CSIRO PUBLISHING

RESEARCH PAPER
https://doi.org/10.1071/WR22022
WILDLIFE RESEARCH

Training and experience increase classification accuracy in white-tailed deer camera surveys

Jace R. Elliott[A,*], Chad H. Newbolt[A], Kelly H. Dunning[A], William D. Gulsby[A] and Stephen S. Ditchkoff[A]

For full list of author affiliations and declarations see end of paper

***Correspondence to:**
Jace R. Elliott
College of Forestry, Wildlife and Environment, Auburn University, Auburn, AL 36849, USA
Email: jace.elliott@dnr.iowa.gov

Handling Editor:
Pablo Ferreras

Received: 9 February 2022
Accepted: 18 August 2022
Published: 16 September 2022

Cite this:
Elliott JR *et al.* (2023)
Wildlife Research, **50**(7), 568–580.
doi:10.1071/WR22022

ABSTRACT

Context. Use of camera trap data in wildlife research is reliant on accurate classification of animals at the species, sex–age category or individual level. One such example is white-tailed deer (*Odocoileus virginianus*) camera surveys, which are often conducted to produce demographic estimates used by managers to establish harvest goals for a population. Previous research suggests that misclassification of deer by sex–age category (e.g. adult male, adult female, fawn) is common in these surveys, and represents a source of bias that could misinform important management decisions. **Aim.** To examine whether training material has an effect on classification accuracy of white-tailed deer and explore other observer-based, experiential factors as they relate to classification accuracy. **Methods.** We developed and tested the efficacy of species-specific training material designed to reduce sex–age misclassifications associated with white-tailed deer images. **Key results.** Exposure to training material resulted in the greatest improvement in classification accuracy of deer images compared with any other respondent-based factors we investigated. Other factors, such as professional experience as a wildlife biologist, field experience viewing white-tailed deer and experience viewing deer images from camera traps, were positively associated with classification accuracy of deer images. **Conclusions.** Our findings suggest that training material has the ability to reduce misclassifications, leading to more accurate demographic estimates for white-tailed deer populations. We also found that prior experience using camera traps and familiarity with target species was positively related to classification accuracy. **Implications.** Species-specific training material would provide a valuable resource to wildlife managers tasked with classifying animals at the species, sex–age category or individual level.

Keywords: camera survey, camera trap, classification accuracy, estimating abundance, observer error, *Odocoileus virginianus*, training material, white-tailed deer.

Introduction

In recent decades, the practice of collecting animal data through camera traps has rapidly grown in popularity within the conservation and ecology fields due to increasingly available and affordable equipment (Rowcliffe and Carbone 2008). Camera traps also offer a relatively non-invasive and passive approach to monitoring elusive species (McCarthy *et al.* 2019). Camera trapping is considered a superior sampling tool when compared with alternative methods, such as live traps or scat surveys, due to their ability to efficiently detect a high number of species and generate a large number of detections for individual species (Wearn and Glover-Kapfer 2019). Camera traps are particularly useful in determining animal occupancy (Gálvez *et al.* 2016), creating species inventories (Silveira *et al.* 2003), estimating abundance indices (Palmer *et al.* 2018), and increasing understanding of population dynamics (Karanth *et al.* 2006). However, these techniques generally depend on reliable and accurate classification of animals at either the species, sex, age, or individual level (Rovero *et al.* 2013).

Classification accuracy is subject to variability from aspects such as image-based constraints (Stevick *et al.* 2001; Meek *et al.* 2015) and vegetation conditions (Wearn and Glover-Kapfer 2019), but observer-based factors, such as experience with target species, can also have an effect on classification accuracy (Newbolt and Ditchkoff 2019). It is generally recognised that species with high visual variation among conspecifics, such as unique natural markings (e.g. spots of a cheetah [*Acinonyx jubatus*]) or secondary sexual traits (e.g. antlers of a white-tailed deer [*Odocoileus virginianus*]), may lead to more reliable classification at the individual, sex, or age–class category (Johansson *et al.* 2020). Conversely, accurately classifying species where individuals may appear visually similar to each other, for example cougars (*Puma concolor*), may provide additional challenges for observers (Kelly *et al.* 2008; Oliveira-Santos *et al.* 2010).

Since a novel approach was developed by Jacobson *et al.* (1997), commonly referred to as the Individual Branched Antlered Method (IBAM), camera surveys have become a widespread method of estimating parameters of white-tailed deer populations (Curtis *et al.* 2009; McCoy *et al.* 2011; Keever *et al.* 2017). All of these methods rely on identifying individual bucks based on unique antler characteristics and creating a sightability (photos/deer) ratio (Jacobson *et al.* 1997). This ratio is then applied to the numbers of doe and fawn images captured with the same camera traps to generate estimates for these sex–age classes. One potential downfall of this method is the assumption of equal detection probability among sex–age categories of deer, which some studies have demonstrated is a false assumption due to differing visitation and feeding patterns (Koerth and Kroll 2000; McCoy *et al.* 2011). More recently, Weckel *et al.* (2011) developed a method attempting to correct for these differences in detection probability.

Of course, these methods primarily rely on an observer's ability to identify individual bucks based on unique antler characteristics (Jacobson *et al.* 1997; Koerth *et al.* 1997); however, resulting estimates also are heavily influenced by correct classification at the sex–age level. For instance, misclassifying a fawn (<12 months) deer as an adult doe would artificially deflate fawn recruitment estimates and inflate doe estimates. Although the antlers of adult bucks provide ubiquitous distinguishing features for this sex–age class, the ability to distinguish a fawn from an adult doe relies on far more subtle determinants. Once fawns molt their neonatal pelage containing definitive spots around 3–4 months after birth (Ditchkoff 2011), observers must base their classification on fairly subjective traits like relative body size or body proportions (Newbolt and Ditchkoff 2019). Additionally, a lack of foundational familiarity with white-tailed deer life history may cause an observer to mistake a yearling buck, particularly small spike-antlered individuals, for a fawn.

In addition to the Jacobson *et al.* (1997) method and its more recent modifications, various other camera survey methods depend on accurate classification, at the species, sex–age, or individual level, to produce reliable population estimates. For instance, spatially explicit capture–recapture (Efford *et al.* 2009; Efford and Fewster 2013) and spatial capture–recapture (Royle *et al.* 2014) are two popular methods that produce density or population size estimates based on an observer's ability to identify individual animals accurately. Similarly, other studies have used occupancy modelling (Duquette *et al.* 2014) and N-mixture modelling (Keever *et al.* 2017) to estimate sex and age structure within a population, which again requires classification at the sex–age category. Even survey methods that may circumvent identifying individual animals or classifying sex–age classes, such as instantaneous, space-to-event, and time-to-event modelling (Moeller *et al.* 2018), can still be subject to bias due to misclassification at the species level.

Misidentification error is a serious concern in wildlife camera surveys, because it reduces the reliability of population estimates. Several studies have demonstrated varying rates of error among observers identifying individual animals within the same set of images (Kelly *et al.* 2008; Oliveira-Santos *et al.* 2010). Further, misclassifications may be a critical source of survey error when comparative sex–age groups lack clear distinctions (Newbolt and Ditchkoff 2019). Additionally, experienced-based factors – such as familiarity with target species and experience conducting camera surveys – likely influence rates of misclassification (Newbolt and Ditchkoff 2019). In attempts to reduce error rates, recommendations have been made to utilise multiple observers to independently classify survey images, evaluate, and monitor observer bias (Kelly *et al.* 2008), as well as conducting camera surveys during seasonal periods that maximise variation among conspecific individuals or classification groups (Newbolt and Ditchkoff 2019).

