Wildlife Society Bulletin 00(0):1–10; 2019; DOI: 10.1002/wsb.985

*Original Article*

# Misidentification Error Associated With Classifications of White-tailed Deer Images

**CHAD H. NEWBOLT,**[1] *School of Forestry and Wildlife Sciences, Auburn University, AL 36849, USA*
**STEPHEN S. DITCHKOFF,** *School of Forestry and Wildlife Sciences, Auburn University, AL 36849, USA*

**ABSTRACT** Camera traps are widely used to monitor wildlife, with important management decisions often relying on interpretation of these data. Animal misidentifications are known to be an important source of error in wildlife surveys that require the identification of unique individuals from camera-trap data; however, the practice of broadly classifying animal images according to sex or age has received less critical attention despite the significant potential for misidentification error under certain circumstances. From 19 January to 1 April 2017, we solicited a group of 726 participants, consisting of both wildlife professionals and nonprofessionals from across the United States, to take an online survey that tested their ability to classify images of known white-tailed deer (*Odocoileus virginianus*) according to sex and age. Our goal was to determine the relative influence of tested observer (i.e., experience and familiarity with classifying deer images) and image-based factors (i.e., distance of deer from camera, day vs. night image) on accuracy of deer classifications. Our results indicated that respondents that were wildlife biologists and those with greater levels of experience viewing deer images were more accurate than others when classifying posthunting season images of deer as adult male, adult female, or fawn. However, the sex–age group of the deer was the most influential predictor of classification reliability, with branched-antlered adult males being classified more accurately by all respondent groups than were adult females and fawns. Our findings emphasized that animal misidentifications may be an important source of survey error not only when identifying unique individuals, but also under any circumstance where comparative groups lack definitive traits. We suggest that those using camera traps to evaluate wildlife populations should select survey periods that maximize differences among classification groups, when possible, and develop species-specific image training for observers to improve the reliability of results. Further, population demographics should be considered when evaluating the overall reliability of survey results for species where classification accuracy varies among sex–age groups. 

**KEY WORDS** camera trap, misidentification error, *Odocoileus virginianus*, survey, trail camera, white-tailed deer.

Camera traps have proven a cost-effective and minimally invasive wildlife monitoring solution; these advantages have led to their widespread use in research, management, and conservation efforts (Cutler and Swann 1999, Burton et al. 2015, Steenweg et al. 2017). Data collected from camera traps are typically classified according to species, sex, age, or unique individuals, and, as such, accurate identifications are central to the reliability of results (Rovero et al. 2013). An observer's ability to correctly classify an image is influenced by many factors, including image quality and observer-based considerations, such as experience with the subject species; however, distinctness of the target animal's appearance relative to surrounding wildlife is fundamental to the accuracy of classifications (Foster and Harmsen 2012). Reliable identifications may be simple to obtain when animals have a highly unique appearance; however, accurate observations may become difficult to achieve when individuals from different classification groups are visually similar (Carbone et al. 2001). For example, a high level of accuracy in classifications could be assumed for observers determining presence or absence of a distinct species, such as African elephant (*Loxodonta africana*), but accuracy likely would be less for those attempting to determine the individual identities of pumas (*Puma concolor*) because that species lacks unique natural markings (Kelly et al. 2008).

In addition to the physical appearance of the animal, image attributes such as resolution, flash type, and distance of animal from the camera, contribute to classification accuracy (Kelly et al. 2008, Yoshizaki et al. 2009, Oliveira-Santos et al. 2010). Observers are reliant upon the 'quality' of the image to make classifications; logic dictates that accuracy of classifications is greater for higher quality images than those of lower quality. Previous research efforts have

Received: 15 January 2019; Accepted: 3 March 2019
Published:

[1]*E-mail: newboch@auburn.edu*

focused on determining the influences of camera specifications (e.g., detection zone, trigger speed, and infrared sensor sensitivity), set-up protocol (e.g., camera mounting height, trigger delay), and sampling design (e.g., camera density, use of attractants) on the reliability of population estimates (Rowcliffe et al. 2011, Rovero et al. 2013, Burton et al. 2015, Meek et al. 2015, Hofmeester et al. 2017). However, there is very little empirical information available regarding the specific image characteristics that define image quality and how they relate to accuracy of classifications.

Observer-based factors, such as field expertise and experience with camera-trap images, have been implied as important considerations influencing accuracy of image classifications (Jacobson et al. 1997, Oliveira-Santos et al. 2010, Després-Einspenner et al. 2017). Researchers often have reported a measure of observer experience with the specific population studied and target species in general as a means to validate the reliability of classifications. Despite emphasis placed on observer-related considerations, currently there is little known regarding the role of these factors on accuracy rates. In fact, results from Snapshot Serengeti, an online citizen-science project, implied that observer expertise was not a prerequisite to obtaining accurate classifications (Swanson et al. 2015). In this study, a large pool of members of the general public were able to identify and count numerous African wildlife species at 96.6% and 90% accuracy, respectively, when compared with 'gold-standard' image-sets that had been classified by experts.

The importance of white-tailed deer (*Odocoileus virginianus;* hereafter, deer) as a game species has resulted in the development of numerous methods of assessing population size and demographics, including the use of camera traps (DeYoung 2011). Jacobson et al. (1997) introduced a method of surveying deer populations using camera traps, commonly referred to as the Individual Branched Antler Method, which has since become one of the most utilized means of estimating deer abundance, sex ratio, and fawn recruitment (Jacobson et al. 1997, McKinley et al. 2006, Curtis et al. 2009). Regardless of the specific method that is used to evaluate deer camera-trap data, observers must, at a minimum, classify images as adult male, adult female, or fawn at an acceptable level of accuracy for population estimates to be reliable. Studies where camera-trap-based population estimates have been compared with known deer abundances or estimates produced by other methods generally have indicated that camera-trap estimating methods produced sound results, which lends support to the accuracy of image classifications (Koerth et al. 1997, McKinley et al. 2006, Moore et al. 2013). However, herd demographics and survey timing are known to influence population estimates because of physical and behavioral variations in deer throughout time (Chitwood et al. 2016). For example, young (<12 weeks) fawns are not as active as their dams; therefore, fawn recruitment estimates will likely be biased low if surveys are conducted during a period when most fawns are young (DeYoung and Miller 2011). Similarly, demographics and timing of survey likely influence reliability of image classifications on account of variations in deer appearance according to age and sex. Koerth et al. (1997), for example, speculated that spike-antlered deer were misidentified as females during helicopter surveys as a result of lack of adequate visual information. These types of errors may also be prevalent when evaluating camera-trap images if spike antlers are not apparent to observers. Further, camera-trap surveys of deer frequently occur after hunting season when fawns aged 5–7 months are of relatively similar size and general appearance as adult females; visual similarities between these groups could result in high rates of misidentification (Gulsby et al. 2015).

