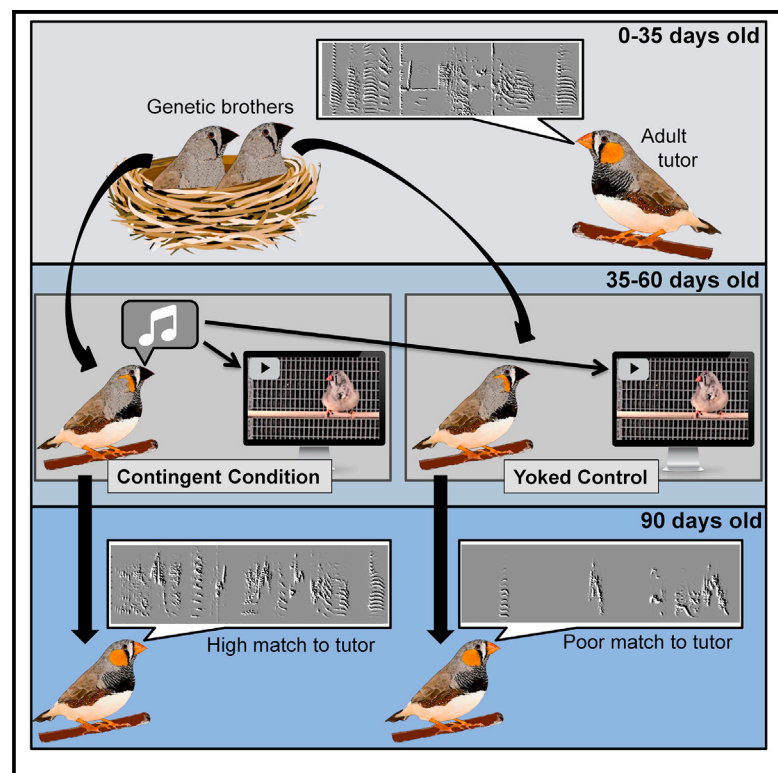


Current Biology

Female Social Feedback Reveals Non-imitative Mechanisms of Vocal Learning in Zebra Finches

Graphical Abstract



Authors

Samantha Carouso-Peck,
Michael H. Goldstein

Correspondence

michael.goldstein@cornell.edu

In Brief

The zebra finch is the most common model of human speech development but, unlike humans, is thought to learn only via imitation. Carouso-Peck and Goldstein show that song learning is affected by non-vocal, visual feedback from females. Young males given feedback contingent on their immature songs developed higher-quality songs than yoked controls.

Highlights

- Imitation, not social factors, is thought to drive zebra finch song learning
- We tested whether non-vocal visual feedback from females affects song learning
- Song learning was facilitated by contingent feedback; yoked controls did not learn
- Socially guided vocal learning is a crucial mechanism of song development



Female Social Feedback Reveals Non-imitative Mechanisms of Vocal Learning in Zebra Finches

Samantha Carouso-Peck¹ and Michael H. Goldstein^{1,2,*}¹Department of Psychology, Uris Hall, Cornell University, Ithaca, NY 14853, USA²Lead Contact*Correspondence: michael.goldstein@cornell.edu<https://doi.org/10.1016/j.cub.2018.12.026>

SUMMARY

Learning of song in birds provides a powerful model for human speech development [1–3]. However, the degree to which songbirds and humans share social mechanisms of vocal learning is unknown. Although it has been demonstrated as a vocal learning mechanism in human infants [3–6], learning via active social feedback is considered rare and atypical among non-human animals [7]. We report here the first evidence that song learning in the zebra finch (*Taeniopygia guttata*), the most common model species of vocal learning and development, utilizes socially guided vocal learning. We demonstrate experimentally that the songs of juvenile zebra finches are guided toward mature vocal forms by real-time visual feedback from adult females that is contingent on their early, immature vocalizations. Using a video playback paradigm, we found that juvenile birds that received non-vocal female feedback contingently on their immature song learned significantly better and more accurate song than did yoked controls that received identical but non-contingent feedback. Both contingent and non-contingent groups sang at similar rates. Thus, we have provided the first evidence suggesting that non-imitative social learning is a crucial, potentially widespread mechanism of vocal development and have established a foundational parallel between humans and our most ubiquitous animal model of vocal learning: the crucial role of social feedback to immature vocalizations in the development of communication.

RESULTS AND DISCUSSION

The role of social influences on vocal learning in non-human animals is poorly understood [1, 2, 8, 9], though social interactions are crucial for early speech learning in human infants [3, 4]. Contingent parental responses to the immature vocalizing of prelinguistic infants facilitate the development of speech and the learning of phonological patterns [5, 6]. Infants who receive non-contingent feedback do not show vocal learning. In contrast to these findings in humans, the immature song of songbirds is generally considered a non-communicative epiphenomenon of motor practice

[10–12]. Little attention has been paid to the social ecology of vocal learners, and few studies have addressed the potentially significant function of immature vocalizations in eliciting social feedback [7, 8, 13]. Early vocal behavior may provide learning opportunities by exploiting information available in the immediate social environment. We investigated the ability of young songbirds to use social feedback to refine their vocal repertoires. Our approach extends traditional models that emphasize imitative learning via memorization and sensorimotor integration.

Traditional models of songbird vocal learning are based on classic work with sparrows: tutor vocalizations are first memorized, and later in development, the learner attempts to match its vocalizations to the memorized template (Figure S1). Sparrows show no evidence of using social feedback to construct their songs, though conspecific behavior plays a role in selecting among previously learned songs [14]. In contrast, vocal learning in a phylogenetically distant songbird, the brown-headed cowbird (*Molothrus ater*) [7], results from social feedback to immature vocalizations. Cowbirds are raised without access to adult song models, due to their reproductive strategy as brood parasites, depositing eggs into nests of other species. Juvenile male cowbirds develop more reproductively potent songs when exposed to non-singing female cowbirds. Females respond selectively to more mature male song elements produced, using a visual cue of a rapid lateral wing movement (a “wing-stroke”) [7]. Juvenile males preferentially incorporate song elements that received wing-strokes into their final songs, resulting in songs more preferred by females. Although cowbirds are the only songbird species in which social cues have been experimentally demonstrated to guide vocal learning, they have not been used as a direct model of human speech acquisition, perhaps because as brood parasites they do not fit the traditional model of learning (Figure S1). Using social feedback to immature vocalizations to guide vocal development—known as “socially guided vocal learning”—is therefore currently considered a rare, atypical learning strategy among songbirds.