Mendoza *et al.* (2011) identify two primary strategies used to overcome misclassifications in wildlife camera-trapping data. First, creating models designed to incorporate rate of misclassification into population estimates can be effective, but only if the magnitude of error is well-known (Yoshizaki *et al.* 2009). Second, automated tools have been created to assist with the process of identifying individuals in the population (Kelly 2001; Speed *et al.* 2007; Azhar *et al.* 2012). However, neither of these strategies deal with the ultimate cause of misclassification, which is human error. Educational material may be a way to eliminate some level of misclassification by training observers to correctly classify wildlife based on objective physical traits or characteristics. For example, practical training and educational materials are frequently provided to respondents of citizen science or volunteer-based projects with the intent of improving data reliability (Newman *et al.* 2003; Cohn 2008; Steger *et al.* 2017; Parsons *et al.* 2018). Recent studies exploring the influence of longer-term (Danielsen *et al.* 2014; van der Wal *et al.* 2016) and single-session (Katrak-Adefowora *et al.* 2020; Perry *et al.* 2021) training programs on identifying wildlife images have generally found training to improve data reliability. However, the majority of previous research in this area has simply required respondents to classify wildlife images to the species level. Although the task of identifying animals to the species level can vary in difficulty among different species (Swanson *et al.* 2016), visual differentiation of unique wildlife species tends to be based on objective morphological criteria. Performing intraspecific sex–age classifications may rely more heavily on subtle, subjective criteria, such as relative size or body proportions (Newbolt and Ditchkoff 2019).

Newbolt and Ditchkoff (2019) found that the sex–age category of a white-tailed deer was the most important predictor of classification accuracy, with branch-antlered bucks classified most accurately, followed by does and fawns, respectively. However, certain observer-based factors, such as professional experience in a wildlife-related field and experience using trail cameras to view deer, had strong associations with classification accuracy as well. The authors postulated that developing species-specific training may improve reliability and accuracy of sex–age classifications for observers (Newbolt and Ditchkoff 2019), and the present study aims to take their findings a step further by introducing species-specific training material to observers designed to reduce sex–age misclassification associated with white-tailed deer images. The specific objectives of this study are to: (1) examine whether training material has an effect on overall classification accuracy; (2) measure how training material affects classification accuracy for each sex–age category of white-tailed deer; and (3) explore other observer-based, experiential factors as they relate to classification accuracy.

Materials and methods

Study area

We collected images of marked, known-age deer for this study at Auburn University's Deer Research Facility, located in the Piedmont region of east-central Alabama, USA. The facility was constructed in October 2007 and consisted of 174 ha enclosed by a 2.6-m steel fence designed to inhibit deer movements. The enclosed deer population consisted of approximately 100 individuals and comprised wild animals (and their descendants) captured during construction. Deer in the facility bred during mid-December to mid-February, with peak conception at approximately 18 January (Neuman *et al.* 2016).

Vegetation within the enclosure was approximately 40% open fields maintained for hay production, 13% bottomland hardwoods (*Quercus* spp.), 26% mature, naturally regenerated mixed oak–hickory–pine forest (oak and hickory [*Carya* spp.], loblolly pine [*Pinus taeda*]), 11% early regenerated thicket areas consisting primarily of *Rubus* spp., sweetgum (*Liquidambar styraciflua*), eastern red cedar (*Juniperus virginiana*), and Chinese privet (*Ligustrum sinense*), and 10% 10–20-year-old loblolly pine forest. A second-order creek bisected the property and provided a stable source of water year-round. Three feeders provided a 16–18% extruded protein feed (Record Rack®, Nutrena Feeds, Abilene, TX, USA) available *ad libitum*. Four timed feeders each provided deer approximately 2 kg/day of corn during October–March each year when we were actively capturing deer as part of additional research objectives.

Deer image collection

We used chemical immobilisation to capture deer in our research facility during eight trapping seasons (~1 October–15 March) from 2007 to 2015 as part of additional research objectives. All methods were approved by the Auburn University Institutional Animal Care and Use Committee (PRNs 2008-1417, 2008-1421, 2010-1785, 2011-1971, 2013-2372, 2016-2964, 2016-2985), and followed the American Society of Mammologists' guidelines (Sikes and Gannon 2011). We gave captured deer a unique 3-digit identification number corresponding with age and order of capture, which was displayed on highly visible ear tags.

We collected images containing marked deer ($n > 100\,000$) using infrared-triggered cameras (Reconyx PC 800 [Reconyx, Holmen, WI, USA]; time-lapse image capture; 1-min delay; factory default image resolution settings) placed at camera-trap sites ($n = 8$) baited with corn in February–March during the years of 2016–2020. This camera model captured full-colour images with no flash during daylight hours and black-and-white images using an infrared flash during low-light periods. Postseason deer surveys in Alabama typically occur from the end of hunting season (10 February) until spring green-up (~15 March–1 April). We selected this period to mirror a typical post hunting-season camera survey in our area. We attached cameras to an adjustable mounting bracket at a height of approximately 132 cm, and placed a 22-kg pile of corn 3.66 m from each camera. We adjusted the vertical angle of cameras such that the lens was focused on a point 72 cm above the centre of the bait pile.

We first classified collected images as adult male, adult female, fawn (i.e. 6–8-month-old deer born during the most recent fawning season), and unknown (i.e. unidentifiable; e.g. deer with obscured top of head, causing inability to determine sex). Images of adult deer used in the survey were of ear-tagged animals for which age and sex were known. Since our deer capture protocol does not include darting fawns, all fawn images used in this survey were of untagged individuals. We feel that the relatively low abundance of untagged adults (<10%), combined with the abundance of visual information provided by 1-min time-lapse imagery, allowed us to minimise instances where we erroneously classified untagged adults as fawns. Images that contained >1 deer were classified according to each individual and placed into multiple categories as needed. For example, an image with an adult female deer and a fawn would be included in both of the two appropriate categories.

Online survey development

We used Qualtrics® survey software (Qualtrics, Provo, UT, USA; accessed February 2020) to develop an online survey that tested the abilities of respondents to accurately classify deer images according to sex and age (i.e. adult vs fawn), and also tested the effect of species-specific training material on classification accuracy. This survey was developed in accordance with Auburn University policies regarding research involving human subjects (Auburn Institutional Review Board protocol #20-485; approved 08 October 2020). We randomly selected images (n = 62 images containing 75 deer) from the pool of sorted images, ensuring that all classification groups were represented in our survey (adult females = 37.3% [female deer 1.5 years and older], adult males = 29.3% [male deer 1.5 years and older], fawns = 28% [male and female deer younger than 1.5 years], unknown = 5.4% [i.e. unidentifiable; e.g. deer with obscured top of head, causing inability to determine sex]). We chose the number of images for the survey to minimise time commitments (<40 min) of respondents while maintaining adequate sample size. Ages of adult male and female deer in the selected images ranged from 1.5 to 6.5 years of age. We edited deer images using Pixlr® photo-editing software (www.pixlr.com; accessed March 2021) to remove all artificial identifying markings given to deer during capture (i.e. ear tags). We added a single-digit identification number to each deer image to link them to specific response areas in our survey.

Adult male images consisted mostly (21 of 22 images) of spike-antlered deer. The training material made available to the test group of respondents was specifically designed to focus on reducing misclassifications of spike-antlered bucks, and so we also intentionally manipulated our image set so that buck images were primarily (95.5%) comprised of spike-antlered individuals. Our justification for this decision was based on findings that branch-antlered bucks were relatively easy to accurately identify due to this conspicuous physical trait (Newbolt and Ditchkoff 2019); therefore, we were less interested in examining the effects of training material on branch-antlered bucks.