The prominent role of camera traps in the management of deer and other wildlife species necessitates a clear understanding of errors that may be found in these data sets and their associated sources. Misidentification of animals likely contributes to overall error in many studies that utilize camera-trap data; however, relatively little is known regarding factors contributing to these errors. Greater understanding of factors that contribute to misidentifications of animals in images will allow for the refinement of current methodologies to improve accuracy in these investigations. We solicited a diverse group of respondents to participate in an online survey that first gathered pertinent individual information (i.e., profession, level of experience using game cameras for deer) and then tested their ability to classify images of known deer according to sex and age. Survey images were from both day and night and of deer at varying distances, which allowed us to also investigate the influence of image-based factors on reliability of classifications. Our objectives were to 1) estimate the accuracy of respondent classifications of deer according to age and sex, and 2) determine the relative influences of tested observer and image-based factors on reliability of classifications.

## 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 2.6-m steel fence designed to inhibit deer movements. The enclosed deer population consisted of approximately 150 individuals and comprised wild animals captured during construction and their descendants. Deer in our 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 hardwoods (oak and hickory [*Carya* spp.]) and 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. 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.

## METHODS

We used chemical immobilization to capture deer in our research facility during 8 trapping seasons (~1 Oct–15 Mar) 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 Mammalogists' 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 ($n = 34{,}728$) 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 during 27 February–7 March 2015. This camera model captured full-color 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 Feb) until spring green-up (~15 Mar–1 Apr). We selected this period of time in an effort to mirror a posthunting-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 at a point 72 cm above the center of the bait pile. We placed 3 wooden stakes 4.57 m from cameras in a radial manner behind bait pile to provide a distance reference.

We first sorted collected images containing deer according to day or night, then further organized images according to the distance of deer from the camera (i.e., inside or outside the wooden stakes). Once images had been characterized by these attributes, we classified deer in images as adult male, adult female, or fawn (i.e., 6–8-month-old deer born during the 2014 fawning season). Images of adult deer used in the survey were of ear-tagged animals for which age and sex were known, but fawn images were of both tagged and untagged animals. We did not capture adequate numbers of fawns to exclusively use images of tagged individuals in our survey; however, we feel that the relatively low abundance of untagged adults (<10%), combined with the abundance of visual information provided by 1-minute time-lapse imagery, allowed us to minimize instances where we erroneously classified untagged adults as fawns. Images that contained >1 deer were sorted according to each individual and placed into multiple categories as needed. For example, an image with an adult female deer inside the stakes and fawn outside the stakes would be included in both of the 2 appropriate categories.

We used Qualtrics® survey software (Qualtrics, Provo, UT, USA; accessed Apr 2016) 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). This survey was developed in accordance with Auburn University policies regarding research involving human subjects (Auburn Institutional Review Board protocol #15-327; approved 22 Sep 2015). We randomly selected images ($n = 55$ images containing 96 deer) from the pool of sorted images, ensuring that all classification groups were represented in our survey (Table 1). We choose the number of images for the survey in an attempt to minimize time commitments (<30 min) of respondents while maintaining adequate sample size. Ages of adult male and female deer in the selected images ranged from 1.5 to 9.5 years. Adult male images consisted mostly (28 of 32 images) of branch-antlered deer. We edited deer images using Pixlr® photo-editing software (www.pixlr.com; accessed Apr 2016) 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 (Fig. 1).

We solicited volunteers from across the U.S. for our survey with assistance from multiple partners and web-based outlets, including web requests for participation distributed by Quality Deer Management Association (www.qdma.com), ALDEER.COM (www.aldeer.com), and Deer and Deer Hunting (www.deeranddeerhunting.com), and posts on Auburn Deer Laboratory social media sites. Adults 19 years of age or older were eligible to participate in our survey. The survey was open access via an anonymous link during 19 January to 1 April 2017. We took precautions to prevent participants 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 an information letter describing the purpose of the research, participation requirements, and privacy information as required by our institution's protocol.

**Table 1.** Total number of deer images by age–sex category, time of day, and distance from the camera used in Qualtrics® white-tailed deer identification survey. The survey was conducted 19 January–1 April 2017, and participants consisted of individuals from across the United States that responded to online solicitations from Auburn University Deer Laboratory social media, Quality Deer Management Association (www.qdma.com), ALDEER.COM (www.aldeer.com), and Deer and Deer Hunting (www.deeranddeerhunting.com).

| | Day image[a] | | Night image[a] | |
|---|---|---|---|---|
| **Group** | **Inside[b]** | **Outside[b]** | **Inside[b]** | **Outside[b]** |
| Adult male[c] | 7 | 9 | 10 | 6 |
| Adult female[c] | 9 | 8 | 12 | 9 |
| Fawn[d] | 6 | 5 | 7 | 6 |

[a] Day images were full-color images taken with no flash. Night images were black-and-white images taken with infrared flash.
[b] Deer were inside or outside wooden stakes placed 4.57 m from camera.
[c] Adult refers to ≥1.5 yr of age.
[d] Fawns were ~6–8 months of age.

