ROYAL SOCIETY OPEN SCIENCE

royalsocietypublishing.org/journal/rsos

CC BY

Research

**Cite this article:** Smith GE, Folkertsma R, Duran E, Shair Ali EG, Jacobs I, Seed AM, Auersperg AMI, Lambert ML. 2025 Physio-behavioural manifestations of 'surprise' in two parrot species: kea (*Nestor notabilis*) and Goffin's cockatoos (*Cacatua goffiniana*). *R. Soc. Open Sci.* **12**: 251264. https://doi.org/10.1098/rsos.251264

Received: 3 July 2025
Accepted: 15 October 2025

**Subject Category:**
Psychology and cognitive neuroscience

**Subject Areas:**
cognition, behaviour, psychology

**Keywords:**
violation of expectations, information seeking, surprise, kea, cockatoo, comparative cognition, infrared thermography

**Author for correspondence:**
Gabriella E. Smith
e-mail: gabriella.smith@vetmeduni.ac.at

Supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.8169408.

THE ROYAL SOCIETY PUBLISHING

# Physio-behavioural manifestations of 'surprise' in two parrot species: kea (*Nestor notabilis*) and Goffin's cockatoos (*Cacatua goffiniana*)

Gabriella E. Smith[1,2], Remco Folkertsma[1], Elif Duran[3], Eva G. Shair Ali[1], Ivo Jacobs[4], Amanda M. Seed[2], Alice M. I. Auersperg[1] and Megan L. Lambert[1]

[1]Messerli Research Institute, Department of Interdisciplinary Life Sciences, University of Veterinary Medicine Vienna, Medical University of Vienna, University of Vienna, Vienna, Austria
[2]School of Psychology and Neuroscience, University of St Andrews, St Andrews, UK
[3]International School for Advanced Studies (SISSA), Trieste, Italy
[4]Department of Cognitive Science, Lund University, Lund, Sweden

GES, 0000-0002-6136-6351; RF, 0000-0003-4839-9703; ED, 0000-0002-9603-0057; EGSA, 0009-0009-3483-3709; IJ, 0000-0002-2182-7686; AMS, 0000-0002-3867-3003; AMIA, 0000-0001-7405-9791; MLL, 0000-0003-3618-7260

How animals react to violations of expectations offers a valuable lens into their perception of the world. Kea (*Nestor notabilis*) and Goffin's cockatoos (*Cacatua goffiniana*) are optimal animal models to examine susceptibility to violations of expectations owing to their sophisticated physical exploration and cognition skills. This study involved a bait-and-switch of food items varying in desirability, followed by an information-seeking phase to examine not only violation susceptibility but also surprise behaviours expressed overtly and covertly in both positive and negative violation directions. We found that the kea, and not the cockatoos, exhibited susceptibility to the violation via likelihood to explore the violating stimulus. Furthermore, while this project serves, to our knowledge, as the first cognitive study using infrared thermography in fully unrestrained avian subjects, no

© 2025 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

conclusive results can be reported as to the effect of violations of expectation on subjects' periorbital temperature, and thus physiological response. Implications for this research include novel perspectives into studying overt and covert responses following violated expectations, as well as expressions of information seeking in non-human species.

# 1. Introduction

Violations of expectation experiments offer valuable insight into an individual's perception of the world. Such studies have been performed in both humans and non-human animals (hereafter, animals) examining susceptibility and reactions to violations including but not limited to object permanence [1,2], physical support [3–5], motion [6,7] and social expectations [8–10]. In studies such as these, expectations are considered violated if subjects express surprise in violation conditions compared to controls. This is recorded in ways such as investigation time ([11–13]; see [14]), pupil dilation [15,16], emotional response [17–19], facial expression [20,21], as well as blood flow and/or brain electrical activity changes [22,23]. Thus, overall, surprise responses can manifest in subtle, overt, and covert ways in response to a violation.

Surprise responses can also vary depending on the direction of the violation (i.e. valence). These violation directions can depend on the value of the stimuli involved, such as in studies of object permanence and individuation in which preferred food is secretly swapped for less-preferred food (negative surprise) and vice versa (positive surprise) (e.g. [24,25]). In the case of negative surprises, behavioural responses tend to fall under the term 'frustration effect' ([26–29]; see also surprise reward omissions [30]), such as delay or refusal to eat the swapped food and/or production of frustration vocalizations [18,19]. Frustration responses such as these are similar to those observed in humans, such as children behaving more intensely towards the negative violating stimulus (e.g. hitting a barrier preventing access to once-accessible toys [17,31,32]). Such frustration effect in animals could be related to loss aversion [33], the communication of which could prove evolutionarily valuable for group cohesion (e.g. [34]).

Behaviours expressed by animals in response to positive surprise, however, are less well-documented. Termed 'elation effect' ([27,28]; but see [29]), in humans, positive surprises are expected to produce behaviours rooted in delight and wonder [35], such as upon discovering unexpected cash in your pocket. In the case of animals, Crespi [27] found that rats ran faster in a runway following trials in which they had been given more food than expected, suggesting an elation effect (see also [36]). Meanwhile, work on dogs and great apes found no difference in emotional responses between positive or negative surprise conditions [24]. A possible explanation for the apparent scarcity of evidence for the elation effect in animals could be that such behaviour may draw the undesirable attraction of competitors to the discovery of a valuable resource (see cache protection in scrub-jays [37]). In fact, some violation studies have found subjects higher in the hierarchy are more likely to express surprise behaviours in response to swapped food than those lower in the hierarchy (e.g. [25]). This suggests that surprise responses may indeed depend on resource security, such as that afforded by high hierarchical status. Emotional responses are thus valuable to study as they offer insight not only into a subject's susceptibility to surprise but also the roles of emotion in response to changes in one's environment in different social and ecological settings.

Emotion can be broadly defined as a state of physiological arousal with a corresponding valence [38]. In the human literature, the nuance of both the arousal intensity and valence direction of an emotion can be explored through various experimental approaches including but not limited to behavioural responses (e.g. facial expressions [39]), physiological activity (e.g. functional magnetic resonance imaging [40]) and descriptive assessments (e.g. self-reports [41]). In the case of animals and the obvious absence of self-reporting, the study of emotions instead includes physiological measures (e.g. blood glucocorticoids [42]) and behavioural proxies. In the case of the latter, such research has included various approaches such as coding of species-specific, emotional behaviours (e.g. tail-wagging versus growling in dogs [43]), measures of playful activity [44] and even cognitive bias tasks for more nuanced emotions (e.g. pessimism [45]). However, disagreement exists as to how comparable the interpretations of emotions are between humans and animals (see [46]), especially those beyond evolutionarily conserved emotions (e.g. fear [47]). For the purposes of this paper, we are considering animal emotion in terms of the arousal following a violation-of-expectation event in either the positive (elation) or the negative (frustration) valence directions.

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

An order of animal species with a wide range of emotional responses is the parrots (*Psittaciformes*). Consisting of four lineages—*Strigopoidea* (New Zealand parrots), *Cacatuoidea* (cockatoos), *Psittaculidae* (Old World parrots) and *Psittacidae* (Afro-American parrots) [48]—parrots are known to exhibit vastly different emotional behaviours [49]. For example, in response to an arousing stimulus, some species are known to blush (e.g. macaws [50]), pin their eyes (e.g. Amazons [51]), fan their tails (e.g. parrotlets [52]) or raise crests (e.g. cockatoos [53,54]). However, even these overt behaviours can be difficult to interpret as they can vary between species, contexts and emotional valence (see [55]). Still other parrot species lack such overt, context-specific emotional behaviours, at least to the human eye [56]—with the exception of exaggerated fear or aggressive responses [57, p. 97]—resulting in the difficulty of their study in particular contexts [58].

The growing use of physiological technology in behavioural research offers potential in addressing this difficulty. Similarly to the techniques described above, infrared thermography (hereafter, IRT) is a promising approach to the study of animal emotion in terms of arousal and valence (see [59] for a recent review). Specifically, emotional events trigger the sympathetic and parasympathetic nervous systems that regulate blood flow, which can be detected by IRT cameras as a change in surface temperature of exposed skin [59–61]. For example, macaques (*Macaca mulatta*) exhibit a decrease in nasal temperature after seeing a threatening human [62], and similarly, the periorbital temperature of blue tits (*Cyanistes caeruleus*) decreases in response to trapping [63]. Such technology has also been used to examine positive valence events, such as the nasal temperature of common marmoset males (*Callithrix jacchus*) increasing following the positive stimulus of food calls ([64]; see also [65]). Thus, IRT holds promise in detecting and quantifying endothermic animals' emotions in response to violations of expectations, whether negative or positive in valence, but to our knowledge has not yet been employed within avian cognition research.

