



Evaluating Mood Changes in Response to Anthropogenic Noise with a Response-Slowing Task in Three Species of Zoo-housed Primates

Katherine A. Cronin¹, Emily J. Bethell², Sarah L. Jacobson¹, Crystal Egelkamp¹,
Lydia M. Hopper¹, and Stephen R. Ross¹

¹ Lester E. Fisher Center for the Study and Conservation of Apes, Lincoln Park Zoo, 2001 N. Clark Street, Chicago, IL 60614

² Research Centre for Brain and Behaviour, School of Natural Sciences and Psychology, Liverpool John Moores University, Byrom Street, Liverpool L3 3AF, UK

*Corresponding author (Email: KCronin@lpzoo.org)

Citation – Cronin, K. A., Bethell, E. J., Jacobson, S. L., Egelkamp, C., Hopper, L. M., & Ross, S. R. (2018). Evaluating mood changes in response to anthropogenic noise with a response-slowing task in three species of zoo-housed primates. *Animal Behavior and Cognition*, 5(2), 209–221. <https://doi.org/10.26451/abc.05.02.03.2018>

Abstract – In the zoo environment, anthropogenic noise is common as sound levels fluctuate due to visitors, construction, habitat design, and special events. In this study, changes in the mood of three species of zoo-housed primates in response to a loud annual event were evaluated with the response-slowing paradigm. In this paradigm, animals experiencing anxiety slow responses on simple cognitive tasks when emotional content is displayed. Following a previously validated approach, we measured latencies to touch potentially threatening (conspecific faces with directed gaze) and non-threatening (conspecific faces with averted gaze) images overlaid on a grey square, relative to neutral control images (grey squares only) on a touchscreen. In Experiment 1, four Japanese macaques (*Macaca fuscata*) were tested in two conditions: during a baseline (non-stressful) period and opportunistically during three days during which loud jets frequently flew overhead. Results indicated a significant effect of condition, with an increase in latency to touch images of conspecific faces relative to control images during the days of the loud event. In Experiment 2, chimpanzees (*Pan troglodytes*, $n = 4$) and western lowland gorillas (*Gorilla gorilla gorilla*, $n = 2$) were tested during the same loud event following a similar methodology. The results revealed subtle changes across conditions; however, this was likely driven by the apes increasing their response speed to face stimuli relative to control stimuli over time (habituation). These findings suggest that the macaques, but not the apes, underwent detectable affective changes during the loud event. With additional development, this relatively simple paradigm may be an effective and feasible way to evaluate real-time changes in the mood of zoo-housed animals.

Keywords: Response slowing, Zoo, Mood, Affect, *Macaca fuscata*, *Gorilla gorilla*, *Pan troglodytes*, Noise

Anthropogenic noise can influence animal behavior, physiology and well-being through effects on the auditory system, interference with sounds important to survival and reproduction, and generation of startle or fear-related responses (Blickley & Patricelli, 2010; Wright et al., 2007). Anthropogenic noises are often more frequent, louder, and less predictable than natural (non-anthropogenic) acoustic stimuli (Kight & Swaddle, 2011). In the zoo environment anthropogenic noise is common, as sound levels may increase due to visitors, construction, habitat design, and special events. Past studies have evaluated how various sources of noise in the zoo environment influence behavioral and physiological indicators of welfare in a variety of species (e.g., Birke, 2002; Chosy, Wilson, & Santymire, 2014; Davey,

2007; Kight & Swaddle, 2011; Orban, Soltis, Perkins, & Mellen, 2017; Quadros, Goulart, Passos, Vecchi, & Young, 2014).

There is growing consensus that one of the essential contributors to an animal's welfare state is their subjective, or affective experience (Dawkins, 2015; McGuire, Vonk, Fuller, & Allard, 2017; Mellor, 2015; Whitham & Wielebnowski, 2013). However, to date, there has been no research directly investigating the relationship between anthropogenic noise and affective state in zoo-housed animals. Several approaches to evaluating affective states that rely on judgement biases have emerged in recent years (reviewed in Bethell, 2015), and most require extensive training of the animals involved (Deakin, Browne, Hodge, Paul, & Mendl, 2016; Harding, Paul, & Mendl, 2004; Mendl, Burman, Parker, & Paul, 2009; but see Brydges & Hall, 2017). However, one recently-developed paradigm, the "response-slowing paradigm," does not require extensive training and, therefore, is more feasible for use in a zoo environment (Bethell, Holmes, MacLarnon, & Semple, 2016). The response-slowing paradigm is grounded in human psychological research, and specifically, the discovery that people experiencing anxiety show an impairment (slower response time) on simple cognitive tasks when emotionally threatening content is displayed, compared to non-anxious individuals (reviewed in Bar-Haim, Lamy, Pergamin, Bakermans-Kranenburg, & Van Ijzendoorn, 2007; Mogg & Bradley, 2016).

Bethell et al. (2016) demonstrated that the response-slowing paradigm can be used to detect anxious mood in laboratory-housed rhesus macaques (*Macaca mulatta*) who had recently undergone veterinary exams. The authors trained the macaques to touch a plain grey square presented on a touchscreen monitor and then measured changes in latency to touch the square when emotional content (images of conspecifics directing gaze at the subject) was added to the square. Direct gaze is a threatening communicative signal for macaques (Maestriperi, 1997), therefore, the authors presumed that the squares containing conspecific direct gaze images had negative emotional valence. Relative to their performance on control trials, the monkeys were slower to touch the direct gaze face when they had recently undergone a presumably stressful veterinary procedure, compared to less stressful, baseline periods. The authors found no such slowing effect following the veterinary exam for images containing averted (submissive) gaze faces. Together, these findings suggest that response latencies to touch single images presented on a touchscreen may provide a measure of changes in affective state without the need for extensive training. In the following experiments we adapted these methods for use with macaques in a zoo setting (Experiment 1) and then applied a similar methodology for use with other primate species in zoos (chimpanzees and western lowland gorillas, Experiment 2).

Experiment 1

In Experiment 1, we applied the response-slowing paradigm to test whether zoo-housed Japanese macaques (*Macaca fuscata*) experience changes in affect corresponding with a noisy, annual public event, the Chicago Air and Water Show (A&W). This event spans three days and is characterized by loud, low-flying jets intermittently passing directly over the macaques' habitat. Previous work has shown that loud noises such as heavy machinery and vehicles, over which macaques have no control, lead to increases in plasma cortisol, aggression, and other behavioral indicators of stress (Hanson, Larson, & Snowdon, 1976). We hypothesized that the monkeys would experience anxious mood during the days of A&W and show an impairment (slower response time) on a simple cognitive task when emotionally threatening content is displayed, compared to a baseline period.