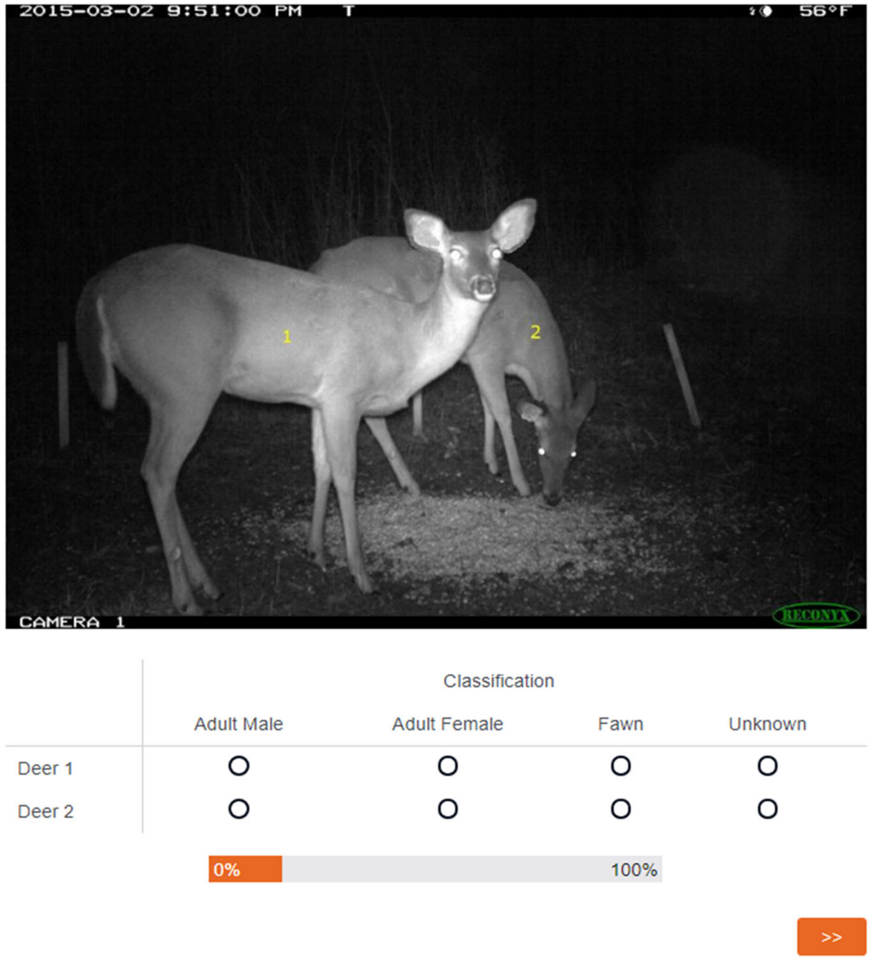


**Figure 1.** Sample question from Qualtrics® white-tailed deer identification survey conducted 19 January–1 April 2017. Survey participants consisted of individuals from across the United States that responded to online solicitations from Auburn University Deer Laboratory social media, Quality Deer Management Association (www.qdma.com), ALDEER.COM (www.aldeer.com), and Deer and Deer Hunting (www.deeranddeerhunting.com).

Respondents were then asked 2 questions focused on general demographic information, followed by 5 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 4 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 randomized 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. Upon completion of the survey, we thanked respondents and asked them submit a valid e-mail address if they were willing to take part in future research projects.

We organized responses into 2 groups for our analyses: 1) unknown responses and 2) adult male, adult female, and fawn responses. Unknown responses are neither correct nor incorrect, and we felt that they were best 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 images. We used generalized mixed-effects regression models with binomial distribution in Program R (R Core Development Team, version 3.4.1 accessed Jul 2017) to examine the influence of professional experience with wildlife (Q3; Table 2), local experience with deer (Q4), general experience with using

**Table 2.** Personal information questions used in Qualtrics® white-tailed deer identification survey followed by number of responses. The survey was conducted 19 January–1 April 2017, and participants consisted of individuals from across the United States that responded to online solicitations from Auburn University Deer Laboratory social media, Quality Deer Management Association (www.qdma.com), ALDEER.COM (www.aldeer.com), and Deer and Deer Hunting (www.deeranddeerhunting.com).

| | | | | |
|---|---|---|---|---|
| Q1– | Please indicate your gender. | | | |
| | Male - 649 | Female - 77 | | |
| Q2– | Please indicate your age classification. | | | |
| | 19–24 - 202 | 35–44 - 142 | 55–64 - 25 | |
| | 25–34 - 272 | 45–54 - 74 | 65 or older - 11 | |
| Q3a– | Do you have any professional/working experience in a wildlife-related field? | | | |
| | Yes - 158 | No - 568 | | |
| Q3b– | If Yes, how would you classify your professional/working experience in a wildlife-related field? Please select all that apply. | | | |
| | Wildlife Biology - 82 | Land Management - 73 | Hunting Guide - 50 | |
| | Outdoor Industry - 34 | Other (Please describe) = 27 | | |
| Q4– | Do you have hunting/field experience viewing white-tailed deer in Alabama or the immediately surrounding states (FL, GA, MS, TN)? | | | |
| | Yes - 421 | No - 305 | | |
| Q5– | In your opinion, what level of experience do you currently have using trail cameras to view white-tailed deer for any purpose? | | | |
| | High - 308 | Moderate - 331 | Low - 77 | None - 10 |
| Q6a– | 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 - 235 | No - 491 | | |
| Q6b– | If Yes, how many of these kind of trail camera surveys have you completed? | | | |
| | 4 or fewer - 162 | 5–10 - 49 | 11 or more - 24 | |
| Q7– | Please provide any other information regarding your experience with white-tailed deer that you feel would further clarify responses and help us place you in the most appropriate subject group. | | | |

trail cameras to view deer (Q5), advanced experience using trail cameras to view deer (Q6), classification of the "known" deer image (i.e., adult male, adult female, fawn), distance of deer from camera (i.e., inside or outside wooden stakes), and type of image (i.e., day or night) on accuracy of survey responses. 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 (i.e., Q1, Q2, Q3a, Q4, Q5, and Q6a) 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).

