Integrative Zoology 2023; 18: 720–735

doi: 10.1111/1749-4877.12670

ORIGINAL ARTICLE

INTEGRATIVE ZOOLOGY

Alteration of the temporal association between courtship audio and visual components affects female sexual response

Clémentine MITOYEN,[1#] Cliodhna QUIGLEY,[1,2,3#] Virginie CANOINE,[1] Silvia COLOMBO,[1] Simon WÖLFL[1] and Leonida FUSANI[1,2,3]

[1]Department of Behavioral and Cognitive Biology, University of Vienna, Austria, [2]Konrad Lorenz Institute of Ethology, University of Veterinary Medicine, Vienna, Austria and [3]Vienna Cognitive Science Hub, University of Vienna, Vienna, Austria

Abstract

Some multimodal signals—that is, occurring in more than one sensory modality—appear to carry additional information which is not present when component signals are presented separately. To understand the function of male ring dove's (*Streptopelia risoria*) multimodal courtship, we used audiovisual playback of male displays to investigate female response to stimuli differing in their audiovisual timing. From natural courtship recordings, we created a shifted stimulus where audio was shifted relative to video by a fixed value and a jittered stimulus, where each call was moved randomly along the visual channel. We presented 3 groups of females with the same stimulus type, that is, control, shifted, and jittered, for 7 days. We recorded their behavior and assessed pre- and post-test blood estradiol concentration. We found that playback exposure increased estradiol levels, confirming that this technique can be efficiently used to study doves' sexual communication. Additionally, chasing behavior (indicating sexual stimulation) increased over experimental days only in the control condition, suggesting a role of multimodal timing on female response. This stresses the importance of signal configuration in multimodal communication, as additional information is likely to be contained in the temporal association between modalities.

Key words: artificial stimulus, audiovisual synchronization, multisensory playback, signal timing, temporal configuration

INTRODUCTION

Courtship displays in many species are multimodal, combining different sensory modalities (Partan & Marler 1999; Higham & Hebets 2013). Typically, empirical studies that attempt to explain the prevalence and adaptive advantages of multimodal signals use cue isolation, comparing responses to single modalities versus full multimodal signal. Exposure to multimodal signals usually results in stronger stimulation compared to unimodal signals (Friedman 1977; Uetz *et al*. 2009), either because different modalities carry separate pieces of information (multi-message hypothesis) or because they improve signal efficiency (redundant signal hypothesis) (Moller & Pomiankowski 1993; Johnstone 1995; Partan & Marler 2005). Modalities can also interact to enhance

Correspondence: Clémentine Mitoyen and Leonida Fusani, Department of Behavioral and Cognitive Biology, University of Vienna, UZA 1, Althanstrasse 14, 1090 Vienna, Austria.
Email: clementine.mitoyen@gmail.com;
leonida.fusani@univie.ac.at
[#]Clementine Mitoyen and Cliodhna Quigley contributed equally to this work.

or suppress their respective effects (Partan & Marler 1999), resulting in a multimodal signal that is "more than the sum of its parts" (Rowe & Guilford 1996; Taylor & Ryan 2013; Hebets *et al.* 2016; Halfwerk *et al.* 2019; Mitoyen *et al.* 2019).

An aspect which has received surprisingly little attention is the temporal configuration of the single modalities within a multimodal signal. Some authors (e.g. Higham & Hebets 2013) have distinguished between "fixed" signals that are intrinsically multimodal and cannot be separated in time during production, such as the visible vocal sac inflation that accompanies acoustic communication in frogs (Starnberger *et al.* 2014), or human speech (Mcgurk & Macdonald 1976), and "fluid" multimodal signals whose timing can be controlled by the emitter. Indeed, multimodal signaling implies spatial and temporal coupling (Halfwerk *et al.* 2019), and for the receiver, multisensory integration is typically facilitated by spatial and temporal coincidence (Meredith & Stein 1986; Meredith *et al.* 1987; Holmes & Spence 2005), although this is not necessarily the case (Spence 2013). Even for species with fixed multimodal signals, such as the bow-call of the male ring dove (*Streptopelia risoria*) which relies on abdominal muscle control of bowing movements and coo vocalization (Gaunt *et al.* 1982), variation in the relative timing of components of multimodal signals is common between (Hutchison *et al.* 1997) and within (Mitoyen *et al.* 2021) individuals. Research into the temporal coordination of signal components has been done in avian courtship displays, which are diverse and often multimodal. In the broad-tailed hummingbird (*Selasphorus platycercus*), male courtship dives show an impressive consistency in their audiovisual timing (Hogan & Stoddard 2018). Similarly, in the Montezuma oropendola (*Psarocolius montezuma*), 2 elements of the visual courtship display (bow and wing spread) co-occur with 2 elements of the auditory courtship display (loudest note and lowest peak frequency, respectively) (Miles & Fuxjager 2018). In superb lyrebirds (*Menura novaehollandiae*), males produce specific songs together with specific visual display types (Dalziell *et al.* 2013). Finally, in zebra finches (*Taeniopygia guttata*), the courtship dance is strongly temporally associated with specific parts of the auditory display (Ullrich *et al.* 2016).

Several authors examined the advantage of producing co-occurring multisensory signals rather than emitting them sequentially (Partan & Marler 1999; Partan 2013; Uy & Safran 2013). The "by-product" hypothesis proposes that one part of the display directly causes another, making them necessarily coincident (Halfwerk *et al.* 2014). Multimodal timing could also be due to mechanical constraints, as in brown-headed cowbirds (*Molothrus ater*) where males may time their visual display during silence to avoid any influence of movement on sound production (Cooper & Goller 2004). Simultaneous multimodal signals could also improve signal efficiency by transmitting information in a shorter time (Hogan & Stoddard 2018). Finally, in the specific case of courtship behavior, multimodal timing could be an indicator of individual quality as it reflects motor performance and neuromuscular ability. Examples include the complex synchronized displays in Montezuma oropendola (Miles & Fuxjager 2018), golden-collared manakin (*Manacus vitellinus*) (Barske *et al.* 2011), or zebra finch (Williams 2001). In addition, coordinating multiple display elements can be cognitively challenging and therefore also under sexual selection (Dalziell *et al.* 2013).

Understanding how the configuration of a multimodal display affects receiver's responses cannot be assessed by cue isolation, but rather by manipulating the relationships between its component signals (see Smith & Evans 2013). One way of investigating the importance of the association between modalities is to experimentally disrupt their spatial or temporal relatedness (Halfwerk *et al.* 2019). For instance, a robotic male túngara frog (*Physalaemus pustulosus*) has been used to present females with different temporal combinations of visual (inflated vocal sac) and auditory (whine and chuck) courtship signals, showing that female response was reduced when calls and sac inflation were temporally interleaved (Taylor *et al.* 2017). Another study in the same species showed that females did not prefer a synchronized over a unimodal signal, but would strongly reject an asynchronous one (Taylor *et al.* 2011). Finally, in the wolf spider (*Schizocosa ocreata*), shifts in the relative timing of visual and vibratory courtship components negatively influenced female receptivity (Kozak & Uetz 2016). However, in all the above studies manipulation of the temporal association between multimodal courtship elements went beyond the natural range of variation. Although this approach is interesting to investigate mechanisms related to cross-modal integration, it tells us little about whether small variations in the timing of multimodal signals affect the response of the receiver. In addition, temporal consistency in the form of low variance in timing between an individual male's displays may play a role in female choice (Byers 2007).

The technical difficulties implicit in this type of study have limited investigations to a restricted number of species. In addition, few studies have investigated the influence of courtship elements' timing on physiological responses, rather employing cue isolation to study physiological changes such as sex steroid concentration

(Crews 1975; Friedman 1977; Kelso & Martins 2008). Playback experiments are powerful tools to study multimodal displays (Chouinard-Thuly *et al*. 2017; Witte *et al*. 2017) as they allow targeted modification of specific aspects of interest. In birds, auditory stimuli have been successfully used in laboratory (Milligan 1966; Zimmer 1982) and field studies (Horning *et al*. 1993; Otter *et al*. 1999). Nowadays, high-speed recording and display devices allow efficient testing of how variation in visual display parameters affects receiver response, even in animals with high temporal visual resolution such as birds (Boström *et al*. 2016), although limitations still persist, namely regarding color and polarization perception (Muheim 2011). Multimodal playback has been used successfully in fowls (*Gallus gallus*) (Smith & Evans 2008) and black capped chickadees (Baker *et al*. 1996), among others. In Columbiformes, Shimizu (1998) presented male pigeons (*Columbia livia*) with female videos and successfully triggered courtship behavior. More recently, female pigeons responded with natural sexual behavior to audiovisual stimuli of courting males, and the response was stronger to multimodal compared to unimodal playback (Partan *et al*. 2005). Again in pigeons, motion quality of displayed videos influenced receiver response (Ware *et al*. 2015), with video frame rate found to be critical for revealing differences in birds' behavioral responses to different stimulus material.