Method

Subjects. Four Japanese macaques (one male and two females, 9-10 years old, and one male, 1-year old), who were members of a troop of 12 housed at Lincoln Park Zoo (Chicago, USA), voluntarily participated in this study.

Materials and testing environment. The full troop was comprised of three adult males, five adult females, one juvenile male and three infant females. The troop inhabited a large naturalistic outdoor

habitat of 685 m² equipped with natural and artificial trees, bushes, large rocks, a pool, grass and mulch during days with additional indoor space (348 m²) available during nights and times of low temperatures or inclement weather. Fresh produce and monkey chow were scattered daily throughout their habitat and monkeys had access to water *ad libitum*.

Testing took place in one of two touchscreen computer booths integrated into the monkeys' outdoor habitat (Figure 1). Stimuli were presented on a 22" Viewsonic TD2240 touch-sensitive monitor connected to a personal computer in the adjacent researcher area. A stimulus consisting of a grey rectangular frame measuring 10.2 X 12.7 cm (width X height on screen) served as the control stimulus. Test stimuli consisted of color photographs of Japanese macaques obtained from the Internet (non-copyrighted images from www.flickr.com and Google image search). Pictures were selected for neutral expressions directing gaze toward the camera (12 pictures and their mirror image resulting in 24 stimuli) or averting their gaze from the camera (12 pictures and their mirror image resulting in 24 stimuli). In Adobe Photoshop CS4, images were trimmed so that only the macaque's head was visible and superimposed onto the grey stimulus used for control trials (resulting in control and test images of equal size). Luminosity and contrast energy were obtained for each image and there were no significant differences in either measure between directed and averted stimuli sets (two-sample t-test, luminosity $t(22) = 0.484, p = .69$; contrast $t(22) = 0.136, p = .89$).

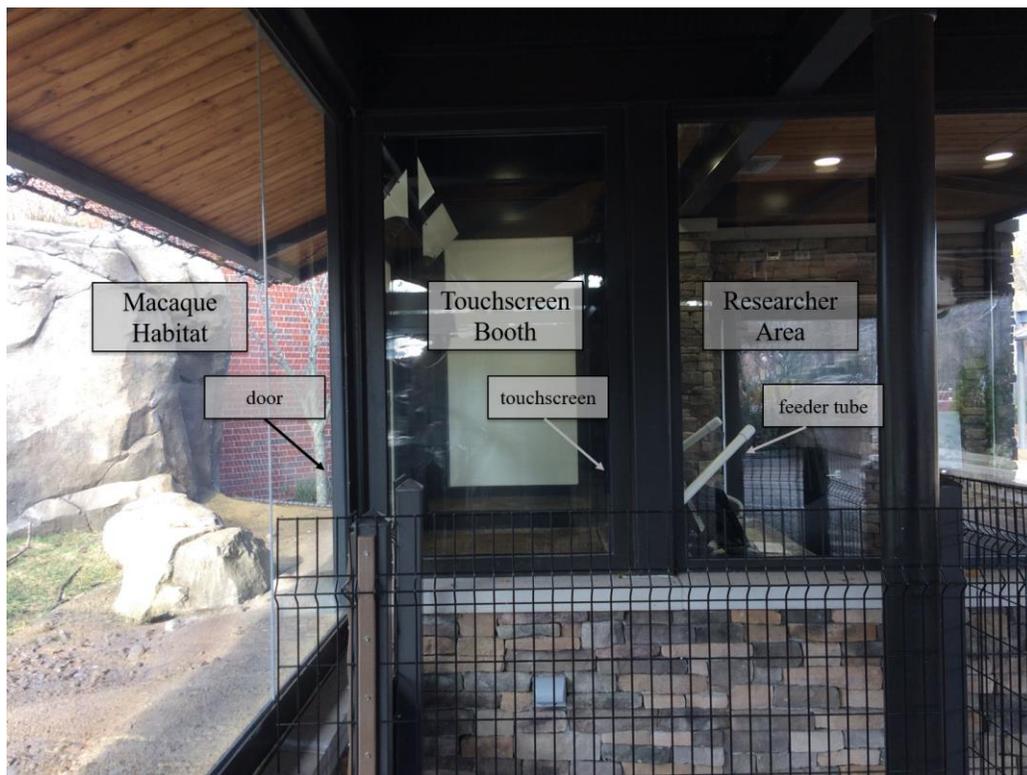


Figure 1. Touchscreen computer booths integrated into the Japanese macaque habitat at the Lincoln Park Zoo.

Procedure. All participation was voluntary; the macaques could enter or exit the booth through a hinged hanging door at any time during the test sessions that took place weekdays between 11:20 a.m. and 1:00 p.m. Sessions were paused when more than one monkey was present in the booth to avoid aggression and allow clear identification of the participant.

Prior to participation in this study, monkeys were trained to touch a single dot when it appeared on the screen in order to receive a food reward (using Zenrichment ApeTouch software). This training began in March 2015. PsychoPy version 1.83.04 (Peirce, 2009) was used to program and run the present experiment (Psychopy software is free; experiment code available upon request).

Following the methods of Bethell et al. (2016), on each trial, monkeys were presented with one image: either the control stimulus (grey square), a grey square containing a direct-gaze stimulus, or a grey square containing an averted-gaze stimulus. The trial types (control, direct gaze, averted gaze) were presented in a ratio of 1:2:2 in a random order such that the control stimulus was shown on average half as frequently as the other trial types. On direct-gaze and averted-gaze trials, exemplars were selected randomly without replacement from the 24 available stimuli of each type. The image was vertically centered on a black screen and randomized and counterbalanced between a central, left and right position (Figure 2). We chose the 1:2:2 ratio to follow the validated methods of Bethell et al. (2016). Different locations were used to safeguard against position biases. The maximum trial length that a single image was on the screen was 60 s and the inter-trial interval was 8 s, during which time a black screen was shown. If a monkey left the booth mid-trial, the aborted trial was excluded from analyses.

