Mechanisms underlying the social enhancement of vocal learning in songbirds

Yining Chen[a], Laura E. Matheson[b], and Jon T. Sakata[a,b,c,d,1]

[a]Integrated Program in Neuroscience, McGill University, Quebec, QC, Canada H1A 2B4; [b]Department of Biology, McGill University, Quebec, QC, Canada H3A 1B1; [c]Centre for Research on Brain, Language, and Music, Quebec, QC, Canada H3G 2A8; and [d]Center for Studies in Behavioral Neurobiology, Quebec, QC, Canada H4B 1R6

Social processes profoundly influence speech and language acquisition. Despite the importance of social influences, little is known about how social interactions modulate vocal learning. Like humans, songbirds learn their vocalizations during development, and they provide an excellent opportunity to reveal mechanisms of social influences on vocal learning. Using yoked experimental designs, we demonstrate that social interactions with adult tutors for as little as 1 d significantly enhanced vocal learning. Social influences on attention to song seemed central to the social enhancement of learning because socially tutored birds were more attentive to the tutor's songs than passively tutored birds, and because variation in attentiveness and in the social modulation of attention significantly predicted variation in vocal learning. Attention to song was influenced by both the nature and amount of tutor song: Pupils paid more attention to songs that tutors directed at them and to tutors that produced fewer songs. Tutors altered their song structure when directing songs at pupils in a manner that resembled how humans alter their vocalizations when speaking to infants, that was distinct from how tutors changed their songs when singing to females, and that could influence attention and learning. Furthermore, social interactions that rapidly enhanced learning increased the activity of noradrenergic and dopaminergic midbrain neurons. These data highlight striking parallels between humans and songbirds in the social modulation of vocal learning and suggest that social influences on attention and midbrain circuitry could represent shared mechanisms underlying the social modulation of vocal learning.

birdsong | catecholamines | attention | social influences | speech

Social interactions are critical for the acquisition of speech and language (1–4). For example, whereas passive exposure to speech leads to minimal vocal learning in infants and children, social and vocal interactions with adults lead to robust sensory and motor learning of phonemes (5–7). Furthermore, developmental disorders that are characterized by deficits in speech and language acquisition are also associated with dysfunctions in social behavior (1, 2, 8, 9). As such, revealing how social processes influence vocal learning is fundamental to understanding general mechanisms underlying vocal learning and the etiology of communicative disorders (1–3).

It has been hypothesized that social interactions promote speech and language learning by modulating attentional, sensory, and sensorimotor mechanisms (1–7). For example, visual and acoustic information provided by adults during interactions with infants, as well as reciprocal vocal interactions between adults and infants, are hypothesized to enhance attention to and acquisition of speech sounds (1, 2, 6, 7, 10). Furthermore, changes to the acoustic and prosodic structure of speech when adults direct speech to infants (e.g., slower speed, greater repetition, higher and more variable pitch) could similarly serve to enhance attention and promote vocal learning (11, 12).

Understanding how social interactions influence vocal learning in nonhuman animals can provide important support for models of speech learning and deeper insight into neurobiological mechanisms underlying vocal learning. Songbirds are ideal for this endeavor

because they, like humans, learn their vocalizations during development (13–17). Furthermore, the strength and trajectory of vocal learning are influenced by social interactions between adult and juvenile songbirds (14, 18, 19). For example, juvenile songbirds that are allowed to interact with an adult throughout development have been found to produce more accurate imitations of the adult's song than socially isolated birds that are only passively exposed to adult songs (14). Despite the importance of social influences on song learning, little is known about the behavioral and neural processes underlying the social enhancement of vocal learning.

Results

Social Interactions for as Little as One Day Enhance Vocal Learning. Songbirds like the zebra finch acquire song during development by learning the acoustic structure of song (sensory learning) and the motor commands to produce the learned sounds (sensorimotor learning). Juveniles learn their song primarily by interacting with and listening to adults, in particular their fathers (20). Juvenile zebra finches typically remain with their parents until they are 50–60 d old and have abundant opportunities to interact with an adult tutor and learn their tutor's song (20).

Previous experiments have reported that social interactions with adult tutors throughout development enhance song learning (14, 21–23). For example, juvenile zebra finches that are allowed to visually and acoustically interact with tutors for 2–3 mo demonstrate more robust vocal learning than juveniles only passively exposed to song (21), and developing white-crowned sparrows learn more from exposure to social tutors than from passive exposure to song (22). However, the degree to which social interactions with tutors

---

Significance

Social interactions are fundamental to the acquisition of speech and language; therefore, it is critical to reveal mechanisms of social influences to understand general mechanisms of vocal learning. Like humans, songbirds learn their vocalizations during development. We demonstrate that social interactions rapidly enhance vocal learning, attention to song, and the activity of catecholamine-synthesizing neurons in juvenile zebra finches. We also reveal that adult finches alter the structure of their vocalizations when interacting with juveniles in ways that resemble how humans alter their speech when interacting with infants, and that could enhance learning. Our data suggest that social influences on attention, acoustic structure, and catecholaminergic activity represent evolutionarily important mechanisms mediating the social enhancement of vocal learning across humans and songbirds.

PSYCHOLOGICAL AND COGNITIVE SCIENCES

NEUROSCIENCE

contribute to the enhancement of learning is unclear from such studies because of the presence of adult females and juveniles and because of group differences in the duration of exposure to tutor song (e.g., refs. 21 and 22). Additionally, we know little about the time course of social influences on song learning, the contribution of social interactions to sensory vs. sensorimotor learning, and the neural circuits that could mediate the social enhancement of song learning.