In this paper, we studied how the temporal association between courtship elements affected female responses in ring doves (*S. risoria*). Ring doves are monogamous, territorial breeders, and different courtship phases occur where both males and females display different behavior during several days before mating (Lovari & Hutchison 1975). During the first phase of courtship, male ring doves perform a highly temporally synchronized audiovisual display, the bow-call, where they repeatedly bow toward the female while simultaneously vocalizing. This courtship phase can last from a few seconds' to a few minutes' duration without interruption. Toward the end of the courtship interaction, behavior usually switches to more copulation-oriented displays such as nest-soliciting display and nest-cooing. Courtship has been associated with changes in physiological response (estradiol and progesterone) in many bird species (Hinde & Steel 1978), including the ring dove (Korenbrot *et al*. 1974; Lovari & Hutchison 1975). Repeated exposure to a courting male is necessary to induce physiological changes in females, resulting in increased blood estradiol (E2) concentration (Cheng 1979), and depends on exposure to multimodal courtship (Friedman 1977). In subsequent stages of courtship, these effects are mediated by the induction of female's own nest-coo vocalizations during nest building (Cheng 2003). It is, therefore, likely that the configuration of the signal impacts female behavior and also female physiological responses.

We hypothesized that female ring doves respond to the temporal association between the visual and acoustic components of the male bowing display. To assess the role of multimodal temporal configuration, we used playback to manipulate the temporal association between acoustic and visual components of male bowing displays. The temporal interval between body movements and vocalizations was either unchanged in its natural variation ("control"), shifted by a fixed amount ("shifted"), or made less temporally consistent ("jittered"), with the temporal parameters chosen to lie within anatomically possible ranges. We predicted that shifted and jittered stimuli would elicit a weaker behavioral and physiological response in female doves.

MATERIALS AND METHODS

Study species

Ring doves were obtained from breeders in Austria and France and were at least one year old to ensure that they were sexually mature (sexual maturation in this species occurs around 6 months of age (Craig 1909; Friedman 1977). Originally, all the doves were housed in mixed-sex colonies and had been exposed to courtship where they could at least hear and see males courting. We housed the 24 females in outdoor aviaries (3.6 × 3.0 × 2.7 m) separated from males 5 weeks prior to testing and transferred them indoors into individual cages (50 × 38 × 60 cm) 3 days before testing, where they remained during the whole experiment. This is a standard pre-experimental procedure in this species to limit the effects of social interaction on sexual receptivity (Hutchison 1970; Lovari & Hutchison 1975). The light regime was 14D:10N.

Ethical note

This work was approved by the local ethics committee of the Faculty of Life Sciences, University of Vienna, and by the national committee of the Austrian Federal Ministry of Education, Science and Research (BMWFW permit 66.006/0042-WF/V/3b/2017) and adheres to the ASAB/ABS guidelines for the Use of Animals in Research, the ARRIVE guidelines and the Directive 2010/63/EU guidelines. Birds had access to seed mix, grit, and water *ad libitum*. Prior to the experiment, doves underwent a habituation period of several weeks with daily handling and exposure to the experimental setup. We monitored birds during playback testing and blood sampling for signs of severe stress, which could include picking feathers, stereotypical behavior, or hissing (Miller & Miller 1958), which was never observed.

Experimental apparatus

The experimental apparatus was composed of 2 compartments (each 50 × 50 × 50 cm) separated by a net, one containing a screen and a speaker, the other containing the tested female (see Supporting Information for detailed specification of playback setup, Fig. S1). A camera recorded a side view of the female at 60 frames per second, and a microphone recorded vocalizations. After placing a female in the setup, the screen displayed an empty setup and she was allowed to habituate. We started audio and video recording of the female and then started playback stimulus display. The testing session lasted for 15 min (see Fig. S4, Supporting Information, for a detailed timeline), after which recording was stopped and the female returned to her individual cage.

Stimulus recording and manipulation

The stimulus recording setup was the same as for testing except that a female was placed behind the camera in a small cage to elicit male courtship and record it from the female's perspective (i.e. courting from the front to obtain a stimulus as similar as possible to natural courtship). We recorded male courtship audio at 48 kHz and video at 120 frames per second to ensure that their movements were perceived as continuous by the female doves. In pigeons (*Colombia livia*), a closely related species, the temporal quality of the video displayed during a playback experiment directly influenced the duration of sexual behavior (Ware *et al*. 2015). In the latter study, videos were displayed between 15 to 60 frames per second. We are therefore confident that our videos, with even higher temporal resolution (120 frames per second), appeared continuous to doves and triggered natural responses. We attempted to record courtship from each of the test males twice a day, and obtained courtship recordings of sufficient duration from 2 males. During recordings, we measured sound amplitude using a sound pressure level meter (Voltcraft 329, Conrad Electronic SE) to calibrate volume of acoustic playback during later testing. For each of the original courtship recordings, we first denoised the audio tracks using the spectral noise gating algorithm from Audacity® (Audacity Team 2017). We then annotated the timing of the bow (Loopy, http://loopb.io, loopbio gmbh, Vienna, Austria) and the bow-call (PRAAT, v.6.0.26) in order to estimate the natural variation in temporal offset between the starts of bows and bow-calls. Male A had an offset of −0.275 s (SD = 0.139 s, n = 133) between bow and bow-call of his display, whereas male B had a mean offset of −0.362 s (SD = 0.127 s, n = 109). We used these values to determine the stimulus modification values (Fig. 1; see Supporting Information and Fig. S2). We created 3 stimulus treatments by varying the relative timing between auditory and visual courtship tracks (Narins *et al*. 2005; Taylor *et al*. 2011; Kozak & Uetz 2016) (Fig. 1). Using Matlab (Mathworks, Natick, MA), each bow-call was extracted from the source track using the annotations mentioned above. For the control stimulus, no change in timing was made and bow-calls were placed at the same onsets in the new audio track (Fig. 1). For shifted courtship, each bow-call was placed 0.32 s earlier in the new audio track (corresponding to a shift in the mean of natural audiovisual timing). In the jittered courtship condition, the new onset of each bow-call was randomly drawn from a normal distribution with mean = 0 and SD = 0.25 s (with the constraint that subsequent bow-calls did not overlap). Thus, for both shifted and jittered treatments, we ensured that the modified stimuli fell into naturally occurring variation (Fig. S2, Supporting Information). The spectral and temporal structure of the calls and the total number of bow-calls were unchanged for the 3 stimulus treatments. Finally, we normalized audio of all stimuli to the same peak amplitude. Videos of the 3 stimulus types can be found in the Supporting Information (Videos S1–S3).