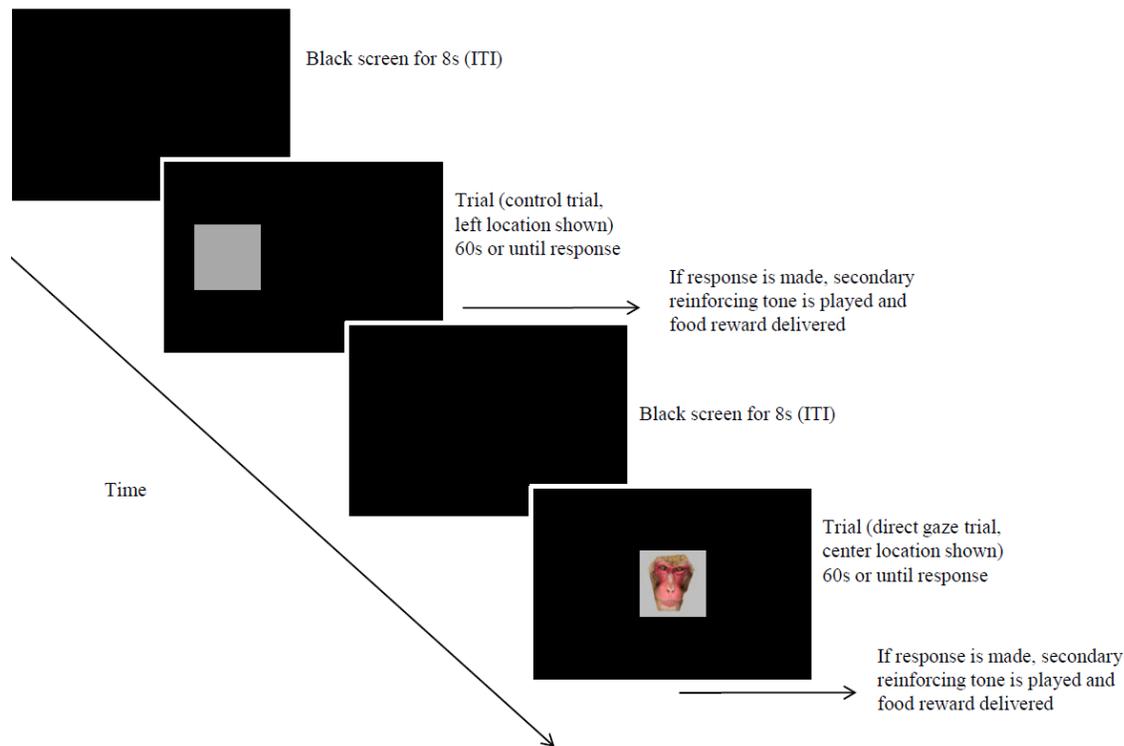


Figure 2. Example of the experimental procedure.

The identification of the participating monkey was manually entered into the software prior to their first trial upon each booth entry. Monkeys were rewarded for touching the stimulus with approximately one eighth of a peanut and a secondary reinforcement tone. Rewards were delivered manually on a 100% reinforcement schedule via a PVC tube that extended from the experimenter area into the touchscreen booths (Figure 1). Response latencies (time elapsed between presentation of stimulus and touch) were recorded automatically by the computer. Given that participation was voluntary and monkeys were free to enter and exit the booth, there was no set number of trials per day, but a 50-trial maximum per day per subject was imposed, after which no additional trials appeared on the screen.

Subjects were tested during a baseline condition and during three days of the 2016 A&W during which loud, low-flying jets passed over the habitat between seven and 80 times each day between the hours of 10:00 a.m. and 4:00 p.m. The monkeys had been exposed to the jet noise only once previously (2015) during which time we observed fleeing, defecating and hiding, leading us to plan to test changes in affect during the 2016 A&W. During the A&W test sessions, monkeys did not voluntarily participate at

the exact times that jets flew overhead, as they tended to gather and seek cover in their habitat during those moments. For three of the four subjects the baseline condition preceded A&W and for one subject the baseline period followed A&W. The number of baseline days experienced by each monkey varied based on their rate of voluntary participation, as we stopped baseline data collection for an individual once 200 trials were obtained (range 5-7 days). There was a minimum of one month (31 days) between baseline data collection and A&W data collection. The identity of the researcher was constant across conditions.

This study was approved by the Lincoln Park Zoo Research Committee, the governing body for all animal research at the institution. No modifications were made to standard animal care routines and the A&W was outside the control of Lincoln Park Zoo. This research adhered to legal requirements in the United States of America and to the American Society of Primatologists' Principles for the Ethical Treatment of Nonhuman Primates.

Statistical analysis. Analyses were conducted in R version 3.3.1 (R Core Team, 2014). Histograms of response latencies were visualized and latency data were trimmed so that responses greater or less than 2.5 standard deviations from each participant's mean were excluded in order to remove outliers (Ratcliff, 1993). Data were also analyzed with trims of 2.0 and 3.0 SD and results did not differ from those reported below. The histogram revealed a positive skew in the distribution; therefore, data were normalized using a \log_{10} transformation. Reaction time ratio scores for each subject were calculated for direct gaze and averted gaze trials following Bethell et al. (2016). Specifically, the reaction time ratio (RTRatio) for each gaze trial was calculated as \log_{10} latency gaze trial/mean \log_{10} latency control trial, where the denominator was calculated separately for each monkey in each condition to account for differences in response speeds that may be due to attention or arousal. RTRatios > 1 reveal slowing of responses toward faces relative to control trials (i.e., the grey square), while ratio scores < 1 reveal speeding of responses toward faces relative to control trials.

Data were analyzed using a linear mixed effects model that included subject and unique image ID as random effects and condition (baseline vs A&W), trial type (directed gaze vs averted gaze), and the interaction between condition and trial type as fixed effects to predict the dependent variable RTRatio. We also included by-subject random slopes for the effect of RTRatio (Barr, Levy, Scheepers, & Tily, 2013). Likelihood ratio tests using the anova function and chi-square distribution were used to compare the full model with null models excluding each variable of interest (condition, trial type, and their interaction). The assumptions of linearity and the absence of heteroscedasticity were examined through plotting residuals, the assumption of normality through visualization of a Q-Q plot, and the absence of collinearity of predictors was assumed from model design. Mixed-effects models were run using the lme4 package (Bates, Maechler, & Bolker, 2014).

Results

Each monkey participated in 5-7 baseline test days (completing 200 trials, with the exception of one monkey that completed only 100 trials due to computer error), and 1-2 days during A&W (completing 23, 27, 55 or 100 trials). No monkey reached the maximum trial length (60 s) on any trial. There was no evidence of habituation over sequential trials during the baseline period (Pearson correlation, trial number and logRT, separately by subject and excluding control trials, all $p > .10$). The full-null model comparison examining the effect of condition (baseline vs A&W) was significant ($\chi^2(3) = 9.32, p = .03$), and the full-null model comparison examining the effect of trial type (averted vs directed) was not significant ($\chi^2(3) = 2.33, p = .13$). Full-null model comparisons revealed no significant interaction between condition and trial type ($\chi^2(1) = 0.05, p = .83$). Complete model results are available in the [Supplemental Information](#). RTRatios were greater during A&W (mean RTRatio = 1.57) than at baseline (mean RTRatio = 0.98), revealing slowing of responses on experimental trials relative to control trials (Figure 3).

