Supplemental Materials

1.1 Data Processing

**1.1.1 GPS data processing**

Data processing and analysis took place using R (version 3.1) and Matlab (version 2017b). Raw GPS data were manually error corrected, first using the functions findLocErr, findLocNA, findTimeDup and findMissing (swaRm package (Garnier 2016)) to identify potential errors. The nature of the identified error was then assessed and manually corrected if it resulted from an obvious writing error, or otherwise omitted. Sequences of NA’s were inserted when data were not available for a given individual for a given second. Sometimes missing values were due to the collar having malfunctioned for a significant period of time. However, sometimes missing values were simply due to the GPS device temporarily losing satellite signal. When a sequence of missing values occurred that was less than or equal to 10 seconds long, the “fixLocNA” function (swaRm package (Garnier 2016)) was used to linearly interpolate the missing values. Sequences of missing values longer than 10 seconds were not interpolated. In addition, the data were cropped from the time the goats left the pen to the time they arrived back at the pen. For a display of all GPS coverage times see Supplementary Figure 1.

**1.1.2 Identifying Calls**

In order to determine the timing of call production and identity of the caller, we had to process the continuous audio streams collected by our dataloggers (See Figure S1 for a display of audio coverage times). We trained a support vector machine algorithm (Kernlab package v0.9-27, (Karatzoglou et al. 2004)) on the set of loud contact calls that were documented in the observational data collected during the study (n = 248 vocalizations). These vocalizations were located in the audio data by visually and auditorally searching for calls around the times documented in the observational data using Raven Pro (version 1.5). Once the calls were located, we annotated the start and end of each call. We also created an equivalent dataset of verified non-calls by randomly selecting sound clips from the same recordings from which each verified call had been recorded. Each randomly selected segment was manually verified to not contain any goat vocalizations by a group of trained listeners. We then extracted a 3-second clip for each manually verified call and non-call, which consisted of 1.5 seconds on either side of the center of the vocalization or the center of the non-call sound clip. A pre-emphasis filter was then applied to these segments which aids signal detection by putting greater emphasis on high frequencies and less on low frequencies, thus enabling greater disambiguation from background noise. We then calculated the amplitude envelope and the fundamental frequency track for each clip and found the average values of each for both calls and non-calls. To examine the ability of these two measures to enable differentiation between calls and non-calls, we fitted a gaussian curve to the average amplitude envelope and fundamental frequency track for calls (Figure S2) and found the correlation with these curves for both calls and non-calls (Figure S3). We then trained a support vector machine algorithm on half of the manually verified data set and validated it with the second half which resulted in the successful detection of 88.7% of observed calls. The machine learning algorithm was then applied to all audio recordings which moved a 3-second window along each recording in 1-second increments and assigned each interval a probability from 0 to 1 that that segment contained a call. We considered all segments with a probability of 0.5 or greater to be a call in order to reduce the number of false negatives. However, since this threshold resulted in many false positives, we proceeded to manually verify all positive identifications to determine whether they were goat vocalizations or not and whether they were produced by the individual wearing the collar or another individual. Since the microphone was positioned very close to the throat of the wearer, the amplitude of vocalizations produced by the wearer was much greater than calls produced by nearby individuals. In our analyses we only consider vocalizations produced by the individual wearing the collar. This resulted in a final dataset of 651 vocalizations, 244 of which fell during the selected periods of movement data.

**1.1.3 Assessing Synchrony of GPS and Audio Data**

To assess the alignment of the audio data with the GPS data, we examined events in which the goats visited a watering trough, since these events were both visually apparent when overlaying the GPS trajectories over a map of the park and also acoustically apparent due to the distinctive sound of the animal drinking (Lynch et al. 2013), which is amplified by the positioning of the microphone at the animal’s throat. Individuals arriving and then departing from a watering trough displayed a sharp trajectory change that aligned with the position of the trough. We identified the time of arrival, to the second, at the change point in the GPS data as well as the time of initiation of drinking sounds in the audio data, with the aid of spectrograms generated in Raven Pro. Based on the GPS-derived start time of the drinking bouts and the amount of time known to have elapsed in the audio data up to the initiation of drinking sounds, we calculated the GPS-derived start time of the audio recording for that day. When we compared the set of audio start times calculated from the time spoken at the beginning of the day with the audio start times inferred from the GPS data (n = 78) we found a range of time differences from -25 to 34 seconds with a median difference of 0 seconds and a mean (SD) difference of -0.1 (6.7) seconds. Based on this assessment, we concluded that there was no need to shift the times relative to one another. We used the audio recording start times derived from the time spoken at the beginning of each day for calculating the times of any calls that occurred that day, since these times were available for the majority of recordings. If a given start time was not available on a given day, for instance, due to a delay in the initiation of recording, we did not include any acoustic data from that device for that day in our analysis (See Figure S1 for an overview of all audio data available for analysis).

References

Garnier, Simon. 2016. “SwaRm.” https://github.com/swarm-lab/swaRm.

Karatzoglou, Alexandros, Alex Smola, Kurt Hornik, and Achim Zeileis. 2004. “Kernlab - An *S4* Package for Kernel Methods in *R*.” *Journal of Statistical Software* 11 (9): 1–20.

Lynch, Emma, Lisa Angeloni, Kurt Fristrup, Damon Joyce, and George Wittemyer. 2013. “The Use of On-Animal Acoustical Recording Devices for Studying Animal Behavior.” *Ecology and Evolution* 3 (7): 2030–37.