Using a yoked experimental design, we assessed the degree to which limited social interactions with an adult zebra finch ("tutor") affected vocal learning in juvenile zebra finches ("pupils"). We compared song learning within cohorts of pupils that were either allowed to visually and acoustically interact with a tutor for only 5 d (socially tutored pupils) or allowed only to passively hear the songs of tutors for the same duration (passively tutored pupils; $n = 10$ cohorts) (SI Appendix, Fig. S1A). This yoked design normalizes the amount and type of song exposure between groups, two factors that can affect learning and were confounded in previous studies (13, 14, 24). After 5 d of tutoring, adult tutors were removed, and pupils were housed individually for 70–80 d to allow them to develop their song. To assess the strength of song learning, we computed the degree to which the spectral structure of each pupil's adult song resembled that of its tutor's song using quantitative and subjective measures; the average of these measures of similarity was used as our index of song learning ("similarity score"). We compared similarity scores not only between socially and passively tutored pupils but also between pupils and birds that were not exposed to song throughout development ("untutored birds") (SI Appendix).

Overall, the adult songs of both socially and passively tutored pupils were significantly more similar to their tutor's songs than were the adult songs of untutored birds (dashed line in Fig. 1A) ($P < 0.05$ for both). Therefore, birds that were socially or passively tutored for 5 d demonstrated significant learning. Moreover, social interactions with adult tutors significantly enhanced song learning. The songs of socially tutored pupils were more similar to their tutor's songs than the songs of passively tutored pupils (Fig. 1A) ($P = 0.0331$). These data indicate that exposure to tutor songs for only 5 d can drive significant vocal learning and that social interactions with live tutors enhance song learning even when song exposure is yoked across conditions.

To further assess the rapidity of song learning and the potency of social interactions to enhance vocal learning, we restricted exposure to tutors to <1 d. Furthermore, to assess the individual contributions of visual and acoustic interactions with tutors to the social modulation of vocal learning, we analyzed variation in vocal learning between juveniles that were allowed to visually and acoustically interact with tutors (V-A pupils), to only acoustically interact with tutors (A pupils), or to only passively hear tutors (P pupils) using a yoked design ($n = 6$ cohorts) (SI Appendix, Fig. S1B). Relative to normally reared zebra finches that hear thousands of tutor song renditions across development, these pupils heard, on average, only ~50 renditions of song. Despite the paucity of song exposure, V-A pupils demonstrated significant song learning because their songs were more similar to their tutor's songs than were the songs of untutored birds (dashed line in Fig. 1B) ($P = 0.0109$). In contrast, neither A nor P pupils demonstrated significant song learning. Additionally, song learning was comparably low in A and P pupils, suggesting that the reduction in learning in P pupils was not simply due to the potential loss of sound fidelity caused by speakers. These data highlight the rapidity with which social interactions promote vocal learning and the importance of visual interactions with tutors. Furthermore, because tutors had minimal opportunities to provide juveniles with social feedback during vocal practice (18), these data strongly suggest that social influences on the sensory acquisition of song are important for the social enhancement of song.
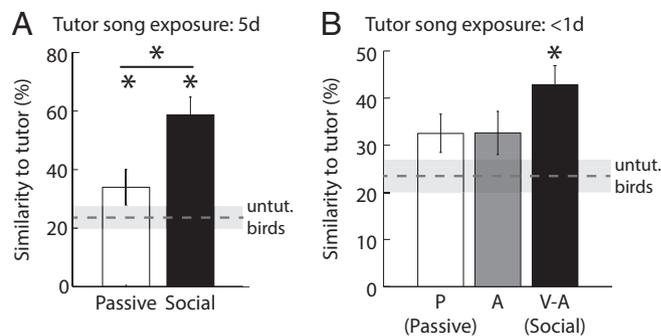

Fig. 1. Social interactions with tutors for as little as 1 d enhance vocal learning. (A) Birds that were socially ($n = 10$) or passively ($n = 9$) tutored (SI Appendix, Fig. S1A) for 5 d produced songs as adults that were significantly more similar to tutor songs than the adult songs of birds that remained untutored throughout development ("untut. birds") (SI Appendix). Additionally, socially tutored pupils produced songs as adults that were significantly more similar to their tutor's song than the adult songs of passively tutored pupils. *$P < 0.05$. (B) In another experiment, pupils were allowed to visually and acoustically interact with the tutor (V-A pupil), to only acoustically interact with the tutor (A pupil), or to passively hear the tutor's song (P pupil) for <1 d (SI Appendix, Fig. S1B). Overall, V-A pupils ($n = 6$) but not A ($n = 5$) or P ($n = 6$) pupils produced songs as adults that were more similar to their songs to tutor songs than the songs of untutored birds ("untut. birds"; *$P < 0.05$). Similarity scores were not significantly different between V-A, A, and P pupils ($P = 0.1394$). Bars ± error bars and dashed lines ± shaded areas represent the mean ± SEM.

Social Influences on Attention Play a Central Role in the Social Enhancement of Vocal Learning. Attention is generally important for learning, and it has been proposed that social interactions promote vocal learning by enhancing attention to communicative sounds (1, 3, 25). Consequently, we analyzed the relationships among social interactions, attention, and song learning in pupils tutored for 5 d (SI Appendix). We scored juveniles as attentive to tutor song if they were awake, quiet, and not engaged in other behaviors (e.g., feeding, drinking, vocalizing, flying) when tutors produced song. Attention can be very difficult to assess behaviorally, but we reasoned that the lack of engagement in other activities is a useful proxy for juvenile attention to song (SI Appendix).

