Identification of African antelope species: using thermographic videos to test the efficacy of real-time thermography

Anne E. Goodenough[1*], William S. Carpenter[1], Lynne MacTavish[2],

Charles Theron[2], Matthew Delbridge[2] and Adam G. Hart[1].

1. School of Natural and Social Science, University of Gloucestershire, Cheltenham, GL50 4AZ, UK
2. Mankwe Wildlife Reserve, Mogwase, Northwest Province, South Africa

*Corresponding author

Running head: Real-time thermography in ecological surveys

Real-time thermography using the live-view function of a thermal camera has considerable potential to improve surveys of nocturnal wildlife relative to traditional spotlighting, while also decreasing disturbance. However, ability to identify species accurately is paramount. We use video as a proxy for real-time thermography to test African antelope identification accuracy among 34 observers of differing experience. Overall accuracy was 41% but there were substantial species-specific differences (e.g. wildebeest (*Connochaetes taurinus*) = 81%; reedbuck (*Redunca arundinum*) = 12%). Observer experience was significantly positively related to accuracy (inexperienced=30%; expert=61%) with experienced observers being better able to use subtle movement and behavioural information to identify animals. However, the effect of experience was inconsistent between species: even experts found some species challenging (e.g. waterbuck (*Kobus ellipsiprymnus*) where coat patterning was invisible thermographically). Self-assessed confidence did not correlate with accuracy. Observers who were good at identifying species were also good at assessing group size. We conclude that real-time thermography is not a “magic bullet” and understanding of species-specific effectiveness is vital. However, for some species and some groups of observers, accuracy can be extremely high (e.g. 100% for expert observers viewing wildebeest). Tailored training is essential for real-time thermography to be a reliable field technique.

Keywords: Antelope identification; real-time thermography; species identification; survey tool; wildlife monitoring

Introduction

Technological imaging approaches are of considerable, and increasing, benefit to ecologists and conservation practitioners seeking accurate and precise animal census data. Such technologies include cameras mounted on unmanned aerial vehicles (UAVs: Anderson & Gaston, 2013; Bryson *et al*., 2014), camera traps (O'Connell, Nichols & Karanth, 2010; Welbourne, 2013), and the use of satellite imagery (Fretwell & Trathan, 2009). One particularly useful image-based approach is infrared thermography, whereby a sensor sensitive to infra-red radiation is used to produce an image in the same way that a conventional digital camera uses visible light (McCafferty, 2007; Cilulko *et al*., 2013). Thermography has been used for a range of different ecological surveys, including the detection of cryptic and burrowing species (Boonstra *et al*., 1994) and the location and surveying of nocturnal mammals and birds (e.g. Dunn, Donnelly & Krausmann 2002; Ganow, Caire & Matlock, 2015; McCafferty *et al*., 1998; Zehnder *et al*., 2001).

The ability to locate species at night with minimal disturbance is one of the major advantages of thermography (Lavers *et al*., 2005). Indeed, thermography has been shown to be almost twice as effective for locating white-tailed deer (*Odocoileus virginianus*; Zimmermann, 1780) as using a spotlight and is also less invasive (Collier *et al*., 2007). However, for it to be useful as a general surveying tool, thermography needs to go beyond simple detection to become an accurate and precise method by which individuals can be identified to species, counted, and even categorized in terms of sex and/or age. Achieving such detailed information can be challenging, especially in complex environments where thermography is hampered by thick vegetation and varied topography, as well as the presence of heat-reflective surfaces (e.g. water) and heat-retaining structures (e.g. rocks) (Graves, Bellis & Knuth, 1972; Naugle, Jenks & Kernohan, 1996). Moreover, it is not always possible to distinguish between co-occurring species using thermography because distinguishing features, such as colour or coat pattern, may not be apparent on thermal images (Franke *et al*., 2012). Such challenges mean that the usefulness of thermography as an ecological survey tool in any specific environment, or for any given species, cannot be assumed without empirical testing, especially given the effect of relatively small levels of species misidentification on ecological inference (Costa *et al*., 2015).

Surveying African antelope often relies on direct observation rather than or in addition to, use of indirect evidence such as footprints. Direct observation is especially important when species community structure or population dynamics of particular species is important, when accurate species identification and robust estimates of abundance are vital. During daytime conditions, driven or walked large mammal transects can often be used to good effect (Plumptre, 2000; Ogutu *et al*., 2006). However, at night, visual surveys are harder to undertake and frequently reply on spotlighting, which can cause disturbance by eliciting flight responses and inducing night-blindness in animals. Moreover, it can still be very difficult to accurately identify animals within spotlighted areas, even when using powerful lights and binoculars. Recent research by Goodenough *et al* (2017) showed that thermography is a useful tool in surveying African antelope, with 38% of identifications being correct to species level and 50% being correct to subfamily level. The technique was particularly useful at night or where visual observation was compromised but the thermal signature remained clear (e.g. animals in long grass). Identification accuracy was positively correlated to observer confidence, skill and experience as well as the number of animals and negatively related to the distance at which the image was taken. These findings were based on thermal photographs (i.e. still images). In actual field conditions, however, thermography is often used in real-time, whereby moving footage on the screen of the thermal camera using "live view" is simply visually inspected rather than photographs/videos being recorded. In this way, using a thermal camera becomes a transitory survey tool to help the observer interpret what is happening in real-time, similar to using binoculars. As noted by Goodenough *et al*. (2017), a useful avenue for future research would be to assess observer accuracy in species identification using moving thermographic footage, when it is possible to assess locomotion and behaviour, rather than still photographs. Here, we use video clips as a proxy for real-time thermography to test whether observers of different experience levels are able to identify various African antelopes to species level, and to determine the number and sex of animals present. We also examine the role of observer parameters (experience, confidence and skill) and ecological parameters (distance, group size, presence of males) on data accuracy.

