

Research



Cite this article: Baciadonna L, Solvi C, La Cava S, Pilenga C, Gamba M, Favaro L. 2021 Cross-modal individual recognition in the African penguin and the effect of partnership. *Proc. R. Soc. B* **288**: 20211463. <https://doi.org/10.1098/rspb.2021.1463>

Received: 28 June 2021

Accepted: 20 September 2021

Subject Category:

Behaviour

Subject Areas:

cognition, behaviour, evolution

Keywords:

bioacoustics, bird cognition, multimodal representations, multimodal signalling, pair bonds

Authors for correspondence:

Luigi Baciadonna

e-mail: luigi.baciadonna@unito.it

Livio Favaro

e-mail: livio.favaro@unito.it

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.5638851>.

Cross-modal individual recognition in the African penguin and the effect of partnership

Luigi Baciadonna¹, Cwyn Solvi², Sara La Cava¹, Cristina Pilenga³,
Marco Gamba¹ and Livio Favaro¹

¹Department of Life Sciences and Systems Biology, University of Turin, Turin, Italy

²Ecology and Genetics Research Unit, University of Oulu, POB 3000, 90014 Oulu, Finland

³Zoomarine Italia, Torvaianica-Pomezia, Rome, Italy

LB, 0000-0003-0011-801X; CS, 0000-0003-2517-6179; CP, 0000-0003-4883-092X; MG, 0000-0001-9545-2242; LF, 0000-0002-8698-472X

An animal's ability to recognize another individual by matching their image to their voice suggests they form internal representations of other individuals. To what extent this ability, termed cross-modal individual recognition, extends to birds other than corvids is unknown. Here, we used an expectancy violation paradigm to determine whether a monogamous territorial seabird (*Spheniscus demersus*) can cross-modally recognize familiar conspecifics (partners or colony-mates). After pairs of penguins spent time together in an isolated area, one of the penguins was released from the area leaving the focal penguin alone. Subsequently, we played contact calls of the released penguin (congruent condition) or a different penguin (incongruent condition). After being paired with a colony-mate, focal penguins' response latency to the auditory stimulus was faster in the incongruent compared to congruent condition, indicating the mismatch violated their expectations. This behavioural pattern was not observed in focal penguins after being paired with their partner. We discuss these different results in the light of penguins' natural behaviour and the evolution of social communication strategies. Our results suggest that cross-modal individual recognition extends to penguins and reveals, in contrast with previously thought, that social communication between members of this endangered species can also use visual cues.

1. Introduction

Humans have the ability to visualize familiar people by simply hearing their voice. This seemingly effortless process, termed cross-modal individual recognition, requires our brain to simultaneously integrate information from different sensory modalities and to identify an individual based on their unique set of multimodal characteristics [1]. The ability to integrate identity cues across senses demonstrates the presence of higher-order cognitive representations in the brain that are independent of modality [2,3]. Determining which animal species possess cross-modal individual recognition and how different sensory modalities are integrated and used by different species will help us better understand the evolution of communication within and across species.

Evidence for cross-modal individual recognition in non-human animals has relied on variations of expectancy violation paradigms. Here, a focal animal first experiences a second animal in one sensory modality (e.g. visually). Then, following the removal of the second individual, information from a different sensory modality (e.g. vocal call) that either belongs to the second animal or a different animal is presented to the focal subject. If the focal subject has a unique internal representation of the second individual, then this should manifest as shorter response times to incongruent sensory information. For example,

when horses were shown a familiar horse and then the call of a different horse was played from a loud speaker, the focal horse responded more quickly and looked significantly longer in the direction of the call than when the call matched the familiar horse they had just seen [3]. This result suggests that the incongruent combination violated the focal horse's expectations and indicates that they recognized the other horse across sensory modalities as a unique individual.

Auditory–visual cross-modal individual recognition of conspecifics has been demonstrated in a few mammals, including African lions (*Panthera leo*) [4], goats (*Capra hircus*) [5], horses (*Equus caballus*) [3] and rhesus macaques (*Macaca mulatta*) [6]. Cats (*Felis catus*) [7], dogs (*Canis familiaris*) [8], horses [9,10] and rhesus macaques [11] have also been shown to recognize individual humans cross-modally using auditory and visual information. In addition, auditory–olfactory representations of conspecifics have been shown in lemurs (*Lemur catta*) [12].

