified with the wrong species. Results are shown for participants with varying backgrounds in biology (none, some, and professional) and for both treatments (just-in-time training [JITT] and no training).

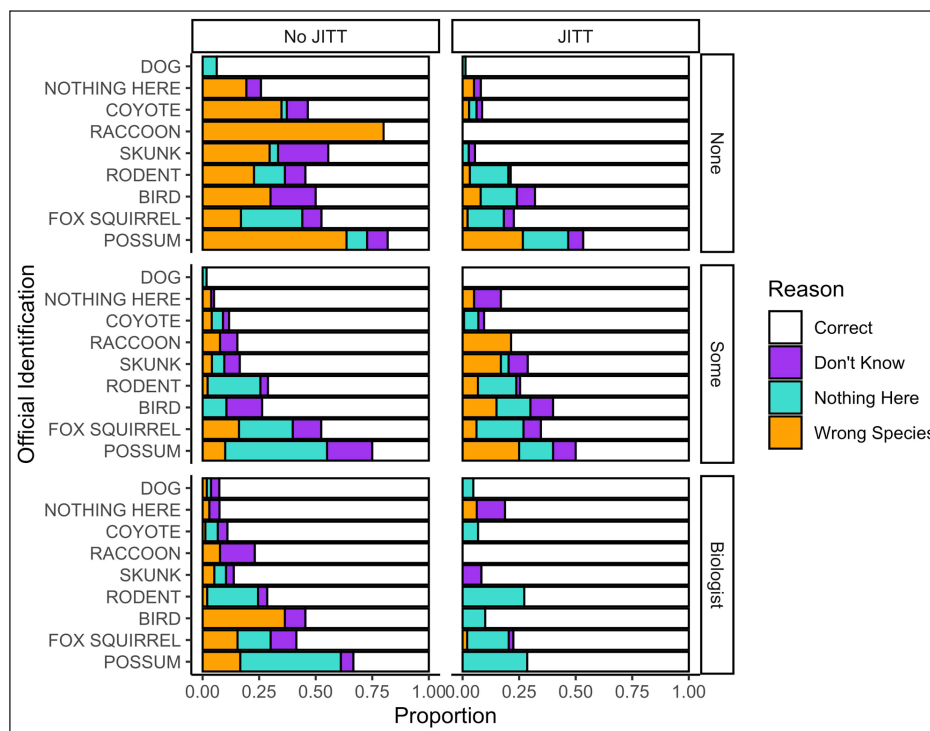


Figure 4: Proportion of correct and incorrect identifications of wildlife photos for each species for participants with and without training. For each of the official identification categories, the proportion of correct and incorrect identifications are shown. Incorrect identifications were split into three categories: 1) “Don’t Know” was assigned to pictures identified as having an organism but the participant was unsure of the species; 2) “Nothing Here” was assigned to pictures identified as having no organisms in them when in fact there were organisms present; and 3) “Wrong Species” was assigned to pictures identified with the wrong species.

Discussion

Citizen scientists are able to contribute high quantities of data to important biological research (Bhattacharjee 2005; Bonney et al. 2009; Silvertown 2009), but these contributions depend on the accuracy of the data produced by the volunteers (Kosmala et al. 2016). Using camera trap images, we assessed the accuracy of wildlife identifications made by volunteers with little to no biology background versus volunteers with professional biology backgrounds, with or without the added assistance of JITT. Our results demonstrate that when provided with relatively modest training materials, volunteers with no biology background can improve the accuracy of their identifications (**Table 1**; **Figure 2**). Based on these results, we conclude that citizen scientists can produce accurate data for scientific research when provided JITT materials.

Our results complement previous studies, which demonstrate that thorough training of volunteers improves accuracy, from tree identification (Ahrends et al. 2011) to visual surveys of fishes (Thompson and Mapstone 1997). Not only did the trainings in our study improve the wildlife photo identification accuracy of the citizen scientists with no biology background in our study, but the rates of accuracy for citizen scientists in this study that received training were also on par with those of previous studies with more intensive training programs (the range of mean accuracy of species identification in our study was 76.3–85.1% compared with 70–95% accuracy in other studies [Delaney et al. 2007; Fuccillo et al. 2015]). Importantly, this finding demonstrates that minimal training, such as JITT, not only improves identification accuracy but

also improves the accuracy just as well as other training methods, including longer-term trainings. Our results add to the existing research by indicating that minimal training can offer large dividends in improving the accuracy of identifications for volunteers with limited backgrounds in biology. Thus, extensive training may not be necessary for some types of studies, particularly with subject identification tasks.