Pupils were scored to be attentive to as few as 22% or as many as 94% of their tutor's songs, and this variation in attentiveness was significantly and positively correlated with the strength of song learning (i.e., similarity score) (Fig. 2A) ($P = 0.0004$). Overall, birds that paid more attention to their tutor's songs demonstrated greater learning. Interestingly, whereas this relationship was significant for both socially and passively tutored birds, increases in attention led to larger increases in learning in socially tutored pupils (slope ± SEM, 97.2 ± 18.7; $P = 0.0013$) than in passively tutored pupils (59.1 ± 14.3; $P = 0.0061$). Furthermore, variation in the degree to which social interactions increased attentiveness to tutor song within a cohort significantly and positively correlated with variation in the degree to which social interactions enhanced learning within a cohort ($P = 0.0291$) (Fig. 2B). Variation in attention or learning was not correlated with variation in hearing sensitivity (SI Appendix, Fig. S4). Taken together, these analyses underscore the importance of attention to vocal learning and to the social modulation of vocal learning and motivate investigations into the factors that influence attention to song.

Human infants differentially attend to speech depending on whether or not speech is directed at them (e.g., refs. 5, 26, and 27). Therefore, we analyzed the degree to which socially and passively tutored pupils differentially attended to songs that tutors directed toward pupils ["pupil-directed" (PD) songs] and to songs that tutors did not direct toward pupils ["undirected" (UD) songs]. Overall, pupils were more attentive to PD songs
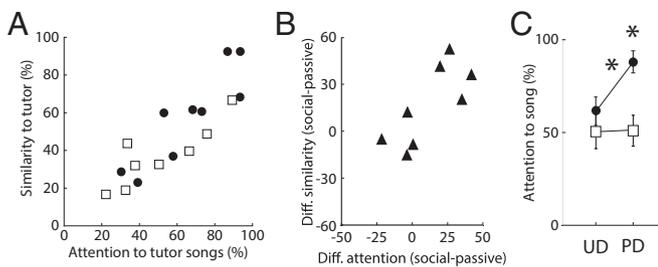
Fig. 2. Attention to song is affected by interactions with tutors and is correlated with song learning. (*A*) Across socially tutored (●; *n* = 10) and passively tutored (□; *n* = 9) birds, there was a significant and positive relationship between attention to tutor's songs and the similarity of the pupil's adult song to his tutor's song. (*B*) There was a significant and positive relationship between the magnitude of difference in attention to song and the magnitude of difference in song learning between socially and passively tutored pupils of a cohort. (*C*) Tutors produced songs that were directed at pupils [pupil-directed (PD) songs] or not directed at pupils [undirected (UD) songs], and, overall, pupils were more attentive to PD songs than UD songs. However, this difference was significant only for socially tutored birds (●; Tukey's HSD; *P < 0.05). Relatedly, socially tutored birds paid significantly more attention than passively tutored birds (□) to PD songs (*P < 0.05) but not to UD songs.

than UD songs (*P* = 0.0435), and socially tutored pupils were more attentive to song than passively tutored birds (*P* = 0.0033) (Fig. 2*C*). The difference in attentiveness to PD vs. UD song varied between socially and passively tutored birds (*P* = 0.0747). Whereas socially tutored pupils attended significantly more to the PD than the UD songs of their tutor (*P* = 0.0362), passively tutored pupils attended to PD and UD songs to the same degree as socially tutored pupils attended to UD songs. This finding suggests that visual signals from tutors could signal the production of PD song and promote attention to song in socially tutored pupils.

In addition to visual signals, there could also exist acoustic modifications to song that allowed socially tutored pupils to differentiate between PD and UD songs. In humans, speech directed toward infants is characterized by longer pauses between utterances, more repetition, and higher pitch than other forms of speech (27, 28). As such, we compared temporal and spectral features of PD and UD songs (Fig. 3 and *SI Appendix*, Fig. S5 and Tables S1 and S2). Zebra finch song bouts are preceded by repetitions of introductory notes and consist of repetitions of a single stereotyped sequence of syllables ("motif") (Fig. 3*A*). Introductory notes were repeated more often before tutors produced PD than UD songs (*P* = 0.0322) (Fig. 3*B*). Furthermore, the intervals between consecutive motifs within a bout were significantly longer when tutors produced PD than UD songs (*P* = 0.0287) (Fig. 3*B*). These data demonstrate that the temporal structure of birdsong changes when adult zebra finches sing to juveniles in a manner similar to how speech changes when adult humans speak to infants. In addition to temporal features of song, the mean frequency and spectral entropy of song syllables decreased, and the "goodness of pitch" (a measure of harmonic pitch periodicity) increased when tutors produced PD songs (*P* < 0.05 for all) (Fig. 3*B*). These data suggest that the acoustic modifications to tutor songs and the visual components of PD songs could act in concert to promote attention to song and song learning in socially tutored pupils.

To further support the notion that acoustic changes during PD song could influence song learning, we analyzed the uniqueness of these acoustic modifications by comparing how song features changed when these same tutors directed songs at adult females ["female-directed" (FD) songs] (Fig. 3*C* and *SI Appendix*, Tables S3 and S4). Similar to PD song, the number of introductory notes preceding FD song tended to be higher than that preceding UD song (*P* = 0.0695). However, in contrast to PD song, FD song was not characterized by longer intervals between motifs or by

changes to mean frequency, spectral entropy, and goodness of pitch. Conversely, the variability of the fundamental frequency of syllables with flat, harmonic structure was significantly lower when tutors produced FD song (*P* = 0.0150), but this variability was not altered when tutors directed songs at pupils. The magnitude of song changes across different audiences was significantly different for some song measures (*SI Appendix*, Fig. S6). For example, the interval between motifs increased by 26.9 ± 11.1% (mean ± SEM) when tutors produced PD songs but decreased by 7.1 ± 4.0% when tutors produced FD songs, and this difference was significant (*P* = 0.0160). These analyses indicate that male zebra finches change their songs in distinct ways when directing songs to pupils and to females and support the notion that changes to the timing of motifs and some spectral features of syllables could promote song learning in developing birds (29).