The only evidence for cross-modal individual recognition in a non-mammalian species thus far has been in crows (*Corvus macrorhynchos*) [13]. In an expectancy violation paradigm, crows were found to look faster and for longer when the visual and auditory cues were incongruent, but only when these cues were from group members. This finding is in line with the idea that individual recognition develops through repeated social interactions in which multiple cues are associated with specific individuals [14]. This aspect of familiarity (the quality of the relationship and the quantity of social interactions) is an understudied component of individual recognition [2,14]. Besides the work on crows just mentioned [13], only one other study tried to address this aspect. Pitcher *et al.* [5] used an auditory–visual preference looking paradigm in goats to test the impact of familiarity on individual recognition. In this type of paradigm, a stronger response towards the congruent match is assumed to indicate cross-modal recognition [10,15]. After hearing the call of a conspecific (stable mate or random herd member), goats responded more strongly towards the visual presence of the goat that matched the previous auditory stimulus [5]. Similar to the crow study, there was no evidence of cross-modal recognition for random heard members, suggesting that subjects needed a certain high level of familiarity with another individual to show evidence of cross-modal recognition in this paradigm.

Evidence of cross-modal recognition in corvids may not be very surprising, given their impressive cognitive and social communication skills. Sometimes referred to as 'feathered apes', corvids' cognitive abilities can often match or outperform those of our primate cousins [16,17]. However, the extent to which any other bird species form cross-modal representations of other individuals is unknown.

Banded penguins (*Spheniscus* spp.) provide an excellent model species to explore the extent to which cross-modal individual recognition extends within the avian taxon. Individuals of this species have a distinctive and unique pattern of black spots on their chests (figure 1a) that appear at the age of three to five months and does not change throughout their entire lives [18,19]. This unique phenotypic feature could provide a salient visual cue necessary for penguins to visually discriminate conspecifics.

Complex forms of individual recognition are expected to emerge when simpler mechanisms do not provide enough information to classify individuals [14]. Extreme environmental

conditions and a unique breeding system have driven the evolution of an efficient vocal coding mechanism in non-nesting penguins that ensures discrimination of mates and offspring [20–23]. However, in contrast with non-nesting penguins, like the king and emperor (*Aptenodytes* spp.), individual identity information encoded in the vocal calls of nesting penguins, like banded penguins, is much weaker [20,22,24,25]. This may have provided nesting penguins with the pressure to have evolved complex modality-independent representations of individuals in order to recognize other individuals reliably.

Banded penguins are colonial, territorial, monogamous, philopatric seabirds [26], and, therefore, have both significant and different levels of familiarity with their partners and colony-mates. Utilizing these characteristics, here we used African penguins, a species of banded penguin, as a model to investigate cross-modal recognition based on the degree of familiarity between individuals (partner versus familiar non-partner colony-mate). Using an expectancy violation paradigm, we expected to find evidence of cross-modal recognition and that this would be stronger for partners than familiar non-partners.

2. Methods

(a) Animals

African penguins were housed at the Zoomarine Marine Park (Torvaianica, Italy). The colony was kept in an outdoor communal exhibit (70 m²) made of sand and pebbles and a 228 m³ saltwater pool. The penguins were hand-fed by the keepers three times a day. The penguin colony was established in 2014 from seven (three females and four males) penguins from the Burger Zoo (The Netherlands) and five (two females and three males) from the Biopac Les Sables D'Olonne (France). Four penguins (two males and two penguins of unknown sex at the time of the study) were born *ex situ*. The colony consisted of 13 adult individuals and three young birds (age range: five months–17 years). Each penguin can be identified by a uniquely coloured plastic ring attached to one of their wings. In addition, the unique dot pattern present on the ventral part of their body can be used to distinguish each individual. The study was performed in an adjacent and connected area to the main exhibit area (figure 1b). This area was usually used as utility space and was not normally accessible to the penguins. Therefore, it was necessary to habituate penguins to the study area before testing commenced. For three months before testing (August–October 2020), the keepers moved the feeding sessions gradually from the main area to the testing area. This habituation allowed the birds to become familiar with the new environment.

For the study, we selected and tested 10 individuals from July 2020 to January 2021, including the habituation phase. These 10 penguins (five females and five males) were stable breeding pairs. According to the veterinary book records, penguins did not have hearing or eyesight problems and were comfortable being around humans. The subjects had not been used in any other cognitive studies.

(b) Contact calls acquisition

Contact calls (distinctive short calls expressing isolation from groups or their mates; [27]) were collected using the focal animal sampling method (cf. [28]) from February to October 2020, for a total of 44 days (time range 7.00–13.00) of data collection and 220 h of recordings. Recordings were collected at a distance of 3–10 m from the vocalizing individuals with a RODE NTG-2 shotgun microphone (frequency response 20 Hz