Zebra finches are the most common model species for human vocal learning but, unlike humans, are thought to acquire their learned vocalizations solely via imitation of a tutor. However, there is mounting evidence that social influences are crucial to zebra finch vocal development. Zebra finches require interaction with a live tutor to develop species-typical song [2, 15, 16]. They attend and learn better when a song is presented by an interactive tutor, even a heterospecific [17], or contingent on their own pressing of a key [18]. Furthermore, visual cues are essential for normal vocal interaction [19] and song learning, as juveniles kept in visual isolation from conspecifics do not develop normal song



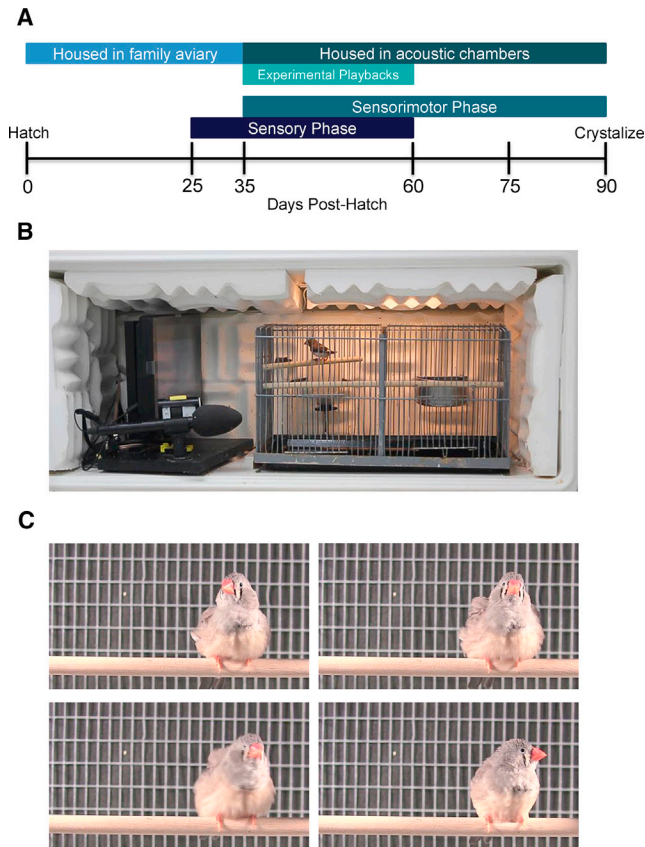


Figure 1. Experimental Apparatus and Stimulus for Video Playback Study

(A) Experimental timeline for all subjects, housed in family aviaries until relocation to acoustic chambers at 35 dph, the start of the sensorimotor period (Figure S1). Playbacks occurred daily until the end of the sensory and sensorimotor overlap phase at 60 dph. Birds were kept in chambers until song crystallization and recording at 90 dph.

(B) Sound attenuation chambers used for playbacks and recording (see STAR Methods).

(C) Stills from the fluff-up video stimulus shown to subjects, contingent on CC male song, sampled every 0.5 s. The video stimulus (3 s total duration) viewed by subjects consisted of a life-sized adult female zebra finch fading into view on a perch over 0.5 s (top left), erecting body feathers over 0.6 s (top right), performing a 0.2 s fluff-up “shake” (bottom left), then returning to a neutral position (bottom right) and fading from view. See Video S1 for the video stimulus as presented to subjects.

even with acoustic interaction [20]. The proximal mechanism driving improved song learning in the presence of a live tutor is unknown. Non-singing female listeners also affect song development, as males raised with deaf females sing more frequently and develop more atypical songs than those raised with hearing females [2], and blindfolded males develop more accurate song when raised with a female sibling [21]. Despite such evidence of social influences on vocal development, socially guided vocal learning has never been experimentally demonstrated in zebra finches or any non-parasitic songbird.

To determine whether zebra finches use socially guided vocal learning, and to discover what cues non-singing females may provide to males as they develop their song, we manipulated

the timing of female responses to juvenile songs. We used video playback of a non-vocal female arousal behavior, presented contingently on juvenile zebra finch song production across vocal development, to influence song learning outcomes.

Using video playbacks of female finches as stimuli enabled us to achieve precise experimental control over displays that could serve as social reinforcement. Zebra finches are known to learn from and sing to videos of conspecifics [22, 23]. Subjects consisted of nine pairs of juvenile zebra finch genetic brothers raised by their respective parents until 35 days post-hatch (dph), the beginning of the sensorimotor song practice phase (Figure 1A), when each brother was placed in a sound attenuated chamber equipped with a video monitor and camera (Figure 1B) and randomly assigned to an experimental contingent condition (CC) or yoked control (YC) condition. For 1 h daily for 25 days (Figure 1A), CC birds were video and audio monitored by an experimenter. Each time the bird sang, the experimenter triggered playback on the monitor of an adult female appearing and performing a “fluff-up,” consisting of erecting her feathers followed by high-frequency side-to-side movements of the upper body (Figure 1C). Like the wing-strokes of cowbirds, fluff-ups are most commonly exhibited in response to complex, attractive song [24]. The video stimulus was shown to the CC subject immediately contingent on their own song production whenever an immature song was produced during the experimental period. The video stimulus appeared simultaneously on the monitors of the CC and YC male siblings, such that playback occurred contingently on the song production of CC birds but unrelated to YC birds’ behavior. Thus, CC and YC subjects received identical and simultaneous amounts of video stimulus presentation, but videos were not contingent on YC subjects’ own song production.

Adult songs of all subjects were recorded after sexual maturity at 90 ± 1 dph (Figures 2A and S2). To assess learning accuracy, songs of each sibling pair were acoustically compared to those of their shared genetic and social father using Sound Analysis Pro 2.0 [25]. CC birds incorporated more of their father’s song material into their motifs than YC birds, as indicated by significantly higher acoustic percent similarity to tutor song ($t(8) = 4.418$, $p = 0.002$) (Figure 2B). CC bird similarity outcomes ($M = 65.266$, $SD = 5.348$) were comparable to those of zebra finches raised with optimal levels of exposure to tutor song [18]. Eight of the nine CC subjects outperformed their YC brothers in song similarity scores (Wilcoxon signed ranks $z = -2.547$, $p = 0.011$) (Figure 2C).

To analyze spectral features of songs, we performed a principal-component analysis (PCA) constructed from average whole-song pitch, frequency modulation, entropy, goodness of pitch, and amplitude modulation (see STAR Methods). The first two components of the PCA had respective eigenvalues of 1.826 and 1.586 and accounted for 36.533% and 31.72% (68.235% cumulative) of the total variance in song. We found significant differences in spectral structure of songs between CC and YC males using PC2 ($t(16) = 2.77$, $p = 0.014$) (Figures 3A and 3B). A Kolmogorov-Smirnov test revealed that YC subjects have a significantly broader distribution of entropy than CC subjects ($z = 1.414$, $p = 0.037$) (Table S1).