The goal of our second analysis was to examine the specific types of error associated with incorrect responses. We first restricted our data to include only incorrect responses, then organized this data set into 3 subgroups according to our classification of the deer image (i.e., adult male, adult female, fawn). We used generalized mixed-effects regression models with binomial distribution in Program R to model each of these subgroups with a conditional response of one of the 2 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 probability of occurrence for the 2 possible incorrect responses respective to each of the 3 deer classification groups.

Our final analysis focused on examining factors contributing to unknown responses. We used our full data set to create a conditional variable based on unknown responses for this analysis. We used generalized mixed-effects regression models with binomial distribution in Program R to examine the influence of professional experience with wildlife (Q3), local experience with deer (Q4), general experience with using trail cameras to view deer, (Q5), advanced experience using trail cameras to view deer (Q6), our classification of the deer image, distance of deer from camera, and type of image on the conditional unknown response variable. Random effects terms for respondent ID and deer image ID were included to account for variation associated with these effects.

# RESULTS

We had 726 respondents complete our survey during the 10-week study period. Respondents were primarily male and from a wide range of age groups (Table 2). Most (78%) did not have professional experience in a wildlife-related field. Respondents that did have professional experience primarily identified as wildlife biologists, and those identifying as OTHER included taxidermist ($n = 2$), outdoor journalist ($n = 2$), agricultural educator ($n = 2$), deer farmer ($n = 4$), wildlife trapper ($n = 1$), game warden ($n = 3$), fishing guide ($n = 2$), and wildlife rehabilitator ($n = 1$). Our pool of respondents mostly (58%) consisted of individuals with some experience viewing deer in Alabama or the immediately surrounding states. Most (88%) indicated they had HIGH or MODERATE level of experience using trail cameras to view deer for any purpose; however, many (68%) had never conducted a camera survey to estimate deer population information. The majority (69%) of individuals with experience conducting camera surveys for deer indicated that they had completed ≤4 surveys.

Estimates of collinearity among predictors related to respondent personal information were low (VIF: Table 2; Q1 = 1.07, Q2 = 1.06, Q3a = 1.08, Q4 = 1.02, Q5 = 1.21,

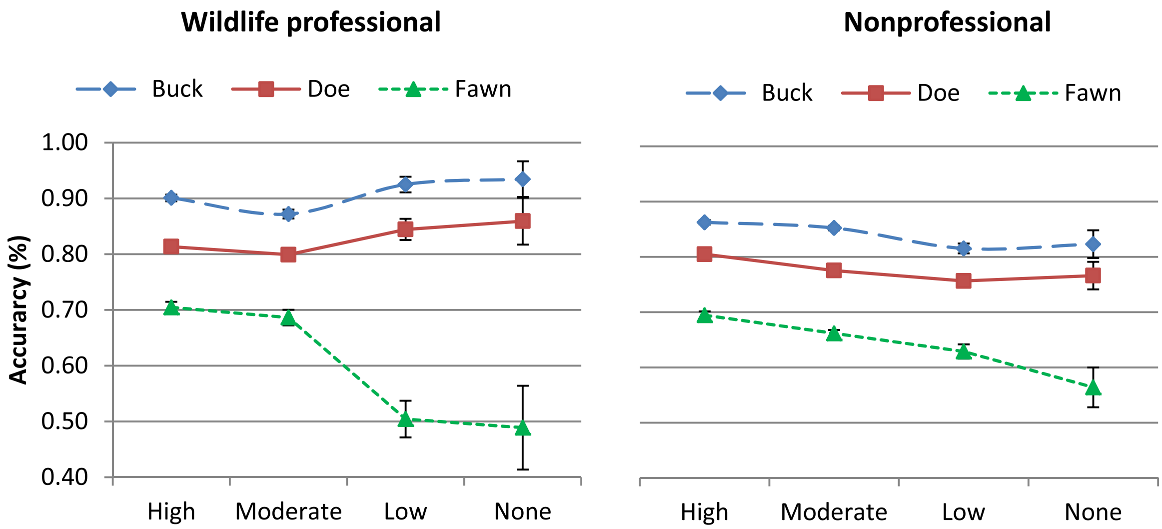


**Figure 2.** Mean (95% CL) accuracy of white-tailed deer image classifications among Qualtrics® white-tailed deer identification survey respondents according to deer classification group, professional experience in wildlife-related field, and level of experience using trail cameras to view deer for any purpose. 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.' The survey was conducted 19 January–1 April 2017, and participants consisted of individuals from across the United States that responded to online solicitations from Auburn University Deer Laboratory social media, Quality Deer Management Association (www.qdma.com), ALDEER.COM (www.aldeer.com), and Deer and Deer Hunting (www.deeranddeerhunting.com).

Q6a = 1.16). Our analysis suggested that accuracy of deer classifications was associated positively with both professional experience in a wildlife-related field and experience using trail cameras to view deer for any purpose (Fig. 2). Respondents with professional experience were 1.22 (95% CL = 1.11–1.34; $P < 0.001$) times more likely to accurately classify deer images than nonprofessionals, with wildlife biologists primarily accounting for the positive effect: Wildlife Biologist, 1.16 (95% CL = 1.01–1.33; $P = 0.04$); Forestry, 1.05 (95% CL = −1.18–1.30; $P = 0.66$); Land Management, 1.07 (95% CL = −1.11–1.28; $P = 0.43$); Hunting Guide, −1.04 (95% CL = −1.25–1.16; $P = 0.69$); Outdoor Industry, 1.07 (95% CL = −1.14–1.30; $P = 0.51$); and Other 1.09 (95% CL = −1.13–1.34; $P = 0.40$). Respondents with HIGH experience using trail cameras to view deer were 1.21 (95% CL = 1.09–1.32; $P < 0.001$), 1.49 (95% CL = 1.30–1.70; $P < 0.001$), and 1.56 (95% CL = 1.13–2.16; $P = 0.01$) times more likely to correctly classify deer than those with MODERATE, LOW, and NONE experience, respectively. Those with MODERATE experience were 1.22 (95% CL = 1.08–1.39; $P < 0.001$) times more likely to accurately classify deer than those indicating LOW. Accuracy of classifications was similar between those with MODERATE and NONE experience (Exp(β) = 1.29 [95% CL = −1.07–1.78; $P = 0.12$]) and LOW and NONE experience (Exp(β) = 1.05 [95% CL = −1.33–1.47; $P = 0.78$]). We did not detect a significant relationship between experience viewing deer in Alabama or surrounding states and classification accuracy (Exp(β) = −1.01 [95% CL = −1.10–1.06; $P = 0.71$]), or experience conducting camera surveys to estimate deer population characteristics and accuracy (Exp(β) = 1.02 [95% CL = −1.07–1.11; $P = 0.68$]).