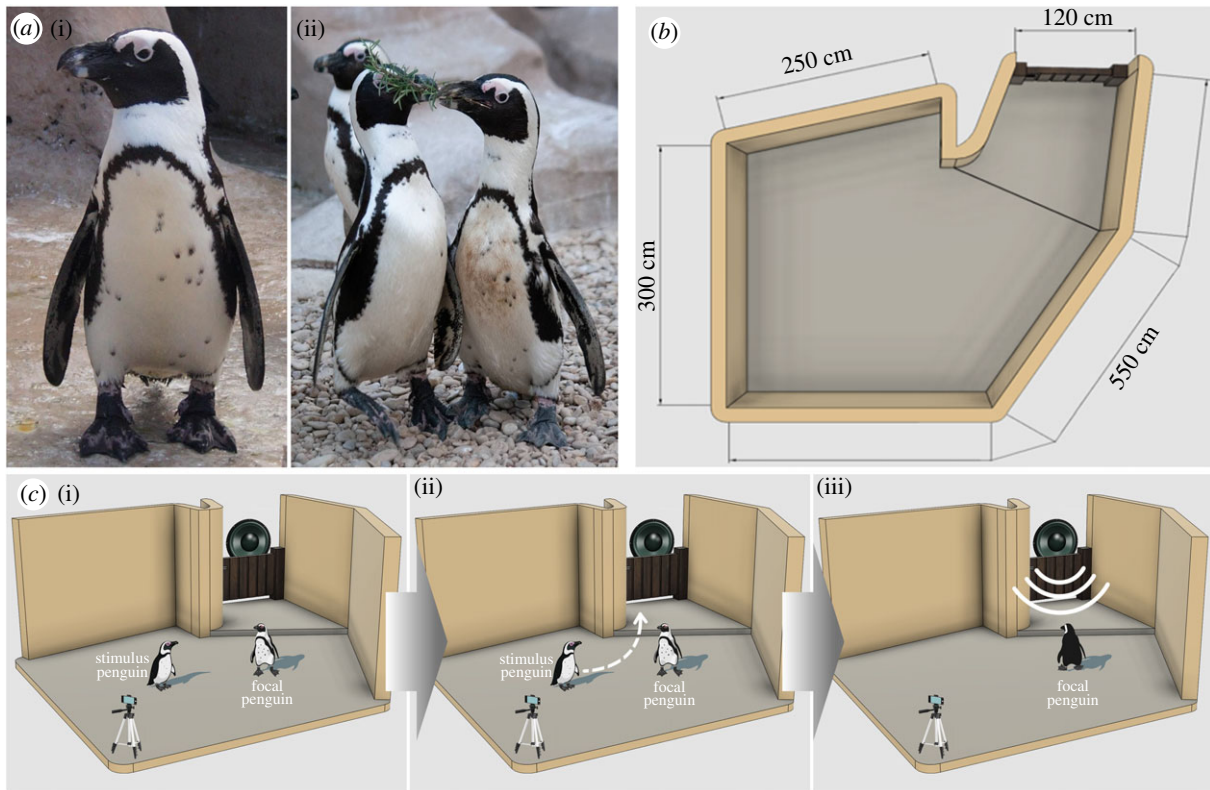


Figure 1. Experimental design. (a) Photos of three different African penguins used in the experiments. Note the distinctive unique pattern of black spots on their chests, which appear at the age of three to five months and remain the same for the rest of their lives. (b) Graphical representation of the test area from an aerial view. (c) Graphical representation of the experimental procedure. First, the stimulus and focal penguins are left alone for 60 s in the test area (i). The stimulus penguin is then led out of the area by the experimenter and the focal penguin remains in the test area alone. This process took on average 21.80 ± 1.2 s (ii). Finally (iii), two contact calls (either congruent or incongruent to the stimulus penguin) separated by 10 s are played through a speaker behind the door where the stimulus penguin exited the area. The test ended 60 s after the offset of the last call in the playback sequence. (Online version in colour.)

to 20 kHz, max SPL 131 dB) connected to a ZOOM H5 handy recorder (48 kHz sampling rate). Audio files were saved in WAV format (16-bit amplitude resolution) and stored on a secure digital (SD) memory card. From the original database, we selected a total of 54 non-temporally consecutive contact calls of good quality, preventing pseudo-replication. Fifty-four calls were extracted from 12 donor penguins (four to six calls per penguin; from the 10 tested penguins and two additional familiar non-partner colony-mate penguins simply to increase the number of call options used for testing).

(c) Playback sequence and procedure

Each playback sequence (54 total sequences) consisted of two identical calls, separated by a 10 s time interval, concatenated using PRAAT [29]. Each vocalization was broadcasted from a Bose Soundlink Mini II loudspeaker connected wirelessly to an Oppo A72 smartphone, at an approximately natural amplitude ($72.40 + 2.47$ dB) measured at 1 m using a Monacor SM-2 sound level meter. The peak amplitude of each call was equalized during the preparation of the sequences, while the duration was kept as in the original recording. Each study subject ($n = 10$) was tested using a within-subject design, with two auditory cues (partner, P; and non-partner, NP), in two conditions (congruent and incongruent). Auditory cues were from the same (S) or opposite sex (O) to the tested subjects. Therefore, each subject was tested a total of six times (two congruent and two incongruent NP, and one congruent and one incongruent P). We carried out a total of 60 trials, and because only 54 playback sequences were available, we used six sequences (11.11%) twice. Each individual was tested using an opportunistic method to comply with the limitations imposed by the hosting facility