Methods

Study site and study species

This study was performed at Mankwe Wildlife Reserve, Northwest Province, South Africa. The site supports >40 species of large mammal within a bushveld matrix of grasses and *Vachellia* scrub. Nine antelope species were included in the study: (1) greater kudu (*Tragelaphus strepsiceros*; Pallas 1766), (2) tsessebe (*Damaliscus lunatus*; Burchell, 1823), (3) impala (*Aepyceros melampus*; Lichtenstein, 1812), (4) blue wildebeest (*Connochaetes taurinus*; Burchell, 1823), (5) waterbuck (*Kobus ellipsiprymnus*; Ogilby, 1833), (6) common reedbuck (*Redunca arundinum*; Boddaert, 1785), (7) eland (*Tragelaphus oryx*; Pallas, 1766), (8) gemsbok (*Oryx gazella*; Linnaeus, 1758), and (9) sable (*Hippotragus niger*; Harris, 1838).

Thermographic videos

The study aimed to assess the value of real-time thermography for surveying African antelope (i.e. visual inspection of images on the screen of a thermal imaging camera while surveying is taking place as a real-time survey tool). In practice, however, it was impossible for multiple (>2) participants to view simultaneously any one camera in real-time, primarily because of the relatively small screen display size of handheld thermal imaging cameras. Thus, a thermographic video camera (Flir T620 with a 30 Hz infrared detector) was used to record a series of video clips, which could ultimately be shown to participants under identical conditions, as a proxy for using the technology in real-time.

The camera had a measurable temperature range of -40 to +650°C with a calibrated thermal sensitivity of 0.04°C. Previous research has demonstrated this can detect mammals at distances up to ≥300m in African bushveld (Hart *et al*., 2015). The camera was set to display in "grey" colour palette, a setting that optimized the clarity of moving animals against an often-complex thermal background (e.g. including rocks or bodies of water) when compared to alternative palettes such as "lava". For the purposes of this study, videos were taken during both the day (when the species being filmed, abundance, and the presence/absence of males, was easy to determine) and at night (only when the species/sex information could be determined unequivocally). The camera was set with an automatic temperature range with a dynamic palette whereby the colours (shades of grey in this case) were adjusted automatically to maximise the difference between hot and cold features. Thus almost all videos had excellent contrast between the animal(s) and the background landscape; any that did not were discarded and not included in the study regardless of whether they had been take at night or in the day. Although there was considerable variation between the videos based on species, distance, and surrounding landscape features, there were no obvious visual differences between those shot diurnally versus nocturnally. Videos were recorded over a 10-day period in May 2017 at the maximum possible resolution of 640*480 pixels using the maximum frame-rate of 30 frames per second.

All videos were shot from the back of an open vehicle, typically in open grassland with a sward height of up to 1m such that the antelope individual(s) were shown clearly and without substantial interference from vegetation. In approximately 15% of cases, individuals were partly concealed visually by tall vegetation but their thermal signature remained clear as per Gill, Thomas & Stocker (1997). The distance from the camera was determined during filming using a Bushnell v3 Laser Rangefinder (Bushnell Performance Optics, Surrey, UK); in cases where there were several individuals at different distances or where individual(s) moved towards or away from the camera, the nearest distance was taken. The distance to the animals varied between 10m and 150m (mean = 76m; s.d. = 43m; n = 60) and data were normally distributed (Kolmogorov-Smirnov test D = 0.130, d.f. = 60, p = 0.140).

Videos were edited to a fixed duration of 8 seconds using Microsoft Movie Maker version 16.4. Each resultant video clip involved animal activity (e.g. locomotion, head movement, or social interaction) and was given a unique reference. In total, there was 60 video clips, each of which showed a single antelope species (9 species; 6-8 videos of each). Figure 1 shows six still images, each taken from a different video clip (videos available as Supplementary Information). The number of individuals present in the video clips ranged from 1 to 12; 33 videos (55%) included at least one male while 17 videos (28%) were of female(s) only. In the remaining 10 videos (17%), it was not possible to determine whether a male was present. This was typically for species such as tsessebe and gemsbok, where both males and females have horns, and where the videos were shot at >80m.

a)

Impala at 15m, one individual, male present

b)

Sable at 20m, one individual, male present

c)

Waterbuck at 45m, two individuals, no male present

d)

Wildebeest at 80m, 9 individuals, male present

e)

Eland at 60m, one individual, no male present

f)

Tsessebe at 76m, one individual, no male present

Figure 1

Volunteer observers and video screening

Thirty-four volunteer observers (19 female and 15 male) took part in this study. Ages ranged from 20 to 76. Observers were allocated to one of three groups based on their experience. Observers with limited experience (10 days' fieldwork at the study site; n=21) formed the inexperienced group. Observers with some previous experience of working with African antelope (6-12 weeks' duration within the previous 2 years plus 10 days' experience of fieldwork immediately prior to this study; n=8) formed the intermediate group. Observers that had worked on animal reserves in Africa for >24 months prior to the study (n=5) formed the experienced group. Observers were also asked to rate their confidence in identifying African antelope species on a scale of 1 (very low confidence) to 10 (very high confidence), while their skill was tested objectively by asking observers to view a series of 28 colour photographic images of African antelope and identify the species shown. Importantly, there was substantial variation among participants in both self-reported confidence (median = 4.5; range = 1-10) and skill (mean percentage accuracy 66%; range 36% to 100%); these parameters constituted independent variables that were used in subsequent analyses (see below).

The videos were shown in a dark room on a large (1.80 m x 1.35 m) screen with participants sitting at a distance of 3-4 m. This study setup was used because it mimicked use of the small screen on the back of the thermal camera (960 mm x 720 mm) under field conditions: viewing the large screen from 3-4 m subtended the same angle on the participants' field of view as viewing the camera's screen when the camera was held 16-21 cm from the eyes (standard posture when using the camera: *pers. obs.*). The aspect ratio of both screens was the same at 4:3. Each observer was asked to identify the species and count the total number of individuals present; observers were asked not to confer to ensure independence of the data. There were two screenings of the videos to ensure that all participants could clearly see the screen and were seated so that they could not inadvertently see other participants' answers. Videos were shown in the same random order at both screenings. The distance between the camera and the animals shown was given to provide a sense of the focal animals' likely size. Before viewing the videos, participants were informed that all the species were antelopes found at the study site and that only one species was shown per video.