Several behavioral results also suggested that differences in learning outcomes may have been caused by differing individual

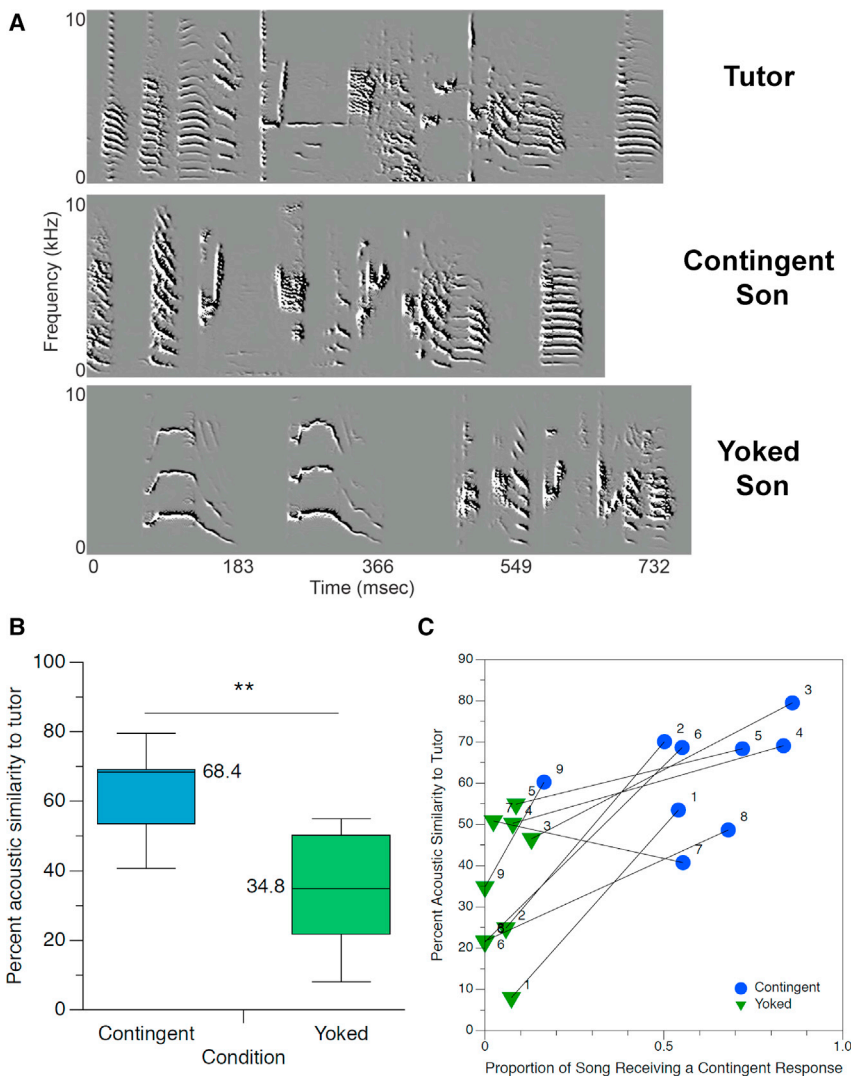


Figure 2. More Accurate Learning of Tutor's Crystallized Song in Subjects Exposed to Contingent Playback Condition

(A) Sample spectrograms of crystallized directed song of an adult tutor (top) and its two male offspring, recorded at 90 dph. Subjects were raised with the tutor until 35 dph and then exposed to the video playback procedure from 35 to 60 dph. As seen in this yoked spectrogram (bottom), YC birds produced atypical songs. Examples of spectrograms from all subjects are shown in Figure S2B. (B) Percent acoustic similarity to tutor of crystallized song by group, contingent ($n = 9$) and yoked ($n = 9$) ($t(8) = 4.418$, $p = 0.002$, two-tailed paired t test). Boxes indicate interquartile range. (C) Relation between final song percent similarity to tutor and proportion of subject-produced songs receiving a video playback within 1 s, Wilcoxon signed ranks $z = -2.547$, $p = 0.011$ ($n = 9$). Lines connect sibling pairs of CC and YC subjects. Sibling pairs are indicated by pair number.

levels of interest in the video across development (Figure S3), indicating that contingent responses increased social feedback salience. Within the CC group, we found a trending positive correlation between average number of arousal behaviors per playback session and final overall similarity to tutor song ($r(8) = 0.660$, $p = 0.053$) (Figure 4A). Overall similarity was significantly higher for CC than YC subjects ($t(8) = 3.074$, $p = 0.015$) (Figure 4B).

Our findings are the first demonstration of a proximal mechanism by which females affect male song development and facilitate socially guided vocal learning in the zebra finch. Juvenile finches that received contingent, non-vocal social feedback from a video of a female performing a fluff-up behavior learned their tutor's song significantly more accurately than a sibling that saw identical videos on a yoked schedule. Several previous studies have found that the presence of non-singing female zebra finches improves song learning in juvenile males [2, 21], the mechanisms of which have never been explained. We hypothesize that contingent social feedback serves to reinforce the memorized representation of the song, indicating to the young learner the social potency of his attempts to reproduce

the song. By receiving generalized positive feedback over the developmental period of song learning, the learner is likely motivated to continue his attempts at producing his memorized song. Such social motivation is known to facilitate song development [9, 26].

These results are also the first experimental demonstration that the song ontogeny of young male zebra finches can be significantly influenced by contingent visual displays from non-singing conspecifics. As human infants also use contingent social feedback to guide vocal development [5, 6], our results provide evidence for a previously unknown parallel between zebra finch song acquisition and human speech development. The presence of socially

guided vocal learning in zebra finches suggests that this active form of learning, which most closely resembles human learning of speech, may be more common than previously thought. This study is a first step in elucidating socially guided vocal learning in the zebra finch and reveals new avenues of research to elucidate proximal mechanisms of this learning strategy. To determine the importance of the form of the feedback, we are currently extending these findings using videos of female finches exhibiting wing-strokes, as well as non-biological stimuli.