regarding handling and separating the individuals involved in the study. The same subject was never tested twice on the same day, and there was always a minimum of 1 day between tests to prevent habituation. Each experimental trial was performed as follows. Before one of the daily feedings, two keepers entered the main exhibit area, called the attention of the entire colony and lead them all into the testing area and began the feeding session as usual. Gradually, the experimenter herded all penguins, except the two to be tested, out of the testing area. The experimenter activated the camera and moved away from the testing area while the focal penguin and stimulus penguin (partner or familiar non-partner) were left alone for 60 s, at which point the keeper entered the area and gently ushered the stimulus penguin from the testing area and therefore out of the view of the focal penguin. The average time between the stimulus penguin leaving the testing area and the broadcast was 21.80 ± 1.2 s. The sound came from behind a wooden door where the stimulus penguin had been seen leaving. The trial ended after 60 s from the offset of the second call. At the end of the trial, the keeper entered the testing area and released the focal penguin (figure 1c).

(d) Video analysis

Experimenters analysed all videos blind to conditions using BORIS software [30]. Two dependent variables were scored for the first and the second call: *response latency at the first and second call*, defined as the time between the onset of the call and when the subject's body rotated towards the direction of the loudspeaker (for subjects that did not respond, a maximum time of 10 and 60 s was assigned, respectively, to the first and second call [3,13]; and *duration of looking at the first and second*

call, defined as the time between when the head/body started to move towards the loudspeaker's direction and the time when the head/body moved away from the speaker. In addition, we scored *total duration of looking* and *total occurrences of looking* (towards the speaker). Twenty per cent of the videos were scored by a second observer blind to the study design and conditions (i.e. all the videos were re-labelled, eliminating the identity and condition in which the subject was tested). The interclass correlation coefficient calculated for all the behaviours considered was never below 0.94.

(e) Statistical analyses

Using the lme4 package [31] in R v. 3.6.1 [32], we created four generalized linear mixed models (GLMMs). The first and the second models included *response latency at the first and second call* and *duration of looking at the first and second call* (square root transformation to improve the normality) as response variables and condition (C = congruent, I = incongruent), familiarity (NP = familiar non-partner, P = partner), sex (S = acoustic sequences from an individual of the same sex of the tested subject, or D = acoustic sequences from an individual of the opposite sex of the tested subject) and call number (1 = first call, 2 = second call) as fixed factors. The interaction between condition and partner was also included as a fixed factor. The third and the fourth models included *total duration of looking* and *total occurrences of looking* as response variables and condition (C, I), partner (NP, P), sex combination (S or O) as fixed factors. The interaction between condition and partner was also included as a fixed factor. The identity of the penguins was included as a random factor to control for repeated measurements of the same subject in all models performed.

The significance of the full model was established by comparing this model with the model that included only the random factor (null model) using the likelihood ratio test. The absence of collinearity was verified among predictors by inspecting the variance inflation factors (*vif* package; [33]). The model fit and the over-dispersion were checked by using R-package DHARMA 0.3.3.0 [34]. For the models, including the interaction between two factors, all pairwise comparisons were performed using *lsmeans* multiple contrast package [35] with a Tukey *post hoc* test. The estimate, *t*-ratio and *p*-value are reported.

3. Results

Using an expectancy violation paradigm, we examined African penguins' responses to matched and unmatched visual and auditory cues from individual conspecifics with different degrees of familiarity. If penguins possess multimodal representations of other individual penguins, then after seeing a familiar conspecific, the presentation of a mismatched auditory cue should violate their expectations and result in differences in their reactions to the incongruent calls compared to the congruent calls. We further predicted that because penguins should have greater familiarity with their own partners compared to other colony members, the observed differences in their reactions should be larger for the partnered situation, i.e. when the focal penguin and stimulus penguin are partners, compared to the non-partnered situation.

Penguins were significantly faster to respond to the auditory stimulus that was incongruent (8.96 ± 2.86 s) with a familiar non-partner (NP) penguin compared to when the auditory stimulus was congruent (20.90 ± 3.60 s) with a familiar non-partner (figure 2; see table 1 for the model parameters; GLMM: estimate = 11.94, s.e. = 4.0, $t = 2.98$, $p =$

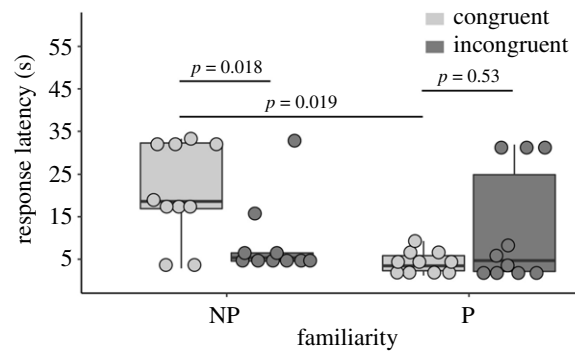


Figure 2. Focal penguins' response latency at the first and second call. The horizontal lines within the boxplot indicate the median. The boxes extend from the lower to the upper quartile and the whiskers indicate the interquartile range above the upper quartile (max) or below the lower quartile (min). Each filled circle represents the average response latency of an individual penguin. NP, non-partner colony-mate; P, partner.