Statistical analysis

A multiple linear regression analysis was used to assess which observer parameters affected identification accuracy. This analysis used the number of correct identifications from each observer as the dependent variable from 0 (no identifications were correct) to 60 (identifications were correct for all 60 videos). There were 34 cases (one case per observer). Five predictor variables were entered: (1) observer gender; (2) observer age; (3) experience level (inexperienced, intermediate, expert); (4) observer-reported confidence in identifying antelope (rank 1-10); and (5) assessed identification skill as measured by photograph test (% correct). Collinearity between the predictors was within acceptable limits (actual VIF <5.888 in all cases; threshold = 10) (Field, 2000).

A second multiple linear regression analysis was used to establish which ecological parameters affected the ability of observers to identify species accurately. Here, the number of times the species in each video was identified correctly out of 34 attempts was the dependent variable, with each video being allocated a score between 0 (no one correctly identified species) and 34 (everyone correctly identified species). There were 60 cases (one case per video). Three predictor variables were entered: (1) the distance in metres between the thermographer and the animal(s); (2) the number of animals in the group; and (3) whether ≥1 male was present (1 or 0). The final variable was added to determine whether readily identified sex-specific features, notably the presence of horns (prominent in males only in greater kudu*,* impala, waterbuck and common reedbuck; thicker or longer in males in eland, gemsbok, sable wildebeest and tsessebe), was of use in identifying species. The presence of ≥1 male was known in 50 of the 60 cases; where this was unknown or unclear a missing value was entered and the analysis was undertaken based on pairwise deletion. Collinearity between the predictors was within acceptable limits (actual VIF <1.235 in all cases).

Additional analyses were undertaking using Pearson's correlation for normally distributed data or Spearman's rank for non-normal data to assess relationships between: (1) the accuracy with which a species could be identified using thermography and the number of species with which it was confused; (2) an observer's accuracy in identifying species and their accuracy at assessing group size; and (3) the order in which the videos were shown and identification accuracy to assess the possible effect of experience.

Results

Baseline results

In total, 2,040 identification attempts were made (34 participants each assessing 60 thermographic videos), of which 834 were identified correctly to species level (41%) and 1,021 (50%) were identified correctly to subfamily level (the latter being determined by the research team by assessing whether identifications erroneous at the level of the species were within the correct subfamily). Thirteen different antelope species were represented in the results, the nine study species (kudu, tsessebe, impala, wildebeest, waterbuck, reedbuck, eland, gemsbok, sable) plus another four species that were erroneously given as identifications: blesbok (*Damaliscus pygargus phillipsi*; Harper, 1939), red hartebeest (*Alcelaphus caama*; Saint-Hilaire, 1803), common duiker, (*Sylvicapra grimmia*; Gray 1871) and steenbok (*Raphicerus campestris*, Thunberg, 1811). In 17 cases (0.8%), observers failed even to detect the presence of animal(s) in the video.

There was substantial variation in identification accuracy between species (Fig. 2). Wildebeest and impala were identifiable with comparatively high accuracy (87% and 62%, respectively). In contrast, the identification accuracy of reedbuck (12%) and tsessebe (15%) was effectively that expected by chance alone (although the later did rise sharply to 44% accuracy in terms of identifying the correct subfamily). It was notable that there was a statistically significant negative relationship (Spearman's rank correlation r_s = -0.714, n = 9, p = 0.047; Fig 2) between the accuracy with which a species could be identified using thermography and the number of species with which it was confused. The species identified with the highest accuracy, and for which the number of confusion species was lowest, were wildebeest and impala. Tsessebe, reedbuck and waterbuck were associated with the lowest accuracy and a high number of confusion species. Gemsbok, sable, eland and kudu were intermediate in both accuracy and the number of confusions species.

