

Chimpanzees (*Pan troglodytes*) Exhibit Gaze Bias for Snakes Upon Hearing Alarm Calls

Yutaro Sato^{1, 2}, Fumihiro Kano^{2, 3}, Naruki Morimura^{1, 2}, Masaki Tomonaga⁴, and Satoshi Hirata^{1, 2}

¹ Wildlife Research Center, Kyoto University

² Kumamoto Sanctuary, Wildlife Research Center, Kyoto University

³ Institute for Advanced Studies, Kyoto University

⁴ Inuyama, Japan

Calls of several species of nonhuman animals are considered to be functionally referential. However, the cognitive mechanisms underlying audience behaviors remain unclear. This study used an audiovisual cross-modal preferential-looking paradigm to examine whether captive chimpanzees spontaneously associated a conspecific call with images of a corresponding item. Chimpanzees were presented with videos of snakes and fruit side-by-side while hearing playbacks of alarm calls, food-associated calls, or no sound (as a baseline condition). Chimpanzees looked at videos of snakes for longer when hearing alarm calls compared with food calls or baseline. However, chimpanzees did not look at videos of fruit for longer when hearing food calls compared with baseline. An additional experiment tested whether chimpanzees' gaze bias to the snake videos was driven by negative affective states in general via affect-driven attention biases. When chimpanzees were presented with the same snake and fruit videos while hearing playbacks of conspecific screams or no sound, they exhibited no gaze bias for snake videos. These results suggest that chimpanzees spontaneously associated alarm calls with images of a potential threat in a preferential-looking experiment and that this response was not simply driven by an affective state matching process. These findings should be interpreted in consideration of a procedural limitation related to pseudoreplication in the experimental stimuli.

Keywords: chimpanzee, cross-modal preferential looking, vocalization, alarm call, snake

Previous studies have suggested that some calls of nonhuman animals are referential, at least functionally, in that audience animals respond as if the vocalizations indicate specific external events or objects (reviewed in Macedonia & Evans, 1993; Seyfarth & Cheney, 2003; but see also Wheeler & Fischer, 2012 for a critical view). For example, in several species of birds and mammals, distinct alarm calls in response to the presence of predators have been observed to induce audience animals to exhibit adaptive evasion responses according to the nature of predator movements (reviewed in Suzuki, 2016; Townsend & Manser, 2013). Moreover, carefully designed playback experiments suggest that audience animals are not simply responding to calls in a fixed manner (Clay & Zuberbühler, 2011; Evans & Evans, 2007; Seyfarth & Cheney, 1990; Slocombe & Zuberbühler, 2005; Suzuki, 2018, 2020; Zuberbühler, 2000; Zuberbühler et al., 1999). For example, in a field experiment (Seyfarth & Cheney, 1990), vervet monkeys (*Cercopithecus aethiops*) repeatedly heard playbacks of either conspecific alarm calls or alarm calls of a sympatric bird species (superb starlings: *Spreo superbus*). Monkeys progressively reduced their responses to playback (i.e., looking toward the speaker), which indicates that they became habituated to it. Monkeys were found to transfer this habituation between playbacks of conspecific calls elicited by eagles and bird alarm calls elicited by raptors as if the two types of call redundantly informed monkeys of similar threats. Furthermore, an experiment with fowls (*Gallus gallus*) reported that individuals exhibited greater search effort for food when they heard food calls compared with when they heard alarm calls but did not differentiate search effort when they had recently discovered food by themselves (Evans & Evans, 2007). A previous study (Slocombe & Zuberbühler, 2005) suggested that chimpanzee (*Pan troglodytes*) food grunts functioned referentially in an audience of chimpanzees in an experimentally induced situation. The chimpanzees first learned that the two locations were baited with cardboard tubes containing either pieces of apple or bread. Subsequently, when the chimpanzees heard a series of food grunts emitted by groupmates in response to apples or bread, the individual exhibited greater search effort for the corresponding food item (for a similar finding with bonobos [*Pan paniscus*], see Clay & Zuberbühler, 2011). These results indicate that audience animals do not just exhibit an adaptive response to each call in a fixed manner but adjust their responses more flexibly according to the types of conspecific calls and the contexts in which they heard those calls.

This body of evidence suggests that animals may recollect representations of relevant items upon hearing the calls (Evans & Evans, 2007; Seyfarth & Cheney, 1990; Zuberbühler et al., 1999). Another line of evidence supporting this claim comes from cross-modal matching experiments, in which participants are sequentially presented with auditory and visual stimuli, and observed for spontaneous responses to the visual referents of the auditory stimuli (Suzuki, 2016). In studies of wild tits, birds approached a wooden stick that simulated snake-like movements only when they heard snake-specific alarm calls of conspecifics (Suzuki, 2018; or heterospecifics: Suzuki, 2020) but not when they heard an irrelevant call, or when the stick exhibited irrelevant movement (Japanese great tits [*Parus minor*]: Suzuki, 2018; coal tits [*Periparus ater*]: Suzuki, 2020). Thus, tits appeared to be primed to detect snakes when they heard snake-specific calls. In a somewhat related study (Shibasaki et al., 2014), captive Japanese macaques (*Macaca fuscata*) were presented with images of a snake and a flower side-by-side while

simultaneously hearing playbacks of either alarm calls or irrelevant contact calls. Monkeys looked at the snake for longer than the flower when hearing the alarm call, although this effect was limited to when the snakes were presented on the left side of the screen. Similarly, human (*Homo sapiens*) infants looked at videos of snakes for longer when hearing frightened human voices than when hearing happy voices (DeLoache & LoBue, 2009).

The current study aimed to contribute to this accumulating evidence by testing whether chimpanzees spontaneously associated conspecific calls with specific visual referents in cross-modal preferential-looking experiments following a design implemented in a previous study of monkeys (Shibasaki et al., 2014) with several modifications, as described in the following text. Chimpanzees are known to flexibly adjust their call utterances and behavioral responses to conspecific calls depending on context (e.g., Kalan et al., 2015; Schel et al., 2013; Slocombe & Zuberbühler, 2005, 2006). For example, chimpanzees have been reported to adjust utterances of alarm calls in response to a realistic model of a snake, possibly according to the knowledge status of the audience chimpanzees (i.e., whether audience members had already seen the model snake and whether audience members had heard alarm calls; Crockford et al., 2012; for a similar trend, see also Girard-Buttoz et al., 2020), or on the basis of whether audience presence was simulated by playbacks of alarm calls, indicative of their awareness of threat, or playbacks of an irrelevant call, indicative of their unawareness of threat (Crockford et al., 2017). Moreover, chimpanzees were more likely to exhibit marking-like behaviors (i.e., repositioning of their body and gaze alteration between audience and snake) when audiences had not seen the snake compared with when audiences had seen the snake (Crockford et al., 2017). Alarm calls in combination with marking behaviors more effectively induced audiences to cautiously approach the snake compared with alarm calls alone.