Table 1. Summary of the generalized linear mixed model examining the influence of the fixed factors on response latency. (Results of the reduced model (full versus null: $\chi^2_5 = 33.25$, $p = 3.347 \times 10^{-6}$). No collinearity was found between the fixed factors (range *vif*_{min} = 1; *vif*_{max} = 2.5). The fixed factor call number and the interaction between condition and partner predicted the response latency, whereas the sex combination did not. The response latency to the first call (5.01 ± 0.47 s) was faster than to the second call (20.27 ± 3.38 s). *Post hoc* analyses revealed that the response latency was significantly faster in the incongruent non-partner condition (8.96 ± 2.86 s) compared to the congruent non-partner condition (20.90 ± 3.60 s; estimate = 11.94, s.e. = 4.0, $t = 2.98$, $p = 0.018$). In addition, the response latency was faster in the congruent partner condition (4.23 ± 0.83 s) compared to the congruent non-partner condition (20.90 ± 3.60 s; estimate = 15.64, s.e. = 5.30, $t = 2.95$, $p = 0.019$). All other comparisons were not statistically significant ($p \geq 0.43$).

	estimate	s.e.	<i>t</i>	<i>p</i> -value
(intercept)	12.24	3.76		
call number	15.26	3.19	4.78	4.977×10^{-6}
sex combination	2.05	3.91	0.52	0.59
condition:	19.63	6.77	2.89	0.0044
partner				

0.018). This result suggests that penguins possess multimodal representations of familiar colony members. Surprisingly, when the stimulus penguin was the partner of the focal penguin, we found no difference in response latency to congruent and incongruent auditory calls (GLMM: estimate = -7.67 , s.e. = 5.66, $t = -1.35$, $p = 0.53$; figure 2). However, when the auditory calls were congruent with the stimulus penguin, focal penguins looked nearly five times faster in response to their partner's call (4.23 ± 0.83) compared to a non-partner's call (20.90 ± 3.60 ; GLMM: estimate = 15.64, s.e. = 5.30, $t = 2.95$, $p = 0.019$; figure 2) suggesting that the partner-related cues may in general prompt stronger behavioural reactions. We discuss these results further below.

Neither condition (congruent or incongruent) nor familiarity (partner or familiar non-partner) predicted duration or occurrences of looking (Methods; *duration of looking at first and second call*—full versus null: $\chi^2_5 = 3.08$, $p = 0.68$; *total duration of looking*—full versus null: $\chi^2_4 = 2.97$, $p = 0.56$; *total*

occurrences of looking—full versus null: $\chi^2_4 = 7.81$, $p = 0.09$, see Methods).

4. Discussion

Cross-modal individual recognition in birds has been previously described only in large-billed crows [13]. In an expectancy violation paradigm, crows were found to associate visual and auditory cues for familiar group members but not for unfamiliar non-group members. To our knowledge, our study is the first to examine how different types of familiar relationships (partner versus familiar non-partner) affect cross-modal recognition, and in general whether a bird species phylogenetically distant from *Oscines* forms cross-modal representations of others. The results of our expectancy violation paradigm suggest that African penguins integrate visual and auditory identity cues of familiar non-partner individuals into a modality-independent cognitive representation [3].

We expected to find evidence of cross-modal recognition in these penguins given their ecology and social behaviour. African penguins are territorial, but partners nest quite close to their colony-mates. Therefore, the ability to identify one's friendly neighbours both visually and vocally may have evolved to help reduce unnecessary conflicts. In addition, during the breeding season, one of each partner pair always stays at the nest with the eggs while the other hunts with the colony. Cross-modal recognition may have proved valuable in the turbulent environment among the waves and rocks, where visual identifiers, e.g. their unique pattern of black spots (figure 1), may not be a reliable salient cue to recognize others. Therefore, to better coordinate and maintain contacts during hunting sessions, other cues were necessary, e.g. vocal calls.

The reason why we observed no differences between conditions in duration and occurrences of looking may be explained by the behavioural significance of contact calls. Contact calls are distinctive short calls expressing isolation from groups or their mates [27]. Ecstatic display calls or mutual displays calls play important roles in long behavioural processes, like mating and bonding [24,36]. Unlike these types of calls, only very brief attention is required to locate and respond appropriately to contact calls. Indeed, allocating longer attention in any way to these calls would probably be pointless in normal conditions or even disadvantageous, given that the caller would normally be moving.