This study aims to examine species differences in overt surprise by exposing a violation of physical expectations to two parrot species: the kea parrot (*Nestor notabilis*) and the Goffin's cockatoo (*Cacatua goffiniana*). By creating a violation of expectation—specifically in which one food item is secretly swapped for another—in positive and negative surprise directions, this study aims to examine exploratory and emotional responses of parrot species. Specifically, as both species exhibit sophisticated physical cognition skills [66–70], we predict both will prove susceptible to the violation via prolonged exploration of the violating stimulus and longer latency to eat food in violation trials. Similarly, with regard to emotional responses, we predict both species to exhibit relatively more 'surprise' behaviours in violating trials than non-violating trials (e.g. vocalizations). Furthermore, in the case that no overt emotional responses are detected in one or either species, this study also employs IRT to examine potential covert responses via periorbital temperature changes and predicts that if subjects are surprised by the switch, we expect a relative change in surface temperature of the periorbital region in violation trials. Owing to the mixed results of prior research into frustration and elation effects, behavioural and physiological surprise relative to violation direction (positive or negative) is harder to predict, but will be explored in this study.

# 2. Material and methods

## 2.1. Subjects

From April to June 2023, 19 captive kea subjects (10 females and nine males; electronic supplementary material, table S1) were individually tested in visually isolated testing compartments (6 × 10 × 4 m), to which they were already habituated. Subjects are kept in an outdoor aviary (52 × 10 × 4 m) at Haidlhof Research Station in Bad Vöslau, Austria, and maintained on a daily diet of seeds, fruits, vegetables and meat, with access to water ad libitum.

In July 2023, 15 captive Goffin's cockatoo subjects (nine males and six females; electronic supplementary material, table S1) were individually tested in a single, visually isolated testing room (3 × 2.5 × 3 m) to which they were already habituated. Subjects are kept in an indoor aviary (45 $m^2$; 3–6 m high) with an outdoor (*ca* 200 $m^2$; 3-4.5 m high) compartment at the Goffin Laboratory in Goldegg, Lower Austria. Subjects' diet consists of egg, starch (e.g. cooked pasta; potatoes), soy yoghurt, zwieback toast, fresh and dry fruit, seeds and bird pellets mixed with mineral supplements. Subjects have access to water ad libitum.

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

While the kea have engaged with violation-of-expectation experiments before (e.g. [71]), and the cockatoos have experienced experiments involving digging in sand (Colbourne *et al*., 2023, unpublished data), neither population has experienced a research set-up of this kind.

## 2.2. Stimuli

The stimuli in the kea experiment included two pairs of wooden towers (13 × 10 × 11 cm each, four total), one white and one black in each pair. One pair had clear plastic shelves inside (to catch dropped reward), the other did not (to allow reward to drop through). In experimental trials, one of these towers (colour and shelf presence determined by trial) was placed atop a wooden block (15 × 15 × 6.5 cm) attached to a wooden plate (35 × 35 × 1.5 cm), the block containing a circular depression (8 × 2.5 cm) into which sand was poured before each trial. The single, clear side of each tower in the kea experiment was an artefact from a previous design, but did not influence the experiment. Plastic bottle caps (2.8 × 1.8 cm) were used in demonstration trials (see figure 1).

The stimuli in the cockatoo experiment included two pairs of PVC towers (8 × 7.5 × 0.3 cm each, four total), one white and one black in each pair. One pair had hard plastic shelves inside (to catch dropped reward), the other did not (to allow reward to drop through). In experimental trials, a single tower (colour and shelf presence determined by trial) was placed atop a terracotta cup (7 × 6 × 0.4 cm) into which sand was poured before each trial, the cup attached to a terracotta plate (28 × 3.5 cm). Yellow, wooden cubes (1.5 × 1.5 × 1.5 cm) to which the birds were already familiar were used in demonstration trials (see figure 2).

Stimuli materials were chosen for each species based on species appropriateness (e.g. size, destructibility) and pre-existing familiarity. For both species, each pair of towers had one white and one black tower so that any exploration exhibited by subjects after the experiment could be deemed generalized or specific with respect to tower colour and thus violated expectations (see §2 for exploration and experimental trials; note also that the tower associated with violation trials was termed 'violation tower,' and the tower associated with non-violation trials was termed 'non-violation tower').

## 2.3. Procedure

The experimental procedure was identical between species.

### 2.3.1. Food preference tests

Prior to testing, subjects were given three sessions of 12 food preference tests in which they were presented with pairings of three common food rewards they were known to eat willingly and allowed to select one of the two options. In kea, options were peanut (high value), pellet (medium value) and apple (low value). In cockatoos, options were cashew (high value), walnut (medium value) and cooked pasta (low value). Chi-square tests determined the high and low preference rewards for each individual before proceeding with experimental trials (see below).

### 2.3.2. Exploration and experimental trials

Subjects experienced two unrewarded, 3 min exploration periods with the shelf-less set of towers (left-right presentation order randomized between subjects), one before and one after experimental trials to examine the effect of the violation on exploration (figure 3). After the initial exploration period, functional demonstrations of the towers were shown to the subjects using the shelved set of towers (clear side turned away from kea subjects to prevent view of the shelf; cockatoos could not see the shelf from demonstration vantage). In these two trials, single, non-food objects (bottle caps for kea; wooden cubes for cockatoos) were dropped into the top of the towers one at a time (tower colour order randomized between subjects), followed by the removal of the tower and reveal of an identical, pre-baited object sitting in the sand below (figure 4). These trials were aimed at establishing the expectation that objects pass through the towers as expected, consistent with the species' understanding of object permanence and individuation (see [70,72]). Subjects had 1 min to access the sand at will following tower removal.

After the demonstration phase, the same procedure was employed with the same-size food in experimental trials (figure 5). The order of these trials was always non-violation/violation/violation/

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

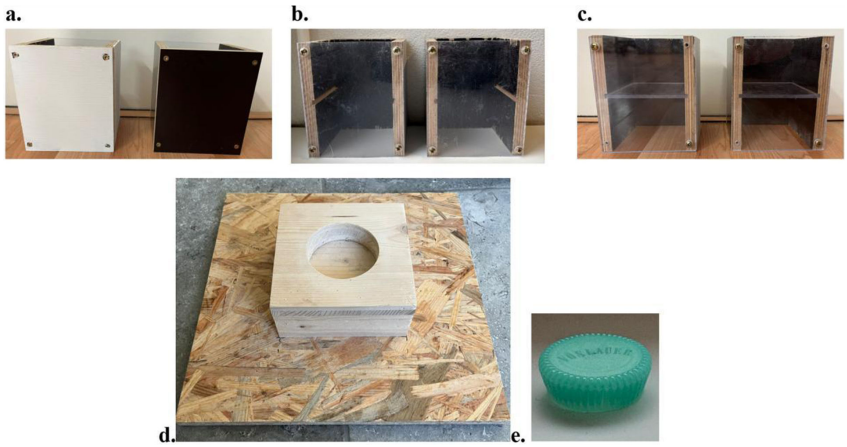


**Figure 1.** (a) Kea tower pairs had one white and one black tower each. (b) Set one (presented in exploration trials) had no shelves. (c) Set two (used in demonstration and test trials) had shelves. (d) In experimental trials, a single tower was placed atop sand in the wooden plate. (e). In demonstration trials, a non-food item was dropped through.

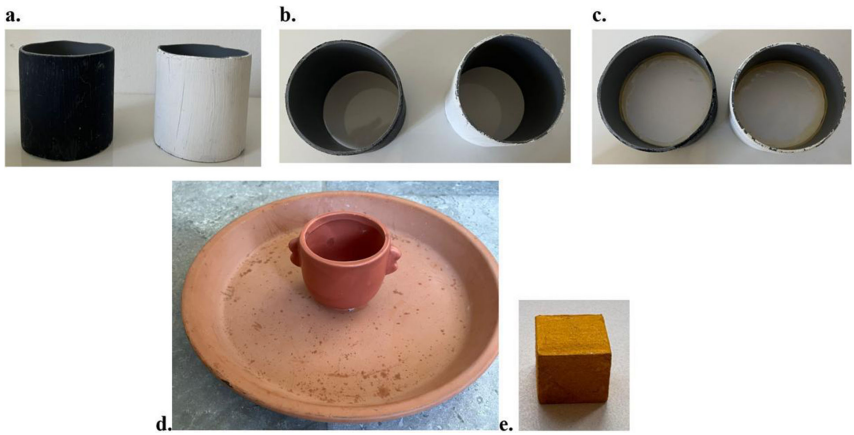


**Figure 2.** (a) Cockatoo tower pairs had one white and one black tower each. (b) Set one (presented in exploration trials) had no shelves. (c) Set two (used during demonstration and test trials) had shelves. (d) In trials, a single tower was placed atop sand in the terracotta plate. (e) In demonstration trials, a non-food item was dropped through.