In contrast to songbirds, the presence of socially guided vocal learning is better characterized in infant marmoset monkeys and humans, which both have the capacity for socially guided vocal learning. Young marmosets that receive more vocal feedback from parents contingent on their immature calls develop mature calls more quickly [13]. Lack of parental interaction during development results in long-term disruptions to the acoustic structure of marmoset vocalizations, suggesting that parental feedback is necessary for proper vocal learning [27]. Similarly, human infants rapidly learn to produce new phonological patterns in response to contingent reactions of caregivers, and their ability to learn

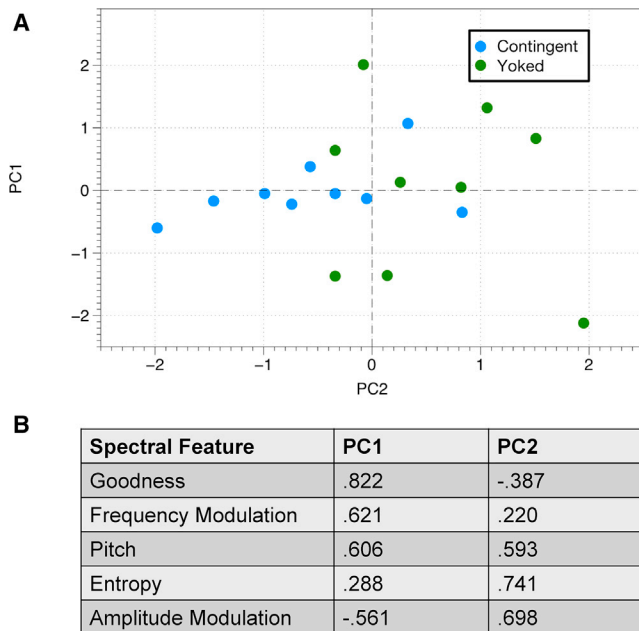


Figure 3. Differences in Spectral Acoustic Features of Crystallized Song between Contingent and Yoked Condition Subjects

(A) Principal-component analysis calculated from spectral features of crystallized songs. Contingent and yoked bird values on PC2, which accounted for 31.72% of variance, were significantly different ($t(16) = 2.77$, $p = 0.014$) (see STAR Methods).

(B) Weighting of spectral features included in each principal component (for values of each acoustic parameter, see Table S1).

phonological rules is not based on imitation [5]. Despite their phylogenetic distance, humans, marmosets, zebra finches, and cowbirds share life history traits that may have given rise to socially guided vocal learning as a solution to the problem of developing communicative competence. First, all three species are socially gregarious, ensuring developmental access to social feedback. Second, they use their learned vocalizations to facilitate and maintain social bonds. Third, they can learn new vocal forms at the same time they are producing immature vocalizations (Figure S1), which may facilitate a role for social feedback in response to immature vocalizations as a means of influencing vocal learning [28]. Furthermore, zebra finches and cowbirds are both non-territorial and use song for attracting mates, meaning they benefit from attending to the song preferences of the opposite sex while developing a vocal repertoire.

Our finding of a novel mechanism that shapes vocal learning in the zebra finch offers a new approach to studying comparative vocal development, as it demonstrates that this ubiquitous model species learns from social contingencies as humans do rather than solely from imitating previous auditory exposure. Why does social feedback have such a robust effect? The motivation and social circuits of the brain are inextricably connected, predisposing gregarious organisms to attach reward value to social partners [29, 30]. Socially guided vocal learning requires additional connections between the social-motivation system and the vocal-learning system [31]. By placing the zebra finch, a species often studied in isolation, into a larger social context, we have uncovered new processes of non-vocal feedback that

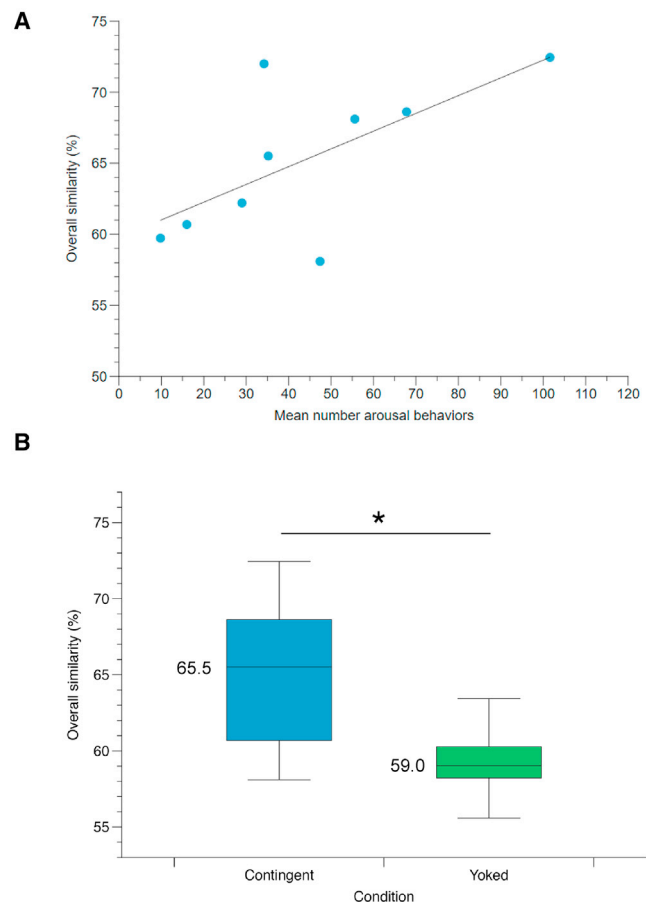


Figure 4. Correlation between Arousal Behaviors during Playback Sessions and Final Overall Song Similarity to Tutor for Contingent Condition Birds

Overall similarity is a composite measure of percent similarity, accuracy, and sequential match to tutor song.

(A) There was a positive trend in correlation ($r(8) = 0.660$, $p = 0.053$) for CC subjects between average number of arousal behaviors (fluff-ups and beak wipes) they exhibited across developmental time (35–60 dph) during experimental video playback periods and final overall similarity of adult crystallized song to tutor song at 90 dph.

(B) Significant difference in final overall acoustic song similarity to tutor between contingent and yoked subjects ($t(8) = 3.074$, $p = 0.015$). Similarity was measured at 90 dph.

require investigation at neural and neuroendocrine levels of organization. Song learning is clearly affected by social factors, but how song system and social reward circuitry are linked is largely unknown. An emerging body of evidence indicates a strong role of nonapeptides such as arginine vasotocin on social motivation and song learning in zebra finches [9, 26]. Learners perform real-time error correction in response to auditory feedback via dopaminergic connections between ventral tegmental area (VTA) and Area X of the song system [32]. Producing song in a social context appears to be highly rewarding, as it elevates dopamine in the striatum of songbirds [33]. Dopamine administration stimulates socially motivated vocalizations in songbirds [34], while dopamine antagonists inhibit vocal behavior [35]. In the zebra finch, EGR-1 expression in catecholaminergic neurons

in VTA is significantly higher in birds that have been tutored socially than in passively tutored birds, suggesting that social interaction, not merely hearing song, leads to activity in VTA [8]. These seemingly disparate neuroendocrine findings can be integrated with our behavioral findings as part of a social-feedback system that guides learning.