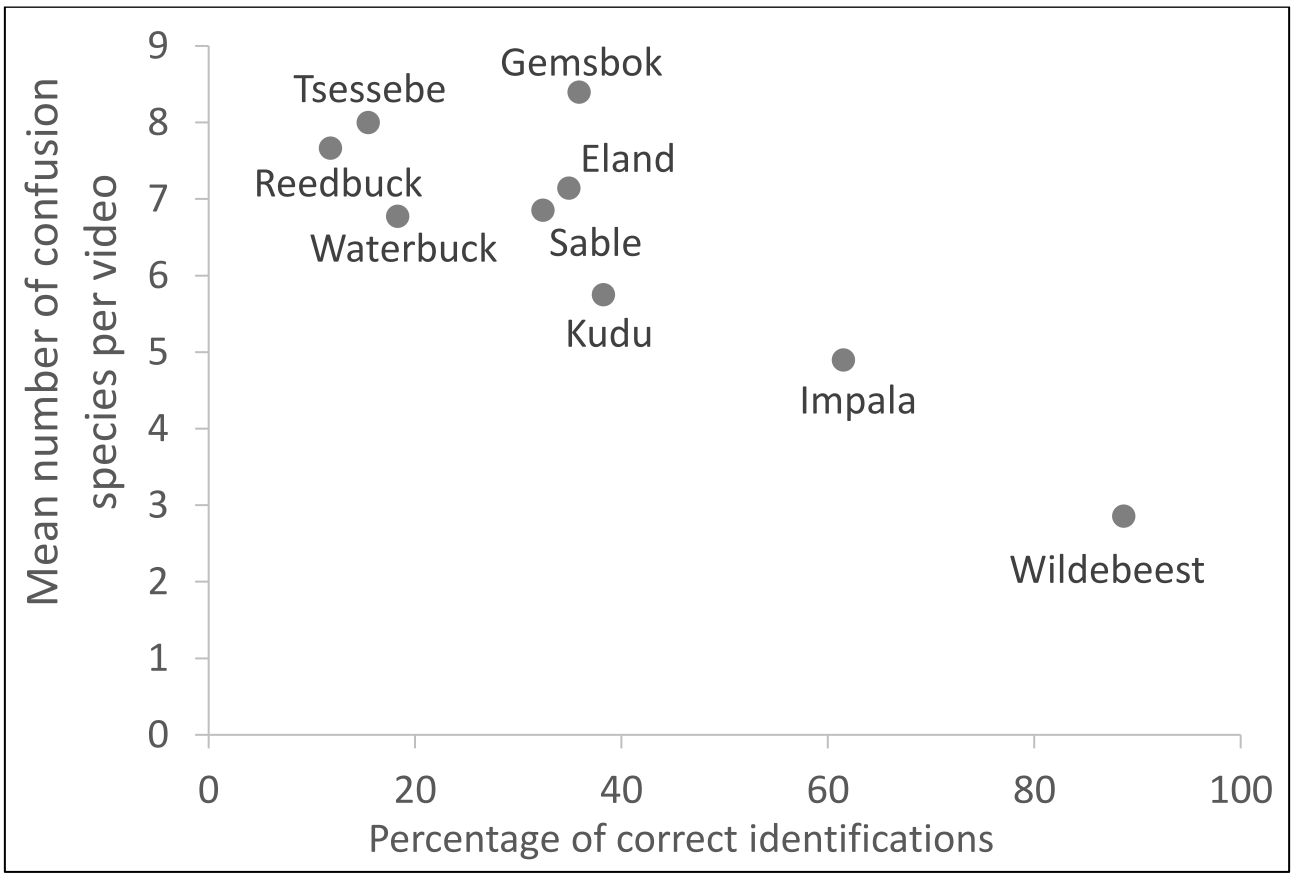

Figure 2

Misidentifications

Common species confusion patterns are shown in Fig. 3. A number of the frequently-confused identifications involved species within the same subfamily (e.g. tsessebe/wildebeest, sable/gemsbok, eland/kudu, waterbuck/reedbuck). Common mistakes were external to subfamily groupings were: (1) reedbuck, kudu, waterbuck and tsessebe all misidentified as impala; and (2) eland misidentified as waterbuck. Some of the common misidentifications occurred for all three experience groups (e.g. reedbuck misidentified as waterbuck), whereas others (e.g. misidentifying tsessebe as kudu) only occurred amongst inexperienced observers.

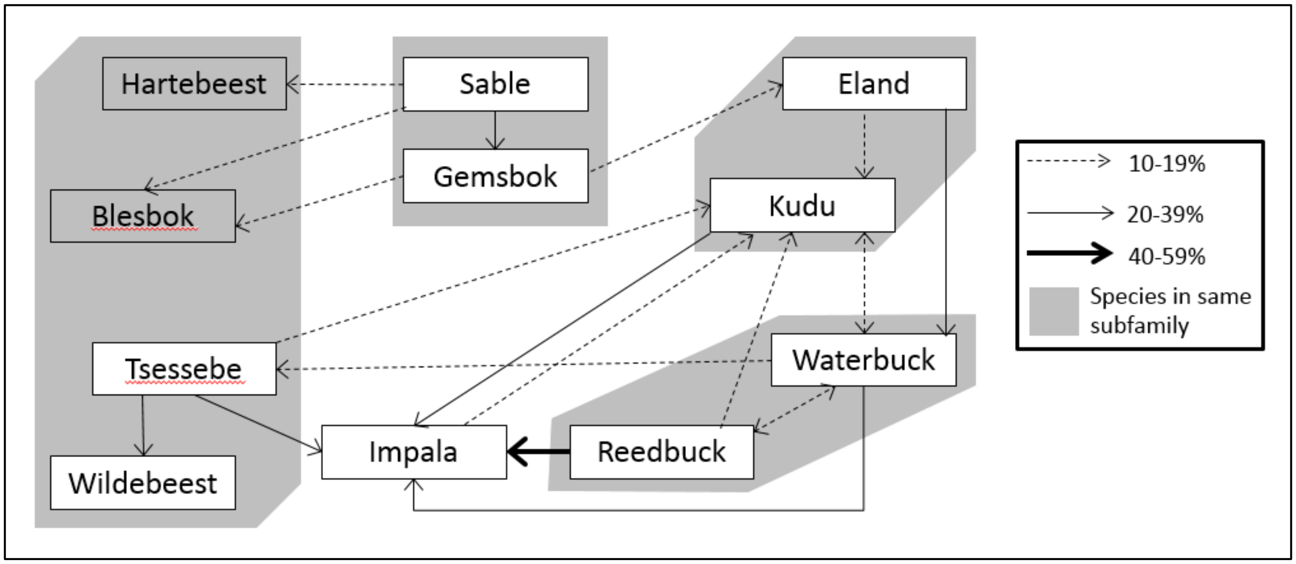

Figure 3

Observer parameter effects on identification accuracy

Across all species, there was a consistent pattern of identification accuracy improving with increasing experience: inexperienced = 43% and 30% accuracy for subfamily and species; intermediate = 57% and 49% accuracy for subfamily and species; experienced = 70% and 61% accuracy for subfamily and species (Fig 4). However, there were species-specific differences in the magnitude of the improvement between experience categories. For example, identification accuracy for tsessebe increased from 4% (inexperienced) to 22% (intermediate) and to 55% (expert) while accuracy also improved substantially for sable (13%→44%→70%) and impala (55%→68%→88%). For other species, such as kudu, the effect of experience was less dramatic (32%→39%→45%). It was also notable that for some species both inexperienced and intermediate groups performed approximately equally in terms of identification accuracy with improvement only occurring at expert level. For example, reedbuck identification accuracy was ~7% among inexperienced and intermediate observers but jumped to 46% among the expert observers (Fig. 4).