In addition to changes to acoustic features of tutor song, the abundance and timing of tutor song could affect pupil attention and learning. We found that the rate at which tutors produced song was inversely related to pupil attention (*P* = 0.0324) (Fig. 4*A*) and song learning (*P* = 0.0592) (Fig. 4*B*). The magnitudes of these relationships were not significantly different between socially and passively tutored birds (*SI Appendix*). The effect of tutor song rate
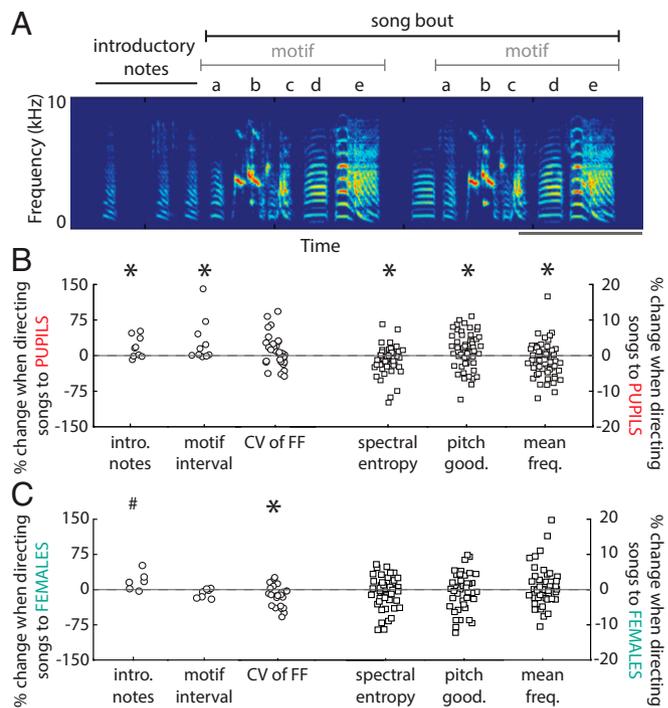
Fig. 3. Tutors changed their songs in distinct ways when directing songs to pupils than when directing songs to females. (*A*) An example of a single bout of zebra finch song. Above the spectrogram are labels that identify individual syllables as well as introductory notes that are repeated before song onset. (Scale bar: 500 ms.) (*B*) Relative to undirected songs, pupil-directed songs were preceded by more introductory notes, contained longer intervals between motifs, and consisted of syllables with lower spectral entropy and mean frequency and higher goodness of pitch (*n* = 10). Introductory notes and intermotif intervals were analyzed per bird whereas spectral features were analyzed per individual syllable using mixed effects models. Although raw data were analyzed, plotted is the percent change from undirected to pupil-directed song. *P < 0.05. (*C*) Relative to undirected songs, female-directed songs (*n* = 7) consisted of syllables that were significantly more stereotyped across renditions [i.e., lower coefficient of variation (CV) for fundamental frequency (FF)] and tended to be preceded by more introductory notes. Plotted is the percent change from undirected to female-directed song. *P < 0.05, #P < 0.10. For both *B* and *C*, the magnitude of change for the first three features (○) scales to the left *y* axis whereas the magnitude of change to the last three features (□) scales to the right *y* axis.
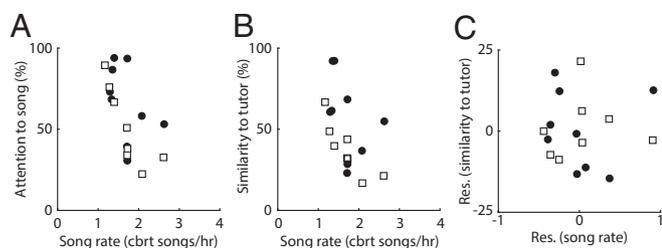
Fig. 4. Tutor song rate affects pupil attention and learning. Pupil attention (A) and song learning (B) were inversely related to the rate at which tutors produced song [songs per hour; cube-root (cbrt)-transformed]. (C) The relationship between tutor song rate and pupil song learning was not significant after controlling for pupil attention. Plotted are the residuals from the analyses of how attention relates to tutor song rate and to song learning.

on song learning seemed to be mediated by the effect of tutor song rate on attention because factoring out attention eliminated the relationship between tutor song rate and learning (Fig. 4C and *SI Appendix*). On the other hand, the relationship between attention and song learning remained significant even after controlling for variation in tutor song rate ($P < 0.005$) (*SI Appendix, Fig. S3*). Importantly, tutor song rate did not affect the degree to which social interactions enhanced attention or song learning ($P > 0.25$), suggesting that tutor song rate and social interactions independently contributed to attention and learning.

The timing of adult vocalizations relative to infant vocalizations has been shown to affect attention to speech and speech development (6, 7, 10). To examine whether contingent vocal interactions also modulate attention and subsequent learning in songbirds, we analyzed various measures of contingency and the relationships among contingent vocalizations, attention, and song learning (*SI Appendix, Figs. S7 and S8*). However, we found little evidence that contingent song production was important for the social modulation of attention or vocal learning (*SI Appendix, Fig. S8 and Table S5*).

Neurochemical Correlates of Social Influences on Vocal Learning. The preceding analyses emphasize that social influences on attention could be important for the rapid social enhancement of vocal learning. Therefore, we measured the expression of a cellular marker of activity (30, 31) in catecholamine-producing neurons implicated in attention, salience, and learning and memory (32–35) and that innervate avian forebrain circuits important for song learning and control (36, 37).