Although the relationships between vocalizations and audience behavior have been researched in some monkey species, it remains unclear whether and to what extent chimpanzees associate calls with relevant objects. Chimpanzees have been reported to emit alarm calls such as alarm barks, screams, and “hoos” in response to various types of threats such as snakes and larger-bodied mammals including leopards and buffalos (Crockford & Boesch, 2003; Goodall, 1986, Chapter 6). In playback experiments (Crockford et al., 2015, 2017), chimpanzees were found to pay more attention to a hidden speaker when hearing playbacks of alarm hoo calls compared with another hoo call emitted in a resting context. However, it is unclear whether and to what extent those calls convey properties of a specific event to the audience (Crockford et al., 2015; Dezechache & Berthet, 2018; Dezechache et al., 2019). Regarding food-associated calls, previous studies suggest that call production is facilitated by larger food resources (Brosnan & de Waal, 2000; Hauser et al., 1993 [divisible food resources in particular]; Hauser & Wrangham, 1987; Leroux et al., 2021 [combinatorial use of calls in particular]; Slocombe et al., 2010; but see O’Bryan, 2015 for a null result) and more preferable types of food (Kalan & Boesch, 2015; Fedurek & Slocombe, 2013; but see O’Bryan, 2015 for an opposite result). Food-associated calls have been reported to attract other chimpanzees to a food patch, at least in some cases (Kalan & Boesch, 2015; O’Bryan et al., 2021; but see O’Bryan, 2015 for a null result). In particular, chimpanzee food grunts have been suggested to exhibit different acoustic profiles depending on the characteristics of food, such as the relative value/preference or size of the food (Hallberg et al., 2003; Kalan et al., 2015 [but only for one tree species]; Slocombe & Zuberbühler, 2005, 2006 [in captivity but not in a wild habitat]; see also O’Bryan, 2015 for a null result). Preliminary evidence suggests that audience chimpanzees may use such acoustic differences to guide subsequent behaviors (Kalan et al., 2015; Slocombe & Zuberbühler, 2005). Similarly, when captive chimpanzees were

presented with four images of different food items on a touchscreen and heard a playback of food barks, chimpanzees were reported to spontaneously select an image depicting a food item of similar value as the one that elicited the food bark more often than expected by chance (Hallberg et al., 2003; Gibbons, 2007). However, this line of inquiry has yielded mixed results (O'Bryan, 2015), and the cognitive underpinnings of audience behavioral responses to those calls remain unclear. In the current study, we sought to examine whether chimpanzees spontaneously match either alarm calls or food grunts to relevant images of snakes and fruit, respectively.

Experiment 1

Chimpanzees were presented with videos of snakes and fruit side-by-side while hearing playbacks of either alarm calls or food grunts. We measured chimpanzees' gaze to each type of video using an eye tracker. If chimpanzees had associations between alarm calls and snakes or between food calls and fruits, they would be expected to look at snake videos for longer when hearing alarm calls than when hearing food grunts and to look at fruit videos for longer when hearing food grunts than when hearing alarm calls.

Method

Participants

Six chimpanzees at Kumamoto Sanctuary (KS), Wildlife Research Center, Kyoto University, participated in Experiment 1 (five females and one male, 10–24 years old; Table 1). Chimpanzees lived in a group comprising these six individuals. Chimpanzees spent the daytime in open-air outdoor enclosures (in total 294 m²) with occasional access to indoor enclosures, in which they received meals and slept at night. These enclosures were equipped with complex structures

such as climbing towers, platforms, hammocks, ropes, and/or small trees to allow chimpanzees to move and climb freely. Chimpanzees received meals comprising fresh vegetables, fruits, and monkey chow three times a day and, in between, small amounts of nuts and seeds in small packages or juice in tubes for enrichment. They had free access to water from the taps located in both outdoor and indoor enclosures. Chimpanzees were not deprived of food or water for the studies. The study depended on their voluntary participation. Animal husbandry and study protocols complied with the Guide for Animal Research Ethics provided by the Wildlife Research Center, Kyoto University (WRC-2018-KS006A).

Stimuli

Visual stimuli were video clips of snakes and fruit (400 px × 400 px) presented on a black screen (1280 px × 720 px) side-by-side. Each video clip of snakes or fruit comprised four 3-s clips depicting different items in each category (12 s in total; Figure 1a), which were presented in the same order across trials and participants. Snake videos depicted four species of snakes found in Japan: the Japanese rat snake (*Elaphe climacophora*), Japanese striped snake (*Elaphe quadrivirgata*), Japanese pit viper (*Gloydius blomhoffii*), and tiger keelback (*Rhabdophis tigrinus*). The chimpanzees in this study were likely to have observed one or more of these species because they are commonly found outdoors in Japan, and the chimpanzees have explored their outdoor compounds freely since their youth (chimpanzees in Experiment 1: Idani & Hirata, 2006; chimpanzees in Experiment 3: Matsuzawa, 2006). Snakes exhibited sinusoidal motions in each of the 3-s snake video clips. Fruit videos depicted pieces of fruit (apple, pineapple, banana, or persimmon) on a silver plate. Chimpanzees were often fed pieces of fruit on this plate for lunch. In the video, a human experimenter picked up and held one piece of fruit. This action was intended to roughly control for the difference in the amount of movement between the fruit and snake videos.

Auditory stimuli were alarm calls (probably those described as “wraaa”: Goodall, 1986, Chapter 6; Marler & Tenaza, 1977; or “Waa Barks”: Schel et al., 2013) and food grunts (described as “rough grunts”: Marler & Tenaza, 1977; Slocombe & Zuberbühler, 2006; or “food grunts”: Goodall, 1986, Chapter 6). For each call type, we used two calls emitted by two of the female chimpanzee participants (Figure 1b). The calls were recorded using a video camera (HDR-CX430V, Sony Marketing Inc., Tokyo, Japan). We recorded alarm calls (0.7 s and 1.0 s of audio were clipped from video recordings and are shown as spectrograms in the left column of Figure 1b) during an incident in which chimpanzees repeatedly emitted alarm calls, and an experienced caretaker suggested that the chimpanzees were likely to have found a snake in an outdoor enclosure, although we did not observe the snake directly during the incident. The 0.7-s and 1.0-s alarm calls were played eleven and nine times, respectively. The calls were distributed over the entire duration of the 12-s video at arbitrary intervals (Video S1 in the online supplemental materials). The same composition was used across trials. We recorded food calls (1.5 s and 3.0 s of audio were clipped from video recordings and are shown as spectrograms in the middle column of Figure 1b) when chimpanzees grunted during feeding times in an experimental booth, during which they could receive pieces of fruit. The 1.5-s and 3.0-s calls were played four and three times, respectively. The calls were distributed over the entire duration of the 12-s video at arbitrary intervals (Video S2 in the online supplemental materials). The same composition was used across trials. The playback volumes of the calls were not equated strictly between call types but were adjusted so that the auditory stimuli could be heard clearly inside the experimental booth (although the exact volume level was not recorded at the time, it was estimated to be approximately 70 dBA by post hoc measurement). We used Adobe Audition CC 2017 to create auditory sequences, Adobe After Effects 2019 to combine the visual and auditory stimuli, and Xmedia Recode to reduce the file sizes.