By contrast, given the results for non-partner colony-mates, we found it surprising that we did not observe similar behavioural responses for partners. Learning and memorizing identity cues are time and energy demanding [14] and is expected to be found only between individuals with necessary regular contacts [5,13,37,38]. African penguins are monogamous and, especially during the breeding season, have frequent daily physical interactions with their partner [26]. Interactions with other colony-mates are less significant and occur much less often. The greater quantity and quality of multimodal interactions between partners should provide more opportunity to make stronger cross-modal associations and form cross-modal representations [14]. One possible explanation of our results is that while nesting penguins do form cross-modal representations of their non-partner colony-mates, a multimodal representation of their partner is not essential. One could imagine that the breeding pairs' interannual fidelity might be based solely on their ecstatic

calls and their dedication to their shared nest site might eliminate the need to locate their partner by more complex means during the breeding season. However, outside of the breeding season, partners will hunt together and, therefore, behave largely as colony-mates. In addition, prior to becoming partners they were colony-mates, and, therefore, would assumedly, based on our results, have already formed individual representations of each other.

A slightly better potential explanation for our results is that these penguins do form cross-modal representations of their partners but do so using different sensory modalities [39]. For example, perhaps olfactory and auditory cues are much more salient for partners and are used to form individual representations. Although the olfactory abilities of penguins and saliency of odours in the wild for individual recognition is unknown for penguins, one study suggests that they can use odours to discriminate between familiar and unfamiliar individuals and between kin and non-kin [40].

However, we speculate that a much more likely explanation has to do with how penguins may react to being separated from their partner. African penguins spend most of their free time very close to their partner compared to their colony-mates, and more importantly their partnerships are monogamous and last their entire lifetime [26]. It may be that in an expectancy violation paradigm where the stimulus penguin is a partner and is removed suddenly, the focal penguin is consequently put in a state that causes them to respond to any call very quickly (e.g. increased attention, vigilance, anxiety, etc.). In normal states, penguins' responses to their partner's calls are stronger than to calls of non-partners (e.g. [41]). In our experiments, within the congruent condition, the latency to respond to a partner's call (mean 4.23 s) was nearly five times faster than to a familiar non-partner's call (20.90 s), highlighting the much greater saliency of a partner's call. Our proposed 'partner separation effect' may be worth investigating on its own. However, for the purposes of evaluating cross-modal recognition, future studies should consider variation in the expectancy violation paradigm or entirely different designs for partnered animals. A preference looking paradigm may be a potential alternative [5,10]. Here, the focal animal can keep their partner, and a non-partner, constantly in view. Their behavioural responses can then be measured to a vocal call that belongs to one of the other two penguins. However, a design like this would be quite difficult to implement in wild animals because it requires capturing and holding three subjects within a small and restricted area. Therefore, clever adjustments to the expectancy violation paradigm or novel more flexible designs are likely to be required for wild animals to assess cross-modal recognition in partners.

Penguins, of the family Spheniscidae, separated from other birds and lost the ability to fly around 65 Ma [42,43]. They are phylogenetically distant and distinct from most other birds, including those most commonly studied in animal cognition research, such as corvids and finches. Extending the study of cognition to such behaviourally distinct species will help reveal how communication has evolved and how and which socio-environmental pressures have shaped differences in cognition through evolution. Our current results suggest that African penguins form internal representations of their colony-mates, suggesting that this ability is much more widespread among the avian taxon than previously thought. Future studies in penguins should investigate how individual

identity information about conspecifics is acquired and how other sensory-perceptual abilities could be used for individual recognition.

Ethics. The Ethics Committee approved the study of the University of Turin (approval number 280324). All procedures follow the guidelines of the Association for the Study of Animal Behaviour for the care and use of animals for research activities (2020) [44].

Data accessibility. The data are provided in the electronic supplementary material [45].

Authors' contributions. L.B. and L.F. conceived and designed the study. L.B. performed all behavioural experiments with the help of S.L.C. and C.P.

L.B. and S.L.C. performed video analyses. L.B. performed all statistical analyses with the help of C.S. and M.G. L.B., C.S. and L.F. wrote the paper with helpful input from all authors.

Competing interests. We declare we have no competing interests.

Funding. During the study, L.B. was supported by the University of Turin through a MIUR postdoctoral fellowship.

Acknowledgements. We thank Andrea Sigismundi, Mattia La Ragione, Giuseppe Monaco, Vanja Marini and all the penguin carers at Zoomarine (www.zoomarine.it) for their excellent support. We are grateful to Chiara Calcarì for helping with the recording of the acoustic stimuli. We also thank the editor and the reviewers for their valuable time and helpful suggestions.