We compared EGR-1 expression between naive juvenile zebra finches that were either socially or passively tutored in a yoked manner (*SI Appendix, Fig. S1A*) (<2.25 h) or kept naive to song (untutored controls; $n = 5$ cohorts of three juveniles). During tutoring sessions, socially and passively tutored birds heard 156 ± 16 s (mean $\pm$ SEM) of song, which was comparable to the amount of song pupils heard when provided <1 d of tutoring (170 ± 13 s; $P = 0.5590$). We found that EGR-1 expression in catecholaminergic neurons in the locus coeruleus (LC) and ventral tegmental area (VTA) was significantly different between socially tutored, passively tutored, and untutored juveniles (LC, $P = 0.0084$; VTA, $P = 0.0077$) (Fig. 5A). On average, ~11% of catecholaminergic neurons in the LC expressed EGR-1 in socially tutored birds, which was significantly greater than the 3–5% of catecholaminergic neurons expressing EGR-1 in passively tutored and untutored birds ($P < 0.05$). Similarly, the percentage of catecholaminergic neurons expressing EGR-1 in the VTA was significantly higher in socially tutored juveniles than in passively tutored or untutored juveniles ($P < 0.05$ for both). The percentage of catecholaminergic neurons in the LC and VTA expressing EGR-1 was not significantly different between passively tutored juveniles and untutored controls, suggesting that hearing song in and of itself

does not lead to significant activation of catecholaminergic LC and VTA neurons.

To assess the degree to which the activation of catecholaminergic neurons in the LC and VTA required both song exposure and social exposure to tutors, we compared EGR-1 in catecholaminergic LC and VTA neurons between juveniles that were socially exposed to a tutor that produced song (social-song), juveniles that were socially exposed to a tutor that did not produce song (social-nosong), and juveniles that were not exposed to a tutor or to song ($n = 5$ cohorts of three juveniles). Similar to the previous experiment, the percentage of catecholaminergic neurons in the LC and VTA that expressed EGR-1 was significantly higher in juveniles that were exposed to a singing tutor (i.e., socially tutored) than in birds that were isolated and not exposed to song (i.e., untutored; $P < 0.05$ for both areas). Across all socially tutored juveniles, individual variation in the extent of EGR-1 activation in response to social tutoring did not significantly correlate with variation in attention to tutor song (*SI Appendix, Fig. S9*). In contrast to social exposure to a singing tutor, social exposure to a nonsinging tutor did not affect EGR-1 expression in catecholaminergic neurons in the LC and VTA ($P > 0.70$ for both). As such, the percentage of catecholaminergic neurons in the LC and VTA expressing EGR-1 was higher in juveniles that were exposed to a singing tutor than in juveniles that were exposed to a nonsinging tutor (Fig. 5B) (LC, $P = 0.0325$; VTA, $P = 0.0754$).

In contrast to the LC and VTA, EGR-1 expression in catecholaminergic neurons in the substantia nigra (SN) was not affected by social tutoring across both experiments ($P > 0.30$ for both) (Fig. 5 A and B).

Discussion

Vocal learning in humans is critically influenced by social interactions, and it is hypothesized that social interactions act on attentional and sensory processes to gate speech and language learning (1–3). Little is known about the degree to which such processes contribute to vocal learning in other species or about the neurobiological mechanisms underlying this modulation. Here, we demonstrated that social interactions with a live tutor rapidly enhanced vocal
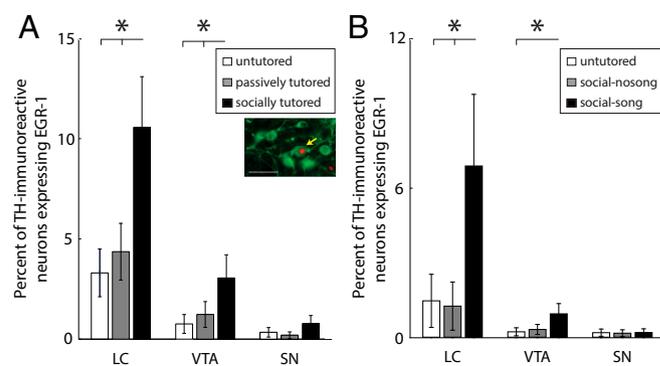
Fig. 5. Social but not passive tutoring significantly increased EGR-1 expression in noradrenergic neurons in the LC and dopaminergic neurons in the VTA. (A) We compared the percentage of tyrosine hydroxylase (TH)-immunoreactive (TH-ir) neurons that expressed EGR-1 across socially tutored pupils, passively tutored pupils, and untutored control pupils. In the LC and VTA but not the SN, a significantly higher percentage of TH-ir neurons expressed EGR-1 in socially tutored birds than in passively tutored birds or untutored birds. Passive tutoring did not increase EGR-1 expression in TH-ir neurons. (Inset) Photomicrograph of midbrain neurons that express TH (green) and EGR-1 (red). Yellow arrow highlights a TH-ir neuron expressing EGR-1. (Scale bar: 40 μm.) (B) Social exposure to a tutor that produced song ("social-song") but not social exposure to an adult tutor that did not produce song ("social-nosong") increased the percentage of TH-ir neurons expressing EGR-1 in the LC and VTA. For both A and B, *$P < 0.05$, and lines indicate significant contrasts.

learning in juvenile zebra finches and that social influences on attentional mechanisms predicted the social modulation of song learning. Juveniles were provided with only limited opportunities to receive social feedback from adults that can guide sensorimotor development (18), and, consequently, our data strongly suggest that social influences on the sensory acquisition of song are central to the social enhancement of song learning. We also documented that adult male zebra finches alter the structure of their song when directing songs to juveniles in ways that resemble how humans modify their speech when speaking to infants. These acoustic changes were distinct from how song is altered when males sing to females and could act in concert with visual cues from tutors to promote attention and learning in socially tutored birds. Finally, we revealed that social interactions that promote song learning activate catecholamine-synthesizing neurons in the midbrain that are hypothesized to influence attention and learning. Together, these experiments reveal striking parallels between humans and songbirds in the social modulation of attention and vocal learning and provide novel insight into the mechanisms underlying the social enhancement of vocal learning.