Apparatus

Chimpanzees were invited into an indoor experimental booth (3 m × 3 m × 2 m), either individually or in a dyad for two

mother–offspring pairs. Chimpanzees sat inside the experimental booth and viewed the stimuli through a transparent 1-cm-thick polycarbonate panel (Figure 2). The experimental equipment was installed outside the booth. Chimpanzees were eye-tracked with an infrared eye tracker (Tobii TX300, Tobii Technology AB, Stockholm, Sweden) while being presented with visual stimuli on a 23-in. LCD monitor at a viewing distance of approximately 70 cm. The sounds were played through speakers located below the monitor. Before conducting daily experiments, we sprayed antifog liquid on the transparent panel to prevent chimpanzees' breath causing condensation on the panel, which can cause noise in eye movement recording. To reduce chimpanzees' head movements, we allowed chimpanzees to sip juice that was slowly dispensed via a nozzle attached to the transparent panel. Another experimenter also held some of the participants' heads gently (for Hat suka, Misaki, and Mizuki) to improve head stabilization during the recording. Chimpanzees' eye movements were never explicitly trained in the present study, in accord with previous studies carried out in this sanctuary (overviewed in Hopper et al., 2021).

Procedures

On the first testing day and whenever necessary, we conducted two-point automated calibration, in which small video clips were presented at each reference point. We used the same calibration record for the same participant across trials (days) in each experiment but manually checked the calibration accuracy before each trial. Specifically, we presented nine small icons on the screen, then guided chimpanzees' gaze onto one of the icons and observed whether the recorded gaze deviated from the icon.

At the start of each trial, we presented a small 3-s video clip in which slideshows of chimpanzees were presented with brown noise (although the exact volume level was not recorded at the time, it was estimated to be approximately 60 dBA by post hoc measurement) to allow chimpanzees to become accustomed to sounds and thereby prevent them from being frightened during the playback of test sounds. A white fixation cross was then presented on a black screen. The test video was played with a key press when the experimenter visually confirmed that chimpanzees appeared to be attending to the screen.

Chimpanzees performed one trial (i.e., 12-s audio visual stimulus) per day. Chimpanzees viewed the visual stimuli in two conditions, in which they heard either alarm calls or food grunts. Each condition comprised two trials presenting the snake and fruit videos on different sides (i.e., left/right). The order of conditions and the side on which the snake/fruit videos were presented were counterbalanced across chimpanzees. We repeated one trial for a female chimpanzee (Misaki) after the other trials were completed because she kept fixating on one location on the screen and appeared to be inattentive to the stimulus content.

Analyses

We used Tobii Studio v.3.4.8 to process the gaze data. The raw gaze recordings were filtered using a default Tobii Fixation Filter. We drew rectangular areas of interest (440 px × 440 px) for snake and fruit videos (Figure 1c). Areas of interest were slightly larger than the video stimuli (400 px × 400 px) to accommodate minor recording errors.

For each trial, we calculated the proportion of time spent looking at snakes, namely, snake gaze time divided by the total gaze time directed at either snake or fruit videos. A higher value indicates a longer gaze time to snake relative to fruit videos and a lower value indicates a longer gaze time to fruit relative to snake videos. A value of 0.5 indicates chance level (i.e., an equal gaze time to snake and fruit videos). Statistical analyses were performed using R v.4.0.0 (R Core Team, 2020). We conducted a repeated-measures analysis of variance with condition (alarm call/food grunt) and location of the snake video (left/right) as independent variables, using the *anovakun* v.4.8.5 R function. For repeated measures, Cousin eau–Morey–Baguley's difference-adjusted normalized confidence intervals (CIs; Baguley, 2012) were calculated using *anovakun* v.4.8.5.

Results and Discussion

Chimpanzees viewed either snake or fruit videos for 10.4 ± 2.37 s ($M \pm SD$) in the alarm call condition and 11.3 ± 1.18 s in the food grunt condition. We found a significant main effect of condition, $F(1, 5) = 7.44$, $p = .041$, $h^2 = .22$, with a higher proportion of snake gaze time during playback of alarm calls ($M = .44$, 95% CI [0.37, 0.51]; two trials were averaged for each participant and the average and 95% CI of the six participants were calculated; Figure 3) compared with playback of food grunts ($M = .28$, 95% CI [0.21, 0.36]). The other effects were not significant, video location: $F(1, 5) = 0.31$, $p = .601$, $h^2 = .01$; Condition × Video:

Location: $F(1, 5) = 2.52$, $p = .173$, $h^2 = .07$. This result suggests that chimpanzees looked at snake videos for longer when they heard alarm calls than when they heard food grunts, and/or that chimpanzees looked at fruit videos for longer when they heard food grunts than when they heard alarm calls. However, it remained unclear whether alarm calls caused gaze bias to snake videos or food grunts caused gaze bias to fruit videos, or both. To test this question, we conducted a follow-up experiment to examine the baseline gaze allocation to the fruit and snake videos when the chimpanzees were not hearing any auditory playbacks (in the online supplemental materials). This follow-up experiment suggested at least that the chimpanzees looked at the fruit videos for longer than the snake videos when there were no auditory playbacks. However, we could not strictly compare this result with that of Experiment 1 because the order of conditions was not counterbalanced across participants. We addressed this issue in Experiment 2.

Experiment 2

We tested another group of chimpanzees by counterbalancing the order of the three conditions (alarm call/food grunt/no-sound base line). Note that we conducted several experiments with some procedural modifications before Experiment 1, all of which produced null results (a brief overview of those experiments is provided in the online supplemental materials). We were concerned that those prior experiences might have somehow affected the results of Experiment 1. Thus, we also aimed to ensure that different chimpanzees (with a less experience of those experiments) yielded the same results with identical procedures in Experiment 2. Chimpanzees would be expected to look at the snake videos for longer when hearing alarm calls than when hearing no sound and to look at the fruit videos for longer when hearing food grunts

than when hearing no sound.