Improvements in accuracy, however, may also vary with the difficulty of the task (Gardiner et al. 2012; Kosmala et al. 2016). Even with training, both just-in-time and long-term, participant accuracy in more difficult tasks may not reach the level required for inclusion in professional research, especially for short-term volunteers. However, for projects that rely on volunteers that are engaged for short periods or irregular intervals, training conducted in parallel with task completion may be the only feasible option. For these reasons, data quality controls, such as multiple identifications for each photo, expert validation, and using standardized equipment should continue to be used to account for inaccuracy in data collected by citizen scientists and to further improve data quality (Kosmala et al. 2016). In addition, assessing the factors that may influence the accuracy of data collected by citizen scientists (e.g., age, level of education, etc.) and applying eligibility requirements can help ensure that researchers understand the reasons for potential variations in accuracy and that citizen scientist participants are able to provide sufficiently accurate data (Delaney et al. 2007).

JITT may increase accuracy because participants misidentify fewer organisms overall, because participants

become more confident in their responses and select “Don’t Know” less frequently, or because participants less frequently select “Nothing Here” since they are better equipped to discern organisms in the photos. Our results suggest that the observed post-training increase in accuracy of participants with no background in biology is due to a combination of the first two mechanisms (**Figure 3**). Participants with no background in biology who received JITT did not show any improvement in accurately identifying photos of hard-to-notice organisms; the proportion of inaccurate selection of “Nothing Here” remained consistent across participants with no background in biology with or without JITT. This was also true for participants with a background in biology (**Figure 3**). In fact, despite the additional training resource, mean accuracy of participants with professional biology backgrounds did not exceed 85%, and the majority of incorrect observations were photos that were incorrectly identified as “Nothing Here.” This implies that additional factors, such as photo quality, may be influencing identification accuracy. After reviewing photos with incorrect identifications, we noted that many of these photos were in fact low quality with difficult-to-distinguish wildlife (e.g., a photo in which an animal is moving out of the field of view). Further, species that are more difficult to distinguish (e.g., fox squirrels) were more likely to be incorrectly identified with the “Nothing Here” option than species that stand out clearly (e.g., dogs; **Figure 4**). Perhaps additional training materials to help participants identify wildlife from non-ideal images would likewise help with improving accuracy overall. Future research should investigate how to improve accuracy for more difficult images.

While our results demonstrate an added benefit of JITT for subject identification tasks, the scale of this benefit may differ depending on the type and difficulty of tasks required for various citizen science projects (Kosmala et al. 2016). For example, while text and image training resources closed the gap in mean accuracy of wildlife photo identifications for participants in this study, advanced participants still outperform novice participants when identifying invasive plants with the assistance of text and image training materials (Starr et al. 2014). In this case, plant identification may present a greater degree of difficulty than the wildlife identification in our study. As a result, it may be ideal for those managing citizen science projects to test different methods of training to determine the best approach for the specific tasks involved in each study (Starr et al. 2014). Our results suggest that including an analysis of the effects of training methods on the accuracy of data in citizen science-based research can help assure data accuracy. Including these quality control data in studies that involve citizen science data may thus improve perceptions of citizen science-based research. Further, this approach can help shed light on which aspects of research citizen scientists can be most helpful and which aspects may be best left to experts (e.g., Casanovas et al. 2014). Future research should therefore focus on determining the success of JITTs for other types of citizen science tasks and at various levels of difficulty for each task.

Although the data suggest a significant boost to citizen science photo-identification accuracy with minimal training, there are some caveats to consider. One of the primary limitations was that participants were able to self-identify their levels of biology experience. Participants were not asked to specify their biology background, though that may have been important in assessing their qualifications (for instance, a background in organismal biology would have likely been more useful in this study than a background in botany or biochemistry). In fact, there was extensive variation in accuracy scores even for participants self-identifying as professional biologists. Despite this limitation, we still see a significant improvement in scores for participants with no biology background who receiving trainings. Future research may be needed to determine how a general background in biology translates to accuracy in citizen science projects. Another possible caveat is that the number of images that each participant identified in Zooniverse varied (e.g., one participant identified 451 photos, whereas another identified 5), which could lead to a bias in the data if accuracy improves with more experience with the task, as has been shown by others (Ratnieks et al. 2016). However, we found no significant difference in the mean ln-transformed number of photos identified regardless of background or training treatment (ANOVA: Background: F -ratio = 0.72, p = 0.50, df = 2; Training: F -ratio = 0.12, p = 0.73, df = 1; Interaction: F -ratio = 3.0, p = 0.06, df = 2, p > 0.05).