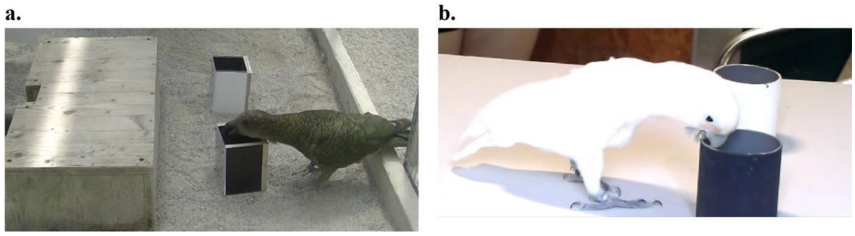


**Figure 3.** (a) Kea and (b) cockatoo exploration of hollow towers.

non-violation with either the white or the black tower colour associated with violation or non-violation trials (colour assignment randomized between subjects), with 1 min to retrieve the food and access the sand at will (trial order was chosen to avoid potential order effects). For example, a kea subject could experience the following trials: (i) an apple piece dropped into the white tower with an apple piece revealed in sand below (non-violation, low-to-low); (ii) a peanut dropped into the black tower with an apple piece revealed in sand below (violation, high-to-low); (iii) an apple piece dropped into the black tower with a peanut revealed in sand below (violation, low-to-high); and (iv) a peanut dropped into the white tower with a peanut revealed in sand below (non-violation, high-to-high). The direction of violation and non-violation trials was randomized within and between subjects.

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

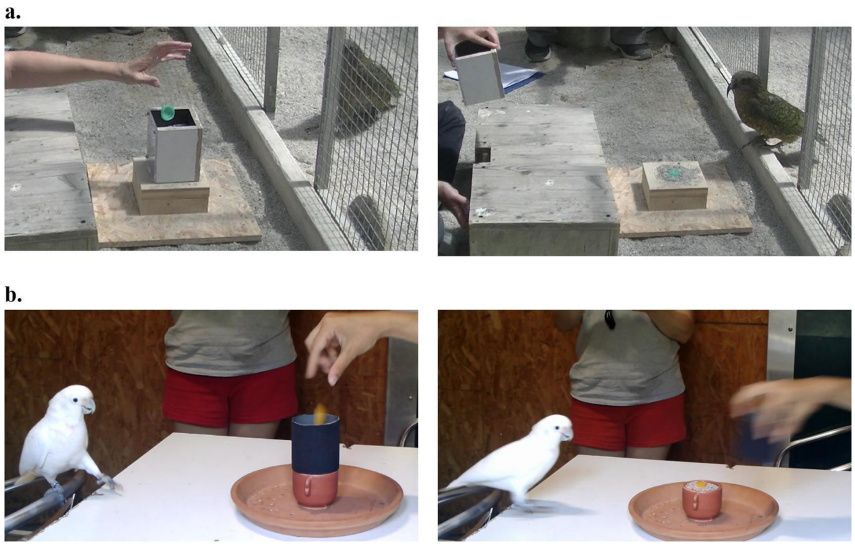


**Figure 4.** (a) Kea and (b) cockatoo demonstration trials.

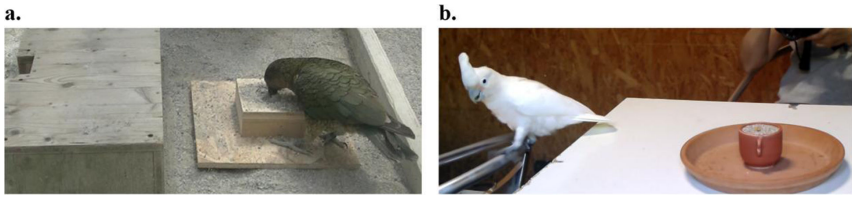


**Figure 5.** (a) Odo the kea digging in sand following a high-to-high non-violation trial. (b) Muki, the cockatoo, crest raising following a low-to-high violation trial.

## 2.4. Data collection and analysis

### 2.4.1. Regular and infrared thermography video recording

Prior to the first exploration trial, we recorded the subject inside the testing compartment for 2 min with an IRT camera (FLIR T530) to obtain their individual baseline temperature. This camera model has an uncooled microbolometer with a resolution of 320 × 240 pixels, thermal sensitivity of <40 mK, accuracy of ±2°C, and spectral range of 7.5–14.0 µm. We used a 24° lens ($f$ = 17 mm), resulting in an instantaneous field of view (spot size) of 1.33 mm at 1 m distance. Videos were recorded at 30 Hz in the −20 to 120°C range and stored in radiometric, lossless format. We adhered to established methodological guidelines (see [73,74]), and the following parameters were inputted beforehand: emissivity was set to 0.95; humidity and ambient temperature were measured with a thermometer positioned in the shade; reflective temperature was considered equivalent to room temperature for the cockatoos. For the kea, it was set as the average cloud temperature on cloudy days, and as −20°C on sunny days. The set distance was 1 m, which was the distance between filmer and subject throughout recording for baseline and test trials. The filmer continuously manually focused (one-shot contrast) the camera on the subject's eye. The camera was held perpendicular to the subject's head as much as possible, with the angle of incidence never exceeding 45°. Birds were tested only in the shade. All test trials were additionally recorded on a JVC Panasonic camcorder.

### 2.4.2. Video coding

Behavioural coding was done using BORIS [75]. For exploration trials, we coded which tower the bird touched first, as well as their duration of interaction with each tower. For test trials, we coded whether subjects ate the food and their latency to do so, the duration subjects spent digging in the sand, vocalization counts and interaction counts with the sand plate (see the electronic supplementary material, table S2, for ethogram). For cockatoos, we additionally recorded counts of crest raising. Inter-observer reliability for durations and latencies on 17% of all subjects' videos (20% of each species'

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

sample size; three individuals per species) was carried out between two independent observers. We calculated the intraclass correlation coefficient using the function 'icc' in the package 'irr', setting the 'model' argument to 'twoway' and the 'type' argument to 'agreement'.

IRT video was coded using FLIR Thermal Studio Standard, with the maximum temperature of the periorbital eye region manually recorded every second for 30 s following the trial start (food drop into tower) (figure 6). The periorbital eye region was chosen because of it being one of the most vascularized and exposed regions of skin in birds, and as it is less affected by environmental conditions than more distal regions such as the feet [63]. Video point quality was also judged (3, in focus; 2, slightly out of focus; 1, out of focus; 0, no data), so that only the in-focus video points (3) were to be considered. Similar inter-observer reliability for thermal coding was also carried out on 17% of all bird subjects' videos (20% of each species' sample size; three individuals per species) between two independent observers, only including data with a video point quality of '3'.

### 2.4.3. Statistical analysis

#### 2.4.3.1. Behavioural

The statistical computing software R (v. 4.2.2; [76]) was used to run a series of generalized linear mixed-effects models. To address the question of whether subjects would change their probability to dig, we fitted a mixed-effects logistic regression including trial condition, final food value and their interaction as test predictors and age as a control predictor. Similarly, we modelled subjects' digging duration as a response (digging duration was log-transformed for kea to improve model fit) using a mixed model with Gaussian error structure and as test predictors, trial condition, final food value and their interaction. The probability of touching the violation tower first was modelled using a mixed-effects logistic regression with exploration trial number as a test predictor. To model the percentage of time exploring the violation tower as well as overall exploration, we fitted a mixed-effects beta regression with exploration trial number as a test predictor and age as a control predictor.

To assess whether eating latency was influenced by the outcome of the trial (i.e. the food value and condition), we fitted a linear mixed-effects model with log-transformed eating latency as the response variable and trial condition (violation or non-violation), final food value (low or high) and their interaction as test predictors. Models with Gaussian and binomial error structures were fitted using the 'lmer' and 'glmer' functions, respectively, of the package 'lme4' [77], and we fitted the beta regression using the function 'glmmTMB' of the package 'glmmTMB' [78]. To avoid pseudoreplication, all of the above models included the random intercept effect of each subject. To keep the type I error rate at the nominal level of 0.05, we included all possible identifiable random slopes [79,80]. We removed correlations among random slopes and random intercepts when they appeared unidentifiable (indicated as being close to +1 or −1) [81]. When model convergence failed, we removed random slopes with the lowest estimated variance until a model converged. For all models, covariates were z-transformed to ease model convergence and achieve easier interpretation of model coefficients [82]. Inter-observer reliability for all continuous responses was excellent (Intraclass Correlation Coefficient (ICC) = 0.936–0.947, $n = 6$).

#### 2.4.3.2. Infrared thermography

We fitted linear mixed models to test whether exposure to a violation resulted in changes in periorbital temperature in the first 30 s of each trial. For each species, we ran four models on different subsets of data. Model 1a and 1b examined the effect of trial type (violation versus non-violation) on temperature in trials where the final food value was the same (i.e. high or low, respectively). Model 2 examined the effect of final food value on temperature in violation trials only. Model 3 examined the overall effects of trial type including all baseline recordings and violation and non-violation trials on temperature. Each model included the random intercept effect of subject. Model 3 also included the random slope of violation type within subjects, which was removed if it was in part unidentifiable. Models were fitted using the function lmer of the package lmerTest. Inter-observer reliability for temperature was good (ICC = 0.86, nobs = 629, $n = 6$).