We developed a species-specific training guide designed to reduce misclassification in white-tailed deer camera surveys (Fig. 1). This training material identified distinguishing physical features among adult females, adult males, and fawns. The training material primarily focused on (1) correctly distinguishing fawns from adult females, (2) correctly distinguishing spike-antlered bucks from fawns, and (3) correctly classifying unknown (i.e. unidentifiable) deer. The distinguishing features related to relative proportions of the head, face, neck, and body used to differentiate adult females from fawns were described. The training material also included definitional information relating to these sex–age categories.

Fawns are male and female deer born during the most recent birthing period. Fawns usually are 6 months of age or younger during most deer camera survey periods. It is unlikely that male fawns will have hard antlers. Typically, they will have soft button-like protrusions on the head. Deer with hard antlers, no matter how small, should be classified as adult bucks. Fawns are born with spots, but usually lose them after a few months. Deer with spots should always be classified as fawns. However, camera surveys often occur during periods where younger fawns will still have spots and older fawns will not. Because of this, it is important to be able to distinguish older fawns from adult deer using characteristics other than spots.

Does are female deer that are at least 1 year old. They can be distinguished from fawns by several physical characteristics:

	Head	Face	Neck	Body
Fawn	Round and stubby	Eyes and nose look exaggerated	Relatively short	Roughly equal length and height (square)
Doe	Long and bottle-shaped	Eyes and nose look proportional	Relatively long	Length is visibly longer than height (rectangle)

Bucks are male deer that are at least 1 year old. They are easily identified by the presence of visible hard antler. Younger bucks frequently have small spike antlers, however, deer with any hard antler are likely adults and should be classified as bucks, not fawns.

Further Considerations

Images that do not provide a clear view of the top of a deer's head should be classified as **unknown**. The purpose of this is to maximise classification errors due to obscured spike antlers or buttons.

If you are uncertain of how to classify an image for any reason, it is best to classify it as **unknown**.

Fig. 1. Training material presented to respondents of Qualtrics® white-tailed deer identification survey. The survey was conducted 26 April–31 May 2021, and respondents consisted of individuals from across the United States that responded to online solicitations from Auburn University Deer Laboratory social media, National Deer Association (www.deerassociation.com) and Deer and Deer Hunting (www.deeranddeerhunting.com).

The training material instructed respondents to classify a deer as 'unknown' when (1) images did not provide a clear view of the top of a deer's head or (2) the respondent was uncertain of how to classify an image for any reason. Finally, a series of nine images followed that reviewed the training material and provided example classification for 14 deer. The training material was made randomly available to approximately half (48.6%) of the respondents (those in the test group) prior to answering the deer classification questions. Training material was not available to 51.4% of respondents (those in the control group).

We solicited volunteers from across the USA for our survey with assistance from multiple partners and web-based outlets, including national deer conservation/hunting organisations and social media. Adults 19 years of age or older were eligible to participate in our survey. The online survey was open for access during 26 April–31 May 2021. We took precautions to prevent respondents from taking the survey more than once by enabling the 'Prevent Ballot Box Stuffing' survey option. This option placed a cookie in the respondent's browser when they submitted a response that aided in restricting them from using the web link for our survey more than once. We first presented respondents with an information letter describing the purpose of the research, participation requirements, and privacy information as required by our institutional review board protocol. Respondents were then asked four questions that focused on general demographic information (Table 1), followed by eight questions that addressed factors we felt might influence an individual's ability to accurately classify deer images (Table 2).

We gave respondents specific information concerning the deer images prior to completing our survey. This information included (1) general geographic and captive facility details, (2) dates the images were taken, and (3) biological information for the captive deer herd (i.e. breeding season dates; approximate ages of fawns; and mass ranges for adult males, adult females, and fawns). We also notified survey respondents that all adult males were in hard antler. We then provided instructions detailing the format of our survey and how to submit responses. Respondents were presented multiple choice boxes corresponding to each numbered deer and asked to classify the image as one of four possible responses. Responses included and were defined as adult male ('male deer that are 1.5 years of age or older'), adult female ('female deer that are 1.5 years of age or older'), fawn ('male or female deer that are younger than 1.5 years of age. These are young-of-the-year deer born during the most recent fawning season. You do not need to determine if these are male or female deer'), and unknown ('not enough visible information to classify'). We randomised the order of questions for each respondent to help prevent sharing of answers, and respondents were notified that images were not in chronological order. Respondents were allowed to take as long as necessary to complete the survey, and we included only completed surveys in our analyses.

Table 1. Demographics of all respondents that took part in online survey conducted 26 April–31 May 2021.

	Frequency	Percentage (%)
Q1 – Please indicate your gender.		
Male	1652	95.27
Female	82	4.73
Prefer not to say	8	–
Q2 – Please indicate your age.		
19–24	77	4.42
25–34	242	13.88
35–44	312	17.89
45–54	365	20.93
55–64	443	25.40
65 or older	305	17.49
Prefer not to say	13	–
Q3 – Which best describes your highest level of education?		
High school degree	231	13.42
Some college	412	23.94
College degree	723	42.01
Graduate degree	355	20.63
Prefer not to say	35	–
Q4 – Which best describes your annual income level? (USD)		
$10 000–25 000	52	3.40
$25 000–50 000	213	13.91
$50 000–75 000	376	24.56
$75 000–100 000	329	21.49
$100 000+	561	36.64
Prefer not to say	195	–

Survey respondents consisted of individuals from across the United States who responded to online solicitations from Auburn University Deer Laboratory social media, National Deer Association (www.deerassociation.com) and Deer and Deer Hunting (www.deeranddeerhunting.com).

Statistical analysis

We organised responses into two groups for our analyses: (1) unknown responses and (2) known responses for adult male, adult female, and fawn images. Although our training material specified conditions that would make 'unknown' a correct response, we recognised there may be multiple reasons for a respondent selecting this response. Therefore, unknown responses were neither correct nor incorrect and were evaluated independently. The goal of our first analysis was to evaluate factors influencing accuracy of responses, without consideration of unknown responses. We determined whether responses were correct or incorrect based upon comparison with our classifications of known deer in

Table 2. Experiential predictors of all respondents that took part in online survey conducted 26 April–31 May 2021.

	Frequency	Percentage (%)
Q5 – Do you have any professional/working experience in a wildlife-related field?		
Yes	279	15.88
No	1478	84.12
Q5a – If yes, how would you classify your professional/working experience in a wildlife related field? Select all that apply		
Wildlife biology	138	49.46
Forestry	118	42.29
Land management	69	24.73
Hunting guide	60	21.51
Outdoor industry	44	15.77
Other	55	19.71
Q6 – Do you have experience hunting white-tailed deer?		
Yes	1713	97.55
No	43	2.45
Q7 – Do you have field experience viewing white-tailed deer?		
Yes	1657	94.42
No	98	5.58
Q8 – Do you have hunting/field experience viewing white-tailed deer in AL or the immediately surrounding states (FL, GA, MS, TN)?		
Yes	548	31.19
No	1209	68.81
Q9 – In your opinion, what level of experience do you currently have using trail cameras to view white-tailed deer for any purpose?		
High	617	35.12
Moderate	834	47.47
Low	241	13.72
None	65	3.70
Q10 – Have you ever conducted a trail camera survey specifically for the purpose of estimating deer population information, such as adult sex ratio, deer density, or fawn recruitment?		
Yes	396	22.54
No	1361	77.46
Q10a – If yes, how many of these kinds of trail camera surveys have you completed?		
4 or less	244	61.62
5–11	86	21.72
11 or more	66	16.67