Some of the most important benefits of citizen science-based research are found in education and outreach, which can be achieved through training. Citizen science not only provides data that can be used in scientific studies, but also helps to educate the public about scientific research (Boudreau and Yan 2004; Newman et al. 2010). By improving training materials for citizen scientists, we can improve the quality of the education volunteers are receiving. In this particular study, JITT was used to educate participants on wildlife identification, which includes having an understanding of the animals’ morphologies, colors, coat patterns, and sizes. As an added benefit, citizen scientists also learned about the types of wildlife living in urban Los Angeles, including organisms rarely seen during the day. Thus, even studies that require little training should consider including JITTs to provide a reciprocal service to citizen scientists. Future research should focus on how JITTs contribute to education and outreach, which provide alternate metrics of success for citizen science research (Freitag and Pfeffer 2013).

Conclusion

Citizen science is a growing and developing field that makes it possible for researchers to collect large sets of data. Citizen science projects, however, are often considered inferior because the nature of these studies requires the involvement of people who have typically not had a significant amount of formal training in a particular scientific field (Kosmala et al. 2016). This study challenges that mentality by demonstrating that when citizen scientists with little to no scientific background are provided with JITT, they can identify wildlife images from camera traps

with as much accuracy as citizen scientists with a professional background in biology. Thus, these results suggest that citizen scientists with no background in the field can contribute accurate and meaningful subject-identification data to scientific research even when provided with only limited JITT and on-demand resources. Future research should focus on how JITTs benefit other types of subject-identification tasks as well as other citizen science-based research.

Data Accessibility Statement

Data can be accessed in the online supplemental material.

Supplementary File

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

- **Supplementary File 1.** Project Data. DOI: <https://doi.org/10.5334/cstp.219.s1>

Ethics and Consent

This study was completed in consultation with the Occidental College Internal Review Board's (IRB) Institutional Animal Care and Use Committee (IACUC) and Human Subjects Research Review Committee (HSRRC) and was declared exempt from IRB approval.

Acknowledgements

We thank two anonymous reviewers for helpful comments on the manuscript. We also thank the participants in our study.

Funding Information

We thank the Occidental Undergraduate Research Center and the Creating Opportunities in Science and Mathematics for Occidental Students (COSMOS) program for supporting this research (NSF S-STEM 1457943).

Competing Interests

The authors have no competing interests to declare.

Authors Contributions

All authors were involved in every aspect of the project.