We found that accuracy of classifications was related to the age–sex of deer (Fig. 2). Images of adult males were 4.73 (95% CL = 1.96–11.43; $P < 0.001$) times more likely to be correctly classified than adult female images and 11.48 (95% CL = 4.27–30.87; $P < 0.001$) times more likely to be correctly classified than fawn images. We did not detect a difference between classification accuracy of adult female and fawn images (Exp(β) = 2.42 [95% CL = −1.06–6.21; $P = 0.06$]). Adult male images that were incorrectly classified were categorized as adult female and fawn at similar rates (Exp(β) = −1.56 [95% CL = −4.23–1.75; $P = 0.38$]). Adult female images that were incorrectly classified were 28.92 (95% CL = −15.98–52.33; $P < 0.001$) times less likely to be classified as adult male than fawn. Fawn images that were incorrectly classified were 11.72 (95% CL = 6.74–20.35; $P = 0.38$) times more likely to be classified as adult female than adult male. Our results indicated that accuracy of classifications did not differ between day versus night images (Exp(β) = 1.14 [95% CL = −1.88–2.42; $P = 0.74$]) or images of deer inside versus outside the wooden stakes (Exp(β) = 1.25 [95% CL = −1.72–2.66; $P = 0.56$]).

Unknown responses comprised 7% of the 68,244 total responses to our survey questions. Our results indicated that unknown responses were not related to any of the investigated factors related to the individual participant. Wildlife professional and nonprofessionals provided similar numbers of unknown responses (Exp(β) = −1.01 (95% CL = −1.40–1.36; $P = 0.93$). Individuals with HIGH level of experience using trail cameras to view deer provided similar numbers of unknown responses as those with MODERATE (Exp(β) = 1.05 [95% CL = −1.26–1.41; $P = 0.71$]), LOW (Exp(β) = 1.20 [95% CL = −1.32–1.90; $P = 0.43$]), and NONE (Exp(β) = −1.56 [95% CL = −4.90–2.02; $P = 0.44$]). Those with MODERATE experience provided similar numbers of unknown responses as those with LOW

($\text{Exp}(\beta) = 1.14$ [95% CL = −1.37–1.77; $P = 0.57$]) and NONE ($\text{Exp}(\beta) = -1.64$ [95% CL = −5.18–1.92; $P = 0.39$]), and unknown responses for those with LOW experience were similar to NONE ($\text{Exp}(\beta) = -1.87$ [95% CL = −6.22–1.78; $P = 0.30$]). Respondents with experience viewing deer in Alabama and surrounding states provided similar numbers of unknown responses as those without local experience ($\text{Exp}(\beta) = -1.21$ [95% CL = −1.57–1.08; $P = 0.16$]), and unknown responses were not different between respondents with and without experience using cameras to conduct deer population surveys ($\text{Exp}(\beta) = 1.32$ [95% CL = −1.02–1.77; $P = 0.06$]).

We found that unknown response rate was related to our age–sex classification group of the deer image. Respondents were 2.36 (95% CL = 1.03–5.39; $P = 0.04$) and 3.49 (95% CL = 1.40–8.82; $P = 0.01$) times more likely to give an unknown response for adult female and fawn images, respectively, than adult male images. Unknown response rates were similar for adult female and fawn images ($\text{Exp}(\beta) = 1.49$ [95% CL = −1.62–3.59; $P = 0.36$]). Our results also indicated that unknown response rate was related to the type (i.e., day vs. night) of image as well as distance of deer from the camera. We observed that night images were 2.15 (95% CL = 1.06–4.37; $P = 0.03$) times more likely to be classified as unknown than day images. Deer outside the wooden stakes were 14.67 (95% CL = 7.22–29.80; $P < 0.001$) times more likely to be considered unknown than those inside the stakes.

## DISCUSSION

Our findings provided evidence that camera-trap images of deer were misidentified by volunteers that attempted to place individual deer into broad age–sex categories. Although experience played a significant role in the accuracy of deer classifications, the factor primarily responsible for misidentification error was the age–sex of the deer in the image. Adult males (hereafter, bucks) in our survey all had visible antlers. The presence of this conspicuous physical trait allowed participants to classify buck images with a greater level of accuracy. In contrast, adult females (hereafter, does) and fawns lacked a definitive trait and were identified less accurately using subjective criteria, such as relative size and body proportions. Does and fawns in our survey were from a range of ages, and age-related variations in appearances likely contributed to the observed classification errors among these groups. Some of the fawns in our survey were younger, smaller, and still had visible spots on their coats (a trait usually lost by 4 months) while others were older, larger, and lacked spots (Ditchkoff 2011). Does were represented by animals 1.5–12.5 years of age, and noticeable differences in body size can exist between yearling (~1.5 yr) and older animals in this population (Neuman et al. 2016). Similarities in appearances between older fawns and younger does likely helped to confound visual criteria used to classify animals and contributed to misclassifications. Further, the presence of spots on some fawns may have led volunteers to believe that all animals in this category should possess similar traits, resulting in the misclassification of unspotted fawns as does.