Method

Participants

Seven chimpanzees at the Primate Research Institute (PRI), Kyoto University, participated in Experiment 2 (six females and one male, 19–43 years old, including estimated ages; Table 1). Chimpanzees lived in two adjacent groups with 11 individuals in total. Chimpanzees spent the daytime in outdoor enclosures (the largest one was open-air, approximately 700 m², accessible every

several days a week) with occasional access to indoor enclosures, in which they received meals and slept at night. These enclosures were equipped with complex structures such as climbing towers, platforms, hammocks, ropes, and/or small trees to allow chimpanzees to move and climb freely. Chimpanzees received meals comprising fresh vegetables, fruits, or nutritionally balanced biscuits three times a day. They had free access to water from the taps located in both outdoor and indoor enclosures. Chimpanzees were not deprived of food or water for the studies. The study depended on their voluntary participation. Animal husbandry and study protocols complied with the Guidelines for the Care and Use of Laboratory Primates by the Primate Research Institute, Kyoto University (2019-001).

Stimuli and Apparatus

Chimpanzees were invited into an indoor experimental booth (1.8 m × 2.15 m × 1.75 m) individually. Chimpanzees viewed stimuli through a transparent 2-cm-thick acrylic panel. We basically used the same stimuli and apparatuses as Experiment 1. Nevertheless, sound volume may have been different because of differences in settings between the facilities such as different computers and experimental booths (the volume level was not recorded).

Procedures

Chimpanzees underwent one trial per day for a total of six trials (3 Auditory Stimuli × 2 Video Locations). Other procedures were the same as those of Experiment 1. We repeated one trial for a female chimpanzee (Cleo) after other trials were completed because she left the eye tracker, presumably because she was afraid of the snake video. An experimenter error occurred in one trial with a male participant (Akira), and this trial was repeated on the same day (this might have happened with another participant [Pan] too).

Results and Discussion

Chimpanzees viewed either video type for 10.9 ± 0.92 s ($M \pm SD$) in the alarm call condition, 10.6 ± 1.56 s in the food grunt condition, and 10.8 ± 2.58 s in the no-sound condition. We found a significant main effect of condition, $F(2, 12) = 5.30$, $p = .023$, $\eta^2 = .09$, whereas the other effects were not significant, video location: $F(1, 6) = 0.37$, $p = .563$, $\eta^2 = .01$; Condition × Video Location: $F(2, 12) = 0.11$, $p = .896$, $\eta^2 = .003$. The observed main effect of condition was explained by a higher proportion of snake gazetime when chimpanzees heard the alarm call ($M = .45$; 95% CI [0.38, 0.53]) compared with when they heard no sound ($M = .32$, 95% CI [0.27, 0.36]), $t(6) = 2.52$, $p = .045$, $d = 0.95$ (paired t test; Figure 4), whereas there was no significant difference between when chimpanzees heard the food grunts ($M = .31$, 95% CI [0.25, 0.36]) and when they heard no sound, $t(6) = 0.34$, $p = .747$, $d = -0.13$. The results suggested that the playbacks of alarm calls caused a gaze bias to the snake videos, whereas playbacks of food grunts did not cause a gaze bias to the fruit videos. Thus, when chimpanzees heard the alarm calls, they may have spontaneously matched the alarm call with the image of snakes over the images of fruit. Alternatively, however, the gaze bias to snakes may have simply resulted from the negative affective state elicited by the calls independently of associations with snakes, for example via emotional contagion (Briefer, 2018; de Waal, 2008; Hatfield et al., 1992). The matched negative affective states may have biased the chimpanzees' attention to potential threats via affect-driven attention biases (reviewed in Crump et al., 2018). We addressed this issue in Experiment 3.

Experiment 3

We examined whether presenting screams, an irrelevant call with negative valence, biased chimpanzees' gaze in the same way as an alarm call. Specifically, we tested the same two groups of chimpanzees using the same visual stimuli as Experiments 1 and 2 coupled with either screams or no sound. Chimpanzees would not be expected to look at snake videos for longer when hearing screams than when hearing no sound.

Method

Participants

The same 13 chimpanzees included in Experiments 1 and 2 participated in Experiment 3 (Table 1).

Stimuli and Apparatus

Auditory stimuli were screams (although a finer categorization [as did in Goodall, 1986, Chapter 6] was not made). We recorded screams (0.8 s and 0.9 s were clipped from video recordings and are shown as spectrograms in the right column of Figure 1b) from two of the female chimpanzee participants at KS during possible agonistic interactions in an outdoor enclosure (in one case, the interaction was not clearly observed). The 0.8-s and 0.9-s calls were played repeatedly 10 times. The calls were distributed over the entire duration of the 12-s video at arbitrary intervals (Video S3 in the online supplemental materials). The same composition was used across trials. We presented the same visual stimuli as Experiments 1 and 2 (although the renditions might have differed slightly from those experiments, and between facilities, due to possible differences in setting and version of Xmedia Recode software) coupled with screams or no sound (although the exact volume

level of playback was not recorded at the time, it was estimated to be approximately 70 dBA by post hoc measurement at KS; not recorded at PRI). As with Experiments 1 and 2, each auditory condition composed of two trials with different location of videos (left/ right), and thus Experiment 3 composed of four trials in total.

Procedures

We used the same procedures as those in Experiments 1 and 2. Two trials for a female chimpanzee (Misaki) were repeated after all other trials were completed because she kept fixating on one location on the screen and appeared to be inattentive to the stimulus content. One trial for a male chimpanzee (Zamba) was also repeated because he substantially changed his posture and viewing angle during the trial.

Results and Discussion

Chimpanzees viewed either video type for 10.8 ± 1.63 s ($M \pm SD$) in the scream condition and 10.7 ± 1.33 s in the no-sound condition. We found no significant effects, condition: $F(1, 12) = 0.37$, $p = .555$, $\eta^2 = .007$; video location: $F(1, 12) = 0.39$, $p = .544$, $\eta^2 = .007$; Condition $\times$ Video Location: $F(1, 12) = 0.06$, $p = .813$, $\eta^2 = .0009$ (scream: $M = .33$, 95% CI [0.28, 0.39]; nosound: $M = .30$, 95% CI [0.25, 0.36]; Figure 5). Note that including group (i.e., KS/PRI) as a between-participants variable did not change the pattern of results.

These findings suggest that chimpanzees did not look at the images of snakes just because the alarm call evoked negative affective states and thereby biased chimpanzees' attention to threatening stimuli via affect-driven attention biases. Nonetheless, we could not completely rule out the possibility that prior experiences with similar stimuli may have affected the results. That is, chimpanzees had experienced other conditions (Experiments 1 for KS chimpanzees; Experiment 2 for PRI chimpanzees) before the scream condition (Experiment 3), and those prior experiences with stimuli might have affected the results of Experiment 3, including the absence of the gaze bias to snake stimuli. However, this possibility seems unlikely to explain all of the results, considering that the playback of alarm calls elicited a gaze bias in KS chimpanzees (Experiment 1), who had experienced similar experiments prior to Experiment 1 (in the online supplemental materials), as well as in PRI chimpanzees (Experiment 2), who had only a few such experiences, which suggests that prior experiences had only a minimal impact.