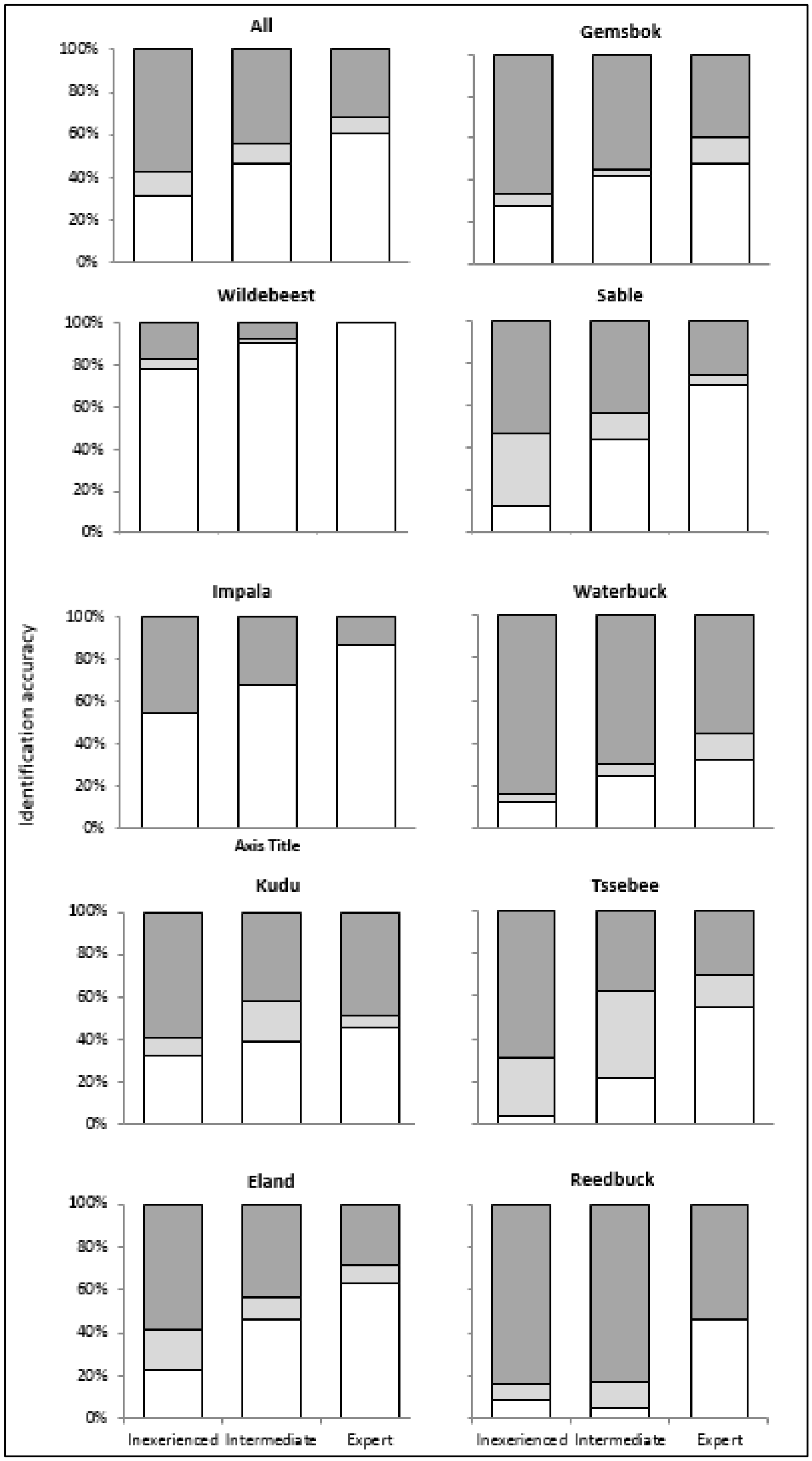
All
Gemsbok
Wildebeest
Sable
Impala
Waterbuck
Kudu
Tsseebee
Eland
Reedbuck
Identification accuracy
Axis Title
0%
20%
40%
60%
80%
100%
Inexerienced
Intermediate
Expert

Figure 4

There was a significant relationship between the number of times an observer correctly identified species in the thermographic videos and: (1) previous experience; and (2) objective identification skill: both relationships were positive. The overall model was highly significant (regression $F_{4,28} = 14.779$; $r^2 = 0.824$; $p < 0.001$); each variable was also significant alone ($p < 0.001$). The best single predictor of identification accuracy when variables were added individually was objective identification skill ($r^2 = 0.620$), but this was closely followed by experience ($r^2 = 0.508$). It was thus possible to explain >60% of the variation in an observer's accuracy by objectively testing their species identification skills using a series of colour images or, more simply and still explaining >50% of the accuracy in variation, by assessing their previous experience. The effect of experience of identification accuracy for individual species can be seen in Fig. 4. There was no significant relationship between identification accuracy and gender ($p = 0.122$) or observer age ($p = 0.565$). Confidence did not significantly relate to identification accuracy ($p = 0.326$), although confidence and objectively-assessed skill, which was a significant predictor of accuracy, were positively correlated (Spearman's rank $r_s = 0.773$, $n = 33$, $p < 0.001$).

There was a significant positive correlation between an observer's accuracy in identifying antelope and their accuracy at assessing group size (Pearson's correlation: $r^2 = 0.193$, $n = 34$, $p = 0.011$). This suggests that some people are simply more proficient at assessing and interpreting thermographic footage than others. There was also a small but significant positive correlation between the order in which the videos were shown (which was consistent between all observers) and identification accuracy (Pearson's correlation: $r^2 = 0.082$, $n = 60$, $p = 0.027$), which suggests that observers quickly become more familiar with assessing thermographic footage.