Camera surveys for deer typically occur during 1 of 2 periods (i.e., prehunting and posthunting season); each of these periods provides benefits over the alternative in terms of the attained information (McCoy et al. 2011). Prehunting season surveys are useful in evaluating the standing crop of bucks and adult sex ratio, but reliable recruitment estimates are difficult to obtain from these surveys because many fawns are relatively young and may not visit camera sites at rates similar to does (McKinley et al. 2006, Chitwood et al. 2016). Posthunting season surveys can provide information regarding the proportion of bucks that survived the hunting season, adult sex ratio estimates, and presumably more stable estimates of fawn recruitment. Additionally, adult deer use of baited sites reportedly is most similar between sexes during the posthunting season period, and camera surveys conducted during this time period may provide the least biased population estimates when bait is used as an attractant (Newbolt et al. 2017). However, previous investigations have made reference to the potential for classification errors among does and fawns during posthunting season surveys; they have provided interobserver reliability estimates of sex ratio and recruitment that support these type of errors (Jacobson et al. 1997). For example, Jacobson et al. (1997) reported fawn recruitment estimates from images obtained during the postseason period that ranged from 0.70 to 1.03 ($n = 5$ observers, $\bar{x} = 0.83$, SE = 0.06). Our results further demonstrate that does and fawns were misclassified when utilizing postseason data, which raises important questions concerning timing of surveys and accuracy of estimates. Although postseason surveys may be the optimal time to conduct camera surveys based upon deer movement patterns and behavior in relation to bait, misclassification of does and fawns may negatively affect accuracy of estimates. The seemingly optimal time to conduct surveys that minimize misclassification error would be when fawns are young enough to be visually distinct, but have adopted behaviors similar to adults. These periods typically coincide with hunting season, which creates difficulties, especially if bait is used at camera sites. The net effect of doe or fawn misclassifications on population estimates from postseason surveys is difficult to ascertain from our data because classification error rates were not found to differ between these groups. However, potential implications of the observed classification errors warrant consideration when selecting camera survey periods for deer and further investigation.

Variation in accuracy of classifications across buck, doe, and fawn images indicates that population demographics may play a role in the reliability of estimates obtained from camera surveys. Deer populations consisting of greater numbers of yearling does may be subject to greater rates of misclassification error than those primarily consisting of older does because of visual similarities between young does and fawns. Further, antler size may influence classification accuracy among bucks as larger antlers are logically easier to see than smaller antlers. Our data set was not adequate to

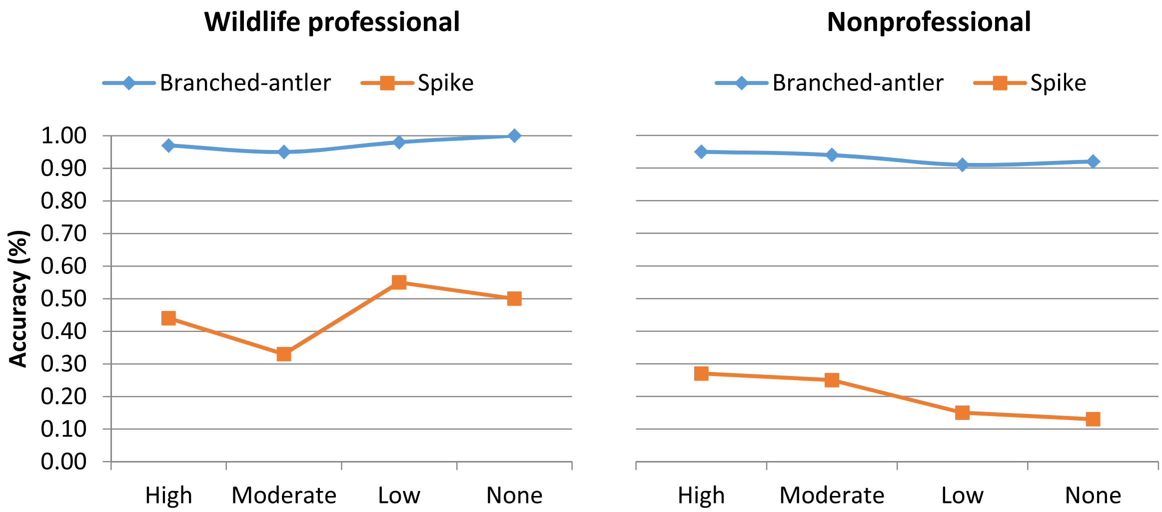


**Figure 3.** Mean (95% CL) accuracy of white-tailed deer image classifications among Qualtrics® white-tailed deer identification survey respondents according to spike and branched-antler males (bucks), professional experience in wildlife-related field, and level of experience using trail cameras to view deer for any purpose. The survey was conducted 19 January–1 April 2017, and participants consisted of individuals from across the United States that responded to online solicitations from Auburn University Deer Laboratory social media, Quality Deer Management Association (www.qdma.com), ALDEER.COM (www.aldeer.com), and Deer and Deer Hunting (www.deeranddeerhunting.com).

specifically address relationships between antler size and accuracy; however, we observed a noteworthy trend in our data where spike-antlered males ($n = 4$) were misidentified more frequently than those with branched antlers (Fig. 3). Spike antlers were, in our opinion, clearly visible in 3 of 4 images, and spikes were most frequently ($\bar{x} = 75\%$ of incorrect responses, SE = 4%) misidentified as fawns. We believe the misclassification of spike-antlered yearling bucks may have been the result of misinformation about antler growth patterns rather than a visual inaccuracy. Few fawns in our facility achieve significant antler growth during their first year, and it is possible that notifying participants of these patterns may have greatly improved accuracy of spike buck classifications.