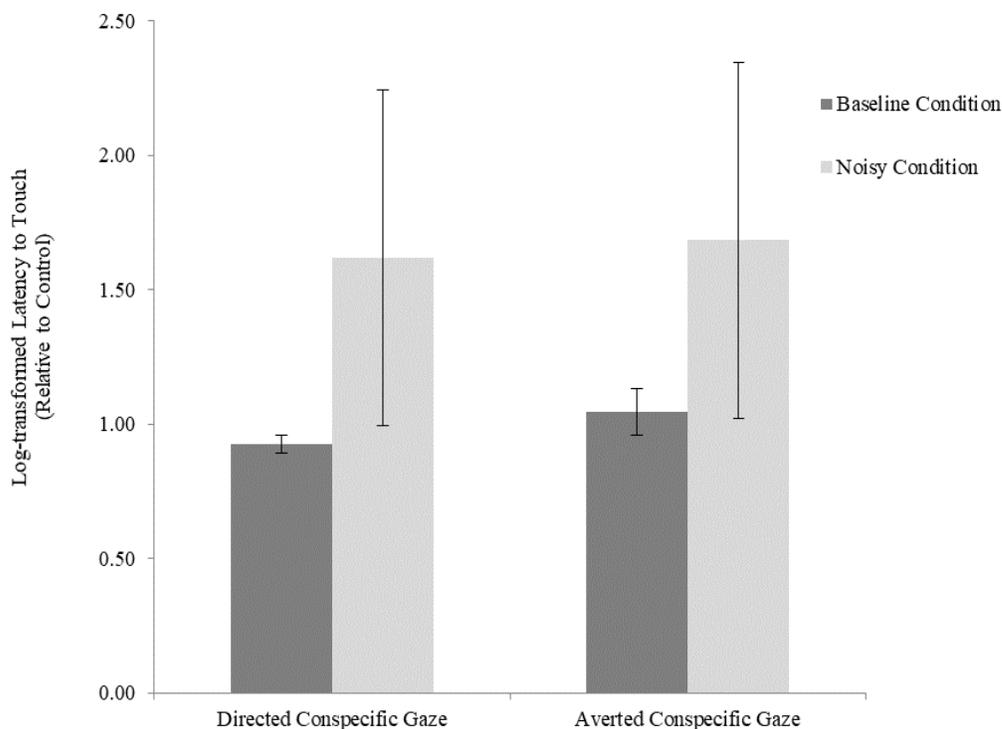


Figure 3. Mean reaction time ratios (\pm SEM) for Japanese macaques across baseline and noisy conditions for both trial types. The reaction time ratios are standardized by the latency to touch control images in each condition, therefore, values > 1 indicate that subjects touched conspecific images more slowly than control images, and values < 1 indicate that subjects touched conspecific images more quickly than control images.

Discussion

In Experiment 1, Japanese macaques demonstrated an increase in their latency to touch stimuli containing images of conspecific faces relative to control stimuli lacking conspecific faces during the days of a loud event. These findings suggest that the macaques experienced changes in affect during A&W. We also predicted that the macaques would respond differently to the directed and averted faces, presuming the first would be emotionally threatening and, therefore, elicit a response-slowing effect while the latter would not. However, in Experiment 1 the comparison between directed and averted faces was not significant. Furthermore, we are limited in our ability to consider expected changes in response times over time given that only one of the four subjects experienced a post-test baseline period. In Experiment 2, we expand the scope of the study to assess response slowing in response to the same loud event in zoo-housed chimpanzees (*Pan troglodytes*) and gorillas (*Gorilla gorilla gorilla*).

Experiment 2

Method

Subjects. Four chimpanzees (one male and three females, range 18-33 years old) and two western lowland gorillas (one female and one male, aged 21 and 28 years) housed at Lincoln Park Zoo in mixed-sex social groups of six and seven, respectively, voluntarily participated in this study. The apes who

participated in this study were the same apes that regularly participate in the Lincoln Park Zoo touchscreen research program.

Materials and testing environment. The gorillas and chimpanzees were housed in naturalistic indoor and outdoor exhibits. Access to the outdoor yard was temperature dependent ($> 5\text{ }^{\circ}\text{C}$) and during the course of the study, all apes had outdoor access at varying times. Indoor spaces ranged from 72 m^2 to 124 m^2 in size; adjacent outdoor yards ranged in size from 116 m^2 to 1127 m^2 . Exhibits incorporated climbing structures, deep-mulch bedding, and additional off-exhibit holding areas. Fresh produce and chow were scattered daily throughout their habitat and apes had access to water *ad libitum*.

Apes were tested on touchscreen monitors attached to a mobile cart adjusted to the height of each animal. During testing, the touchscreen was flush against the mesh ($5.1\text{ cm X }5.1\text{ cm}$) along the perimeter of their indoor habitat (Figure 4). For one female gorilla, social group members were stationed by keepers in several locations simultaneously to allow her to work uninterrupted by conspecifics for approximately five minutes. All other apes were tested freely in their social group without conspecific stationing as they were not interrupted by others. As with the macaques tested in Experiment 1, stimuli were presented on a 22" Viewsonic TD2240 touch-sensitive monitor connected to a personal computer in the adjacent researcher area.