References

- Seyfarth RM, Cheney DL. 2009 Seeing who we hear and hearing who we see. *Proc. Natl Acad. Sci. USA* **106**, 669–670. (doi:10.1073/pnas.0811894106)
- Tibbetts EA, Dale J. 2007 Individual recognition: it is good to be different. *Trends Ecol. Evol.* **22**, 529–537. (doi:10.1016/j.tree.2007.09.001)
- Proops L, McComb K, Reby D. 2008 Cross-modal individual recognition in domestic horses (*Equus caballus*). *Proc. Natl Acad. Sci. USA* **106**, 947–951. (doi:10.1073/pnas.0809127105)
- Gillfillan G, Vitale J, McNutt JW, McComb K. 2016 Cross-modal individual recognition in wild African lions. *Biol. Lett.* **12**, 20160323. (doi:10.1098/rsbl.2016.0323)
- Pitcher BJ, Briefer EF, Baciadonna L, McElligott AG. 2017 Cross-modal recognition of familiar conspecifics in goats. *R. Soc. Open Sci.* **4**, 160346. (doi:10.1098/rsos.160346)
- Adachi I, Hampton RR. 2011 Rhesus monkeys see who they hear: spontaneous cross-modal memory for familiar conspecifics. *PLoS ONE* **6**, e23345. (doi:10.1371/journal.pone.0023345)
- Takagi S, Arahori M, Chijiwa H, Saito A, Kuroshima H, Fujita K. 2019 Cats match voice and face: cross-modal representation of humans in cats (*Felis catus*). *Anim. Cogn.* **22**, 901–906. (doi:10.1007/s10071-019-01265-2)
- Adachi I, Kuwahata H, Fujita K. 2007 Dogs recall their owner's face upon hearing the owner's voice. *Anim. Cogn.* **10**, 17–21. (doi:10.1007/s10071-006-0025-8)
- Lampe JF, Andre J. 2012 Cross-modal recognition of human individuals in domestic horses (*Equus caballus*). *Anim. Cogn.* **15**, 623–630. (doi:10.1007/s10071-012-0490-1)
- Proops L, McComb K. 2012 Cross-modal individual recognition in domestic horses (*Equus caballus*) extends to familiar humans. *Proc. R. Soc. B* **279**, 3131–3138. (doi:10.1098/rspb.2012.0626)
- Sliwa J, Duhamel JR, Pascalis O, Wirth S. 2011 Spontaneous voice-face identity matching by rhesus monkeys for familiar conspecifics and humans. *Proc. Natl Acad. Sci. USA* **108**, 1735–1740. (doi:10.1073/pnas.1008169108)
- Kulahci IG, Drea CM, Rubenstein DI, Ghazanfar AA. 2014 Individual recognition through olfactory-auditory matching in lemurs. *Proc. R. Soc. B* **281**, 20140071. (doi:10.1098/rspb.2014.0071)
- Kondo N, Izawa E, Watanabe S. 2012 Crows cross-modally recognize group members but not non-group members. *Proc. R. Soc. B* **279**, 1937–1942. (doi:10.1098/rspb.2011.2419)
- Wiley RH. 2013 Specificity and multiplicity in the recognition of individuals: implications for the evolution of social behaviour. *Biol. Rev.* **88**, 179–195. (doi:10.1111/j.1469-185X.2012.00246.x)
- Spelke ES. 1985 Preferential looking methods as tools for the study of cognition in infancy. In *Measurement of audition and vision in the first year of postnatal life* (eds G Gottlieb, N Krasnegor), pp. 323–363. Norwood, NJ: Ablex.
- Emery NJ. 2004 Are corvids 'feathered apes'? Cognitive evolution in crows, rooks and jackdaws. In *Comparative analysis of minds* (ed. S Watanabe), pp. 181–213. Tokyo, Japan: Keio University Press.
- Baciadonna L, Cornero FM, Emery NJ, Clayton NS. 2021 Convergent evolution of complex cognition: insights from the field of avian cognition into the study of self-awareness. *Learn. Behav.* **49**, 9–22. (doi:10.3758/s13420-020-00434-5)
- Tilo B, Thomas B, Barham PJ, Janko C. 2004 Automated visual recognition of individual African penguins. In *Fifth Int. Penguin Conf., Ushuaia, Tierra del Fuego, Argentina, 6–10 September 2004*.
- Sherley RB, Burghardt T, Barham PJ, Campbell N, Cuthill IC. 2010 Spotting the difference: towards fully-automated population monitoring of African penguins *Spheniscus demersus*. *Endanger. Species Res.* **11**, 101–111. (doi:10.3354/esr00267)
- Jouventin P, Aubin T. 2002 Acoustic systems are adapted to breeding ecologies: individual recognition in nesting penguins. *Anim. Behav.* **64**, 747–757. (doi:10.1006/anbe.2002.4002)
- Jouventin P, Aubin T, Lengagne T. 1999 Finding a parent in a king penguin colony: the acoustic system of individual recognition. *Anim. Behav.* **57**, 1175–1183. (doi:10.1006/anbe.1999.1086)
- Aubin T, Jouventin P, Hildebrand C. 2000 Penguins use the two-voice system to recognize each other. *Proc. R. Soc. B* **267**, 1081–1087. (doi:10.1098/rspb.2000.1112)
- Aubin T, Jouventin P. 2002 How to vocally identify kin in a crowd: the penguin model. *Adv. Study Behav.* **31**, 243–277. (doi:10.1016/S0065-3454(02)80010-9)
- Favaro L, Gamba M, Alfieri C, Pessani D, McElligott AG. 2015 Vocal individuality cues in the African penguin (*Spheniscus demersus*): a source-filter theory approach. *Sci. Rep.* **5**, 17255. (doi:10.1038/srep17255)
- Robisson P, Aubin T, Brémond J. 1993 Individuality in the voice of emperor penguin *Aptenodytes forsteri*: adaptation to a noisy environment. *Ethology* **94**, 279–290. (doi:10.1111/j.1439-0310.1993.tb00445.x)
- Borboroglu G, Boersma P. 2013 *Penguins: natural history and conservation*. Seattle, WA: University of Washington Press.
- Favaro L, Ozella L, Pessani D. 2014 The vocal repertoire of the African penguin (*Spheniscus demersus*): structure and function of calls. *PLoS ONE* **9**, e103460. (doi:10.1371/journal.pone.0103460)
- Calcarì C, Pilega C, Baciadonna L, Gamba M, Favaro L. In press. Long-term stability of vocal individuality cues in a territorial and monogamous seabird. *Anim. Cogn.* (doi:10.1007/s10071-021-01518-z)
- Boersma P, Weenink D. 2009 Praat: doing phonetics by computer. See <http://www.praat.org/>.
- 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)
- Bates D, Mächler M, Bolker BM, Walker SC. 2015 Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* **67**, 1–48. (doi:10.18637/jss.v067.i01)
- R Development Core Team. 2020 R Foundation for Statistical Computing. See <http://www.R-project.org>.
- Fox J, Weisberg S. 2011 *An R companion to applied regression*, 2nd edn. Thousand Oaks, CA: Sage Publications Inc.
- Hartig F. 2020 DHARMA: Residual diagnostics for hierarchical (multi-level/mixed) regression models. See <https://CRAN.R-project.org/package=DHARMA>.
- Lenth RV. 2016 Least-squares means: the R package lsmeans. *J. Stat. Softw.* **69**, 1–33. (doi:10.18637/jss.v069.i01)