One reason for the lack of attention to social feedback as a driving force of song learning is that zebra finches can learn from taped song when it is triggered by their own key-pressing, though not if played to them passively [e.g., 18]. Key-pressing has long been exploited for vocal learning studies, but its efficacy has never been adequately explained. Our findings suggest that the contingency of social information (song) on the actions of the subject (key-pressing) in traditional paradigms may serve as an artificial proxy for social feedback. Studies in rodents often use key-pressing paradigms to measure social reward value or motivation by having subjects press keys for access to conspecifics [e.g., 36]. Rather than allowing mice to passively experience a social stimulus, such paradigms require effort from subjects to obtain a social reward, potentially rendering it more salient. Similar studies in hamsters have found that subjects work equally hard for social rewards as food rewards [37]. Social feedback appears to be intrinsically rewarding across social species, causing increases in dopamine that mediate appetitive learning via behavioral reinforcement [38]. Social reinforcement only leads to learning when the reward is delivered contingently on the subject's own actions, as is the case for song stimuli in avian key-pressing paradigms.

In conclusion, we found that zebra finches, previously thought to learn only via imitation of an acoustic tutor, utilize the human-like learning strategy of socially guided vocal learning. Juveniles exposed to video playbacks of females performing a fluff-up arousal behavior contingent on their own immature song production developed significantly more accurate songs than their genetic brothers exposed to identical, non-contingent feedback on a yoked schedule. Zebra finches, like cowbirds, are non-territorial and use their song solely for attracting mates, making integration of the preferences of the opposite sex into song a highly adaptive strategy for future reproductive success. Our finding of socially guided vocal learning in the zebra finch suggests the possibility that this learning strategy is more ubiquitous than previously imagined. Given the impact of social influences on zebra finch vocal development and its contrast with the classic model system of sparrows, we emphasize that there is no universal vocal learning strategy across songbird species [39], and the presence of socially guided vocal learning should be investigated across oscines. Incorporating social factors into studies of zebra finch learning will strengthen the species as a model system, as it will uncover new possibilities for drawing parallels with human speech acquisition. Because humans learn to speak via selective social feedback to more advanced vocal forms, the discovery of a similar mechanism in zebra finches presents a new avenue of investigation for more accurately characterizing social mechanisms of human vocal development.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- **KEY RESOURCES TABLE**
- **CONTACT FOR REAGENT AND RESOURCE SHARING**
- **EXPERIMENTAL MODEL AND SUBJECT DETAILS**
 - Subject Rearing and Housing
 - Ethical Note
- **METHOD DETAILS**
 - Video Playback Apparatus
 - Experimental Schedule and Recording
 - Video Stimulus
 - Behavior Recording and Analysis
 - Acoustic Recording and Analysis
- **QUANTIFICATION AND STATISTICAL ANALYSIS**
- **DATA AND SOFTWARE AVAILABILITY**

SUPPLEMENTAL INFORMATION

Supplemental Information includes four figures, two tables, and one video and can be found with this article online at <https://doi.org/10.1016/j.cub.2018.12.026>.

ACKNOWLEDGMENTS

Research was supported by an NSF Graduate Fellowship to S.C.-P. under NSF grant DGE-1650441 and Cornell University Institute for the Social Sciences grant 535-110-233. Animal subject treatment in this study complied with the standards defined by the Cornell Institutional Animal Care and Use Committee. We thank Sarah Bottjer, Barbara Finlay, Asif Ghazanfar, Steve Robertson, and Jennifer Schwade for insightful feedback on the manuscript; Felix Thommes for assistance with statistical analysis; undergraduate assistants Severine Hex, Arielle McDermott-Amos, Annelise Raymond, and Megan Waller for behavioral coding of videos; and Sofia Carrillo for writing our data extraction program.

AUTHOR CONTRIBUTIONS

Both authors designed the study, performed statistical analyses of results, and wrote the manuscript. S.C.-P. performed the experiments and behavioral and acoustic analysis.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: October 11, 2018

Revised: November 20, 2018

Accepted: December 14, 2018

Published: January 31, 2019

REFERENCES

1. Kuhl, P.K. (2003). Human speech and birdsong: communication and the social brain. *Proc. Natl. Acad. Sci. USA* 100, 9645–9646.
2. Williams, H. (2004). Birdsong and singing behavior. *Ann. N Y Acad. Sci.* 1016, 1–30.
3. Goldstein, M.H., King, A.P., and West, M.J. (2003). Social interaction shapes babbling: testing parallels between birdsong and speech. *Proc. Natl. Acad. Sci. USA* 100, 8030–8035.
4. Kuhl, P.K. (2007). Is speech learning 'gated' by the social brain? *Dev. Sci.* 10, 110–120.
5. Goldstein, M.H., and Schwade, J.A. (2008). Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol. Sci.* 19, 515–523.
6. Albert, R., Schwade, J.A., and Goldstein, M.H. (2018). The social functions of babbling: Acoustic and contextual characteristics that facilitate maternal responsiveness. *Dev. Sci.* 21, e12641.