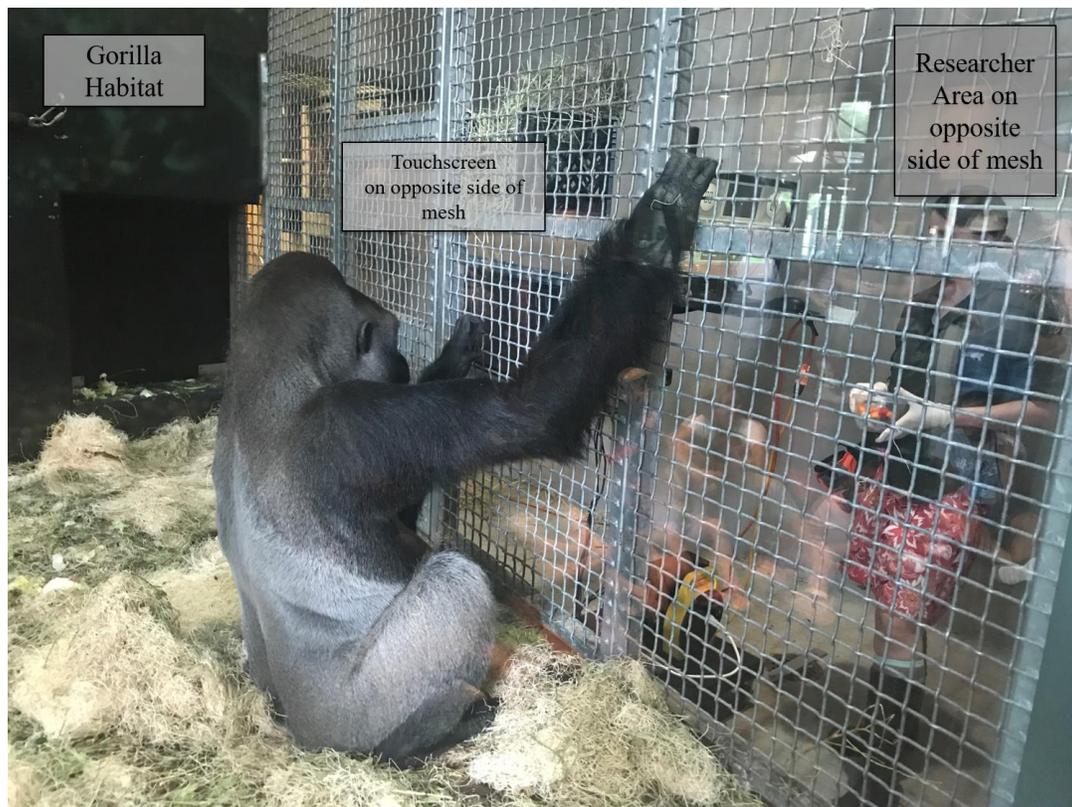


Figure 4. Touchscreen computer session taking place in gorilla habitat at Lincoln Park Zoo.

Experiment 2 was run using ApeTouch Zenrichment software and the stimuli size measured $7\text{ cm X }7\text{ cm}$ on the screen. As in Experiment 1, a grey square served as the control stimulus. Test stimuli consisted of color photographs of chimpanzees or western lowland gorillas unknown to the subjects, obtained from the Internet and selected and processed following the same criteria as Experiment 1. Subjects were shown only stimuli created from faces of conspecifics. Luminosity and contrast energy were obtained for each image and there were no significant differences in either measure between directed and averted stimuli sets (two-sample t-test, luminosity $t(46) = 1.33$, $p = .20$; contrast $t(46) = 1.37$, $p = .18$).

Procedure. Sessions took place between 1:30 and 3:30 p.m. and were voluntary; apes could walk away from the touchscreen at any time. If they did so, the current trial was discarded. Prior to participation in this study, the apes had participated in several touchscreen studies evaluating serial learning and food preferences (e.g., Egelkamp, Hopper, Cronin, Jacobson, & Ross, 2016; Ross, 2009).

The stimuli were presented following the same methods as Experiment 1, with the exception that the location of the image could appear in any location on the screen rather than three pre-set locations, again randomizing locations to safeguard against position biases and maintain interest. The maximum trial length was 60 s and the inter-trial interval was 8 s, during which time a black screen was shown. If an ape walked away from the touchscreen mid-trial, the aborted trial was excluded from analyses.

The identification of the participating ape was manually entered into the software prior to their first trial. Apes were rewarded for touching the stimulus with a single blueberry and a secondary reinforcement tone on a 100% reinforcement schedule via a PVC tube. As in Experiment 1, participation was voluntary with a 50-trial maximum per day per subject imposed, after which no additional trials appeared on the screen.

Subjects were tested during the 2017 Chicago A&W and two baseline periods, one beginning 30 days prior to the onset of A&W and one beginning 18 days after A&W testing was complete. The identity of the researcher was constant across conditions. Baseline data collection for an individual ceased once 200 trials were obtained, and apes completed each baseline period in five (minimum) to ten (maximum) days. Data were collected on three days during A&W. All apes participated on all three days except for one gorilla who participated on only two days. All apes had been exposed to the jet noise annually since they arrived at Lincoln Park Zoo in 2004.

In 2017 we also obtained sound recordings to assess decibel levels on zoo grounds during A&W. Using a sound level meter (SongMeter Model SM2+, Wildlife Acoustics), we recorded 10 min audio samples sequentially and continuously between the hours of 10:00 a.m. and 4:00 p.m. during the three days of A&W testing. We extracted the maximum dBA level per sample using the software Praat (Boersma & Weenink, 2017). For comparison, we repeated sound pressure level sampling during three days of the second baseline phase.

This study was approved by the Lincoln Park Zoo Research Committee, the governing body for all animal research at the institution. No modifications were made to standard animal care routines and the A&W was outside the control of Lincoln Park Zoo. This research adhered to legal requirements in the United States of America and to the American Society of Primatologists' Principles for the Ethical Treatment of Nonhuman Primates.

Statistical analysis. Analyses were conducted following the same methods as Experiment 1, with the exception that there were three levels of condition (pre-baseline, A&W, post-baseline). Given the small sample per species, chimpanzee and gorilla responses were collapsed for analyses (*sensu* Howard, Wagner, Woodward, Ross, & Hopper, 2017).

Results

Each ape voluntarily completed 200 trials during the pre-baseline (pre-BL) and post-baseline (post-BL) periods, with the exception of one female gorilla who completed 191 pre-BL trials and 160 post-BL trials. Apes completed between 57 and 150 trials during A&W (one ape completed 57, one completed 65, one completed 116, and three completed 150). No ape reached the maximum trial length on any trial.

The full-null model comparison examining the effect of condition (pre-baseline, A&W, post-baseline) revealed marginal significance ($\chi^2(2) = 5.61, p < .06$). The direction of the coefficients indicates that the direction of RTRatio change across conditions followed a chronological decrease from pre-baseline to A&W to post-baseline (reference category = A&W, pre-baseline $\beta = 0.004$, SEM = 0.004, post-baseline $\beta = -0.005$, SEM = 0.004; Figure 5). The full-null model comparison examining the effect of trial type (averted vs directed) was not significant ($\chi^2(1) = 2.33, p = .13$). Full-null model comparisons

revealed no significant interaction between condition and trial type ($\chi^2(1) = 2.69, p = .26$). Taken together, the results indicate that the apes exhibited a slowing in their latency to touch faces relative to controls over subsequent conditions and did not treat directed and averted faces differently (Figure 5). Complete model results are available in the [Supplemental Information](#).