36. Favaro L, Gili C, Da Rugna C, Gnone G, Fissore C, Sanchez D, McElligott AG, Gamba M, Pessani D. 2016 Vocal individuality and species divergence in the contact calls of banded penguins. *Behav. Processes* **128**, 83–88. (doi:10.1016/j.beproc.2016.04.010)
37. Injaian A, Tibbetts EA. 2014 Cognition across castes: individual recognition in worker *Polistes fuscatus* wasps. *Anim. Behav.* **87**, 91–96. (doi:10.1016/j.anbehav.2013.10.014)
38. Sheehan MJ, Bergman TJ. 2016 Is there an evolutionary trade-off between quality signaling and social recognition? *Behav. Ecol.* **27**, 2–13. (doi:10.1093/beheco/arv109)
39. Yorzinski JL. 2017 The cognitive basis of individual recognition. *Curr. Opin. Behav. Sci.* **16**, 53–57. (doi:10.1016/j.cobeha.2017.03.009)
40. Coffin HR, Watters JV, Mateo JM. 2011 Odor-based recognition of familiar and related conspecifics: a first test conducted on captive Humboldt penguins (*Spheniscus humboldti*). *PLoS ONE* **6**, e25002. (doi:10.1371/journal.pone.0025002)
41. Clark JA, Boersma PD, Olmsted DM. 2006 Name that tune: call discrimination and individual recognition in Magellanic penguins. *Anim. Behav.* **72**, 1141–1148. (doi:10.1016/j.anbehav.2006.04.002)
42. Vianna JA *et al.* 2020 Genome-wide analyses reveal drivers of penguin diversification. *Proc. Natl Acad. Sci. USA* **117**, 22 303–22 310. (doi:10.1073/pnas.2006659117)
43. Baker AJ, Pereira SL, Haddrath OP, Edge KA. 2006 Multiple gene evidence for expansion of extant penguins out of Antarctica due to global cooling. *Proc. R. Soc. B* **273**, 11–17. (doi:10.1098/rspb.2005.3260)
44. Association for the Study of Animal Behaviour. 2020 Guidelines for the treatment of animals in behavioural research and teaching. *Anim. Behav.* **159**, I–XI.
45. Baciadonna L, Solvi C, La Cava S, Pilenga C, Gamba M, Favaro L 2021 Cross-modal individual recognition in the African penguin and the effect of partnership. Figshare.