7. West, M.J., and King, A.P. (1988). Female visual displays affect the development of male song in the cowbird. *Nature* 334, 244–246.
8. Chen, Y., Matheson, L.E., and Sakata, J.T. (2016). Mechanisms underlying the social enhancement of vocal learning in songbirds. *Proc. Natl. Acad. Sci. USA* 113, 6641–6646.
9. Theofanopoulou, C., Boeckx, C., and Jarvis, E.D. (2017). A hypothesis on a role of oxytocin in the social mechanisms of speech and vocal learning. *Proc. Biol. Sci.* 284, 20170988.
10. Marler, P. (1970). A comparative approach to vocal learning: Song development in white-crowned sparrows. *J. Comp. Physiol. Psychol.* 71, 1–25.
11. Petrinovich, L. (1972). Psychobiological mechanisms in language development. In *Advances in Psychobiology, Volume 1*, G. Newton, and A.H. Riesen, eds. (New York: Wiley).
12. Marler, P. (1997). Three models of song learning: evidence from behavior. *J. Neurobiol.* 33, 501–516.
13. Takahashi, D.Y., Liao, D.A., and Ghazanfar, A.A. (2017). Vocal learning via social reinforcement by infant marmoset monkeys. *Curr. Biol.* 27, 1844–1852.e6.
14. Marler, P., and Nelson, D.A. (1992). Action-based learning: a new form of developmental plasticity in birdsong. *Neth. J. Zool.* 43, 93–103.
15. Eales, L.A. (1989). The influences of visual and vocal interaction on song learning in zebra finches. *Anim. Behav.* 37, 507–508.
16. Derégnaucourt, S., Poirier, C., Kant, A.V., Linden, A.V., and Gahr, M. (2013). Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. *J. Physiol. Paris* 107, 210–218.
17. Immelmann, K. (1969). Song development in the zebra finch and other estrilid finches. In *Bird Vocalizations*, R.A. Hinde, ed. (Cambridge: Cambridge University Press), pp. 61–81.
18. Tchernichovski, O., Lints, T., Mitra, P.P., and Nottebohm, F. (1999). Vocal imitation in zebra finches is inversely related to model abundance. *Proc. Natl. Acad. Sci. USA* 96, 12901–12904.
19. Perez, E.C., Fernandez, M.S.A., Griffith, S.C., Vignal, C., and Soula, H.A. (2005). Impact of visual contact on vocal interaction dynamics of pair-bonded birds. *Anim. Behav.* 107, 125–137.
20. Morrison, R.G., and Nottebohm, F. (1993). Role of a telencephalic nucleus in the delayed song learning of socially isolated zebra finches. *J. Neurobiol.* 24, 1045–1064.
21. Adret, P. (2003). Vocal imitation in blindfolded zebra finches (*Taeniopygia guttata*) is facilitated in the presence of non-singing females. *J. Ethol.* 22, 29–35.
22. Ikebuchi, M., and Okanoya, K. (1999). Male zebra finches and Bengalese finches emit directed song to the video images of conspecific females projected onto a TFT display. *Zool. Sci.* 16, 63–71.
23. Guillette, L.M., and Healy, S.D. (2017). The roles of vocal and visual interactions in social learning zebra finches: A video playback experiment. *Behav. Processes* 139, 43–49.
24. Vyas, A., Harding, C., Borg, L., and Bogdan, D. (2009). Acoustic characteristics, early experience, and endocrine status interact to modulate female zebra finches' behavioral responses to songs. *Horm. Behav.* 55, 50–59.
25. Tchernichovski, O., Nottebohm, F., Ho, C.E., Pesaran, B., and Mitra, P.P. (2000). A procedure for an automated measurement of song similarity. *Anim. Behav.* 59, 1167–1176.
26. Baran, N.M., Peck, S.C., Kim, T.H., Goldstein, M.H., and Adkins-Regan, E. (2017). Early life manipulations of vasopressin-family peptides alter vocal learning. *Proc. Biol. Sci.* 284, 20171114.
27. Gultekin, Y.B., and Hage, S.R. (2018). Limiting parental interaction during vocal development affects acoustic call structure in marmoset monkeys. *Sci. Adv.* 4, eaar2012.
28. Zann, R.A. (1996). *The Zebra Finch: A Synthesis of Field and Laboratory Studies* (Oxford: Oxford University Press).
29. O'Connell, L.A., and Hofmann, H.A. (2011). The vertebrate mesolimbic reward system and social behavior network: a comparative synthesis. *J. Comp. Neurol.* 519, 3599–3639.
30. Newman, S.W. (1999). The medial extended amygdala in male reproductive behavior. A node in the mammalian social behavior network. *Ann. N Y Acad. Sci.* 877, 242–257.
31. Syal, S., and Finlay, B.L. (2011). Thinking outside the cortex: social motivation in the evolution and development of language. *Dev. Sci.* 14, 417–430.
32. Gadagkar, V., Puzerey, P.A., Chen, R., Baird-Daniel, E., Farhang, A.R., and Goldberg, J.H. (2016). Dopamine neurons encode performance error in singing birds. *Science* 354, 1278–1282.
33. Sasaki, A., Sotnikova, T.D., Gainetdinov, R.R., and Jarvis, E.D. (2006). Social context-dependent singing-regulated dopamine. *J. Neurosci.* 26, 9010–9014.
34. Heimovics, S.A., Cornil, C.A., Ball, G.F., and Ritters, L.V. (2009). D1-like dopamine receptor density in nuclei involved in social behavior correlates with song in a context-dependent fashion in male European starlings. *Neuroscience* 159, 962–973.
35. Rauceo, S., Harding, C.F., Maldonado, A., Gaysinkaya, L., Tulloch, I., and Rodriguez, E. (2008). Dopaminergic modulation of reproductive behavior and activity in male zebra finches. *Behav. Brain Res.* 187, 133–139.
36. Martin, L., and Iceberg, E. (2015). Quantifying social motivation in mice using operant conditioning. *J. Vis. Exp.* 102, e53009.
37. Borland, J.M., Frantz, K.J., Aiani, L.M., Grantham, K.N., Song, Z., and Albers, H.E. (2017). A novel operant task to assess social reward and motivation in rodents. *J. Neurosci. Methods* 287, 80–88.
38. Panksepp, J., and Moskal, J. (2008). Dopamine and seeking: Subcortical reward systems and appetitive urges. In *Handbook of Approach and Avoidance Motivation*, A.J. Elliot, ed. (New York: Taylor and Francis), pp. 67–88.
39. Brenowitz, E.A., and Beecher, M.D. (2005). Song learning in birds: diversity and plasticity, opportunities and challenges. *Trends Neurosci.* 28, 127–132.
40. Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., and Sloetjes, H. (2006). ELAN: a professional framework for multimodality research. In *Proceedings of the 5th International Conference on Language Resources and Evaluation*, pp. 1556–1559.

STAR★METHODS

KEY RESOURCES TABLE

| REAGENT or RESOURCE | SOURCE | IDENTIFIER |
|------------------------------|------------|---|
| Deposited Data | | |
| Behavioral and learning data | This paper | https://data.mendeley.com/datasets/63db6zc5fm/2 |
| Other | | |
| Video stimulus | This paper | Video S1 |

CONTACT FOR REAGENT AND RESOURCE SHARING

Further requests for reagents and resources should be directed and will be fulfilled by the Lead Contact, Michael H. Goldstein (michael.goldstein@cornell.edu).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Subject Rearing and Housing

Subjects were eighteen male zebra finches hatched and raised in our laboratory's colony ($n = 9$ pairs). Parents were eight male and eight female unpaired adult zebra finches purchased from Magnolia Bird Farm in Riverside, California, placed in a large flight aviary ($1.2 \times 0.9 \times 0.6$ m) and allowed to pair and breed. All birds in the colony are marked with individually identifying colored leg bands. Birds were provided water, cuttle bone, and Kaytee Forti-Diet Finch Food *ad libitum*, with supplemental fresh spinach, carrots, peas, and hard-boiled egg every other day. Aviaries were equipped with plastic nest boxes lined with coconut fiber and additional loose fiber for nest-building material. Zebra finches are a monogamous, non-territorial, and highly gregarious estrildid finch species native to arid central Australia [28]. They are commonly used in laboratory studies of vocal learning due to their simple song structures, rapid development, and ease of breeding in captivity.