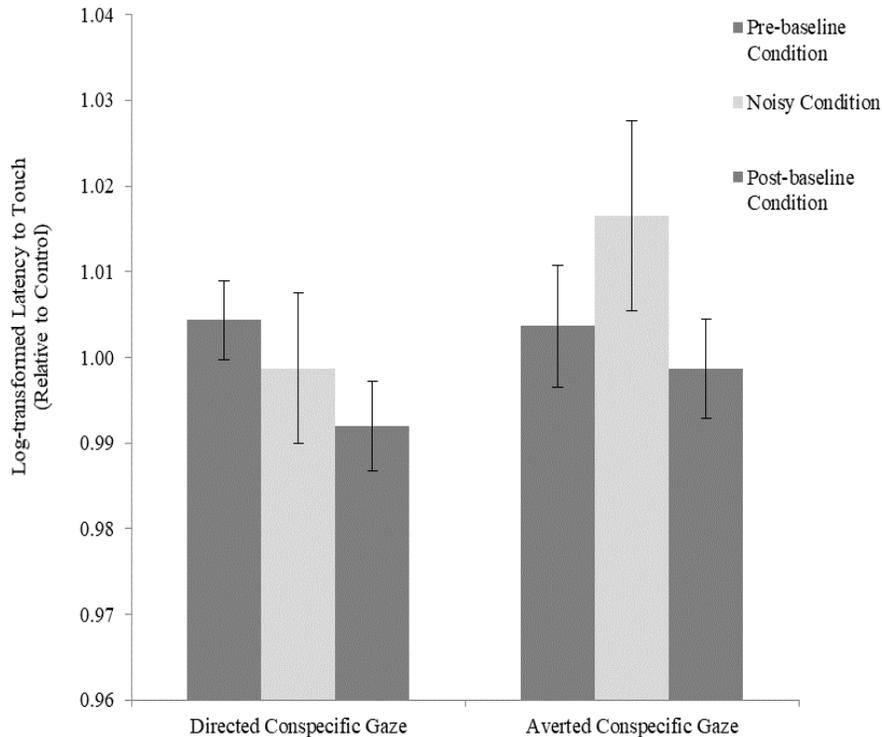


Figure 5. Mean Reaction Time Ratios (\pm SEM) for chimpanzees and gorillas across conditions for both trial types. The reaction time ratios are standardized by the latency to touch control images in each condition, therefore, values > 1 indicate that subjects touched conspecific images more slowly than control images, and values < 1 indicate that subjects touched conspecific images more quickly than control images.

Visual inspection of the spectrograms produced during the 2017 A&W reviewed in conjunction with fly-over times documented by researchers indicated that most flyovers were identifiable in the spectrograms by brief spikes exceeding 90 dBA. During the 2017 A&W show, results revealed that 14.6% of audio samples registered sound pressure levels above 90 dBA (max dBA recorded = 93.5). In contrast, only 1.0% of audio samples registered sound pressure levels above 90 dB during the matched control period.

Discussion

In Experiment 2, we measured whether zoo-housed chimpanzees and gorillas showed a response-slowing effect on a touchscreen task during the days of a loud event to evaluate changes in mood. We obtained sound pressure levels to better characterize the event, and found that levels regularly exceeded 90 dBA producing short, unpredictable bouts of loud noise overhead. The experimental design expanded upon the design of Experiment 1 with all subjects tested in two baseline periods, one preceding and one following the loud event. We predicted that if the event negatively impacted the apes' mood, we would observe a response slowing effect during the A&W condition relative to both baseline conditions. However, when examining the trend toward a significant effect of condition, responses generally

increased in speed relative to control trials over time. Similar to the macaques, the apes did not respond differentially to directed and averted faces. The implications of these findings are discussed together with the results of the Japanese macaques in the General Discussion.

General Discussion

This study evaluated changes in affect resulting from anthropogenic noise in a zoo environment. The response-slowing touchscreen task used here suggests that zoo-housed Japanese macaques underwent changes in mood during a loud annual event, the Chicago Air & Water Show. This interpretation follows from the finding that, during the days of the event, the latency to touch conspecific images, relative to control images, slowed compared to a baseline period. A second experiment was conducted to test whether the mood of chimpanzees and gorillas was affected by the same loud event the following year. The results for the apes were more ambiguous. There was a trend toward significant differences in response latencies to faces compared to controls between baseline days and A&W Show days, but overall, the apes responded to the face stimuli with increasing speed over subsequent conditions, suggesting potential habituation to the stimuli. The full model results, including odds ratios indicating the magnitude and direction of differences, are available in the [Supplemental Information](#).

We tentatively conclude the Japanese macaques experienced a change in affect in response to the loud jets passing overhead whereas the apes did not. The difference in affective response may stem from differences in habitat design and testing location, as the macaques spend the majority of their daytime outdoors and touchscreen tests took place in the periphery of their outdoor habitat. In contrast, the apes choose to spend the majority of their time indoors (Kurtucz, Wagner, & Ross, 2014) and their touchscreen tests also took place inside. The sound level recordings reflect the noise level on zoo grounds outside, and by being inside the apes may have been shielded from some of the noise of the A&W Show. Alternatively or additionally, the difference between macaques and apes may have arisen from different histories of exposure to the noisy event. Specifically, the macaques were tested during their second year of exposure having arrived at the zoo in 2014, whereas the apes had been exposed for 13 consecutive years. The apes may simply have habituated to the noise or learned that it did not pose a threat.

It is also possible that the emotionally threatening stimuli used in this study (direct gaze faces) were not perceived as emotionally threatening by the apes. We tested both sexes of chimpanzees and gorillas with stimuli obtained from conspecifics of both sexes, and it may be that females directing gaze toward males is not a biologically relevant stressor in species with male dominance and extreme sexual dimorphism. Unfortunately, our ape sample is not large enough to consider whether there is an interaction between the subject's sex and the sex of the conspecific stimuli. It is also likely that there are individual differences contributing to the results obtained here, and teasing out how personality or temperament influences responses to social stimuli is an important future direction. Related, there may have been self-selection among the subjects in that the subset of touchscreen trained animals who voluntarily participated during the Air and Water Show may have been willing to do so because they experienced less stress than groupmates (see Morton, Lee, & Buchanan-Smith, 2013; Polgar, Wood, & Haskell, 2017). However, that would have led us to underestimate the impact on mood for Japanese macaques. Finally, it is possible that the difference between the macaques and apes may have arisen due to different evolved predator responses. Macaques are under more threat from aerial predators than apes (Iida, 1999), and noises from above the habitats may elicit a stronger stress response from the macaques than apes. While we speculate about the potential explanations for the differences between the macaques and the apes, we remain cautious in our conclusion of species differences given that the response-slowing paradigm has not yet been validated for apes as it has for macaques.