Attention is defined as a state in which the brain selectively processes a stimulus in the environment (32, 36, 38). Such a state can be difficult to infer through behavioral observation. The orientation of the head or eyes toward a stimulus is frequently used as a behavioral measure of attention (38, 39). We did not, however, use head orientation as a measure of attention because the eyes are laterally displaced in songbirds, indicating that birds can visually attend to an object even when its head is oriented away from that object. Instead, we defined juveniles as attentive to the tutor's song if they were awake, quiet, and not engaged in other behaviors (e.g., eating, drinking, or flying) when the tutor produced song. We reasoned that an individual is much less likely to be attentive to a sound when it is engaged in other behaviors. Further, even if an individual is attentive when engaged in other behaviors, attention is likely to be reduced relative to when the individual is not engaged in other behaviors. Consequently, our measure of attention can be construed as reflecting the minimum requirements for attention to the tutor's song. Despite not knowing precisely what this measure physiologically reflects, our finding that variation in this measure significantly explained variation in song learning suggests that this behavioral state could reflect a heightened state of learning.

Socially tutored pupils were more attentive to songs that tutors directed at them, and it is possible that changes to song structure when adult tutors direct songs at juveniles could promote attention and vocal learning in socially tutored birds. In humans, infant-directed speech is slower in delivery (e.g., longer pauses between word boundaries), more repetitive, and higher and more variable in pitch, and such changes to adult speech have been found to promote vocal learning (2, 27, 28). Relative to undirected songs, adult zebra finch songs directed toward juvenile pupils contained longer pauses between motifs, were preceded by more repetitions of introductory notes, and were composed of syllables that were more structured (e.g., less entropic) and lower in frequency. Changes to the interval between motifs and mean spectral features of syllables were not observed when the same males directed songs at females, indicating that zebra finches, like humans, modulate their vocalizations in distinct ways depending on the audience (29). The specificity of acoustic changes suggests the importance of these modulations to the sensory acquisition of song in juvenile songbirds. Consistent with this notion is that increasing the interval between repeated stimulus presentations enhances long-term memory for sensory stimuli (40, 41). Furthermore, the specificity of the acoustic modulations also suggests the exciting possibility that adult tutors could be actively engaged in teaching song to juvenile songbirds (42, 43).

Another aspect of tutor vocal behavior, namely the rate of tutor song production, was also related to pupil attention and song learning. Interestingly, we found that the rate at which tutors produced song was inversely related to pupil attention and song learning. An inverse relationship between tutor song abundance and pupil learning has previously been observed (44, 45), and our data extend these findings by suggesting that model abundance influences learning through its effects on attention. One interpretation of these data is that pupils become less attentive to song as tutors produce more song. However, another possibility is that pupil behavior could dynamically regulate song production by the tutor; for example, the lack of attention from pupils could promote song production by tutors. These findings suggest that a complex behavioral interplay between pupils and tutors could shape learning outcomes.

How social interactions between tutors and pupils affect the nervous system to modulate attention and learning remains largely unknown. Midbrain neurons that synthesize norepinephrine (NE) or dopamine (DA) could encode social and sensory information and modulate attention and sensory learning (32, 33, 46–49). We discovered that NE-producing neurons in the LC and DA-producing neurons in the VTA were activated by social but not passive song tutoring. Additionally, we documented that the activation of these neurons required the convergence of song and visual stimuli from the tutor. Therefore, just as visual and acoustic interactions were important for the social enhancement of song learning, both seeing and hearing the tutor produce song were necessary for the increase in EGR-1 expression in the LC and VTA. We propose that midbrain catecholaminergic neurons integrate multimodal sensory information provided by social interactions with tutors and release catecholamines into brain areas important for auditory learning to enhance the encoding of tutor songs (33–35, 49, 50). Indeed, NE has been found to enhance the processing of birdsong in auditory areas critical for song learning (48, 51). Additionally, because midbrain neurons habituate after repeated exposures to sensory stimuli (35), the inverse relationship between tutor song abundance and song learning could be mediated by changes to midbrain activity.

Taken together, our experiments in songbirds provide important support for models of speech acquisition in humans. The acquisition of speech is gated by social interactions, and it has been proposed that social interactions promote speech learning by enhancing attentional mechanisms and the sensory learning of communicative sounds (1–7). Our data suggest that social influences on sensory learning, attention, and midbrain circuitry represent evolutionarily important mechanisms underlying the social enhancement of vocal learning. Additionally, our findings that adult songbirds modulate the structure of their song when singing to juvenile pupils and that socially tutored pupils attend more to pupil-directed songs not only extend the parallels between songbirds and humans but also support the notion that changes to the temporal and spectral structure of vocalizations could contribute to the social enhancement of vocal learning. Finally, the sensitivity of catecholaminergic neurons to social information that enhances learning suggests that dysfunctions in these midbrain populations could contribute to the etiology of social and communication disorders.