Offspring were sexed visually at 30 days post-hatch (dph), when males first begin to develop orange cheeks and black-striped chests, which are absent in females. Provided a clutch contained at least two males, the eldest two males became experimental subjects. Offspring were raised in the aviary until 35 ± 1 dph, near the beginning of the sensorimotor learning period when juvenile male zebra finches begin to produce practice song (Figure 1A, Figure S1), and coinciding with independence from parental feeding. At this time, brother pairs were removed from the communal aviary and placed in individual sound attenuated chambers. Because zebra finch chicks usually hatch on sequential days, brothers typically differed in age by one day. We randomized whether the older chick was placed in the Contingent (CC) or Yoked control (YC) condition. Sibling pairs raised simultaneously by the same parents were used to control for potential differences in parental behaviors, complexity of and exposure to tutor song, and genetic effects. Subjects were housed in acoustic chambers until 90 dph (Figure 1A).

Ethical Note

All protocols were approved and overseen by, and animal subject treatment in this study complied with the standards defined by the Cornell Institutional Animal Care and Use Committee. The study was designed to minimize number of birds used, disturbance, stress, and social isolation. No birds were sacrificed for this study, and following final song recordings all subjects were re-homed in other zebra finch labs.

METHOD DETAILS

Video Playback Apparatus

Sound attenuated chambers ($104 \times 48 \times 43$ cm) were internally lined with 'Soundfoam M' attenuating foam from Soundcoat Co Inc. Chambers. Each was equipped with two IKEA Dioder LED lights, on an automatic 12/12 light/dark schedule. A LPH120 linear piston central air pump provided low airflow to all chambers. Cages within the chambers ($46 \times 44 \times 36$ cm) had cuttle bone, water, and seed provided *ad libitum*. At one end of each chamber was a wooden stand mounted with a Sony Actioncam HDR-AS15 digital HD video camera recording audio via a Sennheiser ME62/K6 omni-directional condenser microphone, a Morel MDT 29 speaker, and a 10.4" HP L6010 LED monitor (60 Hz refresh rate, 57% color gamut, 4:3 aspect ratio, 1024×768 resolution) overlaid with a 0.04 mm thick UV filter sheet allowing less than 10% transmission below 390 nanometers. This filter served to prevent excessive UV light produced by the monitors, detectable by the bird visual system but not by humans, from washing out the images on the monitor. The monitor was controlled by an iMac 21.5" 2.7GHz Quad-core Intel Core i5 desktop computer running Microsoft Powerpoint 2010. All audio recordings were run through a high-pass filter at 1100 Hz to exclude the low-frequency noise of the animal facility's air circulation system.

Experimental Schedule and Recording

Experimentation and recording occurred for one hour each day, beginning 30 min after artificial sunrise, from 36 ± 1 dph to 61 ± 1 dph, for 25 total days of recording per subject. During this time, an experimenter monitored the live video and audio from the chamber of the CC bird using an ActionCam wrist monitor and headphones connected to the microphone via an Alesis RA150 Stereo Power Amplifier. The experimenter triggered the video stimulus as soon as a CC bird transitioned from introductory notes into a core motif, which ensured playbacks only occurred when song was produced. The video stimulus was 3 s long, showing a female zebra finch performing a fluff-up (Video S1). New video playbacks were not triggered if the CC subject began singing a new song bout while the video was already playing. This resulted in an average of 60.1% of songs produced by CC birds receiving contingent feedback, while an average of 5% of songs produced by YC subjects received a chance contingent playback. For CC subjects, 81.3% of all contingent feedback videos overlapped with song production. In comparison, 36.5% of the videos that played contingently by chance for YC subjects overlapped with song.

Between sessions, when recordings and playbacks were not occurring, each chamber also housed a female canary as a social partner. Canaries share similar behaviors, diets, and size with zebra finches, but female canaries do not sing and are not attracted to zebra finch song, making them ideal as social partners which would not provide subjects with auditory or visual feedback on their song. To investigate whether final song outcomes were differentially influenced across groups by canary calls or social feedback, we performed an acoustic similarity analysis between the calls of our female canaries and the syllables of the crystallized songs of our subjects. We first recorded twenty canary calls from birds used in the study, and then used Sound Analysis Pro 2011 to perform similarity measures between the calls. We found that all canary calls were highly similar (> 80%) to each other. We chose three calls with the greatest acoustic difference from each other to act as exemplars in further analyses (Figure S2A).

We used Sound Analysis Pro [25] to assess acoustic similarity between the canary calls and zebra finch song syllables. Sound Analysis Pro is commonly used to analyze vocalizations of various finch species, including zebra finches and canaries, and well-suited for comparing the harmonic stack acoustic structures typical of both species. We compared each of the three canary calls with five examples each of every syllable in the songs of each zebra finch subject in both contingent and yoked groups. We then performed paired-sample *t* tests between the contingent and yoked groups on each of sixteen acoustic similarity analyses (Table S2). These included, for each individual canary call and all three averaged, the similarity score between all zebra finch syllables, the similarity score of the single most similar syllable, the average similarity of syllables in each song, and the average similarity of all syllables with non-zero similarity scores (to eliminate the possibility that differences might be driven or washed out by the high number of syllables with similarities of zero). We found no significant difference between contingent and yoked subjects on any of these measures. Thus, yoked birds were no more likely to copy canary calls than were their contingent brothers, and any copying that did occur happened at a very low fidelity.

Video Stimulus

The female fluff-up stimulus shown to subjects lasted 0.8 s, consisting of the female erecting her feathers over 0.6 s and performing a 0.2 s shake of her body. The stimulus fluff-up is highly typical of female fluff-up behaviors in response to male song, which range from 0.37 – 1.12 s in duration, but are most commonly 0.71 – 0.94 s including a 0.15 – 0.33 s shake, based on a sample of 500 fluff-ups from 8 adult females in our colony. Recording was conducted in a cage with a Lexan front and black felt backdrop to reduce glare, and lit with two daylight-balanced lights with aluminum reflectors. Two females were present in the cage simultaneously, as female finches often freeze in place when isolated from conspecifics. The second female was cropped from the video to create the final stimulus. A flight cage with five unpaired adult male zebra finches was placed behind the camera to draw female attention forward and elicit arousal behaviors. The camera was a Canon HD VIXIA HFM31, 3.3 megapixels resolution, filming in 1920 × 1080 resolution. Video playback to subjects was accompanied by background audio of female contact calling to provide a cue to birds facing away from the monitor that playback was occurring. If CC birds did not sing for the first ten minutes of a session, a ‘reminder’ playback was triggered every five minutes until the first instance of song, at which point no more reminder videos would be triggered.