General Discussion

Chimpanzees looked at videos of snakes for longer when they heard alarm calls compared with when they heard food grunts (Experiment 1) or no sound (Experiment 2). These results indicated that when chimpanzees heard alarm calls, they spontaneously matched the call with images of a potential threat over non-threat related images. Experiment 3 showed that playback of screams did not bias chimpanzees' gaze toward the snake video. Therefore, chimpanzees may not have looked at the images of snakes just because the alarm call evoked negative affective states, independently of associations with snakes (e.g., emotional contagion: Briefer, 2018), which biased chimpanzees' attention to threatening stimuli via affect-driven attention biases. These results are in line with the interpretation given by Shibasaki et al. (2014) that audience animals recollected a representation of a snake through an association.

However, several limitations may have affected our results and interpretations. First, chimpanzees did not look at the fruit videos for longer when they heard food grunts compared with baseline (Experiment 2). This finding may have occurred simply because chimpanzees were already biased to look at the fruit videos, as observed in the no-sound baseline condition in Experiment 2 (and in a follow-up study of the findings of Experiment 1, as summarized in the online supplemental materials). An overall bias toward the fruit videos may have made it difficult to detect further gaze bias in the same direction. Second, it remains unclear whether chimpanzees could recall the image of a potential threat without any presented images. We presented participants with two images side-by-side while playing back calls simultaneously. Thus, it is possible that chimpanzees may have compared the presented options and inferred only the most relevant item among them (Adachi et al., 2007; for a relevant finding with human infants, see Wu et al., 2017). Future experiments in which an auditory stimulus is played before showing the visual stimulus may be useful to more strictly examine whether calls evoke a representation of a specific object (Suzuki, 2016, 2018).

Third, although the results revealed that chimpanzees did not match the alarm call with the image of snakes simply based on the negative valence of the stimuli, it remains possible that more subtle differences in affect biased their attention via emotional contagion. That is, alarm calls and screams might have evoked subtly different affective states, if any, which may be closely matched with those of the respective callers, thereby biasing chimpanzees' attention toward the items that elicited the calls. Currently, however, there seems to be little evidence suggesting such fine-tuned state-matching via vocal contagion in nonhuman animals (previous findings are reviewed in Briefer, 2018). Nonetheless, further studies are necessary to test this possibility.

Fourth, it should be noted that we used a single composition for each class of auditory and visual stimuli, although each consisted of several exemplars (i.e., four for snake and fruit videos; two for alarm call, food call, and scream). Thus, this experimental design, with the corresponding statistical analyses, suffered from the issue of pseudoreplication (Hurlbert, 1984; Kroodsma et al., 2001; McGregor, 2000). Thus, we could not infer with confidence to what extent our results are generalizable to the objects, or the call types in general (e.g., Kroodsma, 1989a, 1989b; McGregor, 2000; Wiley, 2003). Ideally, future studies should use a sufficiently large number of exemplars for each class of stimuli. In an additional analysis presented in the online supplemental materials, we examined the effects of different visual exemplars on the observed gaze responses to snakes. Specifically, we divided each trial into four time bins, each approximately corresponding to each 3-s video pair (stills are shown in Figure 1a), and compared the proportion of time spent looking at snake videos across conditions and time bins via repeated-measures analysis of variance. These analyses indicated that the proportion of snake gaze time differed across time bins. However, the differential gaze responses evoked by alarm calls were not due to a particular time bin (and possibly video pair). We did not carry out a similar analysis for the auditory stimuli because the two calls were played with short intervals between each other and sometimes even overlapped, which hindered segmentation of each trial into bins per call sequence (Videos S1 to S3 in the online supplemental materials).

In conclusion, our results revealed that chimpanzees spontaneously matched an alarm call with threat-related over non-threat related stimuli in a cross-modal preferential-looking experiment. It is possible that such audiovisual cross-modal associations indicate the existence of mental representation of the referent that guides flexible adjustments of behavior in both callers and audiences. Further studies will be necessary to test this possibility in depth.