Experimental design

We assigned each female to a treatment group (control, shifted, or jittered stimulus; 8 females per group) and to a stimulus male (1 or 2; 4 females per male per treatment group) (Fig. S3, Supporting Information). Each female was exposed to the playback stimulus during 15-min sessions on each of the 7 consecutive days (Lehrman *et al*. 1961; Barfield 1971). This exposure is sufficient to elicit responses from females, as in this species relatively short exposure to males is enough to trigger behavioral and endocrinal changes (Lehrman 1964). It usually takes 7 to 10 days of male exposure for a female to complete the reproductive cycle and lay an egg (Lehrman *et al*. 1961). Additionally, Barfield (1971) showed that 15-min daily exposure to a live male over 7 days is enough to increase female oviduct size. The playback stimulus was composed of a sequence of

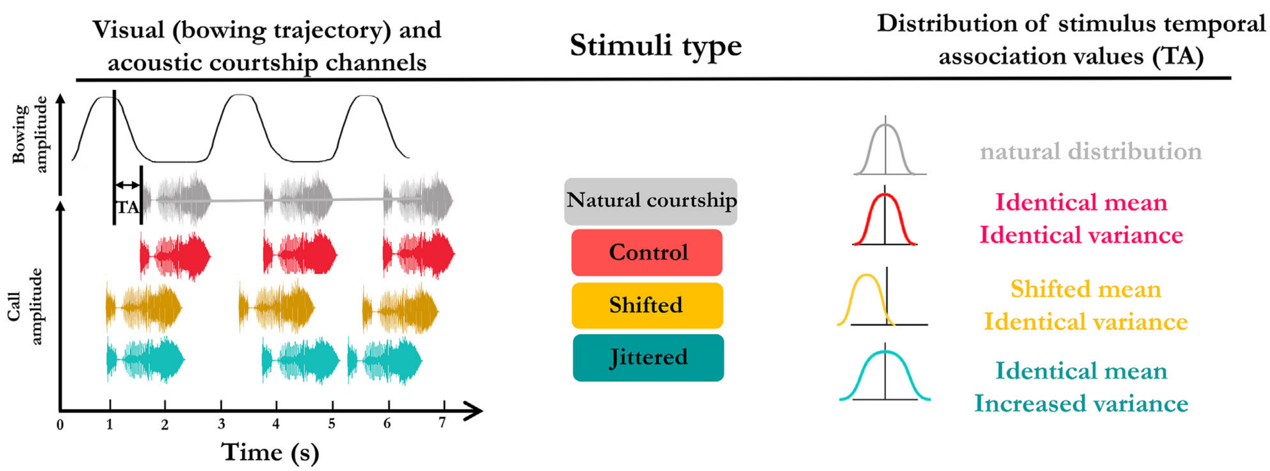

Figure 1 Stimulus creation. Bow-calls were isolated from natural courtship recordings (grey) and placed into new auditory tracks to create modified stimuli. We used unaltered timing between the onset of the bow-call and the onset of the bow (TA = temporal association, corresponding to the interval between the initiation of the descending phase of the bow and the onset of the call) for the control stimulus (red). We built the shifted stimulus (yellow) by shifting the start of each bow-call by a fixed value to change the mean while conserving the variance so that TA remained within the natural variance as depicted on the right (calculated from 242 bow-calls belonging to 2 birds, overall mean = −0.32, SD = 0.14; see Fig. S2, Supporting Information, for individual male data). For the jittered stimulus (green), the mean of the TA distribution was maintained, while the variance was increased.

courtship bouts alternated with videos of the empty setup. The duration of courtship bouts ranged between 30 and 31 s corresponding to 13–14 bow-call displays. In each session, females were presented a series of 2 courtship bouts separated by a short (2–4 s) video of the empty cage, followed by a longer (57 s) video of the empty cage before the next repetition of courtship bouts (Fig. S4, Supporting Information). Courtship videos were each presented 4 times in each session in random order, resulting in 15-min sessions containing 8 min of active bow-call courtship display (see Fig. S4, Supporting Information). The order of presentation of courtship videos was randomized, with the constraint that the same video could not appear twice in a row. As a result, during this 15-min playback, only 8 min displayed a courting male. We used empty video intervals as females may perceive a long and uninterrupted courtship as too aggressive (Mitoyen *et al*. 2021). Because of time constraints that prevented testing all birds in parallel, we tested half of each group (12 females, 4 females of each group) in the second week of August 2019 (testing week a), and half in the third week of August 2019 (testing week b).

Video analysis

Using manual coding software (Loopy, loopbio GmbH, Austria), we coded 3 female behaviors previously associated with sexual behavior in doves and other bird species: tail quivering (Cheng 1973; Witte 2006; Amy *et al*. 2015), preening (Cheng 1973; Partan *et al*. 2005), and chasing. In doves, chasing has not previously been studied in females, but in males, this behavior is typically displayed before the bow-call display. The precise role of the chasing behavior is not known, but because it typically happens during the first phases of courtship interaction, it is thought to be an important component of sexual behavior (Lovari & Hutchison 1975; Cheng 1979). The bird adopts a characteristic horizontal body posture while walking, feathers of the back are often erected, and there is sometimes an approach toward the other individual and a specific call, the kaa-call. Additionally to these 3 sexual behaviors, we also coded the number of steps as a proxy for overall locomotor activity. For each behavior, we recorded its duration but also the time and frame number at which it occurred during the 15 min of testing. As we knew the times when courtship or control videos were played, we could assign behavioral occurrences to either period. If a behavior bridged 2 periods, it was assigned to the first one. Sample videos of coded female behaviors are available in the Supporting Information (Videos S4–S7).

Estradiol assays

One day before the first and one day after the last exposure to playback, we took a blood sample from each female (up to 0.5 mL). Samples were centrifuged for 5 min at 2000 rpm, then plasma was pipetted into Eppendorf tubes and stored at −20°C until assay. Estradiol plasma concentrations were quantified using commercial enzyme-linked immunosorbent assay (ELISA) kits (RE52041; IBL, Hamburg, Germany). Prior to analyzing the samples, validations were conducted to eliminate the interference of the dove plasma matrix that caused a shifting problem (see Supporting Information). Both samples of each individual (before-after) were run on the same assay plate to reduce variability and all samples were analyzed in duplicates. Two assays were performed in total with a sensitivity of 3.6 and 5.81 $\text{pg}\cdot\text{mL}^{-1}$, respectively. Samples below these 2 detection thresholds were marked as left-censored for later analysis. Intra-assay CV% of duplicates was below 5% (mean 0.9%). Inter-assay CV% of the kit control and of an extracted plasma control pool was 4.74% and 0.61%, respectively.

Statistical analysis

We performed all statistical analysis using R (R Core Team 2018). For all female behaviors, number of events and overall duration were highly correlated (Spearman correlation, $r > 0.90$, $P < 0.001$). Therefore, we used only counts for further analysis. For every model, we used a full-null model comparison approach with a likelihood ratio test to investigate the effect of fixed factors (Dobson & Barnett 2018). The null model lacked the respective fixed effects or interaction, but random effects and controls remained.

We used generalized linear mixed models to assess if female behavior differed depending on whether the playback displayed courtship or empty setup (courtship playback present vs. absent). We fitted models using the glmmTMB function and a negative binomial distribution to model number of female behavior events (preening, steps, chasing, tail quivers) and to account for overdispersion. We created the variable "testing week," as we tested 12 females in the second week of August 2020 (testing week a) and 12 more in the third week of the same month (testing week b). We randomly assigned 2 females from each group to each of the 6 available stimuli. We created the variable "testing session" to refer to a given experimental slot for a given female a given day. Full models included female identity nested into testing session as a random effect, playback content (courtship playback present vs. absent), and experimental days (1–7) as fixed effects and testing week (a/b), stimulus male (1/2), and female origin (France/Austria) as controls.

We also used glmmTMB to model the effect of stimulus type (control, shifted, jittered) on female behavioral response. We assumed that female behavior during empty setup videos was influenced by the preceding stimulus and therefore used the full dataset comprising female behavior during the whole 15-min session. Full models included female identity as random effect, interaction between stimulus type and experimental day as fixed effect, and testing week, stimulus male, and female origin as controls.

Some of the concentration values were below the detection threshold (10 out of 24 pre-test samples and 5 out of 24 post-test samples). Where possible, we used methods that take left-censoring into account (Shoari & Dubé 2018) and compare them to standard approaches on the dataset with censored values replaced with the detection threshold. We estimated summary statistics taking left-censoring into account using the Regression on Ordered Statistics approach (functions censummary and censtats in package NADA, Helsel 2012) and plotted censored data using function cenboxplot (Fig. 4). As we did not find an appropriate left-censored alternative and because substituted data were very similar to censored data (see Results for summary statistics), we used linear mixed models (lmer function) and linear models (lm function) with detection-limit substituted data to investigate whether E2 concentration was influenced by stimulus type or sampling phase (pre-exposure/post-exposure). Full models comprised female identity as random factor, stimulus type, and sampling phase as fixed effects, and testing week, female origin and stimulus male as controls.

In addition, we used Kendall's tau correlation for censored data (function cenken in package NADA) to test the association between behavioral variables on the first day and pre-testing E2 blood concentration, and the association between the number of behavioral events on the last day and post-test E2 concentration. We also report the results for Spearman correlation tests (Harrell & Dupont 2020) without taking censoring into account, that is, when non-detects were replaced with detection limit.