Survey respondents consisted of individuals from across the United States who responded to online solicitations from Auburn University Deer Laboratory social media, National Deer Association (www.deerassociation.com) and Deer and Deer Hunting (www.deeranddeerhunting.com).

images. All analyses were conducted in Program R (R Core Development Team, ver. 3.4.1 accessed August 2021). We used generalised mixed-effects regression models with binomial distribution to examine classification accuracy as a function of: (1) the influence of exposure to species-specific training material; (2) professional experience with wildlife; (3) experience hunting deer; (4) field experience viewing deer; (5) local deer hunting experience; (6) general experience with using trail cameras to view deer; (7) experience conducting deer surveys using trail cameras; and (8) classification of the 'known' deer image (adult male, adult female, fawn) on classification accuracy. Random effects terms for respondent identification (ID) and deer image ID were included to account for variation associated with these effects. We calculated variance inflation factors (VIFs) and pairwise correlation coefficients among predictors associated with volunteer responses to personal information questions (Q1–Q12), in addition to exposure to training material, to evaluate collinearity in these data. We determined associations between response and predictor variables using odds ratios. The odds ratio for a predictor variable is the relative amount by which the odds of the outcome increase (odds ratio >1.0) or decrease (odds ratio <1.0) with each unit increase in the predictor variable (Hosmer *et al*. 2013). We calculated overall mean correct response rates for both trained (received training material) and non-trained (did not receive training material) respondents using a data set restricted to only include known responses of known deer images (i.e. adult male, adult female, fawn).

In addition to the previous analysis, we aimed to examine the effects of the training material on classification accuracy of specific sex–age categories of deer images. We restricted our data to include only known responses (excluding 'unknown' response), then organised these data into three subgroups according to our classification of the deer image (i.e. adult male, adult female, fawn). We used generalised mixed-effects regression models with binomial distribution to compare classification accuracy of each of these subgroups and included the same predictor variables and random effects terms used in the previous analysis. This analysis allowed us to determine the specific effect of exposure to training material on each category of deer image (i.e. adult male, adult female, fawn). We also calculated mean correct response rates for each category of known deer image using a data set restricted to only known responses.

We performed a separate analysis to examine the effects of the training material on classification accuracy of deer images we classified as unknown (i.e. unidentifiable). First, we restricted our data to include only images of deer we classified as unknown. Next, we used generalised mixed-effects regression models with binomial distribution to examine the influence of all previous predictor variables on classification accuracy of unknown deer, including previous random effects terms. We calculated mean correct response rates for unknown deer images using the data restricted to only images we classified as unknown.

We also aimed to examine the specific types of error associated with incorrect responses. First, we organised our data into two subgroups according to whether respondents were exposed to training material. Next, we restricted our data to include only incorrect responses, then organised this data set into three subgroups according to our classification of the deer image (i.e. adult male, adult female, fawn). We used generalised mixed-effects regression models with binomial distribution in Program R to model each of these subgroups with a conditional response of one of the two possible incorrect answers. Random effects terms for respondent ID and deer image ID were included to account for variation associated with these effects. This analysis allowed us to determine the likelihood of occurrence for the two possible incorrect responses respective to each of the three known deer classification groups as a function of the respondent being in the test group (received training material) or control group (did not receive training material).

Finally, we focused on examining factors contributing to unknown responses. Rather than examining the effects of our explanatory variables on accuracy of classifying unknown deer, as was described above, this analysis focused on explaining what factors led to respondents selecting 'unknown' as a response, regardless of the known category of deer image. We used our full data set to create a conditional variable based on unknown responses for this analysis. We used generalised mixed-effects regression models with binomial distribution to examine unknown response rate as a function of: (1) exposure to species-specific training material; (2) professional experience with wildlife; (3) experience hunting deer; (4) field experience viewing deer; (5) local hunting experience; (6) general experience with using trail cameras to view deer; (7) experience conducting deer surveys using trail cameras; and (8) classification of the 'known' deer image (adult male, adult female, fawn). Random effects terms for respondent identification and deer image ID were included to account for variation associated with these effects. We also calculated an overall mean unknown response rate for both trained and non-trained respondents using a data set restricted to only unknown responses.

Results

We had 1757 respondents complete our survey during the 5-week study period. We excluded 16 respondents from analysis due to incomplete responses. Respondents were primarily male and from a wide range of age groups, income levels, and education levels (Table 1). Most respondents lacked professional experience in a wildlife-related field, but had experience hunting or viewing white-tailed deer in the field, as well as using trail cameras to view white-tailed deer (Table 2). Estimates of collinearity among predictors related to respondent personal information were low (Q1 = 3.58, Q2 = 4.32, Q3 = 1.62, Q4 = 1.81, Q5 = 1.16, Q6 = 1.47, Q7 = 1.14, Q8 = 1.07, Q9 = 1.43, Q10 = 1.15; Tables 1, 2). Estimates of collinearity for whether respondents received training material was also low (1.03). We did not explore collinearity among predictors that were conditional of a specific response to a separate predictor (Q5a and Q10a).

Our analysis suggested that accuracy of deer classifications was associated positively with professional/working experience in a wildlife-related field, general experience using trail cameras to view deer, and field experience viewing white-tailed deer (Table 3). We also determined that accuracy of deer classifications was associated positively with exposure to training material (Fig. 2). We found an overall correct response rate of 80.5% and 73.4% for trained and non-trained respondents, respectively. Respondents who received training material were 6.42 (95% CL = 5.11–7.92; $P < 0.001$) times as likely to accurately classify images of adult bucks and 1.35 (95% CL = 1.21–1.51; $P < 0.001$) times as likely to accurately classify images of fawns than non-trained respondents. Accuracy of adult female deer classifications was similar between trained and non-trained respondents (Exp(β) = 1.0016 [95% CL = 0.92–1.09; $P = 0.971$).

We found that accuracy of classifications was related to the sex–age category of deer. Images of adult females were 2.43 (95% CL = 1.18–5.03; $P = 0.016$) times as likely to be classified correctly than adult male images. We did not detect a difference between classification accuracy of adult male and fawn images (Exp(β) = 1.58 [95% CL = 0.74–3.4; $P = 0.23$) or between adult female and fawn images (Exp(β) = 1.53 [95% CL = 0.73–3.14; $P = 0.25$). Adult female images that were incorrectly classified were 26.72 (95% CL = 12.14–58.84; $P < 0.001$) and 15.42 (95% CL = 7.61–31.22; $P < 0.001$) times as likely to be misclassified as fawn than adult male for trained and non-trained respondents, respectively. Adult male images that were incorrectly classified were 4.40 (95% CL = 2.22–8.74; $P < 0.001$) and 7.46 (95% CL = 3.26–17.06; $P < 0.001$) times as likely to be misclassified as fawn than adult female for trained and non-trained respondents, respectively. Fawn images that were incorrectly classified were 46.62 (95% CL = 20.28–107.17; $P < 0.001$) and 35.16 (95% CL = 20.62–59.95; $P < 0.001$) times as likely to be misclassified as adult female than adult male for trained and non-trained respondents, respectively.