Unknown responses made up a comparatively low proportion of total responses in our survey, which suggests they may not be an extremely meaningful consideration in deer camera surveys. However, minimizing any level of unknown responses is beneficial because they serve no analytical purpose and increase image processing time. Our results indicated that deer distance from camera and image type (i.e., day vs. night) were not important predictors of classification accuracy, but night images and greater distances of deer from cameras were associated positively with unknown responses. The image type is largely dictated by animal behavior, which leaves little room for improvement outside of technological advancements that improve night image quality. Deer distance from the camera, however, can be easily controlled by placing references, such as our wooden stakes, at a selected distance. Simply eliminating deer outside of the distance references may be an effective way to reduce numbers of unknown responses and image classification effort. Further, placing distance references in the field of view provides the added benefit of allowing for the computation of effective detection distance, which is essential for some population estimating methods such as the Random Encounter Model (Rowcliffe et al. 2008, Hofmeester et al. 2017). We chose the measurement for our distance references based upon familiarity with our camera specifications; however, appropriate distance references will vary according to camera model and wildlife species.

The positive effects of being a wildlife biologist and experience viewing deer in trail camera images suggests that observer-based factors are important to accuracy of animal classifications, even under circumstances that do not require individual identification. Consequently, the overall reliability of information from camera surveys will likely vary according to the experience level of those involved in processing images in situations where animals do not exhibit highly discernable differences in appearances. Although the observed effect size of the significant experience-related factors was relatively small, any significant source of error is important because low rates of misidentification are known to bias estimates (Stevick et al. 2001, Morrison et al. 2011). There currently are inconsistencies and analytical problems associated with methods used to collect and analyze camera-trap data, and researchers have made calls for the refinement of methodologies to address these issues (Foster and Harmsen 2012, Burton et al. 2015). Studies that utilize camera-trap data often provide little information regarding the experience level of those responsible for classifying images, let alone a measure of assumed rate of accuracy associated with each observer. In light of our findings and current shortcomings in camera-trap methods, we suggest that greater emphasis should be placed on providing descriptive information regarding observer experience with the target species and, when possible, analytical information related to the assumed rate of classification accuracy as means for validating results.

We feel that it is important to reiterate that the images in our survey were presented to respondents in a random order, which is obviously not reflective of camera-trap data collected under normal field conditions. Our research required us to present participants a relatively concise set of images, representing various specific classifications of deer, and it was not possible to meet these prerequisites using a single set of chronologically ordered images. Consideration of a chronological series rather than isolated deer images could provide observers additional information that might aid in deer classifications. Consecutive images of individual deer at varying angles would likely influence an observer's decision-making process and possibly accuracy of classifications; however, the myriad of possible time-delay settings across numerous brands of trail cameras, as well as variations in deer behavior in relation to camera traps, make it impossible to determine the exact information that is presented to observers under all scenarios. Our specific reported rates of classification accuracy may not represent those found when evaluating deer camera-trap data that are chronologically ordered, but we believe that our research identifies factors that are important determinants of accuracy, regardless of the interval of time between images.

## MANAGEMENT IMPLICATIONS

Our study demonstrates that observer misclassification of camera-trap data may be an important source of error, not only under conditions where individual animals must be identified, but also in situations where animals from broad age–sex categories lack definitive visual differences. Visual differences between classification groups will vary across seasons for many species because of growth and developmental patterns, and when possible, researchers should strongly consider conducting surveys during periods that maximize the visual differences between animal groups in an effort to reduce error. Further, population demographics should be considered when evaluating the overall reliability of population estimates from camera-trap investigations when classification accuracy varies between animal groups, as is the case for deer. Experience contributed to increased accuracy in our study; we suggest that individuals involved in processing camera-trap data should develop and use species-specific training tools as a means to improve consistency among observers and reduce misclassifications. Although classification errors may be the result of inadequate visual information, misidentifications also may be the result of a simple lack of knowledge regarding the target animal (e.g., adult spike bucks with clearly visible antlers that were misidentified as fawns in our study), which, in theory, would be easy to overcome with training. Image quality played a minimal role in determining the accuracy of classifications in our study, but limiting detection distance with simple point references may reduce occurrences of unknown images and associated wasted processing effort.

## ACKNOWLEDGMENTS

We thank Code Blue Scents–EBSCO Industries and other organizations and individuals who provided private financial support for projects at the Deer Laboratory (complete list of donors available at http://wp.auburn.edu/deerlab/acknowledgements/). We also thank Quality Deer Management Association, ALDEER.com, and *Deer & Deer Hunting Magazine* for their assistance in reaching study participants. Great thanks to the many volunteers who took our survey and made this research possible. Finally, thanks to the Associate Editors of this journal and the *Journal of Wildlife Management,* as well as reviewers, for their valuable insight and contribution to this manuscript.

## LITERATURE CITED

Burton, A. C., E. Neilson, D. Moreira, A. Ladle, R. Steenweg, J. T. Fisher, E. Bayne, and S. Boutin. 2015. REVIEW: wildlife camera trapping: a review and recommendations for linking surveys to ecological processes. Journal of Applied Ecology 52:675–685.

Carbone, C., S. Christie, K. Conforti, T. Coulson, N. Franklin, J. R. Ginsberg, M. Griffiths, J. Holden, K. Kawanishi, M. Kinnaird, R. Laidlaw, A. Lynam, D. W. Macdonald, D. Martyr, C. McDougal, L. Nath, T. O'Brien, J. Seidensticker, D. J. L. Smith, M. Sunquist, R. Tilson, and W. N. Wan Shahruddin. 2001. The use of photographic rates to estimate densities of tigers and other cryptic mammals. Animal Conservation 4:75–79.

Chitwood, M. C., M. A. Lashley, J. C. Kilgo, M. J. Cherry, L. M. Conner, M. Vukovich, H. S. Ray, C. Ruth, R. J. Warren, C. S. DePerno, and C. E. Moorman. 2016. Are camera surveys useful for assessing recruitment in white-tailed deer? Wildlife Biology:wlb.00178.

Curtis, P. D., B. Boldgiv, P. M. Mattison, and J. R. Boulanger. 2009. Estimating deer abundance in suburban areas with infrared-triggered cameras. Human–Wildlife Conflicts 3:116–128.

Cutler, T. L., and D. E. Swann. 1999. Using remote photography in wildlife ecology: a review. Wildlife Society Bulletin 27:571–581.