Ecological parameter effects on identification accuracy

The only video-specific variable to be significantly associated with correct identification to species level was the number of animals in the group (positive; videos with more animals were easier to interpret) (regression $F_{1,58} = 13.573$; $r^2 = 0.190$; $p = 0.001$). This was largely driven by the fact that wildebeest and impala, which are the species that tended to occur in larger groups, were the species that were easiest to identify (Fig. 2). When wildebeest and impala were removed from the analysis, there was no longer a significant relationship between group size and identification accuracy ($F_{1,42} = 0.353$, $p = 0.556$). There was no effect of distance ($p = 0.320$) or whether or not there was a male present ($p = 0.322$) on accuracy of identification.

Discussion

We have shown that, in some circumstances, it is possible to identify antelopes to species level using moving thermographic footage (in this case, using videos as a proxy for real-time thermography). Although overall identification accuracy was only 41%, some species could be identified with a very high level of accuracy (e.g. wildebeest = 87% overall, rising to 100% for the expert group) whereas other species were much more challenging (e.g. tsessebe = 15%, rising to 55% for the expert group). Experience level and objectively-determined identification skill were both important predicators of an observer's ability to identify species.

A previous study on assessment of still thermographic images in identical habitat (Goodenough *et al*. 2017) found a broadly similar level of identification accuracy to that found here for thermographic videos (37% vs 41%, respectively). The aim of this study was not to compare accuracy between photographs and videos *per se*, but rather to establish whether information derived from movement and behaviour (which would be apparent in recorded videos and real-time thermography) resulted in high levels of identification accuracy. In fact, movement information is not as important for correct species identification as might have been assumed, which concurs with Glen *et al*., (2013), who found the ability of researchers to identify species from camera traps did not differ between still or moving footage. It is, however, notable that the accuracy of observers viewing thermographic videos in both the intermediate and expert groups was noticeably higher than for intermediate and expert groups viewing still thermographic images (49% vs 36% and 61% vs 51%, respectively). This situation was reversed for inexperienced observers (30% vs 33%). This suggests that observers with greater experience are able to use movement and interactions in videos to improve identification accuracy, whereas this extra information does not help inexperienced observers.

Identification accuracy increased with experience, rising from 30% (inexperienced) to 61% (experienced). Although this might be expected (e.g. Diefenbach , Brauning & Mattice, 2003; Farmer, Leonard & Horn, 2012; Johnston *et al*., 2017), it cannot be assumed. For example, Tillett *et al.* (2012) found experience did not affect identification accuracy for carcharhinid sharks while a meta-analysis by Lewandowski and Specht (2015) showed that experts were only significantly better at species identification than less-experienced (volunteer) observers in ~50% of cases. Moreover, published evidence is based largely on single-species studies such as bongo antelope (*Tragelaphus eurycerus*) (Gibbon, Bindemann & Roberts 2015) or harbor porpoise (*Phocoena phocoena*) (Hobbs and Waite, 2010)). The multi-species focus of this study showed increases in identification accuracy with experience did not occur uniformly, with higher accuracy of experts being mainly due to increased ability to correctly identify tsessebe, sable and impala. This suggests that the relative difference in identification accuracy between experience categories is species-specific. In a practical context, knowledge of which species can be identified accurately by all experience levels (e.g. wildebeest) means such data can be relied upon across observers. Conversely, knowledge where there is a substantial difference in identification accuracy (e.g. tsessebe) allows for species-specific training. It was notable that there was a small but significant correlation between the order in which the videos were shown and identification accuracy, which suggests that observers require time to "get their eye in" when interpreting thermographic footage. This suggests that training and familiarization activities could improve accuracy relatively quickly (as per Evans *et al*., 2009).

Despite experience being a good predictor identifying accuracy, highly-experienced observers were far from being 100% accurate. The error rate amongst experienced observers found here (39%) is similar to error rates for other hard-to-interpret evidence such as indirect signs (37% for footprints: Evans *et al.* (2009)). It is notable that accuracy in assessment of group size was correlated with accuracy of identification. It would thus seem that, just as some people are better at taking accurate biometric measurements on digital images of animal specimens (Goodenough *et al*., 2012), some people are inherently better at interpreting thermographic footage. As noted by Shea *et al*. (2011) it would be useful to explore what morphological or movement features the most accurate individuals are focusing upon when making their (correct) identifications so this can be incorporated into the training of others.

Species identified with high accuracy tended to have distinctive morphological features that showed well in silhouette, for example, wildebeest have very characteristic posture while impala have characteristic horns. This has been found previously for camera trap footage (Gibbon, Bindemann & Roberts, 2015). Conversely, species with low identification accuracy were typically those identifiable in normal field conditions by coat colour or patterning (e.g. the distinctive white circle around the posterior of waterbuck; the characteristic facial mask of gemsbok). Although prominent coat patterns can sometimes show up on thermal footage (e.g. stripes on zebra Benesch & Hilsberg, 2003; McCafferty, 2007), more subtle patterning is rarely distinguishable thermographically. Anecdotal evidence from participants after they had viewed the videos suggested that the absence of colour and pattern was a barrier to identification, although this was a general comment rather than something that could be linked to specific species. The inherent inability to distinguish some species with similar thermographic signatures has been reported previously (Franke *et al.* (2012) and might be exemplified in the current study by observers misidentifying reedbuck as impala. Videos with larger group sizes were more likely to be identified correctly but this was driven by wildebeest and impala, which are commonly found in large groups with distinctive herd structures. Somewhat surprisingly, distance did not affect identification accuracy, possibly because herd structure, movement and behaviour were easily discernable even at long ranges.

Interestingly, the number of confusion species for each antelope species increased substantially as the accuracy with which that species could be identified decreased. Thus, species such as reedbuck and tsessebe were not simply being confused with a single similar species (e.g. waterbuck and wildebeest, respectively). This multi-species confusion is an important consideration since misidentifications are not easily correctable through a simple equation as per Hobbs and Waite (2010). However, it would be possible to construct a misidentification prior probability matrix giving the proportion of misidentifications for each species relative to every other species (as depicted in Fig. 3) to allow this to be factored into a model. This approach has been used for ice-associated seals (Conn *et al*., 2013). Prior probabilities could be generated by pilot surveys or by having two observers working together as per McClintock *et al.* (2015).