Our results suggested that accuracy of classifying unknown (i.e. unidentifiable) deer was positively associated with both exposure to training material and experience hunting deer in Alabama or surrounding states. However, no other explanatory factors were found to be significant predictors of accurate classification of 'unknown' deer (Table 4). Similarly, our analysis indicated that exposure to training material was the only explanatory variable associated positively with selecting 'unknown' as a response (Table 5). Unknown responses accounted for 8.5% and 4.8% of all responses for trained and non-trained respondents, respectively. We found that unknown response rate was related to our sex–age classification group of the deer

Table 3. The effects of training material and other experiential factors on classification accuracy of 'known' (i.e. adult buck, adult doe, fawn) images of white-tailed deer from camera traps.

		Estimate		95% Confidence limits		Pr(>|z|)
		β	Exp(β)	[0.025	0.975]	
Training material		0.54	1.71	1.60	1.82	<0.001***
Wildlife professionals		0.12	1.13	1.04	1.24	0.005**
Wildlife biologists		0.13	1.14	1.01	1.29	0.030*
Land managers		0.10	1.11	0.95	1.31	0.190
Hunting guides		−0.05	0.95	0.79	1.14	0.560
Outdoor industry		−0.24	0.79	0.65	0.97	0.020*
Other		0.12	1.13	0.96	1.34	0.130
Experience hunting deer		0.17	1.18	0.95	1.48	0.140
Local experience hunting deer		−0.04	0.96	0.90	1.03	0.310
Experience viewing deer in the field		0.18	1.20	1.04	1.38	0.014*
Experience conducting deer surveys with camera traps		0.06	1.06	0.98	1.16	0.100
Viewing deer with trail cameras						
Experience level	Comparison level					
High	Moderate	0.10	1.11	1.03	1.19	0.004**
High	Low	0.17	1.18	1.07	1.31	<0.001***
High	None	0.18	1.20	1.01	1.45	0.050*
Moderate	Low	0.06	1.06	0.96	1.17	0.200
Moderate	None	0.08	1.08	0.90	1.30	0.390
Low	None	0.01	1.01	0.84	1.23	0.860

*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

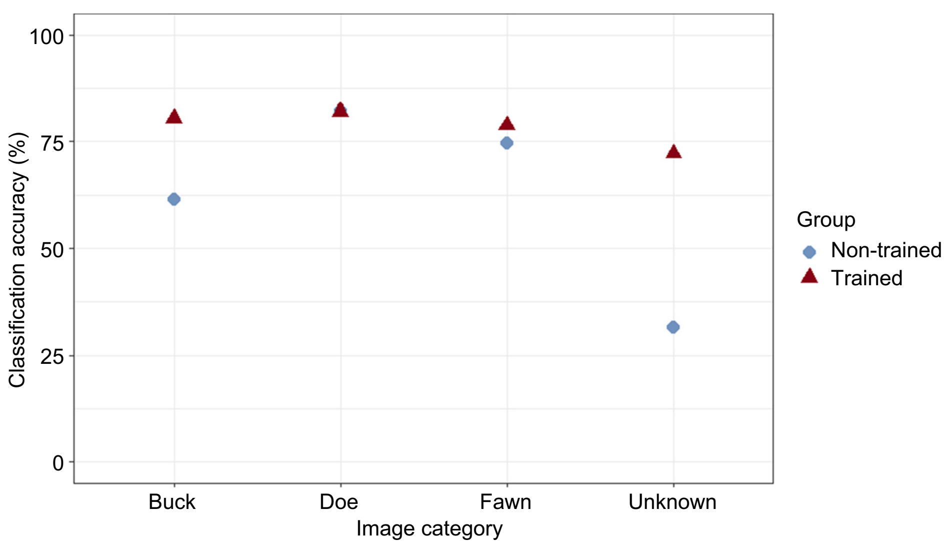

Fig. 2. Mean classification accuracy scores of both trained and non-trained respondents for each sex–age category of deer image. Adult male ≥1.5 years of age: 'buck'; adult female ≥1.5 years of age: 'doe'; juvenile of approximately 6–8 months of age: 'fawn.'; not enough visible information to classify: 'unknown'. The survey was conducted 26 April–31 May 2021, and respondents consisted of individuals from across the USA who responded to online solicitations from Auburn University Deer Laboratory social media, National Deer Association (www.deerassociation.com) and Deer and Deer Hunting (www.deeranddeerhunting.com).

Table 4. The effects of training material and other experiential factors on classification accuracy of 'unknown' (i.e. unidentifiable) images of white-tailed deer from camera traps. We considered deer unidentifiable when the top of head was obscured, causing an inability to determine sex. Accurate classification of these images required respondents selected the 'unknown' response.

		Estimate		95% Confidence limits		Pr(>|z|)
		β	Exp(β)	[0.025	0.975]	
Training material		3.32	27.66	21.37	36.28	<0.001***
Wildlife professionals		0.01	1.01	0.75	1.36	0.950
Experience hunting deer		−0.44	0.64	0.30	1.38	0.260
Local experience hunting deer		0.25	1.28	1.05	1.84	0.036*
Experience viewing deer in field		−0.18	0.83	0.50	1.37	0.470
Experience conducting deer surveys with camera traps		−0.18	0.83	0.70	1.20	0.650
Viewing deer with trail cameras						
Experience level	Comparison level					
High	Moderate	0.05	1.05	0.82	1.34	0.690
High	Low	0.25	1.28	0.89	1.84	0.170
High	None	0.39	1.49	0.79	2.83	0.220
Moderate	Low	−0.21	0.81	0.59	1.14	0.240
Moderate	None	0.34	1.41	0.76	2.66	0.270
Low	None	0.15	1.16	0.60	2.25	0.650

*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

Table 5. Effects of training material and other experiential factors on 'unknown' response rate. This analysis includes all 'unknown' responses regardless of deer image category.

		Estimate		95% Confidence limits		Pr(>|z|)
		β	Exp(β)	[0.025	0.975]	
Training material		1.520	4.570	4.02	5.500	<0.001***
Wildlife professionals		0.004	1.004	0.81	1.250	0.970
Experience hunting deer		−0.450	0.640	0.37	1.100	0.110
Local experience hunting deer		0.160	1.170	0.99	0.139	0.060
Experience viewing deer in field		−0.020	0.980	0.69	1.410	0.950
Experience conducting deer surveys with camera traps		−0.110	0.900	0.74	1.100	0.310
Viewing deer with trail cameras						
Experience level	Comparison level					
High	Moderate	0.100	1.110	0.93	1.330	0.240
High	Low	0.170	1.190	0.93	1.550	0.160
High	None	−0.520	0.590	0.38	0.930	0.02*
Moderate	Low	0.070	1.070	0.85	1.370	0.530
Moderate	None	0.420	1.520	0.97	2.360	0.060
Low	None	0.410	1.510	0.88	2.230	0.150

*$P < 0.05$; **$P < 0.01$; ***$P < 0.001$.

image. Respondents were 2.59 (95% CL = 1.21–5.62; $P = 0.011$) times as likely to select an unknown response for adult female images than adult male images. Respondents were 2.34 (95% CL = 1.08–5.10; $P = 0.031$) times as likely to select an unknown response for adult female images than fawn images. Unknown response rates were similar for adult male and fawn images (Exp(β) = 1.11 [95% CL = 0.39–2.02; $P = 0.79$]).