RESULTS

Female response to audiovisual playback

Playback content had an effect on occurrence of preening ($\chi^2_1 = 109.3$, $P < 0.001$), steps ($\chi^2_1 = 38.206$, $P < 0.001$), chasing ($\chi^2_1 = 55.118$, $P < 0.001$), and tail quivering behaviors ($\chi^2_1 = 25.21$, $P < 0.001$). For every behavior, event number was on average significantly higher when the courtship playback was present than when courtship playback was absent (Fig. 2; Tables 1, 2; Table S1, Supporting Information).

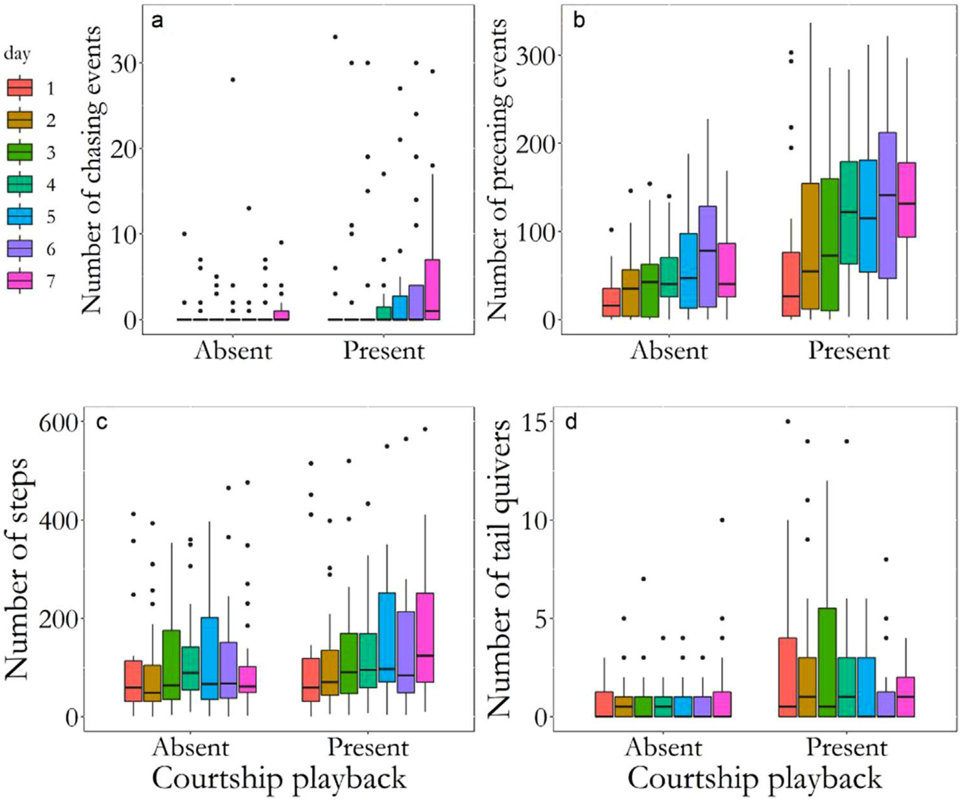

Figure 2 Boxplots of behavioral occurrences displayed by females, grouped by playback content and day of experiment. Playing back courtship stimuli resulted in more behavioral events, and the response increased across the experimental period.

Table 1 Number of behavioral events displayed by females and mean (SD) per female depending on playback content and χ^2 test statistics from the full-null model comparisons

	Empty setup		Courtship		
	Total	Mean (SD)	Total	Mean (SD)	χ^2 test statistics
Chasing	135	0.80 (2.83)	552	3.28 (7.28)	55.11*
Preening	8671	51.61 (49.46)	18 880	112.38 (91.35)	109.3*
Steps	18 409	109.57 (107.79)	22 666	134.91 (127.05)	38.206*
Tail quivers	138	0.82 (1.41)	341	2.02 (3.12)	25.21*

*$P < 0.001$.

Experimental day significantly affected occurrence of preening ($\chi^2_6 = 53.261$, $P < 0.001$), steps ($\chi^2_6 = 18.179$, $P = 0.005$), and chasing ($\chi^2_6 = 24.477$, $P < 0.001$), with number of behavioral events increasing over experimental days. No difference between days was found for tail quiver occurrence ($\chi^2_6 = 7.25$, $P = 0.29$).

Effect of courtship audiovisual timing on female behavioral response

Although females preened more often in the control condition than in shifted and jittered conditions (Table S2, Supporting Information), the effect of audiovisual timing

Table 2 Analyses of deviance table (with Chi-square value, degrees of freedom) from the results of generalized linear models investigating the effect of playback content (PB). Results are presented for all the investigated female number of behavioral events: (a) chasing, (b) preening, (c) steps, and (d) tail quivers

	Analysis of deviance table		
	χ^2	df	*P*-value
(a) Chasing			
PB	73.5667	1	<0.001
Day	17.1165	6	0.008864
Week	0.0167	1	0.897242
Male	1.2083	1	0.271677
Origin	2.5702	1	0.108896
(b) Preening			
PB	156.286	1	<0.001
Day	60.5574	6	<0.001
Week	2.8655	1	0.0905
Male	0.5991	1	0.4389
Origin	1.0782	1	0.2991
(c) Steps			
PB	42.3999	1	<0.001
Day	18.8319	6	0.004457
Week	0.7769	1	0.378082
Male	2.3545	1	0.124924
Origin	3.8125	1	0.05087
(d) Tail quivers			
PB	26.2279	1	<0.001
Day	6.9781	6	0.3229
Week	1.748	1	0.1861
Male	0.0341	1	0.8535
Origin	1.104	1	0.2934

was not significant ($\chi^2_2 = 3.10$, $P < 0.21$). Stimulus type had no effects on number of steps ($\chi^2_2 = 0.52$, $P = 0.76$), tail quivers ($\chi^2_{14} = 17.33$, $P = 0.23$), or chasing behavior ($\chi^2_2 = 1.80$, $P = 0.40$), although similarly to preening, females displayed the most steps and chasing in the control condition. Experimental day impacted preening ($\chi^2_6 = 49.00$, $P < 0.001$), steps ($\chi^2_6 = 22.03$, $P < 0.001$), and chasing ($\chi^2_{14} = 43.33$, $P < 0.001$), with females displaying more of these behaviors as the experimental days passed. However, day did not impact the tail quivering behavior ($\chi^2_{14} = 8.84$, $P = 0.18$).

The interaction between experimental day and stimulus type was not significant for preening ($\chi^2_{12} = 7.38$, $P = 0.83$), steps ($\chi^2_{12} = 14.86$, $P = 0.24$), or tail quivers ($\chi^2_{12} = 14.55$, $P = 0.26$), but was significant for chasing behavior ($\chi^2_{12} = 21.11$, $P = 0.04$). Post hoc tests and pairwise comparisons revealed that chasing significantly increased during the experimental period only in the control condition (Fig. 3; Tables S3, S4, Supporting Information).

Effect of courtship audiovisual timing on female physiological response

E2 plasma concentration ranged from non-detectable (below 5.2 pg·mL^{-1}) to 31.98 pg·mL^{-1} (11.78 ± 8.92 when left-censoring was taken into account; 12.17 ± 8.53 when censoring was not taken into account and non-detects substituted with detection threshold, mean ± SD) for pre-exposure blood sampling, and from non-detectable to 41.88 pg·mL^{-1} (left-censored: 17.88 ± 11.10; detection threshold substitution: 17.92 ± 11.03) for post-exposure blood sampling.

The difference in E2 levels between pre and post-test ranged from −8.52 to 23.44 (5.74 ± 7.51, positive differences reflecting higher post-test concentration; non-detected samples substituted with detection limit, see Materials and Methods), and percentage increase ranged from −28% to 286% (+59% ± 72%). E2 plasma concentration was significantly increased after the playback experiment ($\chi^2_1 = 12.05$, $P < 0.001$, Fig. 4; Tables S5, S6, Supporting Information).

Although plasma E2 concentration was higher in the control condition and stimulus type influenced E2 concentration ($\chi^2_2 = 6.45$, $P = 0.03$), post hoc pairwise comparison tests between stimuli types were not significant (Table S7, Supporting Information). There was no interaction between stimulus type and sampling status ($\chi^2_2 = 1.35$, $P = 0.50$). Stimulus male did not influence E2 concentration ($\chi^2_1 = 0.15$, $P = 0.69$), but females from group 2 (tested 1 week after females from group 1) had lower E2 values than females from group 1 ($\chi^2_1 = 37.37$, $P < 0.001$). However, percentage increase between pre and post-exposure concentration did not differ between groups ($F = 2.26$, $P = 0.14$), condition ($F = 1.48$, $P = 0.25$), or stimulus male ($F = 0.005$, $P = 0.94$).