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026


**Figure 6.** Infrared thermography temperature coding from periorbital region (El1) of kea subject.

#### 2.4.3.3. Behavioural and infrared thermography

After we fitted each model, we checked model assumptions and assessed model stability. We verified the absence of collinearity for each model with multiple predictors by calculating the variance inflation factor for the model excluding any interactions and random effects using the R package 'car' v. 3.0-12 [83], revealing that collinearity was not an issue. We visually inspected and confirmed that the best linear unbiased predictors per level of the random effects were approximately normally distributed [84]. For the models with a Gaussian error structure, we confirmed that the residuals were homogeneous and normally distributed. To estimate the stability of the models [85], we excluded the levels of the random effects one at a time and compared the resulting estimates with those obtained from the model based on all data. This revealed that some of the models were only of moderate stability (see below).

To avoid cryptic, multiple testing in the eating latency and digging models, we compared each full model to its respective null model lacking the test predictors of interest [86]. If this comparison revealed significance, we then continued to test the significance of the individual fixed effects by dropping them from the model one at a time (using the drop1 command) starting with the interaction and comparing the simpler model with the more complex model using likelihood ratio tests [87]. We used the function 'bootMer' of the 'lme4' package to determine 95% confidence intervals of model estimates by means of parametric bootstrapping (1000 bootstraps). Figures were created using the 'ggplot2' package [88].

We also included subjects' count descriptives from violation versus non-violation trials of whether food was eaten, whether the sandplate was roughhoused (dragged and/or slammed), vocalizations and whether cockatoos raised their crests (table 1).

## 2.5. Ethical statement

Subjects were tested individually on a voluntary basis by calling the name of the subject and allowing them to enter the testing compartment/room. If a subject exhibited behaviours indicative of refusing to participate (e.g. ignoring stimuli for more than 30 s; exhibiting neophobia), the subject was returned to the aviary and removed from the experiment. The experiment was approved by the University of Veterinary Medicine Vienna's institutional ethics committee (ETK-119/06/2022 (kea); ETK-054/04/2023 (cockatoos)) in accordance with Good Scientific Practice guidelines and national legislation. All subjects that participated in our experiments were housed in accordance with the Austrian Federal Act on the Protection of Animals (Animal Protection Act-TSchG, BGBl.I Nr.118/2004).

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

**Table 1.** Counts of subjects expressing at least one frustration/elation behaviour.

| behaviour | species | violation | non-violation |
|---|---|---|---|
| food eating | kea | 19 | 18 |
| | cockatoos | 14 | 14 |
| vocalizations | kea | 1 | 0 |
| | cockatoos | 5 | 4 |
| sandplate roughhousing | kea | 4 | 4 |
| | cockatoos | 0 | 1 |
| crest raises | cockatoos | 5 | 2 |

# 3. Results

## 3.1. Food preferences

Of the 19 kea tested, 17 significantly showed the highest preference for peanut (high-value reward) and the lowest preference for apple (low-value reward), while two significantly showed the highest preference for peanut (high-value reward) and the lowest preference for pellet (low-value reward). Of the 15 cockatoos tested, 13 significantly showed the highest preference for cashew (high-value reward) and the lowest preference for pasta (low-value reward), while two significantly showed the highest preference for walnut (high-value reward) and the lowest preference for pasta (low-value reward) (see the electronic supplementary material, table S3).

## 3.2. Behavioural

A full-null model comparison revealed an overall effect of the test predictors on the probability of digging in the kea (electronic supplementary material, tables S4 and S5). The interaction between condition and final food value appeared non-significant. After removing this, we found a significant effect of condition ($\chi^2_1 = 6.03$, $p = 0.014$), but not final food value ($\chi^2_1 = 1.96$, $p = 0.16$) in kea (electronic supplementary material, table S6). This indicates no effect of final food value but shows that the kea were more likely to dig in the violation than in non-violation trials. However, in cockatoos, the full-null model comparison revealed that there was no such clear effect of trial condition or final food value on the probability to dig (electronic supplementary material, tables S4 and S5). Digging duration was not significantly affected by trial condition or final food value in the kea nor in the cockatoos as revealed by both full-null comparisons (electronic supplementary material, tables S4 and S7; figure 7).

The full-null model comparison revealed an overall effect of the test predictors on the kea's, but not the cockatoos', latency to eat the reward (electronic supplementary material, tables S4 and S8; figure 8). Specifically, there was a significant interaction between condition and final food value on the kea's eating latency, with subjects showing the highest latency to eat in the non-violation condition when the final food value was low, and similar latencies in the violation condition regardless of final food value and the non-violation condition when the final food value was high. Quantitatively, there were too few elation or frustration-related behaviours in either species (i.e. vocalizations, sandplate roughhousing, crest raises in cockatoos) to remark on the effect of the violation on these measures (table 1).

The probability of the cockatoos to touch the violation tower first was higher in the first than in the second exploration trial ($\chi^2_1 = 4.06$, $p = 0.044$), but no such effect was found in the kea ($\chi^2_1 = 0.37$, $p = 0.542$) (electronic supplementary material, tables S4 and S9). There was no effect of exploration trial on the percentage of time spent exploring the violation tower in either species (electronic supplementary material, tables S4 and S10). But the cockatoos spent a higher overall percentage exploring the towers in the first compared to the second exploration trial ($\chi^2_1 = 4.53$, $p = 0.033$), while no such effect was found in the kea ($\chi^2_1 = 0.12$, $p = 0.732$) (electronic supplementary material, tables S4 and S11; figure 9).

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

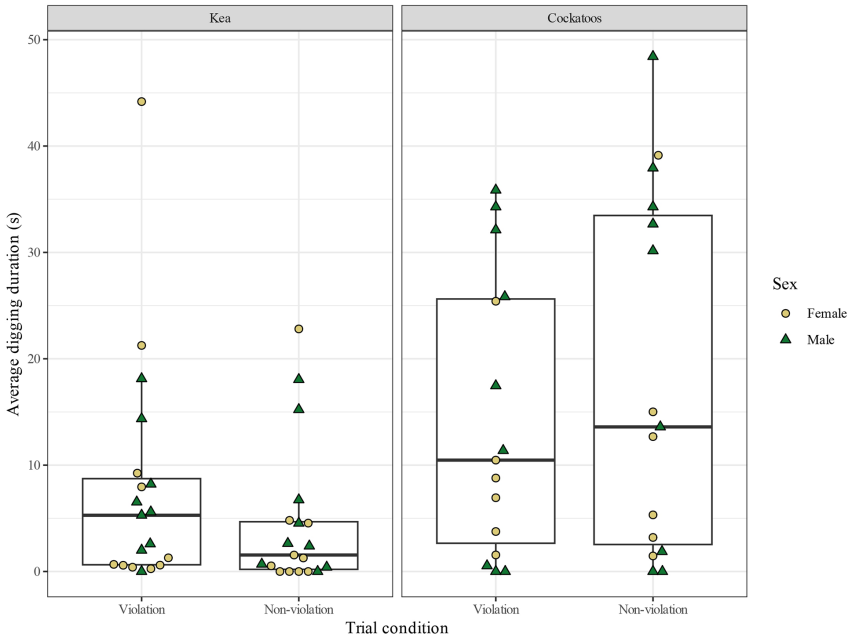


**Figure 7.** Average digging duration in violation and non-violation trials in both species.

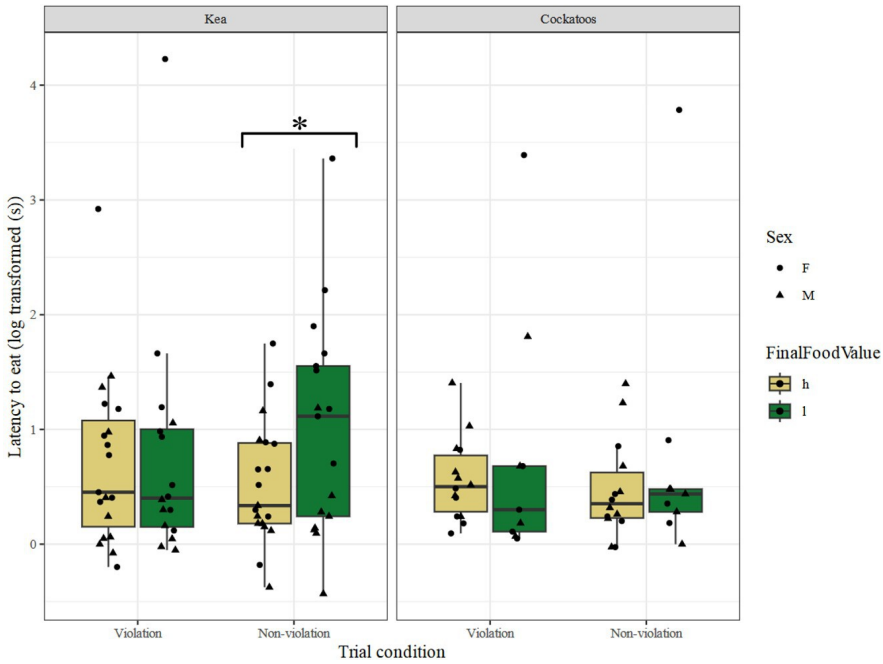


**Figure 8.** Latency to eat food in violation and non-violation trials in both species, split further by final food value.