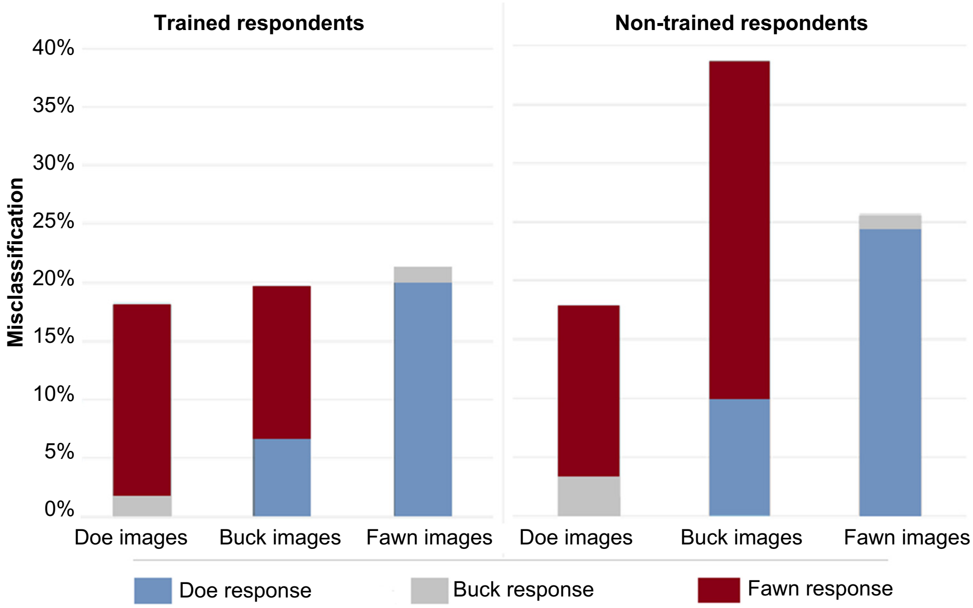

Fig. 3. Misclassification of each known sex–age category of deer image for trained and non-trained respondents. Stacked bars represent proportion of misclassification due to each incorrect response. The survey was conducted 26 April–31 May 2021, and respondents consisted of individuals from across the United States who responded to online solicitations from Auburn University Deer Laboratory social media, National Deer Association (www.deerassociation.com) and Deer and Deer Hunting (www.deeranddeerhunting.com).

Discussion

Of all investigated factors related to the individual respondent, training material resulted in the greatest decrease in misclassification rates. These results are similar to previous studies, which demonstrate that training material can improve quality of wildlife survey data (Ratnieks *et al.* 2016; Katrak-Adefowora *et al.* 2020; Perry *et al.* 2021). Specifically, we saw the greatest reductions in error among adult males and fawns. Newbolt and Ditchkoff (2019) had previously identified spike-antlered males and fawns as being associated with the greatest degree of error during classification of trail camera images.

Our results showed the greatest reduction of misclassification in trained respondents was primarily for adult males, hereafter referred to as bucks, relative to other known sex–age categories (Fig. 3). For both trained and non-trained respondents, most misclassified buck images were mistaken for fawns. Newbolt and Ditchkoff (2019) also found that spike-antlered buck images were most frequently misidentified as fawns, which the authors attributed to misinformation regarding antler growth patterns. The training material we offered respondents specifically addressed this issue by clearly describing the physical distinctions between male fawns and spike-antlered bucks. Not only were respondents who received training more likely to accurately classify buck images, but we also observed a major reduction in the proportion of buck images misclassified as fawns. Based on these findings, we believe the training material was effective for informing respondents on correct antler growth patterns.

We found that trained respondents were also more likely to accurately classify images of fawns. Fawns were most frequently misclassified as adult females, hereafter referred to as does, across both trained and non-trained groups of respondents. Although the reduction in misclassification of fawn images may not have been as drastic as for buck images, most of the reduced error was a result of fewer trained respondents mistaking fawns for does, which we attribute to the multiple strategies for differentiating fawns from does outlined in the training material. Newbolt and Ditchkoff (2019) hypothesised that misclassification between fawns and does results from a lack of distinct physical traits between these two sex–age categories. Rather than looking for relatively obvious physical traits, such as the presence of hard antlers, observers must rely on the ability to use subjective criteria, such as relative size and body proportions to make accurate classifications between fawns and does.

Relative to all the known sex–age categories of deer in the survey images (i.e. buck, doe, fawn), our results showed that training material had the greatest effect of reducing rate of misclassification for unknown (i.e. unidentifiable) deer. Every wildlife camera survey is likely to include images of individual animals that are unidentifiable due to uncertainty caused by distance from the camera, position/orientation of the animal, or a number of other factors. By classifying an unidentifiable animal as unknown, an observer limits the potential error in a survey that may arise from misclassification. If an observer misclassifies an unidentifiable animal at the individual or sex–age category level based on erroneous assumptions, resulting population estimates, such as recruitment or sex ratio, could become biased. Both trained and non-trained respondents were instructed to classify individuals as unknown given 'not enough visible information to identify'; however, trained respondents received advanced instruction regarding unknown classifications in the training material in addition to viewing two example images of unidentifiable deer correctly classified as unknown. We believe that the reported improvement in correct classification of unknown deer can be attributed to exposure to such training material.

Our results also suggested that exposure to training material resulted in respondents being more likely to select 'unknown' when faced with uncertainty in classifying deer images. These findings are contradictory to those of Katrak-Adefowora *et al.* (2020), who reported a lower unknown (i.e. 'Don't Know') response rate for trained than non-trained respondents. However, we believe our study required respondents to perform more difficult classifications (e.g. differentiating sex–age categories based on subjective criteria) than the previously mentioned study (i.e. identifying animals to the 'species' level [bird, cat, dog, skunk, etc.]); therefore, we assume that our respondents faced greater levels of uncertainty in classifying images. Classifying an image as 'unknown' effectively shrinks the pool of usable data in a camera survey, but we maintain our belief that making such a decision in the case of unidentifiable images or general uncertainty will tend to result in net benefits for the reliability of resulting population estimates.

Newbolt and Ditchkoff (2019) also found that the sex–age category of deer was a major predictor of classification accuracy; however, differences between the image sets used in their study and this current study produced contrasting rates of classification accuracy among sex–age categories – specifically, the fact that this study included buck images that were vastly comprised of spike-antlered individuals (21 of 22) was different from Newbolt and Ditchkoff (2019), which primarily used images of branch-antlered bucks (28 of 32). We believe this difference in image sets resulted in greater misclassification of buck images compared with Newbolt and Ditchkoff (2019), who found that spike-antlered bucks were misclassified at a far greater rate than branch-antlered bucks. We believe the difference in rate of misclassified buck images between trained and non-trained respondents reported in this study supports the hypothesis of Newbolt and Ditchkoff (2019) that much of the misclassification of spike-antlered bucks they observed in their study resulted from misinformation about antler growth patterns rather than visual inaccuracy.

We found experiential factors to be important predictors of classification accuracy in this study. These conclusions are corroborated by findings from Newbolt and Ditchkoff (2019), who also found that experience played a role in the accuracy of deer classifications. Specifically, professional experience as a wildlife biologist, field experience viewing deer, and experience viewing deer using trail cameras were important determinants of accurately classifying deer images. Although findings suggest experiential factors may only account for a low rate of error in wildlife camera surveys, any significant source of error must be critically assessed to maximise reliability of survey output (Newbolt and Ditchkoff 2019). Multiple studies have reported that low rates of error when identifying camera images can contribute to considerable biases in survey estimates (Gunnlaugsson and Sigurjónsson 1990; Stevick *et al.* 2001; Morrison *et al.* 2011). Of respondents that indicated professional experience in a wildlife-related field, wildlife biologists were most accurate at classifying. Several ecological studies have revealed an advantage in the reliability of data collected by professional researchers compared with non-professional volunteers (Darwall and Dulvy 1996; Garel *et al.* 2005; Lovell *et al.* 2009; Ahrends *et al.* 2011; Lewandowski and Specht 2015). Receiving training material remained a stronger predictor of classification accuracy than professional experience, even for wildlife biologists; however, we acknowledge that the overall proportion of respondents that indicated experience as a wildlife biologist (<8%) was likely too small to responsibly draw definitive conclusions from this comparison.