The only significant correlation between E2 concentration and behavior was between post-exposure concentration and tail quiver number on the seventh day (Kendell's tau = −0.39, $P = 0.007$; Spearman $r = -0.52$,

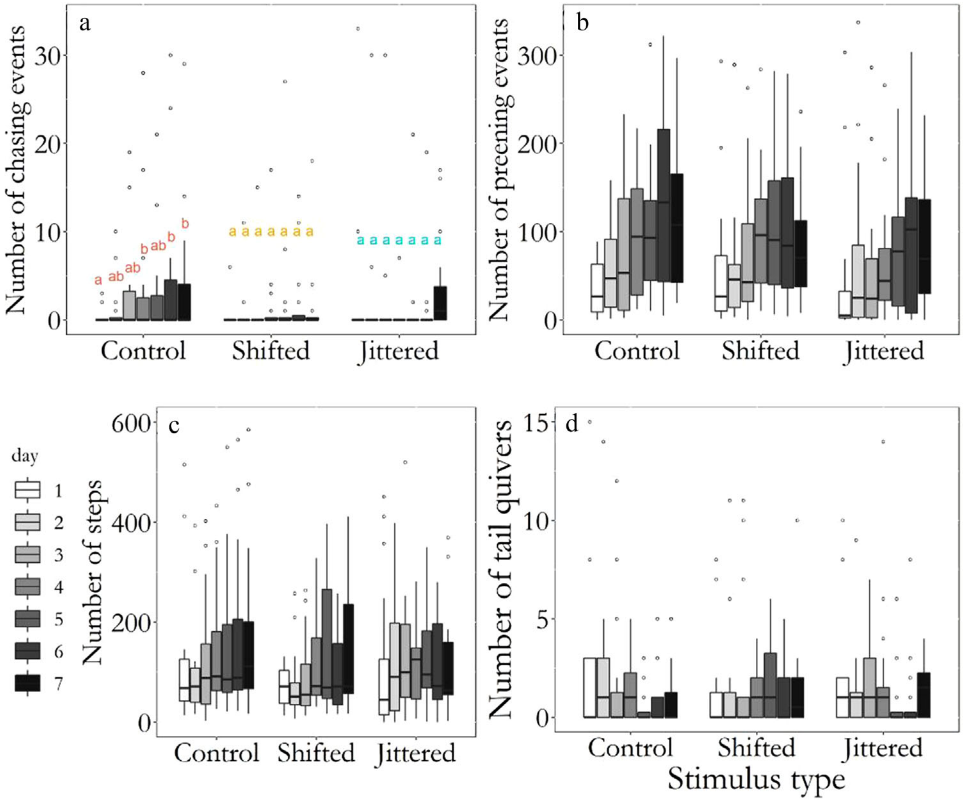

Figure 3 Boxplots (median, box spans lower and upper quartiles, whiskers show values within 1.5 times the interquartile range and dots depict outliers) of behavioral occurrences displayed by females, grouped by stimulus type and day of experiment: (a) chasing, (b) preening, (c) steps, and (d) tail quivers. There was no significant difference between conditions for any behavior. We displayed significance letters between experimental days for chasing only; as for this behavior, we found a significant interaction between stimulus type and days.

$P = 0.008$), with females showing higher number of tail quivers on the last experimental day having lower E2 concentration (Table S8, Supporting Information).

DISCUSSION

We investigated behavioral and physiological responses of female ring doves to playback stimuli of male courtship differing in the audiovisual timing between the visual and acoustic components of the bowing display. In addition, we examined how female behavior changed over the 7 consecutive experimental days during playback of courtship stimuli or of videos showing a silent and empty cage. We found that females responded more to courtship playback than to videos showing an empty cage. This is likely due to the courtship display, and not just the presence of a male on the screen, as Mitoyen *et al*. (2021) showed that female sexual behavioral responses differed when females were exposed to a live courting male compared to a male that was not courting. As both the behavioral and physiological responses of female doves changed during and after exposure to our stimuli, we are confident that they responded to a courting male in the playback, and not simply to an object moving on the screen. We also showed that sexual and non-sexual behaviors increased as the experiment progressed. We additionally documented an increase in E2 concentration at the end of the experiment compared to pre-experimental values, reflecting female sexual stimulation. An increased female physiological response after repeated exposure to courtship has already been documented in doves and canaries (Cheng 1973; Amy *et al*. 2015). Finally, although we could not convincingly show that manipulation of audiovisual timing affected all aspects of female behavioral response, we found that only those females presented with the unaltered, control stimulus significantly

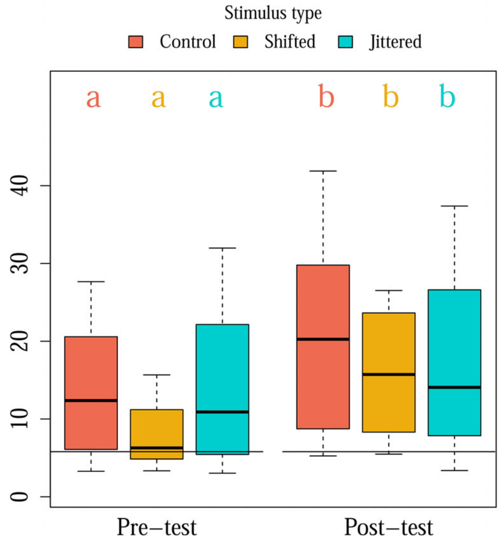

Figure 4 Boxplots (median, lower, and upper quartiles and values within 1.5 times the interquartile range) of E2 concentration (pg·mL^{-1}) of females grouped by pre- and post-exposure and stimulus types and taking into account left-censored data (package NADA). The horizontal line represents the highest censoring threshold. Pre- and post-exposure E2 concentrations are significantly different and letters indicate lack of differences between stimulus types within blood sampling phase.

increased their chasing behavior across the experimental period. Additionally, plasma E2 was higher in females exposed to the control stimulus. This suggests that the natural temporal configuration of audiovisual courtship plays a role in female evaluation of male courtship.

Ring dove courtship is a dynamic phenomenon, with behavior and physiological responses of males and females developing through different stages (Lovari & Hutchison 1975). Before testing, females were separated from males to bring them into a basal stimulation state. In females, E2 production significantly increases after a few days of courtship interaction (Korenbrot *et al*. 1974), which was confirmed by our E2 results. Females' behavioral responses to male bow-calls (Cheng 2008; Mitoyen *et al*. 2021) and their subsequent shift to nest-oriented behavior after a few days of courtship (Cheng 1973) are also well documented. This shift could explain the correlation patterns we found between E2 levels and number of tail quivers. Indeed, if tail quiver is a sign of sexual interest displayed in the first phase of the courtship interaction, we could expect a negative correlation of this behavior with post-exposure E2 concentration, which was measured on the subsequent day. Additionally to physiological stimulation, our results show that audiovisual playback increased females' sexual and non-sexual behavioral responses, as steps, preening, and chasing increased over experimental days.

We designed this study to examine the importance of multimodal configuration in courtship. The amount of courtship (number of bows and bow-calls) in both modalities was the same for all 3 audiovisual timing conditions. A similar response to the control and the shifted condition in contrast to the jittered condition would therefore indicate that more than the average relative audiovisual timing, it is rather the repetition rate and variance (i.e. consistency) in timing of the multimodal courtship that are important for females. In contrast, a similar response to the 2 modified stimuli compared to control could indicate a negative effect of a signal whose parts are too strongly desynchronized. Our results suggest the latter is the case. Consistent shifting of the acoustic track (shifted condition), as well as randomly placing calls along the bowing sequence (jittered condition), negatively affected female sexual response, but not their non-sexual response. The number of chasing events increased from the beginning to the end of the experiment only in females presented with control stimuli, although the overall number was not significantly influenced by stimulus type (although numerically greater in the control condition, Fig. 2). These behavioral results, in addition to the moderate effect of stimulus type on plasma E2 concentration, therefore, provide support to the hypothesis that multimodal timing and signal configuration of courtship influence female response. The lack of post hoc significant differences regarding stimulus type effect on plasma E2 could be explained by an overall weak effect or a small sample size. Some information triggering an increase of chasing and plasma E2 appears to be present in control stimuli, but not in modified videos.