Materials and Methods

Song Tutoring and Video and Acoustic Analyses. We tutored naive juvenile zebra finches (~40 d old) that were raised by their mother or a foster mother from 5 to 7 d of age (*SI Appendix*). Birds were socially or passively tutored in sound-attenuating chambers ("soundboxes") for 5 d (n = 10 cohorts of two birds) or for <1 d (n = 6 cohorts of three birds) using a yoked paradigm (*SI Appendix*, Fig. S1). A majority of cohorts consisted of siblings from the same nest, thus controlling for genetic background to a large extent (*SI Appendix*). For both experiments, passively tutored birds could also hear the developing vocalizations of the pupils housed next to the tutors (*SI Appendix*, Fig. S1). After tutoring sessions, tutors were removed, and juveniles developed their song individually in soundboxes for 70–80 d until they reached sexual maturity (~120 d of age). The adult songs of tutored birds were

recorded and compared with the songs of the tutor to assess the strength of vocal learning.

We analyzed how interactions between tutors and pupils influenced vocal learning in birds tutored for 5 d. We analyzed the occurrence of tutor song, whether the tutor song was directed at the pupil [pupil-directed (PD) song] or not [undirected (UD) song], whether the pupil was attentive to the tutor's song (i.e., awake, still, and not engaged in feeding, flying, calling, singing, or drinking when the tutor was singing), and contingent vocal responses between pupils and tutors (*SI Appendix*).

We analyzed differences in various temporal and spectral features of song between PD and UD song that were produced at comparable times of the day to control for circadian variation in song. To assess the specificity of vocal changes accompanying PD song, we also housed male tutors next to adult females for 5 d and compared temporal and spectral features of female-directed (FD) and UD songs (*SI Appendix*).

Immunocytochemistry and Brain Imaging. We analyzed how social interactions during song tutoring influenced the expression of EGR-1, a cellular marker of neural activity (30, 31), in midbrain catecholaminergic neurons. Juveniles were allowed to visually and acoustically interact with a tutor, allowed to passively hear the songs of tutors, or kept in isolation without song tutoring for <2.25 h. As in the song-learning experiments, socially and passively tutored juveniles were tutored in a yoked manner (*SI Appendix*, Fig. S1A). After tutoring, lights were turned off, and animals were killed using an overdose of isoflurane, and then transcardially perfused. Untutored control birds remained in isolation without song exposure before being killed. All procedures were approved by the McGill University Animal Care and Use Committee in accordance with the guidelines of the Canadian Council on Animal Care.

We also analyzed the degree to which visual and song information was integrated by midbrain catecholaminergic neurons. We compared EGR-1 expression in catecholaminergic neurons between juveniles that were allowed to socially interact with a tutor that produced song, allowed to socially interact with a tutor that did not produce song, or kept in isolation without exposure to song for <2.25 h (*SI Appendix*).

Brains were processed for tyrosine hydroxylase (TH) and EGR-1 immunocytochemistry using established protocols (*SI Appendix*). We quantified EGR-1 expression in the LC, VTA, and SN.

Statistical Analyses. We used mixed effects models to analyze the effect of social interactions on vocal learning, context-dependent changes to the structure of tutor songs, and EGR-1 expression (*SI Appendix*). For example, to examine the effect of social interactions on vocal learning, we analyzed the effect of "CONDITION" (e.g., social vs. passive) on similarity scores. Because juveniles were organized into cohorts that were tutored by the same adult, we also included "COHORT" as a random variable in the model. Tukey's honest significant difference (HSD) test was used for all post hoc contrasts. Data that were not normally distributed were transformed in various ways to improve normality before analyses (*SI Appendix*). All statistics were performed using JMP v11.2 (SAS).