Després-Einspenner, M. L., E. J. Howe, P. Drapeau, and H. S. Kuhl. 2017. An empirical evaluation of camera trapping and spatially explicit capture-recapture models for estimating chimpanzee density. American Journal of Primatology 79:e22647.

DeYoung, C. A. 2011. Population dynamics. Pages 147–180 *in* D. G. Hewitt, editor. Biology and management of white-tailed deer. CRC Press, Boca Raton, Florida, USA.

DeYoung, R. W., and K. V. Miller. 2011. White-tailed deer behavior. Pages 311–354 *in* D. G. Hewitt, editor. Biology and management of white-tailed deer. CRC Press, Boca Raton, Florida, USA.

Ditchkoff, S. S. 2011. Anatomy and physiology. Pages 43–74 *in* D. G. Hewitt, editor. Biology and management of white-tailed deer. CRC Press, Boca Raton, Florida, USA.

Foster, R. J., and B. J. Harmsen. 2012. A critique of density estimation from camera-trap data. Journal of Wildlife Management 76:224–236.

Gulsby, W. D., C. H. Killmaster, J. W. Bowers, J. D. Kelly, B. N. Sacks, M. J. Statham, and K. V. Miller. 2015. White-tailed deer fawn recruitment before and after experimental coyote removals in central Georgia. Wildlife Society Bulletin 39:248–255.

Hofmeester, T. R., J. M. Rowcliffe, and P. A. Jansen. 2017. A simple method for estimating the effective detection distance of camera traps. Remote Sensing in Ecology and Conservation 3:81–89.

Hosmer, D. W., S. Lemeshow, and R. X. Sturdivant. 2013. Applied logistic regression, John Wiley and Sons, Hoboken, New Jersey, USA.

Jacobson, H. A., J. C. Kroll, R. W. Browning, B. H. Koerth, and M. H. Conway. 1997. Infrared-triggered cameras for censusing white-tailed deer. Wildlife Society Bulletin 25:547–556.

Kelly, M. J., A. J. Noss, M. S. Di Bitetti, L. Maffei, R. L. Arispe, A. Paviolo, C. D. De Angelo, and Y. E. Di Blanco. 2008. Estimating puma densities from camera trapping across three study sites: Bolivia, Argentina, and Belize. Journal of Mammalogy 89:408–418.

Koerth, B. H., C. D. McKown, and J. C. Kroll. 1997. Infrared-triggered camera versus helicopter counts of white-tailed deer. Wildlife Society Bulletin 25:557–562.

McCoy, J. C., S. S. Ditchkoff, and T. D. Steury. 2011. Bias associated with baited camera sites for assessing population characteristics of deer. Journal of Wildlife Management 75:472–477.

McKinley, W. T., S. Demarais, K. L. Gee, and H. A. Jacobson. 2006. Accuracy of the camera technique for estimating white-tailed deer population characteristics. Proceedings of the Southeastern Association of Fish and Wildlife Agencies 60:83–88.

Meek, P. D., G.-A. Ballard, and P. J. S. Fleming. 2015. The pitfalls of wildlife camera trapping as a survey tool in Australia. Australian Mammalogy 37:13–22.

Moore, M. T., A. M. Foley, C. A. DeYoung, D. G. Hewitt, T. E. Fulbright, and D. A. Draeger. 2013. Evaluation of population estimates of white-tailed deer from camera survey. Journal of the Southeastern Association of Fish and Wildlife Agencies 1:127–132.

Morrison, T. A., J. Yoshizaki, J. D. Nichols, and D. T. Bolger. 2011. Estimating survival in photographic capture–recapture studies: overcoming misidentification error. Methods in Ecology and Evolution 2:454–463.

Neuman, T. J., C. H. Newbolt, S. S. Ditchkoff, and T. D. Steury. 2016. Microsatellites reveal plasticity in reproductive success of white-tailed deer. Journal of Mammalogy 97:1441–1450.

Newbolt, C. H., S. Rankin, and S. S. Ditchkoff. 2017. Temporal and sex-related differences in use of baited sites by white-tailed deer. Journal of the Southeastern Association of Fish and Wildlife Agencies 4:109–114.

Oliveira-Santos, L. G. R., C. A. Zucco, P. C. Antunes, and P. G. Crawshaw 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.

Rovero, F., F. Zimmermann, D. Berzi, and P. Meek. 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.

Rowcliffe, J. M., C. Chris, P. A. Jansen, R. Kays, and B. Kranstauber. 2011. Quantifying the sensitivity of camera traps: an adapted distance sampling approach. Methods in Ecology and Evolution 2:464–476.

Rowcliffe, J. M., J. Field, S. T. Turvey, and C. Carbone. 2008. Estimating animal density using camera traps without the need for individual recognition. Journal of Applied Ecology 45:1228–1236.

Sikes, R. S., and W. L. Gannon. 2011. Guidelines of the American Society of Mammalogists for the use of wild mammals in research. Journal of Mammalogy 92:235–253.

Steenweg, R., M. Hebblewhite, R. Kays, J. Ahumada, J. T. Fisher, C. Burton, S. E. Townsend, C. Carbone, J. M. Rowcliffe, J. Whittington, J. Brodie, J. A. Royle, A. Switalski, A. P. Clevenger, N. Heim, and L. N. Rich. 2017. Scaling-up camera traps: monitoring the planet's biodiversity with networks of remote sensors. Frontiers in Ecology and the Environment 15:26–34.

Stevick, P. T., P. J. Palsbøll, T. D. Smith, M. V. Bravington, and P. S. Hammond. 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.

Swanson, A., M. Kosmala, C. Lintott, R. Simpson, A. Smith, and C. Packer. 2015. Snapshot Serengeti, high-frequency annotated camera trap images of 40 mammalian species in an African savanna. Scientific Data 2:150026.

Yoshizaki, J., K. H. Pollock, C. Brownie, and R. A. Webster. 2009. Modeling misidentification errors in capture–recapture studies using photographic identification of evolving marks. Ecology 90:3–9.

*Associate Editor: Jacques.*