Based on analysis of videos as a proxy for real-time thermography, we conclude that real-time thermography can be a useful tool for surveying African antelope species. This concurs with previous work using still photographs as a proxy (Goodenough *et al*., 2017). For some species, and for some groups of observers, identification accuracy using thermography is extremely high (even up to 100% - e.g. expert observers viewing wildebeest). Thermography also has advantages over other survey methods such as spotlighting in allowing crepuscular and nocturnal species to be detected with minimal interference (Lavers *et al*., 2005; Collier *et al*., 2007) and can be useful in habitats where normal visibility is compromised by vegetation (Gill, Thomas & Stocker, 1997). However, it is vital to realise that real-time thermography is not a "magic bullet" and overall accuracy, at ~40%, is still fairly low. We thus conclude that: (1) there needs to be an understanding of species-specific effectiveness to establish when the technique can be relied upon and also to inform training; (2) observers need experience to maximise the potential of thermography (accuracy was 61% for expert observers versus 30% for inexperienced observers); (3) the use of expensive thermal imaging equipment is only valuable if identification accuracy is higher than would be possible otherwise, which will depend on species community, site-specific conditions and what other techniques are available. Empirical assessment of real-time thermography in relation to other specific techniques, such as use of night-vision goggles, would be a useful area for future research.

Acknowledgements

We thank all the participants for their willingness to be involved in identifying images; D. MacTavish, L. Levitt, J. Martin and R. Leeman for field assistance and useful insights; and S. P. Lodge and M. Ickey for logistical support.

REFERENCES

ANDERSON, K. & GASTON, K.J. (2013) Lightweight unmanned aerial vehicles will revolutionize spatial ecology. *Front. Eco. Environ*. 11, 138-146.

BENESCH, A.R. & HILSBERG, S. (2003) Infrared thermographic study of surface temperature in zebras. *Zool. Garten*. 73, 74–82.

BOONSTRA, R., KREBS, C. J., BOUTIN, S. & EADIE, J. M. (1994) Finding Mammals Using Far-Infrared Thermal Imaging. *J Mammal*. 75, 1063-1068.

BRYSON, M., REID, A., HUNG, C., RAMOS, F. T. & SUKKARIEH, S. (2014) Cost-effective mapping using unmanned aerial vehicles in ecology monitoring applications. In: *Experimental Robotics* (Eds. O. KHATIB & V. KUMAR). Springer, Berlin, Heidelberg.

CILULKO, J., JANISZEWSKI, P., BOGDASZEWSKI, M. & SZCZYGIELSKA, E. (2013) Infrared thermal imaging in studies of wild animals. *Eur. J. Wildl. Res*. 59, 17-23.

COLLIER, B. A., DITCHKOFF, S. S., RAGLIN, J. B. & SMITH, J. M. (2007) Detection probability and sources of variation in white-tailed deer spotlight surveys. *J. Wildlife Manage*. 71, 277-281.

CONN, P.B., McCLINTOCK, B.T., CAMERON, M.F., JOHNSON, D.S., MORELAND, E.E. & BOVENG, P.L., (2013) Accommodating species identification errors in transect surveys. *Ecology. 94*, 2607-2618.

COSTA, H., FOODY, G.M., JIMÉNEZ, S. & SILVA, L. (2015) Impacts of species misidentification on species distribution modeling with presence-only data. *Int. J. Geo-Inf*. 4, 2496-2518.

DIEFENBACH, D.R., BRAUNING, D.W. & MATTICE, J.A. (2003) Variability in grassland bird counts related to observer differences and species detection rates. *Auk*. 120, 1168-1179.

DUNN, W. C., DONNELLY, J. P. & KRAUSMANN, W. J. (2002) Using thermal infrared sensing to count elk in the southwestern United States. *Wildlife Society Bulletin*. 30, 963-967.

EVANS, J.W., EVANS, C.A., PACKARD, J.M., CALKINS, G. & ELBROCH, M. (2009) Determining observer reliability in counts of river otter tracks. *J. Wildl. Manag.* 73, 426-432.

FARMER, R.G., LEONARD, M.L. & HORN, A.G (2012) Observer effects and avian-call-count survey quality: rare-species biases and overconfidence. *Auk*. 129, 76-86.

FIELD, A. P. (2000) *Discovering Statistics Using SPSS for Windows: Advanced Techniques for Beginners*. Sage Publications, Inc. Thousand Oaks, CA, USA.

FRANKE, U., GOLL, B., HOHMANN, U. & HEURICH, M. (2012) Aerial ungulate surveys with a combination of infrared and high-resolution natural colour images. *Anim. Biodivers. Conserv*. 35, 285-293.

FRETWELL, P. T. & TRATHAN, P. N. (2009) Penguins from space: faecal stains reveal the location of emperor penguin colonies. *Global Ecol. Biogeogr.* 18, 543-552.

GANOW, K. B., CAIRE, W. & MATLOCK, R. S. (2015) Use of thermal imaging to estimate the population sizes of Brazilian free-tailed bat, *Tadarida brasiliensis*, maternity roosts in Oklahoma. *Southwest Nat*. 60, 90-96.

GIBBON, G. E. M., BINDEMANN, M. & ROBERTS, D. L. (2015) Factors affecting the identification of individual mountain bongo antelope. *PeerJ*. 3: e1303 https://doi.org/10.7717/peerj.1303

GILL, R. M. A., THOMAS, M. L. & STOCKER, D. (1997) The use of portable thermal imaging for estimating deer population density in forest habitats. *J. Appl. Ecol.* 34, 1273-1286.