## 3.3. Infrared thermography

For average temperatures between species, sex and trial types, see table 2 and figure 10; see the electronic supplementary material, table S12, for model comparisons between baseline, violation and non-violation temperatures. Overall, there were no significant effects of trial type nor violation direction on average periorbital temperatures in either species (table 3; see the electronic supplementary material, table S13, for percentages of video quality points of experimental trials).

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

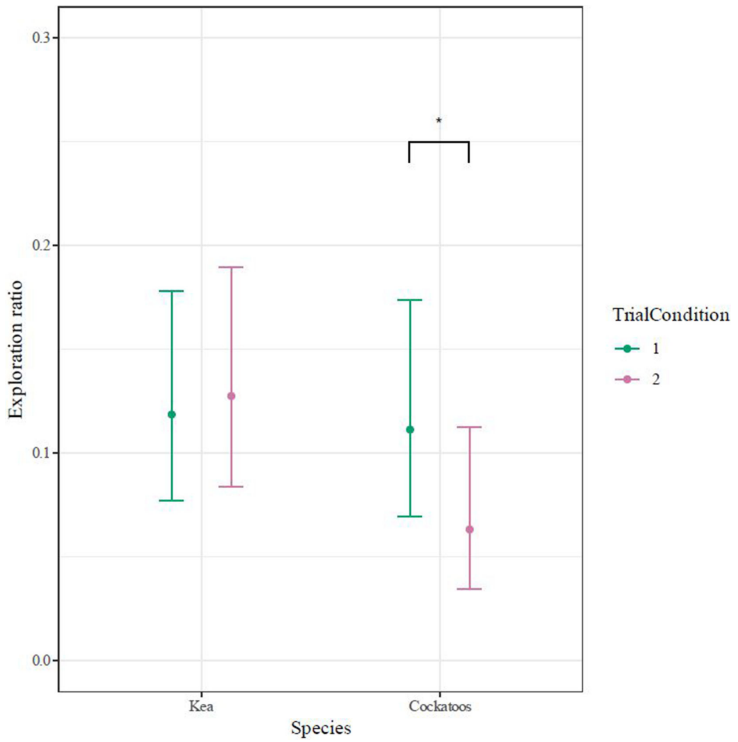


**Figure 9.** Percentage overall tower exploration before and after experimental trials. The exploration ratio was determined as the exploration duration of the towers as part of the total trial time.

**Table 2.** Average temperatures between species, sexes and trial types.

| trial type | species | sex | mean. temp (°C) |
|---|---|---|---|
| baseline | kea | M | 36.49 |
| | | F | 36.76 |
| | cockatoos | M | 38.63 |
| | | F | 38.39 |
| violation | kea | M | 36.98 |
| | | F | 37.14 |
| | cockatoos | M | 38.77 |
| | | F | 38.77 |
| non-violation | kea | M | 36.8 |
| | | F | 37.02 |
| | cockatoos | M | 38.72 |
| | | F | 38.85 |

# 4. Discussion

In this study, we examined susceptibility and physio-behavioural surprise to violations of expectations in kea and Goffin's cockatoos. We found that kea, but not cockatoos, demonstrated susceptibility to the illusion as exhibited through overt behavioural differences between violation and non-violation conditions (probability to dig), though neither species showed significant differences in physiological responses (periorbital temperature) between conditions.

Specifically, while digging duration was not affected by final food value or trial type in either species, we found that only the kea and not the Goffin's were more likely to dig in violation versus non-violation trials (regardless of violation direction). This behaviour likely reflects a continued search

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

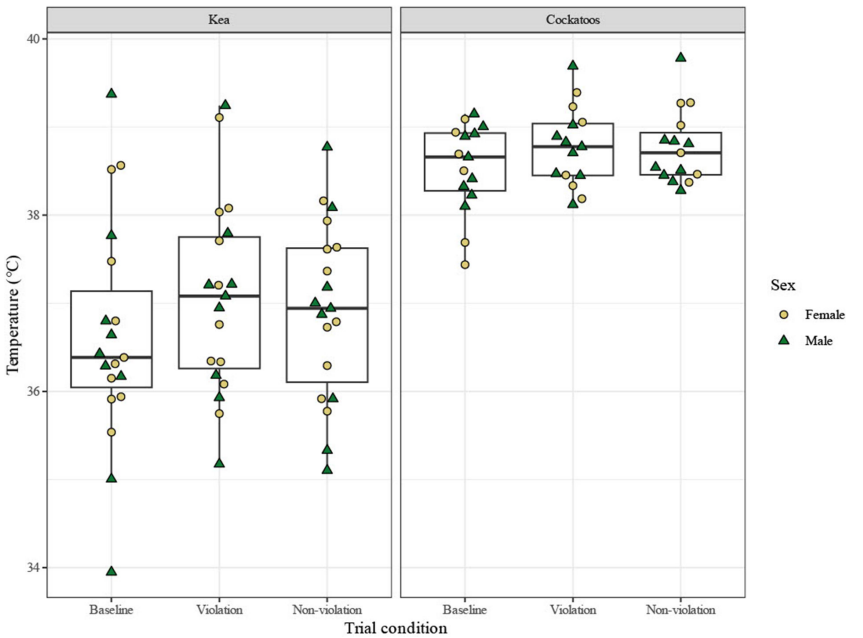


**Figure 10.** Average temperature per trial condition in both species.

**Table 3.** Statistical models of differences in periorbital temperature between trial types and violation directions in both species.

| **kea** | | | |
|---|---|---|---|
| **model** | **test predictor** | **estimate** | **significance** |
| 1a | trial type (violation versus non-violation, high final food value) | -0.134 | $t_{18}$: −0.916; $p = 0.372$ |
| 1b | trial type (violation versus non-violation, low final food value) | -0.063 | $t_{18}$: −0.47; $p = 0.644$ |
| 2 | violation direction | 0.01 | $t_{18}$: 0.059; $p = 0.954$ |
| **cockatoos** | | | |
| **model** | **test predictor** | **estimate** | **significance** |
| 1a | trial type (violation versus non-violation, high final food value) | -0.007 | $t_{14}$: −0.112; $p = 0.912$ |
| 1b | trial type (violation versus non-violation, low final food value) | 0.014 | $t_{14}$: 0.188; $p = 0.854$ |
| 2 | violation direction | -0.091 | $t_{14}$: −1.936; $p = 0.073$ |

for the expected food item, suggesting that the kea indeed perceived the violation. Not only does this indicate that they were susceptible to the violation, but it supports claims of their sophisticated physical cognition (e.g. object permanence and individuation, see [70]). Importantly, it is unlikely that this occurred owing to the species's propensity for digging, as although kea are extractive foragers [69,89], the cockatoos dug more frequently and for longer than the kea in this experiment.

The kea showed a significantly higher latency to eat in the non-violation condition when the final food value was low, but similar latencies in both types of violation trials and in the non-violation condition when the final food value was high. This could be interpreted as the kea approaching food more quickly when they either expected high-value food (violation trials, high-to-low) and/or actually encountered it (violation or non-violation trials, high final food value). By contrast, the Goffin's cockatoos showed no such difference in eating latency between conditions. These results contrast directly with comparable studies of other bird species. For example, Eurasian jays show an increased latency to consume a low-value reward when expecting a high-value reward, compared to a control condition in which they expect and receive a low-value reward [25].

Our overall results are therefore not completely consistent with prior research into violation of expectation susceptibility in animals. For example, great apes, monkeys, corvids, grey parrots and dogs are more likely to exhibit behaviours (e.g. exploration; food rejection) towards the negative violating stimulus consistent with the violation susceptibility (see [13,18,19,25]). It is unclear why the kea, and

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

not the Goffin's cockatoo, would exhibit susceptibility to the violation. The cockatoos' willingness to interact with the materials, as shown through their prolonged digging, suggests that this difference is not attributable to species differences in neophobia or neophilia. Given that Goffin's cockatoos readily pass visible displacement tasks (i.e. stage 5 object permanence tasks, [72]), they should have held an expectation about the food item they would discover. One possibility, then, is that our food-swapping violation was simply not arousing enough to elicit a behavioural response in this species.