References

- Adachi, I., Kuwahata, H., & Fujita, K. (2007). Dogs recall their owner's face upon hearing the owner's voice. *Animal Cognition*, 10(1), 17–21. <https://doi.org/10.1007/s10071-006-0025-8>
- Baguley, T. (2012). Calculating and graphing within-subject confidence intervals for ANOVA. *Behavior Research Methods*, 44(1), 158–175. <https://doi.org/10.3758/s13428-011-0123-7>
- Briefer, E. F. (2018). Vocal contagion of emotions in non-human animals. *Proceedings of the Royal Society B: Biological Sciences*, 285(1873), Article 20172783. <https://doi.org/10.1098/rspb.2017.2783>
- Brosnan, S. F., & de Waal, F. B. M. (2000). Regulation of vocal output by chimpanzees finding food in the presence or absence of an audience. *Evolution of Communication*, 4(2), 211–224. <https://doi.org/10.1075/eoc.4.2.05bro>
- Clay, Z., & Zuberbühler, K. (2011). Bonobos extract meaning from call sequences. *PLoS ONE*, 6(4), Article e18786. <https://doi.org/10.1371/journal.pone.0018786>
- Crockford, C., & Boesch, C. (2003). Context-specific calls in wild chimpanzees, *Pan troglodytes verus*: Analysis of barks. *Animal Behaviour*, 66(1), 115–125. <https://doi.org/10.1006/anbe.2003.2166>
- Crockford, C., Wittig, R. M., Mundry, R., & Zuberbühler, K. (2012). Wild chimpanzees inform ignorant group members of danger. *Current Biology*, 22(2), 142–146. <https://doi.org/10.1016/j.cub.2011.11.053>
- Crockford, C., Wittig, R. M., & Zuberbühler, K. (2015). An intentional vocalization draws others' attention: A playback experiment with wild chimpanzees. *Animal Cognition*, 18(3), 581–591. <https://doi.org/10.1007/s10071-014-0827-z>
- Crockford, C., Wittig, R. M., & Zuberbühler, K. (2017). Vocalizing in chimpanzees is influenced by social-cognitive processes. *Science Advances*, 3(11), Article e1701742. <https://doi.org/10.1126/sciadv.1701742>
- Crump, A., Arnott, G., & Bethell, E. J. (2018). Affect-driven attention biases as animal welfare indicators: Review and methods. *Animals*, 8(8), Article 136. <https://doi.org/10.3390/ani8080136>
- Deloache, J. S., & Lobue, V. (2009). The narrow fellow in the grass: Human infants associate snakes and fear. *Developmental Science*, 12(1), 201–207. <https://doi.org/10.1111/j.1467-7687.2008.00753.x>
- de Waal, F. B. M. (2008). Putting the altruism back into altruism: The evolution of empathy. *Annual Review of Psychology*, 59, 279–300. <https://doi.org/10.1146/annurev.psych.59.103006.093625>
- Dezecache, G., & Berthet, M. (2018). Working hypotheses on the meaning of general alarm calls. *Animal Behaviour*, 142, 113–118. <https://doi.org/10.1016/j.anbehav.2018.06.008>
- Dezecache, G., Crockford, C., & Zuberbühler, K. (2019). The development of communication in alarm contexts in wild chimpanzees. *Behavioral Ecology and Sociobiology*, 73(8), Article 104. <https://doi.org/10.1007/s00265-019-2716-6>
- Evans, C. S., & Evans, L. (2007). Representational signalling in birds. *Biology Letters*, 3(1), 8–11. <https://doi.org/10.1098/rsbl.2006.0561>
- Fedurek, P., & Slocombe, K. E. (2013). The social function of food-associated calls in male chimpanzees. *American Journal of Primatology*, 75(7), 726–739. <https://doi.org/10.1002/ajp.22122>
- Gibbons, C. (2007). *The referentiality of chimpanzee vocal signaling: behavioral and acoustic analysis of food barks* [Electronic Thesis or Dissertation, Ohio State University].
- Girard-Buttoz, C., Surbeck, M., Samuni, L., Tkaczynski, P., Boesch, C., Fruth, B., Wittig, R. M., Hohmann, G., & Crockford, C. (2020). Information transfer efficiency differs in wild chimpanzees and bonobos, but not social cognition. *Proceedings of the Royal Society Biological Sciences*, 287(1929), Article 20200523. <https://doi.org/10.1098/rspb.2020.0523>
- Goodall, J. (1986). *The chimpanzees of Gombe: Patterns of behavior* (pp. 117–145). Harvard University Press.
- Hallberg, K. I., Nelson, D. A., & Boysen, S. T. (2003). Representational vocal signaling in the chimpanzee. In F. B. M. de Waal & P. L. Tyack (Eds.), *Animal social complexity: Intelligence, culture, and individualized societies* (pp. 317–321). Harvard University Press. <https://doi.org/10.4159/harvard.9780674419131.c23>
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1992). Primitive emotional contagion. In M. S. Clark (Ed.), *Review of personality and social psychology: Emotion and social behavior* (Vol. 14, pp. 151–177). Sage, Inc.
- Hauser, M. D., Teixidor, P., Field, L., & Flaherty, R. (1993). Food-elicited calls in chimpanzees: Effects of food quantity and divisibility. *Animal Behaviour*, 45(4), 817–819. <https://doi.org/10.1006/anbe.1993.1096>
- Hauser, M. D., & Wrangham, R. W. (1987). Manipulation of food calls in captive chimpanzees: A preliminary report. *Folia Primatologica*, 48(3–4), 207–210. <https://doi.org/10.1159/000156298>
- Hopper, L. M., Gulli, R. A., Howard, L. H., Kano, F., Krupenye, C., Ryan, A. M., & Paukner, A. (2021). The application of noninvasive, restraint-free eye-tracking methods for use with nonhuman primates. *Behavior Research Methods*, 53(3), 1003–1030. <https://doi.org/10.3758/s13428-020-01465-6>
- Howell, D. C. (2010). *Statistical methods for psychology* (7th ed., pp. 461–513). Cengage Learning.
- Hurlbert, S. H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological Monographs*, 54(2), 187–211. <https://doi.org/10.2307/1942661>
- Idani, G., & Hirata, S. (2006). Studies at the Great Ape Research Institute. In D. A. Washburn (Ed.), *Primate perspectives on behavior and cognition* (pp. 29–36). American Psychological Association.
- Kalan, A. K., & Boesch, C. (2015). Audience effects in chimpanzee food calls and their potential for recruiting others. *Behavioral Ecology and Sociobiology*, 69(10), 1701–1712. <https://doi.org/10.1007/s00265-015-1982-1>
- Kalan, A. K., Mundry, R., & Boesch, C. (2015). Wild chimpanzees modify food call structure with respect to tree size for a particular fruit species. *Animal*