References

- Ahrends, A, Rahbek, C, Bulling, MT, Burgess, ND, Platts, PJ, Lovett, JC, Kindemba, VW, Owen, N, Sallu, AN, Marshall, AR, Mhoro, BE, Fanning, E and Marchant, R.** 2011. Conservation and the botanist effect. *Biological Conservation*, 144(1): 131–140. DOI: <https://doi.org/10.1016/j.biocon.2010.08.008>
- Austen, GE, Bindemann, M, Griffiths, RA and Roberts, DL.** 2016. Species identification by experts and non-experts: Comparing images from field guides. *Scientific Reports*, 6(August): 1–7. DOI: <https://doi.org/10.1038/srep33634>
- Bhattacharjee, Y.** 2005. Citizen scientists supplement work of Cornell researchers, *Science*, 308(5727): 1402–1403. DOI: <https://doi.org/10.1126/science.308.5727.1402>
- Bonney, R, Cooper, CB, Dickinson, J, Kelling, S, Phillips, T, Rosenberg, KV and Shirk, J.** 2009. Citizen science: a developing tool for expanding science knowledge and scientific literacy. *BioScience*, 59(11): 977–984. DOI: <https://doi.org/10.1525/bio.2009.59.11.9>
- Bonney, R, Shirk, JL, Phillips, TB, Wiggins, A, Ballard, HL, Miller-Rushing, AJ and Parrish, JK.** 2014. Citizen science: Next steps for citizen science, *Science*, 343(6178): 1436–1437. DOI: <https://doi.org/10.1126/science.1251554>
- Boudreau, S and Yan, N.** 2004. Auditing the accuracy of a volunteer-based Invader, surveillance program for an aquatic Bythotrephes. *Environmental Monitoring and Assessment*, 91: 17–26. DOI: <https://doi.org/10.1023/B:EMAS.0000009228.09204.b7>
- Casanovas, P, Lynch, HJ and Fagan, WF.** 2014. Using citizen science to estimate lichen diversity. *Biological Conservation*, 171: 1–8. DOI: <https://doi.org/10.1016/j.biocon.2013.12.020>
- Crall, AW, Newman, GJ, Stohlgren, TJ, Holfelder, KA, Graham, J and Waller, DM.** 2011. Assessing citizen science data quality: An invasive species case study. *Conservation Letters*, 4(6): 433–442. DOI: <https://doi.org/10.1111/j.1755-263X.2011.00196.x>
- Danielsen, F, Jensen, PM, Burgess, ND, Altamirano, R, Alviola, PA, Andrianandrasana, H, Brashares, JS, Burton, AC, Coronado, I, Corpuz, N, Enghoff, M, Fjeldså, J, Funder, M, Holt, S, Hübertz, H, Jensen, AE, Lewis, R, Massao, J, Mendoza, MM, Ngaga, Y, Pippert, CB, Poulsen, MK, Rueda, RM, Sam, MK, Skielboe, T, Sørensen, M and Young, R.** 2014. A multicountry assessment of tropical resource monitoring by local communities. *BioScience*, 64(3): 236–251. DOI: <https://doi.org/10.1093/biosci/biu001>
- Delaney, DG, Sperling, CD, Adams, CS and Leung, B.** 2007. 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>
- Freitag, A and Pfeffer, MJ.** 2013. Process, not product: investigating recommendations for improving citizen science 'success'. *PLoS ONE*, 8(5): 1–5. DOI: <https://doi.org/10.1371/journal.pone.0064079>
- Fuccillo, KK, Crimmins, TM, de Rivera, CE and Elder, TS.** 2015. Assessing accuracy in citizen science-based plant phenology monitoring. *International Journal of Biometeorology*, 59(7): 917–926. DOI: <https://doi.org/10.1007/s00484-014-0892-7>
- Gardiner, MM, Allee, LL, Brown, PMJ, Losey, JE, Roy, HE and Smyth, RR.** 2012. Lessons from lady beetles: Accuracy of monitoring data from US and UK citizen-science programs. *Frontiers in Ecology and the Environment*, 10(9): 471–476. DOI: <https://doi.org/10.1890/110185>
- Gollan, J, De Bruyn, LL, Reid, N and Wilkie, L.** 2012. Can volunteers collect data that are comparable to professional scientists? A study of variables used in monitoring the outcomes of ecosystem rehabilitation. *Environmental Management*, 50(5): 969–978. DOI: <https://doi.org/10.1007/s00267-012-9924-4>