The pattern of response slowing that emerged when we introduced a second baseline period in Experiment 2 suggests that habituation to face stimuli is a possibility that may limit the utility of this paradigm. This limitation may be especially pronounced given the frequent presentation of the face stimuli in the present design. However, habituation cannot explain the pattern of results obtained for the Japanese macaques given that three of the four subjects were tested in a baseline period that *preceded* the

Air and Water Show, yet the responses to face stimuli *slowed* during the Air and Water Show. Moving forward, determining ideal ratios for displaying face and control stimuli that generate a sufficient amount of repeated samples from individuals to provide enough statistical power to test hypotheses about changes in mood, while simultaneously minimizing habituation, will be essential to creating a useful method. Researchers may also consider the possibility of changing face stimuli throughout the experiment to minimize the potential for habituation, although this strategy would come at the cost of introducing an additional source of variation across conditions.

The response-slowing paradigm is motivated by the finding that emotional content slows response times on simple tasks for humans experiencing anxiety. However, the relationship between anxiety and biases in cognitive and attentional processes is complex with several potential mechanisms at play (Mogg & Bradley, 2016). The pattern of results observed here for Japanese macaques, and for rhesus macaques in Bethell et al. (2016), is consistent with a number of (non-mutually exclusive) interpretations involving attention to threat and a subtle cognitive freeze response. Under stressful conditions, animals may invest additional resources to maintain a state of high social vigilance (Ebitz, Watson, & Platt, 2013). In the response-slowing paradigm, the monkeys may be responding slower under stress because their attention is more strongly captured by the conspecific faces (e.g., Bethell, Holmes, MacLarnon, & Semple, 2012; Bradley, Mogg, & Miller, 2000; Garner, Mogg, & Bradley, 2006). An alternative explanation is an enhanced freeze response to conspecific faces in macaques during the presumably stressful Air and Water Show (Bethell et al., 2016). In order to experimentally test whether heightened vigilance is in fact the mechanism responsible for the slowing, future work could disentangle the behavioral response (touching the image) from the removal of the threatening stimuli (disappearance of the image). Alternatively, adapting spatial cueing tasks (e.g., Kalin, Shelton, Rickman, & Davidson, 1998; Parr, Modi, Siebert, & Young, 2013) in which the spatial location of threatening stimuli is either consistent or inconsistent with the location of a required response (e.g., a screen touch) to be used under stressful and non-stressful conditions could be a fruitful way forward, as this approach would measure which stimuli attract attentional investment.

Although we found an effect of condition that indicates monkeys responded differently during A&W compared to baseline, and these findings were standardized by condition-specific response speeds to control trials to account for variability across conditions due to arousal or practice, we did not find a difference in response slowing between averted and directed faces in either experiment. This pattern of results differs from those of Bethell et al. (2016) that showed slowing in response to directed but not averted faces in veterinary-stressed rhesus macaques. There are many potential explanations for this difference including differences in the sex of subjects and stimuli (Bethell et al. (2016) used all males for both), differences in laboratory and zoo housing conditions, and potential differences in baseline levels of anxiety. Furthermore, we may have lacked the statistical power necessary to detect a difference between directed and averted faces or the distinction between directed and averted faces may have been less pronounced in our study. There may also be species or individual age or personality differences influencing our findings (Adams et al., 2015; Wright et al., 2007). However, if the lack of a significant interaction represents that the primates truly respond similarly to direct and averted faces, then either both image types are being interpreted as threatening or the primates are not affected by whether the image content is threatening as we assume in this paradigm. If the latter interpretation is true, we still maintain that something about the macaques' affective state has likely changed between the baseline and noise conditions given that their responses to social versus non-social stimuli changed. What this reveals about the state of the animal remains to be determined with future work aimed at unpacking the mechanisms underlying the effect.

This study investigated the impact of noisy, unpredictable, repeated events on the mood of zoo-housed Japanese macaques, chimpanzees and gorillas. The results suggested a negative impact of the events on the mood of the Japanese macaques through a change in the macaques' behavior on a response-slowing task consistent with an anxious state. Whether other loud events that are common to zoo environments, such as special events for donors or concerts, have an effect on the mood of zoo animals remains to be determined. This report also demonstrates the feasibility of voluntary cognitive testing in

three primate species without isolating subjects from their social group, which can increase the feasibility and validity of cognitive testing (Cronin, Jacobson, Bonnie, & Hopper, 2017). With further development, the response-slowing paradigm used here may be an effective and feasible way to evaluate real-time changes in the mood of zoo-housed animals under a variety of circumstances.

Acknowledgments

The authors thank the animal keepers in the Primate Department, Maureen Leahy, Michael Brown-Palsgrove and Jill Moyse. We thank Liza Lehrer of the Urban Wildlife Institute for assistance with recording sound levels and Fisher Center interns for assistance. We thank Dan Acheson for assistance with PsychoPy and statistics, Chris Martin for Zenrachment ApeTouch programming, and the Leo S. Guthman Fund and Chauncey and Marion Deering McCormick Foundation for support.