We feel it is important to note that the pool of respondents in this study may not accurately represent the demographics and experience level of the individuals who typically conduct wildlife surveys, primarily based on the fact that a vast majority of our respondents lacked professional or working experience in a wildlife-related field. We also acknowledge that our respondents were not representative of the general public. Rather, our pool of respondents was an artifact of the methods used to elicit participation, primarily through deer hunting-based media platforms. Our survey images were collected during a post-hunting season timeframe, which is a common time of year to conduct white-tailed deer surveys in our region (Newbolt and Ditchkoff 2019). Deer camera surveys conducted earlier in the year would offer easier differentiation of adult does and fawns, due to a more apparent difference in body size as well as the higher likelihood of fawn pelage containing spots. However, biologists and managers must be aware of a potentially lower detection probability of fawns during this time, leading to underestimations of fawn abundance (McKinley 2002; McCoy *et al.* 2011). The deer images presented to our respondents were intended to reflect those collected in an actual camera survey. However, the integrity of our data required presenting images in a random order, rather than in a chronological series common under normal field situation. This departure from real-world conditions must be considered when interpreting our reported rates of classification accuracy, which may have been greater had images been presented chronologically due to deer potentially being captured more than once and from multiple angles. Regardless, we believe that our research reveals important trends and factors that contribute to misclassification in wildlife camera surveys.

Our study demonstrates that misclassification of sex–age categories may be a surprisingly widespread source of error in wildlife camera surveys. Oftentimes, important management decisions are informed by survey estimates, and resulting management actions may be misinformed if operating on biased population estimates. Our findings suggest that training material has the ability to improve population estimates from camera surveys by reducing rates of misclassification. The training material in this study was extremely concise and simplistic, yet still significantly increased classification accuracy at the sex–age level. We suggest that similar tools be readily accessible and frequently utilised, even for experienced practitioners, to minimise potential bias resulting from sex–age misclassification.

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, Marchant R (2011) Conservation and the botanist effect. *Biological Conservation* **144**, 131–140. doi:10.1016/j.biocon.2010.08.008

Azhar MAHB, Hoque S, Deravi F (2012) Automatic identification of wildlife using local binary patterns. In 'IET Conference on Image Processing'. (IET Digital Library: London, UK)

Cohn JP (2008) Citizen science: can volunteers do real research? *BioScience* **58**, 192–197. doi:10.1641/B580303

Curtis PD, Boldgiv B, Mattison PM, Boulanger JR (2009) Estimating deer abundance in suburban areas with infrared-triggered cameras. *Human–Wildlife Conflicts* **3**, 116–128.

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, Pipper CB, Poulsen MK, Rueda RM, Sam MK, Skielboe T, Sørensen M, Young R (2014) A multicountry assessment of tropical resource monitoring by local communities. *BioScience* **64**, 236–251. doi:10.1093/biosci/biu001

Darwall WRT, Dulvy NK (1996) An evaluation of the suitability of non-specialist volunteer researchers for coral reef fish surveys. Mafia Island, Tanzania – a case study. *Biological Conservation* **78**, 223–231. doi:10.1016/0006-3207(95)00147-6

Ditchkoff SS (2011) Anatomy and physiology. In 'Biology and Management of White-tailed Deer'. (Ed. DG Hewitt) pp. 43–73. (CRC Press, Boca Raton, FL, USA)

Duquette JF, Belant JL, Svoboda NJ, Beyer DE Jr., Albright CA (2014) Comparison of occupancy modeling and radiotelemetry to estimate ungulate population dynamics. *Population Ecology* **56**, 481–492. doi:10.1007/s10144-014-0432-7

Efford MG, Fewster RM (2013) Estimating population size by spatially explicit capture–recapture. *Oikos* **122**, 918–928. doi:10.1111/j.1600-0706.2012.20440.x

Efford MG, Borchers DL, Byrom AE (2009) Density estimation by spatially explicit capture–recapture: likelihood-based methods. In 'Modeling Ddemographic Processes in Marked Populations, Environmental and Ecological Statistics. Vol. 3'. (Eds DL Thomson, EG Cooch, MJ Conroy) pp. 255–269. (Springer Science & Business Media, LLC: Berlin, Germany)

Gálvez N, Guillera-Arroita G, Morgan BJT, Davies ZG (2016) Cost-efficient effort allocation for camera-trap occupancy surveys of mammals. *Biological Conservation* **204**, 350–359. doi:10.1016/j.biocon.2016.10.019

Garel M, Cugnasse J-M, Gaillard J-M, Loison A, Santosa Y, Maublanc M-L (2005) Effect of observer experience on the monitoring of a mouflon population. *Acta Theriologica* **50**, 109–114. doi:10.1007/BF03192623

Gunnlaugsson T, Sigurjónsson J (1990) A note on the problem of false positives in the use of natural marking data for abundance estimation. *Report of the International Whaling Commission* **12**, 143–145.

Hosmer DW, Lemeshow S, Sturdivant RX (2013) 'Applied Logistic Regression.' (John Wiley and Sons: Hoboken, NJ, USA)

Jacobson HA, Kroll JC, Browning RW, Koerth BH, Conway MH (1997) Infrared-triggered cameras for censusing white-tailed deer. *Wildlife Society Bulletin* **25**, 547–556.

Johansson O, Samelius G, Wikberg E, Chapron G, Mishra C, Low M (2020) Identification errors in camera-trap studies result in systematic population overestimation. *Scientific Reports* **10**, 6393. doi:10.1038/s41598-020-63367-z

Karanth KU, Nichols JD, Kumar NS, Hines JE (2006) Assessing tiger population dynamics using photographic capture–recapture sampling. *Ecology* **87**, 2925–2937. doi:10.1890/0012-9658(2006)87[2925:ATPDUP]2.0.CO;2

Katrak-Adefowora R, Blickley JL, Zellmer AJ (2020) Just-in-time training improves accuracy of citizen scientist wildlife identifications from camera trap photos. *Citizen Science: Theory and Practice* **5**, 8. doi:10.5334/cstp.219

Keever AC, McGowan CP, Ditchkoff SS, Acker PK, Grand JB, Newbolt CH (2017) Efficacy of N-mixture models for surveying and monitoring white-tailed deer populations. *Mammal Research* **62**, 413–422. doi:10.1007/s13364-017-0319-z

Kelly MJ (2001) Computer-aided photograph matching in studies using individual identification: an example from serengeti cheetahs. *Journal of Mammalogy* **82**, 440–449. doi:10.1644/1545-1542(2001)082<0440:CAPMIS>2.0.CO;2

Kelly MJ, Noss AJ, DiBitetti MS, Maffei L, Arispe RL, Paviolo A, DeAngelo CD, DiBlanco YE (2008) Estimating puma densities from camera trapping across three study sites: Bolivia, Argentina, and Belize. *Journal of Mammalogy* **89**, 408–418. doi:10.1644/06-MAMM-A-424R.1

Koerth BH, Kroll JC (2000) Bait type and timing for deer counts using cameras triggered by infrared monitors. *Wildlife Society Bulletin* **28**, 630–635.

Koerth BH, McKown CD, Kroll JC (1997) Infrared-Triggered camera versus helicopter counts of white-tailed deer. *Wildlife Society Bulletin* **25**, 557–562.