- Gooliaff, TJ and Hodges, KE.** 2018. Measuring agreement among experts in classifying camera images of similar species. *Ecology and Evolution*, 8(22): 11009–11021. DOI: <https://doi.org/10.1002/ece3.4567>
- Harmsen, BJ, Foster, RJ, Sanchez, E, Gutierrez-González, CE, Silver, SC, Ostro, LET, Kelly, MJ, Kay, E and Quigley, H.** 2017. Long term monitoring of jaguars in the Cockscomb Basin Wildlife Sanctuary, Belize; Implications for camera trap studies of carnivores, Boyce, M.S. (ed.). *PLOS ONE*, 12(6): e0179505. DOI: <https://doi.org/10.1371/journal.pone.0179505>
- Jiguet, F.** 2009. Method learning caused a first-time observer effect in a newly started breeding bird survey. *Bird Study*, 56(2): 253–258. DOI: <https://doi.org/10.1080/00063650902791991>
- Jones, MJ.** 2001. Just-in-time training. *Advances in Developing Human Resources*, 3(4): 480–487. DOI: <https://doi.org/10.1177/15234220122238409>
- Kelling, S, Johnston, A, Hochachka, WM, Iliff, M, Fink, D, Gerbracht, J, Lagoze, C, La Sorte, FA, Moore, T, Wiggins, A, Wong, WK, Wood, C and Yu, J.** 2015. Can observation skills of citizen scientists be estimated using species accumulation curves? *PLoS ONE*, 10(10): 1–20. DOI: <https://doi.org/10.1371/journal.pone.0139600>
- 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>
- Newman, G, Crall, A, Laituri, M, Graham, J, Stohlgren, T, Moore, JC, Kodrich, K and Holfelder, KA.** 2010. Teaching citizen science skills online: Implications for invasive species training programs. *Applied Environmental Education and Communication*, 9(4): 276–286. DOI: <https://doi.org/10.1080/1533015X.2010.530896>
- Prysbly, MD and Oberhauser, KS.** 2004. Temporal and geographic variation in monarch densities: Citizen scientists document monarch population patterns. In: *The monarch butterfly: biology and conservation*, 9–20.
- Ratnieks, F, Schrell, F, Sheppard, R, Brown, E, Bristow, O and Garbuzov, M.** 2016. Data reliability in citizen science: learning curve and the effects of training method, volunteer background and experience on identification accuracy of insects visiting ivy flowers. *Methods in Ecology and Evolution*, 7(10): 1226–1235. DOI: <https://doi.org/10.1111/2041-210X.12581>
- Roy, HE, Baxter, E, Saunders, A and Pocock, MJO.** 2016. Focal plant observations as a standardised method for pollinator monitoring: Opportunities and limitations for mass participation citizen science. *PLoS ONE*, 1–14. DOI: <https://doi.org/10.1371/journal.pone.0150794>
- Sauer, JR, Link, WA, Fallon, JE, Pardieck, KL and Ziolkowski, DJ.** 2013. The North American Breeding Bird Survey 1966–2011: Summary analysis and species accounts. *North American Fauna*, 79(79): 1–32. DOI: <https://doi.org/10.3996/nafa.79.0001>
- Silvertown, J.** 2009. A new dawn for citizen science. *Trends in Ecology & Evolution*, 24(9): 467–471. DOI: <https://doi.org/10.1016/j.tree.2009.03.017>
- Starr, J, Schweik, CM, Bush, N, Fletcher, L, Finn, J, Fish, J and Barger, CT.** 2014. Lights, camera...citizen science: Assessing the effectiveness of smartphone-based video training in invasive plant identification. *PLoS ONE*, 9(11): e111433. DOI: <https://doi.org/10.1371/journal.pone.0111433>
- Sullivan, BL, Aycrigg, JL, Barry, JH, Bonney, RE, Bruns, N, Cooper, CB, Damoulas, T, Dhondt, AA, Dietterich, T, Farnsworth, A, Fink, D, Fitzpatrick, JW, Fredericks, T, Gerbracht, J, Gomes, C, Hochachka, WM, Iliff, MJ, Lagoze, C, La Sorte, FA, Merrifield, M, Morris, W, Phillips, TB, Reynolds, M, Rodewald, AD, Rosenberg, KV, Trautmann, NM, Wiggins, A, Winkler, DW, Wong, W-K, Wood, CL, Yu, J and Kelling, S.** 2014. The eBird enterprise: An integrated approach to development and application of citizen science. *Biological Conservation*, 169: 31–40. DOI: <https://doi.org/10.1016/j.biocon.2013.11.003>
- 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>
- Thompson, AA and Mapstone, BD.** 1997. Observer effects and training in underwater visual surveys of reef fishes. *Marine Ecology Progress Series*, 154: 53–63. DOI: <https://doi.org/10.3354/meps154053>
- van der Wal, R, Sharma, N, Mellish, C, Robinson, A and Siddharthan, A.** 2016. The role of automated feedback in training and retaining biological recorders for citizen science. *Conservation Biology*, 30(3): 550–561. DOI: <https://doi.org/10.1111/cobi.12705>
- von Ahn, L.** 2009. Human computation. *Proceedings of the 46th Annual Design Automation Conference – DAC '09*, 418–419. DOI: <https://doi.org/10.1145/1629911.1630023>
- White, RL, Sutton, AE, Salguero-Gómez, R, Bray, TC, Campbell, H, Cieraad, E, Geekiyanage, N, Gherardi, L, Hughes, AC, Jørgensen, PS, Poisot, T, DeSoto, L and Zimmerman, N.** 2015. The next generation of action ecology: novel approaches towards global ecological research. *Ecosphere*, 6(8): art134. DOI: <https://doi.org/10.1890/ES14-00485.1>

How to cite this article: 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, pp.1–11. DOI: <https://doi.org/10.5334/cstp.219>

Submitted: 12 November 2018 **Accepted:** 11 October 2019 **Published:** 04 March 2020

Copyright: © 2020 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 <https://creativecommons.org/licenses/by/4.0/>.

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

OPEN ACCESS 