References

- Adams, M. J., Majolo, B., Ostner, J., Schülke, O., De Marco, A., ... Weiss, A. (2015). Personality structure and social style in macaques. *Journal of Personality and Social Psychology, 109*, 338–353.
- Bar-Haim, Y., Lamy, D., Pergamin, L., Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2007). Threat-related attentional bias in anxious and nonanxious individuals: A meta-analytic study. *Psychological Bulletin, 133*, 1–24.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.
- Bates, D., Maechler, M., & Bolker, B. (2014). lme4: Linear mixed-effects models using Eigen and S4. R Package Version 1.7.
- Bethell, E. J. (2015). A “how-to” guide for designing judgment bias studies to assess captive animal welfare. *Journal of Applied Animal Welfare Science, 18*, S18–S42.
- Bethell, E. J., Holmes, A., MacLarnon, A., & Semple, S. (2012). Evidence that emotion mediates social attention in rhesus macaques. *PLOS One, 7*, e44387.
- Bethell, E. J., Holmes, A., MacLarnon, A., & Semple, S. (2016). Emotion evaluation and response slowing in a non-human primate: New directions for cognitive bias measures of animal emotion? *Behavioral Sciences, 6*, 2.
- Birke, L. (2002). Effects of browse, human visitors and noise on the behaviour of captive orangutans. *Animal Welfare, 11*, 189–202.
- Blickley, J. L., & Patricelli, G. L. (2010). Impacts of anthropogenic noise on wildlife: research priorities for the development of standards and mitigation. *Journal of International Wildlife Law & Policy, 13*, 274–292.
- Boersma, P., & Weenick, D. (2017). Praat: Doing phonetics by computer [Computer program]. Version 6.0.33, retrieved 26 September 2017 from <http://www.praat.org>.
- Bradley, B. P., Mogg, K., & Millar, N. H. (2000). Covert and overt orienting of attention to emotional faces in anxiety. *Cognition and Emotion, 14*, 789–808.
- Brydges, N. M., & Hall, L. (2017). A shortened protocol for assessing cognitive bias in rats. *Journal of Neuroscience Methods, 286*, 1–5.
- Chosey, J., Wilson, M., & Santymire, R. (2014). Behavioral and physiological responses in Felids to exhibit construction. *Zoo Biology, 33*, 267–274.
- Cronin, K. A., Jacobson, S. L., Bonnie, K. E., & Hopper, L. M. (2017). Studying primate cognition in a social setting to improve validity and welfare: A literature review highlighting successful approaches. *PeerJ 5*: e3649.
- Davey, G. (2007). Visitors' effects on the welfare of animals in the zoo: A review. *Journal of Applied Animal Welfare Science, 10*, 169–183.
- Dawkins, M. (2015). Animal welfare and the paradox of animal consciousness. *Advances in the Study of Behavior, 47*, 5–38.
- Deakin, A., Browne, W. J., Hodge, J. J., Paul, E. S., & Mendl, M. (2016). A screen-peck task for investigating cognitive bias in laying hens. *PLOS One, 11*, e0158222.
- Ebitz, R. B., Watson, K. K., & Platt, M. L. (2013). Oxytocin blunts social vigilance in the rhesus macaque. *Proceedings of the National Academy of Sciences, 110*, 11630–11635.

- Egelkamp, C. L., Hopper, L. M., Cronin, K. A., Jacobson, S. L., & Ross, S. R. (2016). Using touchscreens to explore the welfare and cognition of zoo-housed primates. *PeerJ Preprints* 4: e2312v1.
- Garner, M., Mogg, K., & Bradley, B. P. (2006). Orienting and maintenance of gaze to facial expressions in social anxiety. *Journal of Abnormal Psychology*, 115, 760–770.
- Hanson, J. D., Larson, M. E., & Snowdon, C. T. (1976). The effects of control over high intensity noise on plasma cortisol levels in rhesus monkeys. *Behavioral Biology* 16, 333–340.
- Harding, E. J., Paul, E. S., & Mendl, M. (2004). Animal behaviour: Cognitive bias and affective state. *Nature*, 427, 312–312.
- Kalin, N. H., Shelton, S. E., Rickman, M., & Davidson, R. J. (1998). Individual differences in freezing and cortisol in infant and mother rhesus monkeys. *Behavioral Neuroscience*, 112, 251–254.
- Kight, C. R., & Swaddle, J. P. (2011). How and why environmental noise impacts animals: An integrative, mechanistic review. *Ecology Letters*, 14, 1052–1061.
- Kurtycz, L. M., Wagner, K. E., & Ross, S. R. (2014). The choice to access outdoor areas affects the behavior of great apes. *Journal of Applied Animal Welfare Science*, 17, 185–197.
- Maestriperieri, D. (1997). Gestural communication in macaques: Usage and meaning of nonvocal signals. *Evolution of Communication*, 1, 193–222.
- McGuire, M. C., Vonk, J., Fuller, G., & Allard, S. (2017). Using an ambiguous cue paradigm to assess cognitive bias in gorillas (*Gorilla gorilla gorilla*) during a forage manipulation. *Animal Behavior and Cognition*, 4, 91–104.
- Mellor, D. J. (2015). Positive animal welfare states and reference standards for welfare assessment. *New Zealand Veterinary Journal*, 63, 17–23.
- Mendl, M., Burman, O. H., Parker, R., & Paul, E. S. (2009). Cognitive bias as an indicator of animal emotion and welfare: Emerging evidence and underlying mechanisms. *Applied Animal Behaviour Science*, 118, 161–181.
- Mogg, K., & Bradley, B. P. (2016). Anxiety and attention to threat: Cognitive mechanisms and treatment with attention bias modification. *Behaviour Research and Therapy*, 87, 76–108.
- Morton, F. B., Lee, P. C., & Buchanan-Smith, H. M. (2013). Taking personality selection bias seriously in animal cognition research: A case study in capuchin monkeys. *Animal Cognition*, 16, 677–684.
- Orban, D. A., Soltis, J., Perkins, L., & Mellen, J. D. (2017). Sound at the zoo: Using animal monitoring, sound measurement, and noise reduction in zoo animal management. *Zoo Biology*, 36, 231–236.
- Parr, L. A., Modi, M., Siebert, E., & Young, L. J. (2013). Intranasal oxytocin selectively attenuates rhesus monkeys' attention to negative facial expressions. *Psychoneuroendocrinology*, 38, 1748–1756.
- Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, 2, 1–8.
- Polgar, Z., Wood, L., & Haskell, M. J. (2017). Individual differences in zoo-housed squirrel monkeys' (*Saimiri sciureus*) reactions to visitors, research participation, and personality ratings. *American Journal of Primatology*, 79, e22639.
- Quadros, S., Goulart, V. D., Passos, L., Vecci, M. A., & Young, R. J. (2014). Zoo visitor effect on mammal behaviour: Does noise matter? *Applied Animal Behaviour Science*, 156, 78–84.
- R Core Team. (2014). *R: A language and environment for statistical computing*. Vienna, Austria. Retrieved from <http://www.R-project.org>
- Ratcliff, R. (1993). Methods for dealing with reaction time outliers. *Psychological Bulletin*, 114, 510–532.
- Ross, S. R. (2009). Sequential list-learning by an adolescent lowland gorilla (*Gorilla gorilla gorilla*) using an infrared touchframe apparatus. *Interaction Studies*, 10, 115–129.
- Whitham, J. C., & Wielebnowski, N. (2013). New directions for zoo animal welfare science. *Applied Animal Behaviour Science*, 147, 247–260.
- Wright, A. J., Soto, N. A., Baldwin, A. L., Bateson, M., Beale, C. M., ... Hatch, L. T. (2007). Anthropogenic noise as a stressor in animals: A multidisciplinary perspective. *International Journal of Comparative Psychology*, 20, 250–273.