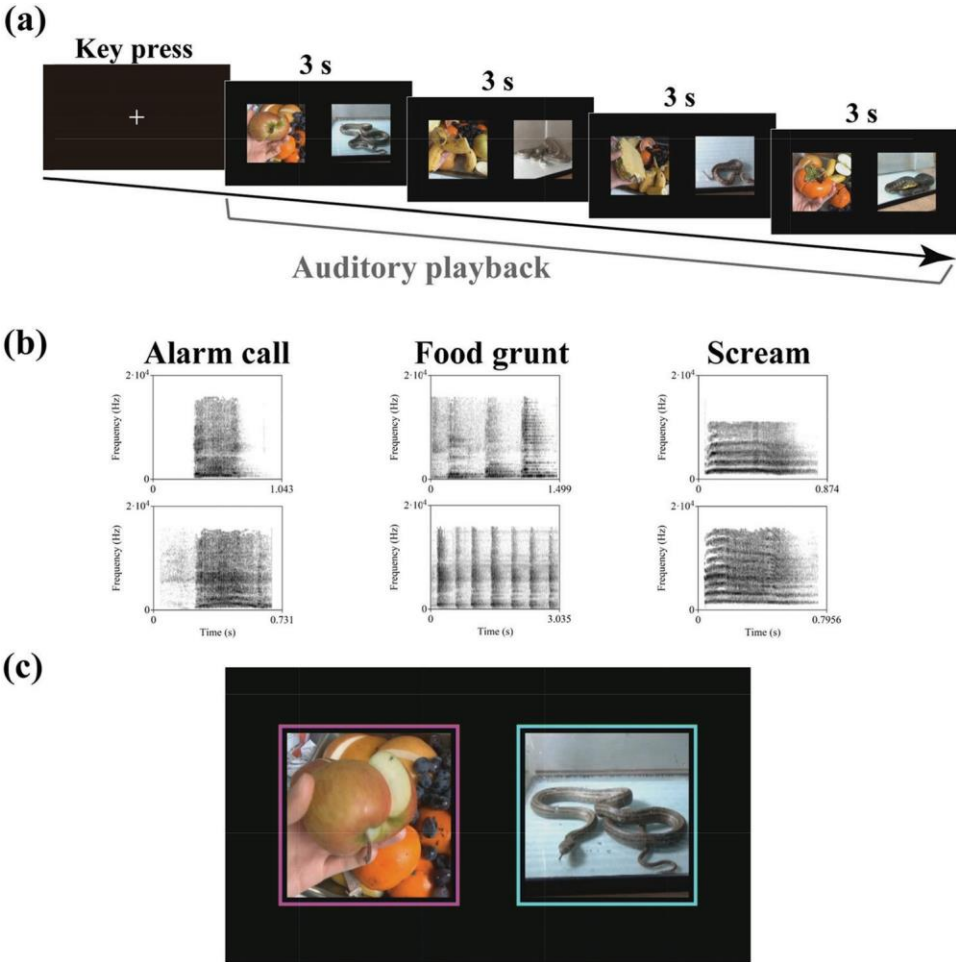
- Behaviour*, 101, 1–9. <https://doi.org/10.1016/j.anbehav.2014.12.011>
- Kroodtsma, D. E. (1989a). Inappropriate experimental designs impede progress in bioacoustic research: A reply. *Animal Behaviour*, 38(4), 717–719. [https://doi.org/10.1016/S0003-3472\(89\)80020-X](https://doi.org/10.1016/S0003-3472(89)80020-X)
- Kroodtsma, D. E. (1989b). Suggested experimental designs for song play backs. *Animal Behaviour*, 37(4), 600–609. [https://doi.org/10.1016/0003-3472\(89\)90039-0](https://doi.org/10.1016/0003-3472(89)90039-0)
- Kroodtsma, D. E., Byers, B. E., Goodale, E., Johnson, S., & Liu, W. C. (2001). Pseudoreplication in playback experiments, revisited a decade later. *Animal Behaviour*, 61(5), 1029–1033. <https://doi.org/10.1006/anbe.2000.1676>
- Leroux, M., Bosshard, A. B., Chandia, B., Manser, A., Zuberbühler, K., & Townsend, S. W. (2021). Chimpanzees combine pant hoots with food calls into larger structures. *Animal Behaviour*, 179, 41–50. <https://doi.org/10.1016/j.anbehav.2021.06.026>
- Macedonia, J. M., & Evans, C. S. (1993). Essay on contemporary issues in ethology: Variation among mammalian alarm call systems and the problem of meaning in animal signals. *Ethology*, 93(3), 177–197. <https://doi.org/10.1111/j.1439-0310.1993.tb00988.x>
- Marler, P., & Tenaza, R. (1977). Signaling behavior of apes with special reference to vocalization. In T. A. Sebeok (Ed.), *How animals communicate* (pp. 965–1033). Indiana University Press.
- Matsuzawa, T. (2006). Sociocognitive development in chimpanzees: A synthesis of laboratory work and fieldwork. In T. Matsuzawa, M. Tomonaga, & M. Tanaka (Eds.), *Cognitive development in chimpanzees* (pp. 3–33). Springer. https://doi.org/10.1007/4-431-30248-4_1
- McGregor, P. K. (2000). Playback experiments: Design and analysis. *Acta Ethologica*, 3(1), 3–8. <https://doi.org/10.1007/s102110000023>
- O'Bryan, L. R. (2015). *The socioecology of chimpanzee foraging and food-associated calling behavior* [Doctoral dissertation, University of Minnesota].
- O'Bryan, L. R., Lambeth, S. P., Schapiro, S. J., & Wilson, M. L. (2021). Playbacks of food-associated calls attract chimpanzees towards known food patches in a captive setting. *Primates*. Advance online publication. <https://doi.org/10.1007/s10329-021-00936-4>
- R Core Team. (2020). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing.
- Schel, A. M., Townsend, S. W., Machanda, Z., Zuberbühler, K., & Slocombe, K. E. (2013). Chimpanzee alarm call production meets key criteria for intentionality. *PLoS ONE*, 8(10), Article e76674. <https://doi.org/10.1371/journal.pone.0076674>
- Seyfarth, R. M., & Cheney, D. L. (1990). The assessment by vervet monkeys of their own and another species' alarm calls. *Animal Behaviour*, 40(4), 754–764. [https://doi.org/10.1016/S0003-3472\(05\)80704-3](https://doi.org/10.1016/S0003-3472(05)80704-3)
- Seyfarth, R. M., & Cheney, D. L. (2003). Meaning and emotion in animal vocalizations. *Annals of the New York Academy of Sciences*, 1000(1), 32–55. <https://doi.org/10.1196/annals.1280.004>
- Shibasaki, M., Nagumo, S., & Koda, H. (2014). Japanese monkeys (*Macaca fuscata*) spontaneously associate alarm calls with snakes appearing in the left visual field. *Journal of Comparative Psychology*, 128(3), 332–335. <https://doi.org/10.1037/a0036049>
- Slocombe, K. E., Kaller, T., Turman, L., Townsend, S. W., Papworth, S., Squibbs, P., & Zuberbühler, K. (2010). Production of food-associated calls in wild male chimpanzees is dependent on the composition of the audience. *Behavioral Ecology and Sociobiology*, 64(12), 1959–1966. <https://doi.org/10.1007/s00265-010-1006-0>
- Slocombe, K. E., & Zuberbühler, K. (2005). Functionally referential communication in a chimpanzee. *Current Biology*, 15(19), 1779–1784. <https://doi.org/10.1016/j.cub.2005.08.068>
- Slocombe, K. E., & Zuberbühler, K. (2006). Food-associated calls in chimpanzees: Responses to food types or food preferences? *Animal Behaviour*, 72(5), 989–999. <https://doi.org/10.1016/j.anbehav.2006.01.030>
- Suzuki, T. N. (2016). Semantic communication in birds: Evidence from field research over the past two decades. *Ecological Research*, 31(3), 307–319. <https://doi.org/10.1007/s11284-016-1339-x>
- Suzuki, T. N. (2018). Alarm calls evoke a visual search image of a predator in birds. *Proceedings of the National Academy of Sciences of the United States of America*, 115(7), 1541–1545. <https://doi.org/10.1073/pnas.1718884115>
- Suzuki, T. N. (2020). Other species' alarm calls evoke a predator-specific search image in birds. *Current Biology*, 30(13), 2616–2620.e2. <https://doi.org/10.1016/j.cub.2020.04.062>
- Townsend, S. W., & Manser, M. B. (2013). Functionally referential communication in mammals: The past, present and the future. *Ethology*, 119(1), 1–11. <https://doi.org/10.1111/eth.12015>
- Wiley, R. H. (2003). Is there an ideal behavioural experiment? *Animal Behaviour*, 66(3), 585–588. <https://doi.org/10.1006/anbe.2003.2231>
- Wheeler, B. C., & Fischer, J. (2012). Functionally referential signals: A promising paradigm whose time has passed. *Evolutionary Anthropology*, 21(5), 195–205. <https://doi.org/10.1002/evan.21319>
- Wu, Y., Muentener, P., & Schulz, L. E. (2017). One- to four-year-olds connect diverse positive emotional vocalizations to their probable causes. *Proceedings of the National Academy of Sciences of the United States of America*, 114(45), 11896–11901. <https://doi.org/10.1073/pnas.1707715114>
- Zuberbühler, K. (2000). Causal cognition in a non-human primate: Field playback experiments with Diana monkeys. *Cognition*, 76(3), 195–207. [https://doi.org/10.1016/S0010-0277\(00\)00079-2](https://doi.org/10.1016/S0010-0277(00)00079-2)
- Zuberbühler, K., Cheney, D. L., & Seyfarth, R. M. (1999). Conceptual semantics in a nonhuman primate. *Journal of Comparative Psychology*, 113(1), 33–42. <https://doi.org/10.1037/0735-7036.113.1.33>