With regard to exploration trials, we found that kea did not show a significant difference in which tower they touched first nor the duration of exploration towards the towers between their first and second exploration trials. By contrast, the cockatoos were significantly more likely to explore the violating tower first in exploration 1 than in exploration 2. Why this occurred is unclear, as there was no reason why subjects should have been attuned to their assigned violation tower before observing any experimental trials. Between the two exploration trials in the cockatoos, there was also a non-significant trend towards overall tower exploration in exploration 1, potentially owing to the novelty of the stimuli at the beginning of the experiment. These results contrast with work on other species, as after observing an event that violated their expectations about occlusion, dogs were more likely to explore the object involved in the violation [13]. As each subject in our study experienced only four experimental trials, it is possible they were not yet attuned to the colour of the violation tower, and so would not explore it specifically. Future research could address this by providing the opportunity to explore directly following violations (as in [13]), rather than in a choice scenario after both violation and control trials. While this would not offer the opportunity to examine specific versus generalized exploration, such a set-up might provide a more sensitive recording of subjects' immediate reaction to the violation rather than relying on a memory of the events.

Regarding elation or frustration effects, there were no significant effects of violation trials on behaviour or periorbital temperature in either species. Furthermore, neither the kea nor the cockatoos exhibited noticeable differences in species-specific emotional behaviours (e.g. vocalizing, sandplate roughhousing, crest raising) depending on violation direction. This contrasts with other bird species such as jays that show a greater likelihood to reject a food reward if it is of a lower value than expected [25] and grey parrots expressing frustration in negative violation trials [18]. While there was weak evidence towards Goffin's cockatoos exhibiting a decreased temperature in high-to-low violation trials, suggesting the potential presence of a frustration effect, further research is needed to differentiate this from general frustration at discovering low-quality food.

To date, the thermal response curves of parrots, along with animals' thermal responses to violations of expectations, remain largely unexplored. Our study explores the feasibility of applying IRT in cognitive research on unrestrained subjects and offers several suggestions for its implementation in future research. While we recorded periorbital temperatures of each bird every second for the first 30 s of each baseline and experimental trial, we nonetheless did not find any significant changes over or between conditions. Descriptively, we found that temperatures varied more among the kea than among the cockatoos, potentially as a result of their outdoor testing enclosure and greater mobility. Finally, although we included a baseline session prior to testing, we suggest that future research incorporate baselines prior to every trial to account for effects of previous trials and fluctuations in temperature over time.

While the stress of restraint has been documented to decrease body temperature in birds [63,90], allowing subjects full rein of their surroundings in this study came with its own issues. During the study, the mobility of the birds coupled with the size of the compartment resulted in brief periods of time with poor or absent data. Future research could also construct smaller arenas appropriately sized for each species (e.g. [63,91]), so as to avoid stress-related temperature changes stemming from restraint, but with constant data recording during trials. On the other hand, an alternative explanation is that the violation was simply not arousing enough for either species to elicit a perceptible temperature change. Therefore, future research in this area may benefit from other violation-of-expectation events varying in fundamental saliency (e.g. gravity [12,13]).

In summary, this comparative study offers a novel approach to examining susceptibility to violations of expectations via physio-behavioural measures of surprise in parrot species. While, to our knowledge, the first of its kind to employ thermography in a cognitive context in unrestrained birds, future research could benefit from minimizing the range of the subjects (e.g. smaller arenas), taking baseline measurements before each trial (for more accurate detection of temperature changes), and/or offering more salient violations of expectation. In this way, research may be afforded an improved view into the emotional states of animals following violation-of-expectation events.

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

**Ethics.** The experiment was approved by the University of Veterinary Medicine Vienna's institutional ethics committee (ETK-119/06/2022 (kea); ETK-054/04/2023 (cockatoos)) in accordance with Good Scientific Practice guidelines and national legislations. All subjects that participated in our experiments were housed in accordance with the Austrian Federal Act on the Protection of Animals (Animal Protection Act-TSchG, BGBl.I Nr.118/2004).

**Data accessibility.** This paper presents new data. Data/R files have been uploaded in addition to a corresponding title/caption legend word doc.

Supplementary material is available online [92].

**Declaration of AI use.** We have not used AI-assisted technologies in creating this article.

**Authors' contributions.** G.E.S.: conceptualization, data curation, investigation, methodology, project administration, visualization, writing—original draft, writing—review and editing; R.F.: data curation, formal analysis, validation, visualization, writing—original draft, writing—review and editing; E.D.: data curation; E.G.S.A.: investigation; I.J.: conceptualization, data curation, formal analysis, investigation, methodology, validation, writing—original draft, writing—review and editing; A.M.S.: conceptualization, investigation, methodology, supervision, writing—original draft, writing—review and editing; A.M.I.A.: conceptualization, funding acquisition, methodology, resources, supervision, writing—original draft, writing—review and editing; M.L.L.: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

**Conflict of interest declaration.** We declare we have no competing interests.

**Funding.** This research was funded in whole or in part by the Austrian Science Fund (FWF): P34533 to M.L.L. and Y1309 to A.M.I.A. For open access purposes, the author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission.

**Acknowledgements.** We thank the kea keepers at Haidlhof Research Station and head keeper Heidelinde Wanas at the Goffin Laboratory. Thanks also go to András Péter for the kea apparatus, as well as Lucrezia Lonardo and Tabea Teichmann for reliability coding. Special thanks to Dr Sylvain Giroud for the loan of his infrared thermography camera. Thanks also to Roger Mundry for providing functions to perform model diagnostics.

# References

1. Baillargeon R, Spelke ES, Wasserman S. 1985 Object permanence in five-month-old infants. *Cognition* **20**, 191–208. (doi:10.1016/0010-0277(85)90008-3)
2. Kersken V, Zhang D, Gomez JC, Seed A, Ball D. 2020 Capuchin monkeys individuate objects based on spatio-temporal and property/kind information: evidence from looking and reaching measures. *Anim. Behav. Cogn*. **7**, 343–364. (doi:10.26451/abc.07.03.06.2020)
3. Baillargeon R, Hanko-Summers S. 1990 Is the top object adequately supported by the bottom object? Young infants' understanding of support relations. *Cogn. Dev*. **5**, 29–53. (doi:10.1016/0885-2014(90)90011-h)
4. Bird CD, Emery NJ. 2010 Rooks perceive support relations similar to six-month-old babies. *Proc. R. Soc. B* **277**, 147–151. (doi:10.1098/rspb.2009.1456)
5. Cacchione T, Krist H. 2004 Recognizing impossible object relations: intuitions about support in chimpanzees (*Pan troglodytes*). *J. Comp. Psychol*. **118**, 140–148. (doi:10.1037/0735-7036.118.2.140)
6. Spelke ES, Katz G, Purcell SE, Ehrlich SM, Breinlinger K. 1994 Early knowledge of object motion: continuity and inertia. *Cognition* **51**, 131–176. (doi:10.1016/0010-0277(94)90013-2)
7. Völter CJ, Huber L. 2021 Expectancy violations about physical properties of animated objects in dogs. *Proc. Annu. Meet. Cogn. Sci. Soc*. **43**, 43. (doi:10.31234/osf.io/3pr9z)
8. Cheney DL, Seyfarth RM, Silk JB. 1995 The responses of female baboons (*Papio cynocephalus ursinus*) to anomalous social interactions: evidence for causal reasoning? *J. Comp. Psychol*. **109**, 134–141. (doi:10.1037//0735-7036.109.2.134)
9. Garcia-Pelegrin E, Miller R, Wilkins C, Clayton NS. 2023 Manual action expectation and biomechanical ability in three species of New World monkey. *Curr. Biol*. **33**, 1803–1808.(doi:10.1016/j.cub.2023.03.023)
10. Mendes WB, Blascovich J, Hunter SB, Lickel B, Jost JT. 2007 Threatened by the unexpected: physiological responses during social interactions with expectancy-violating partners. *J. Personal. Soc. Psychol*. **92**, 698–716. (doi:10.1037/0022-3514.92.4.698)
11. Bräuer J, Belger J. 2018 A ball is not a Kong: odor representation and search behavior in domestic dogs (*Canis familiaris*) of different education. *J. Comp. Psychol*. **132**, 189–199. (doi:10.1037/com0000115)
12. Stahl AE, Feigenson L. 2015 Observing the unexpected enhances infants' learning and exploration. *Science* **348**, 91–94. (doi:10.1126/science.aaa3799)
13. Völter CJ, Tomašić A, Nipperdey L, Huber L. 2023 Dogs' expectations about occlusion events: from expectancy violation to exploration. *Proc. R. Soc. B* **290**, 20230696. (doi:10.1098/rspb.2023.0696)
14. Wilson VA, Bethell EJ, Nawroth C. 2023 The use of gaze to study cognition: limitations, solutions, and applications to animal welfare. *Front. Psychol*. **14**, 1147278. (doi:10.3389/fpsyg.2023.1147278)
15. Gredebäck G, Melinder A. 2010 Infants' understanding of everyday social interactions: a dual process account. *Cognition* **114**, 197–206. (doi:10.1016/j.cognition.2009.09.004)

Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

16. Völter CJ, Huber L. 2021 Dogs' looking times and pupil dilation response reveal expectations about contact causality. *Biol. Lett*. **17**, 20210465. (doi:10.1098/rsbl.2021.0465)
17. Block J, Martin B. 1955 Predicting the behavior of children under frustration. *J. Abnorm. Soc. Psychol*. **51**, 281–285. (doi:10.1037/h0047628)
18. Pepperberg IM, Willner MR, Gravitz LB. 1997 Development of Piagetian object permanence in grey parrot (*Psittacus erithacus*). *J. Comp. Psychol*. **111**, 63–75. (doi:10.1037//0735-7036.111.1.63)
19. Tinklepaugh OL. 1928 An experimental study of representative factors in monkeys. *J. Comp. Psychol*. **8**, 197–236. (doi:10.1037/h0075798)
20. Camras LA *et al*. 2002 Observing emotion in infants: facial expression, body behavior, and rater judgments of responses to an expectancy-violating event. *Emotion* **2**, 179–193. (doi:10.1037//1528-3542.2.2.179)
21. Visser M, Krahmer E, Swerts M. 2014 Contextual effects on surprise expressions: a developmental study. *J. Nonverbal Behav*. **38**, 523–547. (doi:10.1007/s10919-014-0188-y)
22. Berger A, Tzur G, Posner MI. 2006 Infant brains detect arithmetic errors. *Proc. Natl Acad. Sci. USA* **103**, 12649–12653. (doi:10.1073/pnas.0605350103)
23. Cook PF, Prichard A, Spivak M, Berns GS. 2016 Awake canine fMRI predicts dogs' preference for praise vs food. *Soc. Cogn. Affect. Neurosci.* **11**, 1853–1862. (doi:10.1093/scan/nsw102)
24. Bräuer J, Call J. 2011 The magic cup: great apes and domestic dogs (*Canis familiaris*) individuate objects according to their properties. *J. Comp. Psychol*. **125**, 353–361. (doi:10.1037/a0023009)
25. Schnell AK, Loconsole M, Garcia-Pelegrin E, Wilkins C, Clayton NS. 2021 Jays are sensitive to cognitive illusions. *R. Soc. Open Sci*. **8**, 202358. (doi:10.1098/rsos.202358)
26. Amsel A, Roussel J. 1952 Motivational properties of frustration. I. Effect on a running response of the addition of frustration to the motivational complex. *J. Exp. Psychol*. **43**, 363–366. (doi:10.1037/h0059393)
27. Crespi LP. 1942 Quantitative variation of incentive and performance in the white rat. *Am. J. Psychol*. **55**, 467–517. (doi:10.2307/1417120)
28. Flaherty CF. 1982 Incentive contrast: a review of behavioral changes following shifts in reward. *Anim. Learn. Behav*. **10**, 409–440. (doi:10.3758/bf03212282)
29. Spence KW. 1956 *Behavior theory and conditioning,* vol. 128. New Haven, CT: Yale University Press. (doi:10.1037/10029-000)
30. Papini MR, Dudley RT. 1997 Consequences of surprising reward omissions. *Rev. Gen. Psychol*. **1**, 175–197. (doi:10.1037//1089-2680.1.2.175)
31. Fagen JW, Ohr PS. 1985 Temperament and crying in response to the violation of a learned expectancy in early infancy. *Infant Behav. Dev.* **8**, 157–166. (doi:10.1016/S0163-6383(85)80003-5)
32. Longstreth LE. 1960 The relationship between expectations and frustration in children. *Child Dev.* **31**, 667–671. (doi:10.1111/j.1467-8624.1960.tb04981.x)
33. McGraw AP, Larsen JT, Kahneman D, Schkade D. 2010 Comparing gains and losses. *Psychol. Sci*. **21**, 1438–1445. (doi:10.1177/0956797610381504)
34. Adriaense JE, Martin JS, Schiestl M, Lamm C, Bugnyar T. 2019 Negative emotional contagion and cognitive bias in common ravens (*Corvus corax*). *Proc. Natl Acad. Sci. USA* **116**, 11547–11552. (doi:10.1073/pnas.1817066116)
35. Betzel RF, Satterthwaite TD, Gold JI, Bassett DS. 2017 Positive affect, surprise, and fatigue are correlates of network flexibility. *Sci. Rep*. **7**, 520. (doi:10.1038/s41598-017-00425-z)
36. Herrnstein RJ, Morse WH. 1957 Some effects of response-independent positive reinforcement on maintained operant behavior. *J. Comp. Physiol. Psychol*. **50**, 461–467. (doi:10.1037/h0041506)
37. Dally JM, Emery NJ, Clayton NS. 2004 Cache protection strategies by western scrub-jays (*Aphelocoma californica*): hiding food in the shade. *Proc. R. Soc. B* **271**, S387–S390. (doi:10.1098/rsbl.2004.0190)
38. Russell JA. 1980 A circumplex model of affect. *J. Personal. Soc. Psychol*. **39**, 1161–1178. (doi:10.1037/h0077714)
39. Ekman P. 1993 Facial expression and emotion. *Am. Psychol*. **48**, 384–392. (doi:10.1037//0003-066x.48.4.384)
40. Teasdale JD, Howard RJ, Cox SG, Ha Y, Brammer MJ, Williams SC, Checkley SA. 1999 Functional MRI study of the cognitive generation of affect. *Am. J. Psychiatry* **156**, 209–215. (doi:10.1176/ajp.156.2.209)
41. Barrett LF. 2004 Feelings or words? Understanding the content in self-report ratings of experienced emotion. *J. Personal. Soc. Psychol*. **87**, 266–281. (doi:10.1037/0022-3514.87.2.266)
42. Iyasere OS, Durosaro SO, Oyeniran VJ, Daramola JO. 2023 Is an increase of glucocorticoid concentrations related to the degree of arousal or valence experienced by an animal to a stimulus? *Domest. Anim. Endocrinol.* **81**, 106774. (doi:10.1016/j.domaniend.2022.106774)
43. de Winkel T, van der Steen S, Enders-Slegers MJ, Griffioen R, Haverbeke A, Groenewoud D, Hediger K. 2024 Observational behaviors and emotions to assess welfare of dogs: a systematic review. *J. Vet. Behav*. **72**, 1–17. (doi:10.1016/j.jveb.2023.12.007)
44. Held SD, Špinka M. 2011 Animal play and animal welfare. *Anim. Behav*. **81**, 891–899. (doi:10.1016/j.anbehav.2011.01.007)
45. Harding EJ, Paul ES, Mendl M. 2004 Cognitive bias and affective state. *Nature* **427**, 312–312. (doi:10.1038/427312a)
46. Zipple MN, Hazelwood C, Webster MF, Benítez ME. 2024 Animal emotions and consciousness: a preliminary assessment of researchers' perceptions and biases and prospects for future progress. *R. Soc. Open Sci*. **11**, 241255. (doi:10.1098/rsos.241255)
47. Panksepp J. 1998 *Affective neuroscience: the foundations of human and animal emotions*. Oxford, UK: Oxford University Press.
48. Selvatti AP, Galvão A, Mayr G, Miyaki CY, Russo CADM. 2022 Southern hemisphere tectonics in the Cenozoic shaped the pantropical distribution of parrots and passerines. *J. Biogeogr*. **49**, 1753–1766. (doi:10.1111/jbi.14466)
49. Schuck-Paim C, Alonso WJ, Ottoni EB. 2008 Cognition in an ever-changing world: climatic variability is associated with brain size in neotropical parrots. *Brain Behav. Evol*. **71**, 200–215. (doi:10.1159/000119710)


Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

50. Bertin A, Beraud A, Lansade L, Blache MC, Diot A, Mulot B, Arnould C. 2018 Facial display and blushing: means of visual communication in blue-and-yellow macaws (*Ara Ararauna*)? *PLoS ONE* **13**, e0201762. (doi:10.1371/journal.pone.0201762)
51. Gregory R, Hopkins P. 1974 Pupils of a talking parrot. *Nature* **252**, 637–638. (doi:10.1038/252637a0)
52. Waltman JR, Beissinger SR. 1992 Breeding behavior of the green-rumped parrotlet. *Wilson Bull*. **104**, 65–84.
53. Bertin A, Beraud A, Lansade L, Mulot B, Arnould C. 2020 Bill covering and nape feather ruffling as indicators of calm states in the sulphur-crested cockatoo (*Cacatua galerita*). *Behav. Process*. **178**, 104188. (doi:10.1016/j.beproc.2020.104188)
54. Kaplan G. 2015 *Bird minds: cognition and behaviour of Australian native birds*. Victoria, Australia: CSIRO Publishing. (doi:10.1071/9781486300198)
55. Casey R, Heath S, Zulch H (eds). 2022 *Companion animal behaviour problems: prevention and management of behaviour problems in veterinary practice*. Wallingford, UK: CABI.
56. Engebretson M. 2006 The welfare and suitability of parrots as companion animals: a review. *Anim. Welf*. **15**, 263–276. (doi:10.1017/s0962728600030475)
57. Darwin C. 1872 Means of expression in animals. In *The expression of the emotions in man and animals* (ed. ET Carlson), pp. 83–115. London, UK: John Murray. (doi:10.1093/oso/9780195112719.001.0001)
58. Papini MR, Penagos-Corzo JC, Pérez-Acosta AM. 2018 Avian emotions: comparative perspectives on fear and frustration. *Front. Psychol*. **9**, 2707. (doi:10.3389/fpsyg.2018.02707)
59. Travain T, Valsecchi P. 2021 Infrared thermography in the study of animals' emotional responses: a critical review. *Animals* **11**, 2510. (doi:10.3390/ani11092510)
60. Clay-Warner J, Robinson DT. 2015 Infrared thermography as a measure of emotion response. *Emot. Rev*. **7**, 157–162. (doi:10.1177/1754073914554783)
61. Liehrmann O, Ollila A, Lummaa V, Lansade L, Seltmann MW. 2024 Enhancing stress assessment in sledge reindeer (*Rangifer tarandus*): a pilot study on infrared thermal imaging and its opportunities for advancement as a welfare assessment tool. *Anim. Behav. Cogn*. **11**, 293–304. (doi:10.26451/abc.11.03.04.2024)
62. Nakayama K, Goto S, Kuraoka K, Nakamura K. 2005 Decrease in nasal temperature of rhesus monkeys (*Macaca mulatta*) in negative emotional state. *Physiol. Behav*. **84**, 783–790. (doi:10.1016/j.physbeh.2005.03.009)
63. Jerem P, Herborn K, McCafferty D, McKeegan D, Nager R. 2015 Thermal imaging to study stress non-invasively in unrestrained birds. *J. Vis. Exp.* **105**, e53184. (doi:10.3791/53184-v)
64. Ermatinger FA, Brügger RK, Burkart JM. 2019 The use of infrared thermography to investigate emotions in common marmosets. *Physiol. Behav*. **211**, 112672. (doi:10.1016/j.physbeh.2019.112672)
65. Brügger RK, Willems EP, Burkart JM. 2021 Do marmosets understand others' conversations? A thermography approach. *Sci. Adv*. **7**, eabc8790. (doi:10.1126/sciadv.abc8790)
66. Auersperg AMI, Von Bayern AM, Gajdon GK, Huber L, Kacelnik A. 2011 Flexibility in problem solving and tool use of kea and New Caledonian crows in a multi access box paradigm. *PLoS ONE* **6**, e20231. (doi:10.1371/journal.pone.0020231)
67. Auersperg AMI, Szabo B, Von Bayern AM, Kacelnik A. 2012 Spontaneous innovation in tool manufacture and use in a Goffin's cockatoo. *Curr. Biol*. **22**, R903–R904. (doi:10.1016/j.cub.2012.09.002)
68. Auersperg AMI, Van Horik JO, Bugnyar T, Kacelnik A, Emery NJ, von Bayern AM. 2015 Combinatory actions during object play in psittaciformes (*Diopsittaca nobilis, Pionites melanocephala, Cacatua goffini*) and corvids (*Corvus corax, C. monedula, C. moneduloides*). *J. Comp. Psychol*. **129**, 62–71. (doi:10.1037/a0038314)
69. Diamond J, Bond AB. 1999 *Kea, bird of paradox: the evolution and behavior of a New Zealand parrot*. Berkeley, CA: University of California Press. (doi:10.1525/9780520920804)
70. Bastos AP, Taylor AH. 2019 Kea (*Nestor notabilis*) represent object trajectory and identity. *Sci. Rep*. **9**, 19759. (doi:10.1038/s41598-019-56380-4)
71. Smith GE, Lambert ML, Swindell E, Engelmann JM, Völter CJ. 2025 Epistemic curiosity in kea parrots and human children. *Open Mind* **9**, 1528–1542. (doi:10.1162/OPMI.a.34)
72. Auersperg AMI, Szabo B, von Bayern AMP, Bugnyar T. 2014 Object permanence in the Goffin cockatoo (*Cacatua goffini*). *J. Comp. Psychol*. **128**, 88–98. (doi:10.1037/a0033272)
73. Tattersall GJ. 2016 Infrared thermography: a non-invasive window into thermal physiology. *Comp. Biochem. Physiol.* **202**, 78–98. (doi:10.1016/j.cbpa.2016.02.022)
74. Playà-Montmany N, Tattersall GJ. 2021 Spot size, distance and emissivity errors in field applications of infrared thermography. *Methods Ecol. Evol*. **12**, 828–840. (doi:10.1111/2041-210x.13563)
75. Friard O, Gamba M. 2016 BORIS: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods Ecol. Evol*. **7**, 1325–1330. (doi:10.1111/2041-210x.12584)
76. R Core Team. 2022 *R: a language and environment for statistical computing*. See https://www.R-project.org/.
77. Bates D, Mächler M, Bolker B, Walker S. 2015 Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48. (doi:10.18637/jss.v067.i01)
78. Brooks M, Kristensen K, van Benthem K, Magnusson A, Berg CW, Nielsen A, Skaug H, Mächler M, Bolker B. 2017 glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R J*. **9**, 378–400. (doi:10.32614/rj-2017-066)
79. Barr DJ, Levy R, Scheepers C, Tily HJ. 2013 Random effects structure for confirmatory hypothesis testing: keep it maximal. *J. Mem. Lang*. **68**, 255–278. (doi:10.1016/j.jml.2012.11.001)


Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026

80. Schielzeth H, Forstmeier W. 2009 Conclusions beyond support: overconfident estimates in mixed models. *Behav. Ecol*. **20**, 416–420. (doi:10.1093/beheco/arn145)
81. Matuschek H, Kliegl R, Vasishth S, Baayen H, Bates D. 2017 Balancing type I error and power in linear mixed models. *J. Mem. Lang*. **94**, 305–315. (doi:10.1016/j.jml.2017.01.001)
82. Schielzeth H. 2010 Simple means to improve the interpretability of regression coefficients. *Methods Ecol. Evol*. **1**, 103–113. (doi:10.1111/j.2041-210x.2010.00012.x)
83. Fox J, Weisberg S. 2018 *An R companion to applied regression*. Thousand Oaks, CA: Sage Publications. (doi:10.32614/CRAN.package.carData)
84. Baayen RH, Davidson DJ, Bates DM. 2008 Mixed-effects modeling with crossed random effects for subjects and items. *J. Mem. Lang*. **59**, 390–412. (doi:10.1016/j.jml.2007.12.005)
85. Nieuwenhuis R, Grotenhuis MT, Pelzer B. 2012 Influence.ME: tools for detecting influential data in mixed effects models. *R J*. **4**, 38. (doi:10.32614/RJ-2012-011)
86. Forstmeier W, Schielzeth H. 2011 Cryptic multiple hypotheses testing in linear models: overestimated effect sizes and the winner's curse. *Behav. Ecol. Sociobiol*. **65**, 47–55. (doi:10.1007/s00265-010-1038-5)
87. Dobson AJ, Barnett AG. 2018 *An introduction to generalized linear models*. New York, NY: Chapman and Hall/CRC.
88. Wickham H. 2016 *ggplot2: elegant graphics for data analysis*. New York, NY: Springer-Verlag New York.
89. Huber L, Gajdon GK. 2006 Technical intelligence in animals: the kea model. *Anim. Cogn*. **9**, 295–305. (doi:10.1007/s10071-006-0033-8)
90. Gray DA, Maloney SK, Kamerman PR. 2008 Restraint increases afebrile body temperature but attenuates fever in Pekin ducks (*Anas platyrhynchos*). *Am. J. Physiol. Regul. Integr. Comp. Physiol.* **294**, R1666–71. (doi:10.1152/ajpregu.00865.2007)
91. Burton S, Perrin MR, Downs CT. 2008 Thermal biology of African lovebirds and Australian grass parakeets. *J. Therm. Biol*. **33**, 355–362. (doi:10.1016/j.jtherbio.2008.05.003)
92. Smith GE, Folkertsma R, Duran E, Shair Ali EG, Jacobs I, Seed AM, Auersperg AMI, Lambert ML. 2025 Supplementary material from: Physio-behavioral manifestations of 'surprise' in two parrot species: kea (*Nestor notabilis*) and Goffin's cockatoos (*Cacatua goffiniana*). Figshare. (doi:10.6084/m9.figshare.c.8169408)


Downloaded from http://royalsocietypublishing.org/rsos/article-pdf/doi/10.1098/rsos.251264/4439440/rsos.251264.pdf
by guest
on 15 May 2026