Signal configuration, particularly regarding the timing of constituent elements of multimodal signals, is likely to play an important role in sexual interactions, and our results add to existing data suggesting that signal configuration itself is a target of inter- and intra-sexual selection. Our results are in line with work on courtship synchronization in wolf spiders (Kozak & Uetz 2016) and túngara frogs (Taylor *et al.* 2011), where females were more attracted to males displaying natural courtship timing and rejected shifted or jittered stimuli. Although ring dove courtship is less complex and elaborated than the displays of some tropical birds, the capacity to control the timing of multimodal courtship could reflect good motor and neural control and might positively affect female sexual response (Barske *et al*. 2011; Miles & Fuxjager 2018). Whether the by-product hypothesis or the mechanical constraint hypothesis is relevant here is unknown. Male doves can produce calls without simultaneously performing the bow (perch-call and nest-call (Hutchison *et al*. 1997), but abdominal muscles are involved in both behaviors (Gaunt *et al*. 1982). In doves, we know that esophagus inflation amplifies calls (Fletcher *et al*. 2004; Riede *et al*. 2004), and we cannot exclude that bowing further enhances this inflation. The visual bowing component could thereby increase call amplitude, and higher sound pressure level is more attractive for females in many bird species (Ritschard *et al*. 2010; Zollinger & Brumm 2015).

A previous behavioral study by our group showed that tail quiver occurrence in female ring doves was associated with lower fundamental frequency calls but had no relation to average audiovisual timing of courtship (Mitoyen *et al*. 2021). In this previous study, chasing behavior was not investigated, and we therefore cannot exclude that it was also correlated with audiovisual timing. Our present results indicate that manipulations of audiovisual timing also do not affect tail quiver behavior. Rather, the increasing number of tail quivers during playback along the course of the experiments could be linked to repeated exposure to male courtship. In doves, females' own calls are physiologically self-stimulating and increase E2 levels (Cheng 2008). We were not able to measure female vocal behavior as they did not vocalize during experimental sessions, but the effect of male courtship on female stimulation was possibly mediated by females' own acoustic stimulation in their home cages. However, the higher plasma E2 of females in the control stimulus condition (even though this effect would need to be confirmed in further experiments) supports the notion that audiovisual synchronization enhances sexual stimulation of females. The goal of the experiment was to investigate the effect of the audiovisual timing on female response, and not the effect of the presence or absence of a given modality. To formally test this hypothesis, we would need to compare female response to unimodal acoustic and visual stimuli versus multimodal stimuli in a future playback experiment. In túngara frogs, females preferred asynchronous multimodal signals over unimodal signals (Taylor *et al*. 2017) and this could be true for doves as well.

The moderate effects on plasma E2 levels, where we found an effect of stimulus type but did not obtain significant post hoc differences between stimuli, suggest that our stimulus manipulations were possibly not sufficiently large to yield strong differences in the hormonal response between stimulus types. Indeed, E2 concentrations in this experiment were lower than previous studies where average peak values of 85 pg·mL^{-1} were documented for females paired with a male for several days (Korenbrot *et al*. 1974). It is possible that 8 min of active courtship a day for 7 days were not enough to fully stimulate females (Cheng *et al*. 1981), or that the post-exposure blood sampling missed the typically short E2 peak. Moreover, it is possible that the noninteractive nature of playback negatively impacted female sexual stimulation, and that physical interaction with a live male for the same amount of time would have triggered higher hormonal changes and stronger differences between displays differing in their audiovisual configuration.

Remarkably, tail quivering was the only behavior that occurred consistently over experimental days, suggesting that it is an immediate response to courtship in ring doves, showing momentary interest of a female. In contrast, chasing behavior seems to reflect increased sexual stimulation over a longer time scale, meaning courtship likely triggers both immediate and delayed behavioral responses. Our results also showed a peculiar increase in chasing on the 7th day in the jittered condition. Apart from issues related to sample size or individual differences, it could also be that jittered stimuli need a longer period of exposure to trigger sexual response. At this time, we do not have a convincing interpretation for this unexpected aspect of the study.

This study is to our knowledge the first to show that female doves respond to audiovisual playback of courtship displays. The higher E2 concentration at the end of the experiment, and increased behavioral response across testing days and during courtship playback compared to empty setup, are strong evidence that females responded to conspecific courtship in the presented videos. In birds, acoustic-only playback has been used extensively to successfully trigger responses. However, the higher flicker fusion rate of avian visual systems (Boström *et al*. 2016), as well as differences between avian and human color and polarized light perception (Muheim 2011) makes it challenging to create natural-looking visual stimuli for birds. As visual and auditory stimuli were systematically displayed together, the increased behavioral and hormonal response at the end of the experimental period could be due solely to the acoustic part of the playback. However, visual stimuli have been used successfully in pigeons (Partan *et al*. 2005), a closely related species, to trigger natural responses to courtship, and the temporal quality of the video is known to directly influence the duration of sexual behavior displayed by test subjects (Ware *et al*. 2015). We are therefore confident that our videos, with even higher temporal fidelity (120 frames per second), triggered natural responses in doves. Additionally, when males were presented with silent videos of a female dove, they regularly displayed courtship (personal observation). Finally, and most importantly, females from different experimental groups reacted differently (stronger chasing behavior), confirming that both visual and acoustic stimuli were assessed and integrated by females.

This study provides important advances in our understanding of the importance of the configuration of multimodal signals. Audiovisual playback was used for the first time to study how targeted variation in male behavior can influence both female sexual behavior and physiological state. The demonstrated success of the playback technique opens up many possibilities for future studies. Here, we targeted the temporal association between sensory channels, but the spatial configuration of the signal is likely to be important and could also be investigated using multimodal playback. More importantly, our results support the hypothesis that multicomponent signals contain information that is related not only to the presence of multiple components ("summation") but also to the relationship between them as they elicit different responses depending on the way unimodal signals are temporally configured. The nature of this integrated information is still to be determined, and further quantitative studies comparing responses to differently manipulated stimuli need to be performed. In the long term, our aim is to understand if the integration of multiple signals in multimodal courtship is associated with an evaluation of the whole display by the receiver that relies on mechanisms that go beyond the evaluation of single components. This approach, initially used in controlled laboratory conditions and for species that show relatively simple courtship, could later be extended to the elaborate and bizarre displays of many species of arthropods and vertebrates.

ACKNOWLEDGMENTS

This work was supported by Leonida Fusani's Startup funding of the University of Vienna, Austria, by the Austrian Science Fund (FWF): W1262-B29, and by the Vienna Science and Technology Fund (WWTF) CS18-021. We thank Elisabeth Heiderich for her assistance with blood sampling. We thank the animal keepers Alexandra Bohmann and Barbara Timmler from the Department of Cognitive Biology, and Nathalie Skupa from the Konrad Lorenz Institute of Ethology for taking good care of the animals.

All relevant data are available from Phaidra, the University of Vienna data repository (https://phaidra.univie.ac.at/o:1126739)

REFERENCES

Amy M, Salvin P, Naguib M, Leboucher G (2015). Female signalling to male song in the domestic canary, *Serinus canaria*. *Royal Society Open Science* **2**, 140196.

Audacity Team (2017). Audacity® (2.2.1), computer software. [Accessed 1 Oct 2020]. Available from URL: https://audacityteam.org/.

Baker MC, Tracy TT, Miyasato LE (1996). Gargle vocalizations of black-capped chickadees: Test of repertoire and video stimuli. *Animal Behaviour* **52**, 1171–5.

Barfield RJ (1971). Gonadotrophic hormone secretion in the female ring dove in response to visual and auditory stimulation by the male. *Journal of Endocrinology* **49**, 305–10.

Barske J, Schlinger BA, Wikelski M, Fusani L (2011). Female choice for male motor skills. *Proceedings of the Royal Society B: Biological Sciences* **278**, 3523–8.

Boström JE, Dimitrova M, Canton C, Håstad O, Qvarnström A, Ödeen A (2016). Ultra-rapid vision in birds. *PLoS ONE* **11**, e0151099.

Byers BE (2007). Extrapair paternity in chestnut-sided warblers is correlated with consistent vocal performance. *Behavioral Ecology* **18**, 130–6.