GLEN, A. S., COCKBURN, S., NICHOLS, M., EKANAYAKE, J. & WARBURTON, B. (2013) Optimising camera traps for monitoring small mammals. *PLoS ONE*. *8*, e67940

GOODENOUGH, A. E., SMITH, A. L., STUBBS, H., WILLIAMS, R. L., & HART, A. G. (2012) Observer variability in measuring animal biometrics and fluctuating asymmetry when using digital analysis of photographs. *Ann. Zool. Fennici*. 49, 81-92.

GOODENOUGH, A. E., CARPENTER, W. S., MACTAVISH, L., MACTAVISH, D., THERON, C. & HART, A G. (2017) Empirically testing the effectiveness of thermal imaging as a tool for identification of large mammals in the African bushveldt. *Afr. J. Ecol*. http://doi:10.1111/aje.12416

GRAVES, H. B., BELLIS, E. D. & KNUTH, W. M. (1972) Censusing white-tailed deer by airbourne thermal infrared imagery. *J. Wildlife Manage*. 36, 875-884.

HART, A. G., ROLFE, R. N., DANDY, S., STUBBS, H., MACTAVISH, D., MACTAVISH, L., & GOODENOUGH, A. E. (2015) Can handheld thermal imaging technology improve detection of poachers in African bushveldt? *PLoS ONE*. 10, e0131584.

HOBBS, R. C. & WAITE, J. M. (2010) Abundance of harbor porpoise (*Phocoena phocoena*) in three Alaskan regions, corrected for observer errors due to perception bias and species misidentification, and corrected for animals submerged from view. *Fish. Bull.* 108, 251-267.

JOHNSTON, A., FINK, D., HOCHACHKA, W.M. & KELLING, S. (2017) Estimates of observer expertise improve species distributions from citizen science data. *Methods Ecol. Evol*. In Press.

LAVERS, C., FRANKS, K., FLOYD, M. & PLOWMAN, A. (2005) Application of remote thermal imaging and night vision technology to improve endangered wildlife resource management with minimal animal distress and hazard to humans. *J. Phys. Conference Series*. 15, 207-212.

LEWANDOWSKI, E. & SPECHT, H. (2015) Influence of volunteer and project characteristics on data quality of biological surveys. *Conserv. Biol*. 29, 713-723.

MCCAFFERTY, D.J. (2007) The value of infrared thermography for research on mammals: previous applications and future directions. *Mammal Rev*. 37, 207-223.

MCCAFFERTY, D., MONCRIEFF, J., TAYLOR, I. & BODDIE, G. (1998) The use of IR thermography to measure the radiative temperature and heat loss of a barn owl (*Tyto alba*). *J. Therm. Biol*. 23, 311-318.

MCCLINTOCK, B.T., MORELAND, E.E., LONDON, J.M., DAHLE, S.P., BRADY, G.M., RICHMOND, E.L., YANO, K.M. & BOVENG, P.L. (2015) Quantitative assessment of species identification in aerial transect surveys for ice-associated seals. *Marine Mam. Sci*. *31*, 1057-1076.

NAUGLE, D. E., JENKS, J. A. & KERNOHAN, B. J. (1996) Use of thermal infrared sensing to estimate density of white tailed deer. *Wildlife Soc, B*. 24, 37-43.

O'CONNELL, A.F., NICHOLS, J.D. & KARANTH, K.U. (Eds.) (2010) *Camera Traps in Animal Ecology: Methods and Analyses*. Springer. Tokyo, Dordrecht, Heidelberg, London, New York.

OGUTU J.O., BHOLA, N., PIEPHO, H. P. & REID, R. (2006) Efficiency of strip-and line-transect surveys of African savanna mammals. *J. Zool*. 269, 149-160.

PLUMPTRE A. J. (2000) Monitoring mammal populations with line transect techniques in African forests. *J. App. Ecol*. 37, 356-368.

SHEA, C. P., PETERSON, J. T., WISNIEWSKI, J. M. & JOHNSON, N. A. (2011) Misidentification of freshwater mussel species (Bivalvia: Unionidae): contributing factors, management implications, and potential solutions. *J. N. Am. Benthological Soc.* 30, 446-458.

TILLETT, B. J., FIELD, I. C., BRADSHAW, C. J., JOHNSON, G., BUCKWORTH, R. C., MEEKAN, M. G. & OVENDEN, J. R. (2012) Accuracy of species identification by fisheries observers in a north Australian shark fishery. *Fisheries Res. 127*, 109-115.

WELBOURNE, D.J. (2013) A method for surveying diurnal terrestrial reptiles with passive infrared automatically triggered cameras. *Herpetol. Rev*. 44, 247-250.

ZEHNDER, S., AKESSON, S., LIECHTI, F. & BRUDERER, B. (2001) Nocturnal autumn bird migration at Falsterbo, south Sweden. *J. Avian Bio.* 32, 239-248.

Figure legends

Figure 1: Stills from thermographic videos of African antelope at different distances: (a) impala; (b) sable; (c) waterbuck; (d) wildebeest; (e) eland; and (f) tsessebe; the videos have been uploaded as Supplementary Information. In some species, the presence of horns as a distinct morphological feature may aid accurate identification (e.g. horns of the male impala and sable stand out clearly at close range).

Figure 2: The percentage of correct identifications for all video clips of each study species plotted against the mean number of species that acted as confusion species (i.e. that were misidentified as the focal species).

Figure 3: Common misidentification mistakes made by observers studying thermographic videos of African antelope. Line thickness shows prevalence and arrow shows direction of mistakes. Species in the same subfamily are grouped. The two greyed out species – hartebeest and blesbok - were not part of this study directly (i.e. not filmed) but were present at the study site and were frequently confused with sable and gemsbok, which were part of the study.

Figure 4: Accuracy of species identification from thermographic videos expressed as a percentage for a) inexperienced observers; b) intermediate observers; and c) experienced observers for all species combined and then for each specific species (given in order of overall accuracy). For all species except impala, there were other species within the same subfamily so the percentage occurrence of recorders misidentifying animals at species level but ascribing them to the correct subfamily is also shown.