Table 1
Chimpanzee Participant Information

Site	Name	Sex	Age	GAIN ID ^a
Kumamoto Sanctuary	Hatsuka	F	11	0704
(Experiments 1 and 3)	Iroha	F	10-11	0708
	Misaki	F		0593
	Mizuki	F	22	0559
Primate Research Institute (Experiments 2 and 3)	Natsuki	F	14	0677
	Zamba	M	24	0543
	Ai	F	43 ^b	0434
	Akira	M	43 ^b	0435
	Chloe	F	38	0441
	Cleo	F	19	0609
	Pal	F	19	0611
	Pan	F	35	0440
	Pendesa	F	42	0095

Note. In the column Sex, “F” and “M” denote “Female” and “Male,” respectively.
^a GAIN ID is the identification from the Great Ape Information Network (retrieved from <https://shigen.nig.ac.jp/gain/index.jsp>).
^b Estimated ages.

Figure 1
 Illustrations for Materials and Method

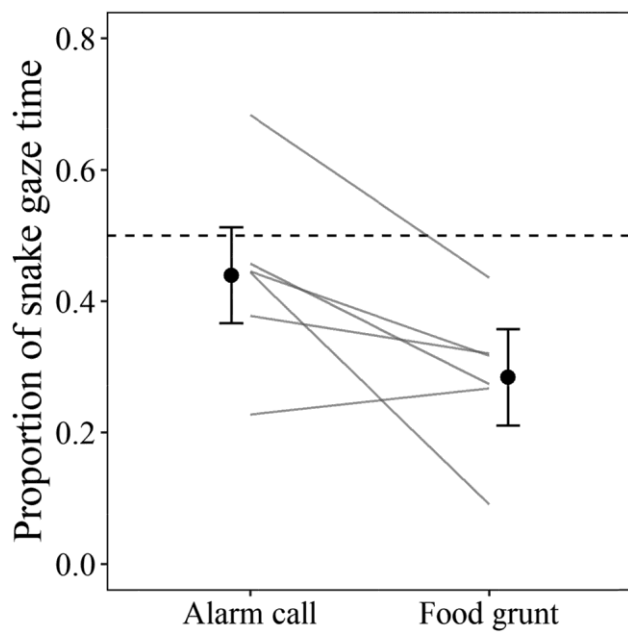


Note. (a) Trial flow. (b) Spectrograms of auditory stimuli. We used PRAAT 6.1.04 to construct spectrograms with the following parameters (window length = 0.005 s, maximum frequency = 20,000 Hz, time step = 0.002 s, frequency step = 20 Hz, window shape = Gaussian, dynamic range = 50 dB). (c) The magenta (dark gray) and cyan (light gray) squares show the areas of interest in analysis and are included in the figure for illustrative purposes only.

Figure 2
Experimental Setting at Kumamoto Sanctuary

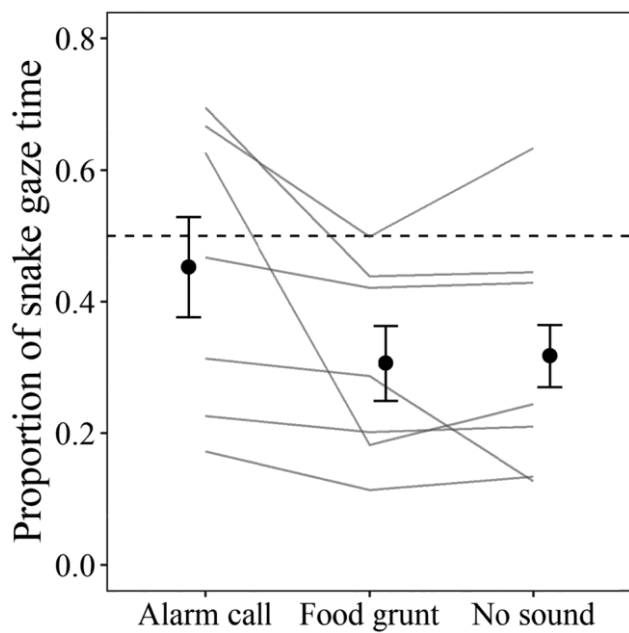


Figure 3
Proportion of Time Spent Looking Toward the Snake Video Stimulus in Experiment 1 (n = 6)



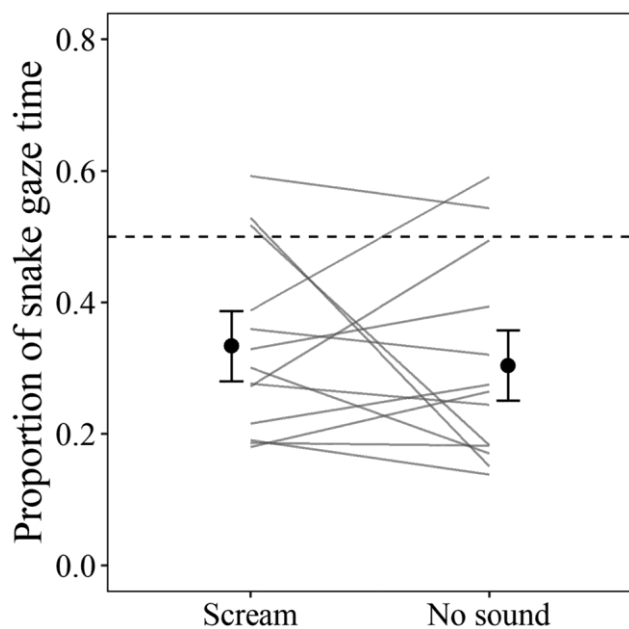
Note. The gray lines represent the individual data (average of two trials). The black dots represent the mean, and error bars represent the 95% CI.

Figure 4
Proportion of Time Spent Looking Toward the Snake Video Stimulus in Experiment 1 (n = 7)



Note. The gray lines represent the individual data (average of two trials). The black dots represent the mean, and error bars represent the 95% CI.

Figure 5
Proportion of Time Spent Looking Toward Snake Video in Experiment 3 ($n = 13$, Data From Two Groups Were Pooled [Kumamoto Sanctuary and Primate Research Institute])



Note. The gray lines represent the individual data (average of two trials). The black dots represent the mean, and the error bars represent the 95% CI.