Cheng MF (1973). Effect of ovariectomy on the reproductive behavior of female ring doves (*Streptopelia risoria*). *Journal of Comparative and Physiological Psychology* **83**, 221–33.

Cheng MF, Porter M, Ball G (1981). Do ring doves copulate more than necessary for fertilization? *Physiology and Behaviour* **27**, 659–62.

Cheng MF (1979). Progress and prospects in ring dove research: A personal view. In: Rosenblatt JS, Hinde RA, Beer C, Busnel MC, eds. *Advances in the Study of Behavior*, Vol. 9. Academic Press, Cambridge, MA, pp. 97–129.

Cheng MF (2008). The role of vocal self-stimulation in female responses to males: Implications for state-reading. *Hormones and Behaviour* **53**, 1–10.

Chouinard-Thuly L, Gierszewski S, Rosenthal GG *et al*. (2017). Technical and conceptual considerations for using animated stimuli in studies of animal behavior. *Current Zoology* **63**, 5–19.

Cooper BG, Goller F (2004). Multimodal signals: Enhancement and constraint of song motor patterns by visual display. *Science* **303**, 544–6.

Craig W (1909). *The Expressions of Emotion in the Pigeons: I. The Blond Ring Dove (Turtur risorius)*. University of Chicago, Chicago, IL.

Crews D (1975). Effects of different components of male courtship behavior on environmentally induced ovarian recrudescence and mating preferences in the lizard, *Anolis carolinensis*. *Animal Behaviour* **23**, 349–56.

Dalziell AH, Peters RA, Cockburn A, Dorland AD, Maisey AC, Magrath RD (2013). Dance choreography is coordinated with song repertoire in a complex avian display. *Current Biology* **23**, 1132–5.

Dobson AJ, Barnett AG (2018). *An Introduction to Generalized Linear Models*. CRC Press, Boca Raton, FL.

Fletcher NH, Riede T, Beckers GJ, Suthers RA (2004). Vocal tract filtering and the “coo” of doves. *The Journal of the Acoustical Society of America* **116**, 3750–6.

Friedman MB (1977). Interactions between visual and vocal courtship stimuli in the neuroendocrine response of female doves. *Journal of Comparative and Physiological Psychology* **91**, 1408–16.

Gaunt AS, Gaunt SLL, Casey RM (1982). Syringeal mechanisms reassessed: Evidence from *Streptopelia*. *Auk* **99**, 474–94.

Halfwerk W, Jones PL, Taylor RC, Ryan MJ, Page RA (2014). Risky ripples allow bats and frogs to eavesdrop on a multisensory sexual display. *Science* **343**, 413–6.

Halfwerk W, Varkevisser J, Simon R, Mendoza E, Scharff C, Riebel K (2019). Toward testing for multimodal perception of mating signals. *Frontiers in Ecology and Evolution* **7**, 124.

Harrell FE, Dupont C (2020). Hmisc: Harrell Miscellaneous, R package version 4.3-1. Available from URL: https://cran.r-project.org/package=Hmisc.

Hebets EA, Barron AB, Balakrishnan CN, Hauber ME, Mason PH, Hoke KL (2016). A systems approach to animal communication. *Proceedings of the Royal Society B: Biological Sciences* **283**, 20152889.

Helsel DR (2012). *Statistics for Censored Environmental Data Using Minitab and R*. John Wiley & Sons, Hoboken, NJ.

Higham JP, Hebets EA (2013). An introduction to multimodal communication. *Behavioural Ecology and Sociobiology* **67**, 1381–8.

Hinde RA, Steel E (1978). The influence of daylength and male vocalizations on the estrogen-dependent behavior of female canaries and budgerigars, with discussion of data from other species. In: Rosenblatt JS, Hinde RA, Beer C, Busnel MC, eds. *Advances in the Study of Behaviour*, Vol. 8. Academic Press, Cambridge, MA, pp. 39–73.

Hogan BG, Stoddard MC (2018). Synchronization of speed, sound and iridescent color in a hummingbird aerial courtship dive. *Nature Communications* **9**, 5260.

Holmes NP, Spence C (2005). Multisensory integration: space, yime and superadditivity. *Current Biology* **15**, R762–4.

Horning CL, Beecher MD, Stoddard PK, Campbell SE (1993). Song perception in the song sparrow: Importance of different parts of the song in song type classification. *Ethology* **94**, 46–58.

Hutchison JB (1970). Differential effects of testosterone and oestradiol on male courtship in barbary doves (*Streptopelia risoria*). *Animal Behaviour* **18**, 41–51.

Hutchison RE, Hutchison JB, Fusani L (1997). Vocal-postural co-ordination of a sexually dimorphic display in a monomorphic species: the Barbary dove. *Behaviour* **134**, 321–35.

Johnstone RA (1995). Honest advertisement of multiple qualities using multiple signals. *Journal of Theoretical Biology* **177**, 87–94.

Kelso EC, Martins EP (2008). Effects of two courtship display components on female reproductive behavior and physiology in the sagebrush lizard. *Animal Behaviour* **75**, 639–46.

Korenbrot CC, Schomberg DW, Erickson CJ (1974). Radioimmunoassay of plasma estradiol during the breeding cycle of ring doves (*Streptopelia risoria*). *Endocrinology* **94**, 1126–32.

Kozak EC, Uetz GW (2016). Cross-modal integration of multimodal courtship signals in a wolf spider. *Animal Cognition* **19**, 1173–81.

Lehrman DS, Brody PN, Wortis RP (1961). The presence of the mate and of nesting material as stimuli for the development of incubation behavior and for gonadotropin secretion in the ring dove (*Streptopelia risoria*). *Endocrinology* **68**, 507–16.

Lehrman DS (1964). The reproductive behavior of ring doves. *Scientific American* **211**, 48–55.

Lovari S, Hutchison JB (1975). Behavioral Transitions in the reproductive cycle of barbary doves (*Streptopelia risoria L.*). *Behaviour* **53**, 126–50.

Mcgurk H, Macdonald J (1976). Hearing lips and seeing voices. *Nature* **264**, 746–8.

Meredith MA, Nemitz JW, Stein BE (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *Journal of Neuroscience* **7**, 3215–29.

Meredith MA, Stein BE (1986). Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Research* **365**, 350–4.

Miles MC, Fuxjager MJ (2018). Animal choreography of song and dance: A case study in the Montezuma oropendola, *Psarocolius montezuma*. *Animal Behaviour* **140**, 99–107.

Miller WJ, Miller LS (1958). Synopsis of behaviour traits of the ring neck dove. *Animal Behaviour* **6**, 3–8.

Milligan M (1966). Vocal responses of white-crowned sparrows to recorded songs of their own and another species. *Animal Behaviour* **14**, 356–61.

Mitoyen C, Quigley C, Boehly T, Fusani L (2021). Female behavior is differentially associated with specific components of multimodal courtship in ring doves. *Animal Behaviour* **173**, 21–39.

Mitoyen C, Quigley C, Fusani L (2019). Evolution and function of multimodal courtship displays. *Ethology* **125**, 503–15.

Moller AP, Pomiankowski A (1993). Why have birds got multiple sexual ornaments? *Behavioral Ecology and Sociobiology* **32**, 167–76.

Muheim R (2011). Behavioral and physiological mechanisms of polarized light sensitivity in birds. *Philosophical Transactions of the Royal Society B: Biological Sciences* **366**, 763–71.

Narins PM, Grabul DS, Soma KK, Gaucher P, Hödl W (2005). Cross-modal integration in a dart-poison frog. *PNAS* **102**, 2425–9.

Otter K, McGregor PK, Terry AMR, Burford FRL, Peake TM, Dabelsteen T (1999). Do female great tits (*Parus major*) assess males by eavesdropping? A field study using interactive song playback. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **266**, 1305–9.

Partan SR (2013). Ten unanswered questions in multimodal communication. *Behavioral Ecology and Sociobiology* **67**, 1523–39.

Partan SR, Marler P (2005). Issues in the classification of multimodal communication signals. *The American Naturalist* **166**, 231–45.

Partan SR, Yelda S, Price V, Shimizu T (2005). Female pigeons, Columba livia, respond to multisensory audio/video playbacks of male courtship behavior. *Animal Behaviour* **70**, 957–66.

Partan SR, Marler P (1999). Communication goes multimodal. *Science* **283**, 1272–3.