Behavior Recording and Analysis

All subjects were recorded for the duration of every recording session, for one hour a day for 25 days. To obtain a detailed longitudinal cross-section of behavior over time, videos from every 5th trial (trials 5, 10, 15, 20, and 25) were coded for each bird. Videos were coded for onset and offset of singing and video playback, and the arousal and motivation behaviors of fluff-ups, beak wipes, and landings on and departures from the perch nearest to the video monitor [24]. Behavioral coding was conducted using ELAN Linguistic Annotation software [40]. All videos were independently coded by two research assistants, and any disparities in coding resolved by a third coder. Videos were first visually coded for arousal behaviors with the sound turned off to ensure coders were blind to experimental condition, before being coded for song and video playbacks.

Acoustic Recording and Analysis

At 90 dph, all subjects were video- and audio-recorded in the acoustic attenuation chambers in the presence of an unfamiliar female zebra finch to obtain directed song recordings. Subjects were taped for as long as required to obtain 20 high-quality recordings of motifs not occluded by background noise or female calls. The motifs used for analysis were therefore the first 20 unoccluded songs produced by each subject following song crystallization, to avoid any differential selection of motif across conditions. Recordings of genetic fathers of the subject were obtained using the same method, though these tutors were > 150 dph at the time of recording.

Each motif from each subject was compared to that of their biological and social father using the Similarity module of Sound Analysis Pro 2011 [25]. The experimenter performing the acoustic analysis was blind to the condition of the bird from which each motif was obtained. Similarity scores generated by SAP2011 have three major components: the percent similarity score is computed over longer intervals (typically 50 - 70 msec) and reflects the amount of song material included from the tutor's song in the juvenile's motif. Accuracy is computed across shorter time windows (5 - 10 msec) and indicates how well the juvenile's song matches the tutor's song across these shorter segments. Sequential match incorporates the temporal order (syntax) of the component syllables, and refers to the similarity of temporal order of final sections (as defined in the SAP 2011 manual) between the reference (tutor) song and the second (juvenile) sound. Overall similarity score is a calculated average of these three similarity components. For analysis, similarity results for all motifs were averaged for each subject into a single score for each similarity component score.

QUANTIFICATION AND STATISTICAL ANALYSIS

All statistical analyses were performed in IBM SPSS Statistics Software, version 22. Unless otherwise stated, all analyses were two-tailed, significance level was $p < 0.05$, and $n = 9$ subjects per group (18 birds total). In total, 20 males completed the experiment, but one sibling pair had to be excluded from analysis due to an equipment failure. A formal sample size calculation could not be conducted prior to experimentation, as the preliminary data necessary to perform such a calculation did not exist prior to this study. We chose to analyze 9 Contingent and 9 Yoked Control subjects, as 4-12 subjects per group is typical in animal studies of learning and performance. For each sibling pair, which brother was placed in the CC condition and which in the YC condition was determined at random by flipping a coin.

Our primary test for learning was based on the acoustic similarity between the subjects and their song tutor, which was also their social and genetic father. Because we used a paired-subjects design, with each Contingent subject paired with an age-matched genetic Yoked brother, we first performed paired subjects *t* tests between the Contingent and Yoked groups on overall similarity, percent similarity, accuracy, and sequential match data obtained from Sound Analysis Pro. Significant results from the *t* test on percent similarity can be seen in Figure 2B and overall similarity in Figure 4B. We then performed a Wilcoxon Signed Ranks test to determine if, within these brother pairs, the Contingent subject was significantly more likely to develop superior song than its paired Yoked subject (see Figure 2C).

To determine which spectral features of the song contributed to differences in similarity, and whether Contingent and Yoked birds vary according to particular aspects of song structure, we first used paired-samples *t* tests to compare mean values of average whole-song pitch, frequency modulation, entropy, goodness of pitch, and amplitude modulation. The means did not vary between groups (all p s > 0.05) (Table S1). We compared the distributions of each spectral component across groups using a Kolmogorov-Smirnov test. Only entropy differed between groups. YC subjects had a significantly broader distribution of entropy than CC subjects ($z = 1.414$, $p = 0.037$).

We then performed a principal components analysis constructed from the mean values of the same five primary spectral components (Figure 3A). All factors were normalized using Kaiser normalization and orthogonally rotated with Varimax. The groups differed significantly on PC2, which was weighted primarily by entropy, amplitude modulation and pitch (Figure 3B).

Using behavioral data acquired from video coding in ELAN, we performed linear regression analyses to determine the correlation between number of arousal behaviors (fluff-ups and beak wipes) performed by subjects across development and song learning outcomes, both within group and pooled across groups. We found a positive trend for CC birds between arousal behaviors and crystallized song similarity (Figure 4A). To determine if the groups differed in arousal behaviors at different time-points in development, we then conducted a 2 (Condition: CC, YC) \times 5 (Age: 40, 45, 50, 55, 60dph) repeated-measures ANOVA on mean number of arousal behaviors (Figure S3). To determine whether our results were driven by differences in motivation to sing, leading to CC birds practicing song more than YC birds, we used video and audio recordings across development, sampled every five days from 35-60 dph, to count total number of song bouts per hour and total time spent singing for each subject. Using paired *t* tests, we found no significant differences between CC and YC groups (Figure S4A). We then performed 2 (Condition: CC, YC) \times 5 (Age: 40, 45, 50, 55, 60dph) repeated-measures ANOVAs on the total duration of time spent singing each session (in seconds) and the number of song bouts per session. We found no significant main effect of Condition and no significant interaction effect. We found a significant effect of Age on song duration, $F(4,32) = 3.32$, $p = 0.022$. A Tukey HSD post hoc test revealed a significant difference in song duration only between 40-50dph ($p < 0.05$). We also found a significant effect of Age on number of songs $F(2,32) = 3.23$, $p = 0.025$. A Tukey HSD post hoc test revealed a significant difference in number of songs only between 40-50 dph ($p < 0.05$) (Figures S4B and S4C).

DATA AND SOFTWARE AVAILABILITY

All behavioral and learning data generated and analyzed during this study are publicly available on Mendeley Data at <https://data.mendeley.com/datasets/63db6z5fm/2>. The custom-written Python code used in this study to extract behavioral measures from ELAN video coding software into a spreadsheet format and calculate contingencies between behaviors is available from the corresponding author upon request.