1. Kuhl PK (2007) Is speech learning 'gated' by the social brain? *Dev Sci* 10(1):110–120.
2. Kuhl PK (2010) Brain mechanisms in early language acquisition. *Neuron* 67(5):713–727.
3. Meltzoff AN, Kuhl PK, Movellan J, Sejnowski TJ (2009) Foundations for a new science of learning. *Science* 325(5938):284–288.
4. Tomasello M (2003) *Constructing a Language: A Usage-Based Theory of Language Acquisition* (Harvard Univ Press, Cambridge, MA), pp 69–88.
5. Kuhl PK, Tsao FM, Liu HM (2003) Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proc Natl Acad Sci USA* 100(15):9096–9101.
6. Goldstein MH, King AP, West MJ (2003) Social interaction shapes babbling: Testing parallels between birdsong and speech. *Proc Natl Acad Sci USA* 100(13):8030–8035.
7. Goldstein MH, Schwade JA (2008) Social feedback to infants' babbling facilitates rapid phonological learning. *Psychol Sci* 19(5):515–523.
8. Adolphs R (2010) Conceptual challenges and directions for social neuroscience. *Neuron* 65(6):752–767.
9. Baron-Cohen S, Tager-Flusberg H, Cohen D, eds (1993) *Understanding Other Minds: Perspectives from Autism* (Oxford Univ Press, Oxford).
10. Bloom K, Russell A, Wassenberg K (1987) Turn taking affects the quality of infant vocalizations. *J Child Lang* 14(2):211–227.
11. Fernald A, Kuhl P (1987) Acoustic determinants of infant preference for motherese speech. *Infant Behav Dev* 10(3):279–293.
12. Liu HM, Kuhl PK, Tsao FM (2003) An association between mothers' speech clarity and infants' speech discrimination skills. *Dev Sci* 6(3):F1–F10.
13. Doupe AJ, Kuhl PK (1999) Birdsong and human speech: Common themes and mechanisms. *Annu Rev Neurosci* 22(1):567–631.
14. Catchpole CK, Slater PJ (2008) *Bird Song: Biological Themes and Variations* (Cambridge Univ Press, Cambridge, UK).
15. Brainard MS, Doupe AJ (2002) What songbirds teach us about learning. *Nature* 417(6886):351–358.
16. Mooney R (2009) Neurobiology of song learning. *Curr Opin Neurobiol* 19(6):654–660.
17. Kojima S, Doupe AJ (2007) Song selectivity in the pallial-basal ganglia song circuit of zebra finches raised without tutor song exposure. *J Neurophysiol* 98(4):2099–2109.
18. West MJ, King AP (1988) Female visual displays affect the development of male song in the cowbird. *Nature* 334(6179):244–246.
19. Snowdon CT, Hausberger M, eds (1997) *Social Influences on Vocal Development* (Cambridge Univ Press, Cambridge, UK).
20. Zann RA (1996) *The Zebra Finch: A Synthesis of Field and Laboratory Studies* (Oxford Univ Press, New York).
21. Eales LA (1989) The influences of visual and vocal interaction on song learning in zebra finches. *Anim Behav* 37(3):507–508.
22. Baptista LF, Petrinovich L (1986) Song development in the white-crowned sparrow: Social factors and sex differences. *Anim Behav* 34(5):1359–1371.
23. Derégnaucourt S, Poirier C, Kant AV, Linden AV, Gahr M (2013) Comparisons of different methods to train a young zebra finch (*Taeniopygia guttata*) to learn a song. *J Physiol Paris* 107(3):210–218.
24. Nelson DA (1998) External validity and experimental design: The sensitive phase for song learning. *Anim Behav* 56(2):487–491.
25. Posner MI, ed (2011) *Cognitive Neuroscience of Attention* (Guilford, New York).
26. Trainor LJ, Clark ED, Huntley A, Adams BA (1997) The acoustic basis of preferences for infant-directed singing. *Infant Behav Dev* 20(3):383–396.
27. Werker JF, Pegg JE, McLeod PJ (1994) A cross-language investigation of infant preference for infant-directed communication. *Infant Behav Dev* 17(3):323–333.
28. Fernald A, et al. (1989) A cross-language study of prosodic modifications in mothers' and fathers' speech to preverbal infants. *J Child Lang* 16(3):477–501.
29. Burnham D, Kitamura C, Vollmer-Conna U (2002) What's new, pussycat? On talking to babies and animals. *Science* 296(5572):1435.
30. Guzowski JF, et al. (2005) Mapping behaviorally relevant neural circuits with immediate-early gene expression. *Curr Opin Neurobiol* 15(5):599–606.
31. O'Donovan KJ, Tourtellotte WG, Millbrandt J, Baraban JM (1999) The EGR family of transcription-regulatory factors: Progress at the interface of molecular and systems neuroscience. *Trends Neurosci* 22(4):167–173.
32. Aston-Jones G, Cohen JD (2005) An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annu Rev Neurosci* 28:403–450.
33. Fields HL, Hjelmstad GO, Margolis EB, Nicola SM (2007) Ventral tegmental area neurons in learned appetitive behavior and positive reinforcement. *Annu Rev Neurosci* 30:289–316.
34. Roozendaal B, McGaugh JL (2011) Memory modulation. *Behav Neurosci* 125(6):797–824.
35. Sara SJ, Bouret S (2012) Orienting and reorienting: The locus coeruleus mediates cognition through arousal. *Neuron* 76(1):130–141.
36. Castelino CB, Schmidt MF (2010) What birdsong can teach us about the central noradrenergic system. *J Chem Neuroanat* 39(2):96–111.
37. Matheson LE, Sakata JT (2015) Catecholaminergic contributions to vocal communication signals. *Eur J Neurosci* 41(9):1180–1194.
38. Knudsen EI (2007) Fundamental components of attention. *Annu Rev Neurosci* 30:57–78.
39. Langton SR, Watt RJ, Bruce I (2000) Do the eyes have it? Cues to the direction of social attention. *Trends Cogn Sci* 4(2):50–59.
40. Domjan M (2014) *The Principles of Learning and Behavior* (Cengage Learning, Stamford, CT).
41. Picton TW, Hillyard SA, Galambos R (1976) Habituation and attention in the auditory system. *Auditory System*, eds Keidel WD, Neff WD (Springer, Berlin), pp 343–389.
42. Caro TM, Hauser MD (1992) Is there teaching in nonhuman animals? *Q Rev Biol* 67(2):151–174.
43. Kline MA (2015) How to learn about teaching: An evolutionary framework for the study of teaching behavior in humans and other animals. *Behav Brain Sci* 38:e31.
44. Tchernichovski O, Lints T, Mitra PP, Nottebohm F (1999) Vocal imitation in zebra finches is inversely related to model abundance. *Proc Natl Acad Sci USA* 96(22):12901–12904.
45. Houx AB, Ten cate C (1999) Song learning from playback in zebra finches: Is there an effect of operant contingency? *Anim Behav* 57(4):837–845.
46. Ungless MA (2004) Dopamine: The salient issue. *Trends Neurosci* 27(12):702–706.
47. Nordeen EJ, Holtzman DA, Nordeen KW (2009) Increased Fos expression among midbrain dopaminergic cell groups during birdsong tutoring. *Eur J Neurosci* 30(4):662–670.
48. Velho TA, et al. (2012) Noradrenergic control of gene expression and long-term neuronal adaptation evoked by learned vocalizations in songbirds. *PLoS One* 7(5):e36276.
49. Bao S, Chan VT, Merzenich MM (2001) Cortical remodelling induced by activity of ventral tegmental dopamine neurons. *Nature* 412(6842):79–83.
50. Grant SJ, Aston-Jones G, Redmond DE, Jr (1988) Responses of primate locus coeruleus neurons to simple and complex sensory stimuli. *Brain Res Bull* 21(3):401–410.
51. Ikeda MZ, Jeon SD, Cowell RA, Remage-Healey L (2015) Norepinephrine modulates coding of complex vocalizations in the songbird auditory cortex independent of local neuroestrogen synthesis. *J Neurosci* 35(25):9356–9368.