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available from URL: http://www.R-project.org/

Riede T, Beckers GJL, Blevins W, Suthers RA (2004). Inflation of the esophagus and vocal tract filtering in ring doves. *Journal of Experimental Biology* **207**, 4025–36.

Ritschard M, Riebel K, Brumm H (2010). Female zebra finches prefer high-amplitude song. *Animal Behaviour* **79**, 877–83.

Rowe C, Guilford T (1996). Hidden colour aversions in domestic chicks triggered by pyrazine odours of insect warning displays. *Nature* **383**, 520–2.

Shoari N, Dubé JS (2018). Toward improved analysis of concentration data: Embracing nondetects. *Environmental toxicology and chemistry* **37**, 643–56.

Shimizu T (1998). Conspecific recognition in pigeons (*Columba livia*) using dynamic video images. *Behaviour* **135**, 43–53.

Smith CL, Evans CS (2008). Multimodal signaling in fowl, *Gallus gallus*. *Journal of Experimental Biology* **211**, 2052–7.

Smith CL, Evans CS (2013). A new heuristic for capturing the complexity of multimodal signals. *Behavioural Ecology and Sociobiology* **67**, 1389–98.

Spence C (2013). Just how important is spatial coincidence to multisensory integration? Evaluating the spatial rule. *Annals of the New York Academy of Sciences* **1296**, 31–49.

Starnberger I, Preininger D, Hödl W (2014). The anuran vocal sac: A tool for multimodal signalling. *Animal Behaviour* **97**, 281–8.

Taylor RC, Klein BA, Stein J, Ryan MJ (2011). Multimodal signal variation in space and time: How important is matching a signal with its signaler? *Journal of Experimental Biology* **214**, 815–20.

Taylor RC, Page RA, Klein BA, Ryan MJ, Hunter KL (2017). Perceived synchrony of frog multimodal signal components is influenced by content and order. *Integrative and Comparative Biology* **57**, 902–9.

Taylor RC, Ryan MJ (2013). Interactions of multisensory components perceptually rescue túngara frog mating signals. *Science* **341**, 273–4.

Uetz GW, Roberts JA, Taylor PW (2009). Multimodal communication and mate choice in wolf spiders: Female response to multimodal versus unimodal signals. *Animal Behaviour* **78**, 299–305.

Ullrich R, Norton P, Scharff C (2016). Waltzing *Taeniopygia*: Integration of courtship song and dance in the domesticated Australian zebra finch. *Animal Behaviour* **112**, 285–300.

Uy JAC, Safran RJ (2013). Variation in the temporal and spatial use of signals and its implications for multimodal communication. *Behavioural Ecology and Sociobiology* **67**, 1499–511.

Ware E, Saunders DR, Troje NF (2015). The influence of motion quality on responses towards video playback stimuli. *Biology Open* **4**, 803–11.

Williams H (2001). Choreography of song, dance and beak movements in the zebra finch (*Taeniopygia guttata*). *Journal of Experimental Biology* **204**, 3497–506.

Witte K (2006). Time spent with a male is a good indicator of mate preference in female zebra finches. *Ethology Ecology & Evolution* **18**, 195–204.

Witte K, Gierszewski S, Chouinard-Thuly L, Editors G (2017). Virtual is the new reality. *Current Zoology* **63**, 1–4.

Zimmer UE (1982). Birds react to playback of recorded songs by heart rate alteration. *Zeitschrift Für Tierpsychologie* **58**, 25–30.

Zollinger SA, Brumm H (2015). Why birds sing loud songs and why they sometimes don't. *Animal Behaviour* **105**, 289–95.

SUPPLEMENTARY MATERIALS

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 Experimental setup used to present audio-visual playback stimuli to female ring doves.

Figure S2 Distribution of the audio-visual timing differences between the start of the coo and the start of the bow for the 2 stimulus males used in the experiment. To plot the red curve, we used the values derived directly from the data: stimulus male 1 had mean offset value of -0.2753 (SD $= 0.1393$, $n = 133$); stimulus male 2 had a mean of -0.3620 (SD $= 0.1274$, $n = 109$). The yellow curve depicts the distribution resulting from the mean shift applied to create the shifted condition, whereas the blue curve represents the random distribution that we used for generating offsets used in the jittered condition.

Figure S3 Experimental design. We used courtship from 2 stimulus males (1 and 2) to create the stimuli. For each male, we recorded and extracted four courtship bouts of 30 seconds. We then used those bouts to create 3 different stimulus types as in Figure 1 of the main text. A control, a shifted and a jittered stimulus were created for each male, resulting in 6 unique stimulus types, each containing 4 videos. We created 2 female groups depending on which week they were tested (a or b), each containing 12 females. Females from week a were tested in the 2$^{\text{nd}}$ week of August 2020 and females from week b in the 3rd week of the same month. We randomly assigned 2 females from each week to each of the 6 available stimuli. We ran all playback sessions between 0800 and 1300. Every day, we pseudo-randomized the order of testing so that every female was tested in every possible time slot during the experiments.

Figure S4 Playback timeline. Timeline of Video Stimuli used in the experiment. Courtship videos (green) were alternated with empty cage videos to limit the aggressive character of uninterrupted male courtship (long durations, light grey) and to avoid strange effects when subsequent videos had featured the male appearing in different screen locations (short durations, dark grey). This resulted in a 15 minutes playback testing session, which included 8 minutes of courtship videos.

Table S1 Results of generalized linear models (with estimates, standard error (SE), *z* and *P*-values) investigating the effect of playback content (PB). Fixed effects are as follow: PB_courtship: playback displaying courtship behavior; day2: second day of the experiment; week_b: second group tested in August; male_2: second male used as stimulus; origin_France: dove coming from France

Table S2 Total number of behavioral events and mean (SD) per female displayed by females for each of the 3 stimulus types

Table S3 Results of generalized linear models (with estimates, standard error (SE), *z* and *P*-values) investigating the effect of stimulus type (stimulus A: control, B: shifted and C: jittered), as well as analyses of deviance table (with Chi-square value, degrees of freedom (df) and *P*-values). Results are presented for all the investigated female number of behavioral events: (a) chasing, (b) preening, (c) steps, (d) tail quivers

Table S4 Tukey post-hoc test comparing difference in chasing behavior between days, for each of the stimulus type. (a) Estimated marginal means with their standard errors, lower and upper confidence intervals for days and stimulus type. Results are averaged over the levels of week, male and origin (b). Result of the pairwise comparisons with estimates, degrees of freedom (df), *t*-value and *P*-values.

Table S5 Mean (standard deviation) of E2 plasma concentration (pg·mL^{-1}) depending on stimulus type and treatment.

Table S6 (a) Results of linear mixed models (with estimates, standard error (SE), degrees of freedom (df), *t* and *P*-values) investigating the effect of stimulus type (A: control, B: shifted, C: jittered) on E2 plasma concentration. (b) Analyses of deviance table (Sum- and mean-square values, degrees of freedom, *F* values and *P*-values).

Table S7 Tukey post-hoc test test comparing difference in E2 plasma concentration between stimulus types (A: control, B: shifted, C: jittered) (a) Estimated marginal means with their standard errors, lower and upper confidence intervals for stimulus type. Results are averaged over the levels of treatment, week, male and origin. (b) Result of the pairwise comparisons with estimates, degrees of freedom (df), *t*-value and *P*-values

Table S8 Correlation tables with *r* (corr coefficient) and *P*-values with the corrplot function (corrplot package), between (a) pre-test E2 plasma concentration and behavior displayed the first day, and (b) post-test E2 plasma concentration and behavior displayed the last day. The coefficients and *P*-values from the columns "Pre [E2] (censored)" and "Post [E2] (censored)" were obtained taking into account the left-censored data using the function cenken from the NADA package

Video S1 Sample video of the control courtship stimulus.

Video S2 Sample video of the shifted courtship stimulus.

Video S3 Sample video of the jittered courtship stimulus.

Video S4 Sample video of the female chasing behavior.

Video S5 Sample video of the female normal step behavior.

Video S6 Sample video of the female preening behavior.

Video S7 Sample video of the female tail quivers behavior.

Cite this article as:

Mitoyen C, Quigley C, Canoine V, Colombo S, Wölfl S, Fusani L (2023). Alteration of the temporal association between courtship audio and visual components affects female sexual response. *Integrative Zoology* **18**, 720–35. https://doi.org/10.1111/1749-4877.12670