Lewandowski E, Specht H (2015) Influence of volunteer and project characteristics on data quality of biological surveys. *Conservation Biology* **29**, 713–723. doi:10.1111/cobi.12481

Lovell S, Hamer M, Slotow R, Herbert D (2009) An assessment of the use of volunteers for terrestrial invertebrate biodiversity surveys. *Biodiversity and Conservation* **18**, 3295–3307. doi:10.1007/s10531-009-9642-2

McCarthy MS, Després-Einspenner M-L, Farine DR, Samuni L, Angedakin S, Arandjelovic M, Boesch C, Dieguez P, Havercamp K, Knight A, Langergraber KE, Wittig RM, Kühl HS (2019) Camera traps provide a robust alternative to direct observations for constructing social networks of wild chimpanzees. *Animal Behaviour* **157**, 227–238. doi:10.1016/j.anbehav.2019.08.008

McCoy JC, Ditchkoff SS, Steury TD (2011) Bias associated with baited camera sites for assessing population characteristics of deer. *The Journal of Wildlife Management* **75**, 472–477. doi:10.1002/jwmg.54

McKinley WT (2002) Evaluating infrared camera and other census techniques for white-tailed deer in Mississippi. M.Sc. Thesis, Mississippi State University, MS, USA.

Meek PD, Ballard G-A, Fleming PJS (2015) The pitfalls of wildlife camera trapping as a survey tool in Australia. *Australian Mammalogy* **37**, 13–22. doi:10.1071/AM14023

Mendoza E, Martineau PR, Brenner E, Dirzo R (2011) A novel method to improve individual animal identification based on camera-trapping data. *The Journal of Wildlife Management* **75**, 973–979. doi:10.1002/jwmg.120

Moeller AK, Lukacs PM, Horne JS (2018) Three novel methods to estimate abundance of unmarked animals using remote cameras. *Ecosphere* **9**, e02331. doi:10.1002/ecs2.2331

Morrison TA, Yoshizaki J, Nichols JD, Bolger DT (2011) Estimating survival in photographic capture–recapture studies: overcoming

misidentification error. *Methods in Ecology and Evolution* **2**, 454–463. doi:10.1111/j.2041-210X.2011.00106.x
Neuman TJ, Newbolt CH, Ditchkoff SS, Steury TD (2016) Microsatellites reveal plasticity in reproductive success of white-tailed deer. *Journal of Mammalogy* **97**, 1441–1450. doi:10.1093/jmammal/gyw087
Newbolt CH, Ditchkoff SS (2019) Misidentification error associated with classifications of white-tailed deer images. *Wildlife Society Bulletin* **43**, 527–536. doi:10.1002/wsb.985
Newman C, Buesching CD, Macdonald DW (2003) Validating mammal monitoring methods and assessing the performance of volunteers in wildlife conservation—"Sed quis custodiet ipsos custodies?". *Biological Conservation* **113**, 189–197.
Oliveira-Santos LGR, Zucco CA, Antunes PC, Crawshaw PG Jr (2010) Is it possible to individually identify mammals with no natural markings using camera-traps? A controlled case-study with lowland tapirs. *Mammalian Biology* **75**, 375–378. doi:10.1016/j.mambio.2009.08.005
Palmer MS, Swanson A, Kosmala M, Arnold T, Packer C (2018) Evaluating relative abundance indices for terrestrial herbivores from large-scale camera trap surveys. *African Journal of Ecology* **56**, 791–803. doi:10.1111/aje.12566
Parsons AW, Goforth C, Costello R, Kays R (2018) The value of citizen science for ecological monitoring of mammals. *PeerJ* **6**, e4536. doi:10.7717/peerj.4536
Perry JR, Sumner S, Thompson C, Hart AG (2021) 'Citizen identification': online learning supports highly accurate species identification for insect-focused citizen science. *Insect Conservation and Diversity* **14**, 862–867. doi:10.1111/icad.12528
Ratnieks FLW, Schrell F, Sheppard RC, Brown E, Bristow OE, 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**, 1226–1235. doi:10.1111/2041-210X.12581
Rovero F, Zimmermann F, Berzi D, Meek P (2013) "Which camera trap type and how many do I need?" A review of camera features and study designs for a range of wildlife research applications. *Hystrix* **24**, 148–156. doi:10.4404/hystrix-24.2-8789
Rowcliffe JM, Carbone C (2008) Surveys using camera traps: are we looking to a brighter future? *Animal Conservation* **11**, 185–186. doi:10.1111/j.1469-1795.2008.00180.x
Royle JA, Chandler RB, Sollmann R, Gardner B (Eds) (2014) 'Spatial capture–recapture.' (Elsevier: Waltham, MA, USA)
Sikes RS, Gannon WL (2011) Guidelines of the American Society of Mammalogists for the use of wild mammals in research. *Journal of Mammalogy* **92**, 235–253.
Silveira L, Jácomo ATA, Diniz-Filho JAF (2003) Camera trap, line transect census and track surveys: a comparative evaluation. *Biological Conservation* **114**, 351–355. doi:10.1016/S0006-3207(03)00063-6
Speed CW, Meekan MG, Bradshaw CJA (2007) Spot the match – wildlife photo-identification using information theory. *Frontiers in Zoology* **4**, 2. doi:10.1186/1742-9994-4-2
Steger C, Butt B, Hooten MB (2017) Safari science: assessing the reliability of citizen science data for wildlife surveys. *Journal of Applied Ecology* **54**, 2053–2062. doi:10.1111/1365-2664.12921
Stevick PT, Palsbøll PJ, Smith TD, Bravington MV, Hammond PS (2001) Errors in identification using natural markings: rates, sources, and effects on capture–recapture estimates of abundance. *Canadian Journal of Fisheries and Aquatic Sciences* **58**, 1861–1870. doi:10.1139/cjfas-58-9-1861
Swanson A, Kosmala M, Lintott C, Packer C (2016) A generalized approach for producing, quantifying, and validating citizen science data from wildlife images. *Conservation Biology* **30**, 520–531. doi:10.1111/cobi.12695
van der Wal R, Sharma N, Mellish C, Robinson A, Siddharthan A (2016) The role of automated feedback in training and retaining biological recorders for citizen science. *Conservation Biology* **30**, 550–561. doi:10.1111/cobi.12705
Wearn OR, Glover-Kapfer P (2019) Snap happy: camera traps are an effective sampling tool when compared with alternative methods. *Royal Society Open Science* **6**, 181748. doi:10.1098/rsos.181748
Weckel M, Rockwell RF, Secret F (2011) A modification of Jacobson *et al.*'s (1997) individual branch-antlered male method for censusing white-tailed deer. *Wildlife Society Bulletin* **35**, 445–451. doi:10.1002/wsb.64
Yoshizaki J, Pollock KH, Brownie C, Webster RA (2009) Modeling misidentification errors in capture–recapture studies using photographic identification of evolving marks. *Ecology* **90**, 3–9. doi:10.1890/08-0304.1

Data availability. The data that support this study will be shared upon reasonable request to the corresponding author.

Conflicts of interest. The authors declare no conflicts of interest.

Declaration of funding. All funding for this study was provided by the Agricola Wildlife Foundation.

Acknowledgements. We thank the National Deer Association and *Deer & Deer Hunting Magazine* for their assistance in reaching study participants. Additional thanks to the many participants who volunteered their time to take our survey and made this research possible. This paper forms part of the MSc thesis of Elliott (2021), titled Translocating Captive Female White-tailed Deer and Training and Experience Increase Classification Accuracy in White-tailed Deer Camera Surveys.

Author affiliation
[A]College of Forestry, Wildlife and Environment, Auburn University, Auburn, AL 36849, USA.