Evolution and Human Behavior 31 (2010) 354–364

Evolution and Human Behavior

Original Article

Joint music making promotes prosocial behavior in 4-year-old children☆,☆☆

Sebastian Kirschner*, Michael Tomasello

Department of Developmental and Comparative Psychology, Max Planck Institute for Evolutionary Anthropology, 04103 Leipzig, Germany

Initial receipt 25 April 2009; final revision received 24 April 2010

Abstract

Humans are the only primates that make music. But the evolutionary origins and functions of music are unclear. Given that in traditional cultures music making and dancing are often integral parts of important group ceremonies such as initiation rites, weddings or preparations for battle, one hypothesis is that music evolved into a tool that fosters social bonding and group cohesion, ultimately increasing prosocial in-group behavior and cooperation. Here we provide support for this hypothesis by showing that joint music making among 4-year-old children increases subsequent spontaneous cooperative and helpful behavior, relative to a carefully matched control condition with the same level of social and linguistic interaction but no music. Among other functional mechanisms, we propose that music making, including joint singing and dancing, encourages the participants to keep a constant audiovisual representation of the collective intention and shared goal of vocalizing and moving together in time — thereby effectively satisfying the intrinsic human desire to share emotions, experiences and activities with others.

Keywords: Music; Dance; Gene–culture co-evolution; Human behavior; Child development; Social cognition; Joint action; Shared intentionality; Altruism; Cooperation; Rhythmic entrainment; Gender differences; Interpersonal synchrony

1. Introduction

The evolutionary origins of music are a puzzle, since music lacks any obvious adaptive function (Darwin, 1871; Wallin, Merker, & Brown, 2000). Some theorists have speculated that it actually has no adaptive function, but rather music was invented as a pure pleasure stimulant and all components of human musicality originally evolved for non-musical purposes, like for language, fine motor-control or emotional communication (James, 1890; Pinker, 1997). This by-product hypothesis provides a plausible theoretical rationale for the initial step in music evolution at the point where human ancestors produced the first music-like behaviors, but it does not preclude the possibility that music could have later acquired some adaptive function(s), either biological or cultural.

As music is an omnipresent behavior across all cultures (Merriam, 1964), with deep roots in human ontogeny (Trehub, 2001), an ancient history of at least 40,000 years (Conard, Malina, & Munzel, 2009) and powerful psychological effects on mood and emotions (Sloboda & Juslin, 2001), other theorists have proposed various adaptive functions of music-like behaviors — at least at some stage of human biological and cultural evolution. Such adaptive theories are not necessarily mutually exclusive, since each may account for certain aspects of human music today which could have evolved at different evolutionary periods. In addition, it is important to be clear about whether music in its biological or cultural dimension is at issue: particular forms of music are products of cultural evolution, whereas the innate and universal components of human musicality are products of biological evolution.

Darwin (1871) proposed that musical behaviors once had an adaptive function that they no longer have (i.e., music is an evolutionary vestige). In this view, the major components

☆ Author contributions*:* S. Kirschner and M. Tomasello designed the study and wrote the paper. S. Kirschner conducted the experiments and analyzed the data.

☆☆ The research of S. Kirschner is supported by a grant from the German National Academic Foundation.

* Corresponding author. Tel.: +49 341 3550 474; fax: +49 341 3550 444.

E-mail address: kirschner@eva.mpg.de (S. Kirschner).

1090-5138/$ – see front matter © 2010 Elsevier Inc. All rights reserved.
doi:10.1016/j.evolhumbehav.2010.04.004

of human musicality originated in an ancient, pre-linguistic, songlike communication system comprising learned and complex acoustic signals (Brown, 2000b; Mithen, 2005; Richman, 1993). At a later stage in human evolution, this communication system was upgraded by a more efficient one — human language — leaving our species with the innate predisposition to create today's music. Analogous to bird song during courtship, Darwin suggested that music-like behaviors first evolved by means of sexual selection, as individuals advertised for mates (for an extended argument see Miller, 2000).

The idea that music originally evolved as a display signal was also put forward by Merker (2000), who proposed that synchronous chorusing by hominid males served as an indicator of coalition strength, helping to defend territory and at the same time attract migrating females. Similarly, Hagen and Bryant (2003) suggested that group music making and dancing evolved as between-group displays signaling internal stability and the group's ability to act collectively, thereby establishing meaningful relationships — whether cooperative or hostile — between groups. These group display theories, however, have a hard time explaining how such a signaling system could be invented and stabilized in the first place within large groups of often non-related individuals, since it appears vulnerable to cheating. For example, individuals might participate in the musical group performance, but only pretend to share the group's coalition agreement, later taking personal advantage of the others' commitment. Furthermore, these signal theories are supported by rather sparse ethnomusicological evidence and do not account for the majority of musical encounters observed across cultures today where music is part of peaceful within-group ceremonies outside any sexual or competitive context (Clayton, 2009).

Another group of theories treat music not as a signal but as a tool — thereby circumventing the problem of cheating. For example, Dissanayake (2000) and Falk (2004) have advocated a kin-selected function for an ancient musical communication system in mother–infant bonding: prosodic utterances might have served to keep mothers and their infants in psychological contact when they were physically separated: e.g., while mothers prepared food or manufactured tools. Indeed, the use of lullabies to soothe infants is considered a human universal (Trehub, 2001) and when it comes to communicating emotion through infant-directed speech, "the melody is the message" (Fernald, 1989).

Another related hypothesis is that music and dance, once invented, turned out to be effective tools to establish and maintain social bonds and prosocial commitment among the members of social groups, ultimately increasing cooperation and prosocial in-group behavior (Huron, 2001; McNeill, 1995; Roederer, 1984). Unlike the many modern Western examples of individual music consumption (e.g., via iPod or car radio), in traditional small-scale societies music is typically performed for pragmatic reasons (Bohlman, 2000), integrated into ritual ceremonies such as worship, weddings, funerals or preparations for hunt or combat (Clayton, 2009; Dissanayake, 2006). These ceremonies are usually considered to be essential for the maintenance of the group's identity, with the music being an indispensable part of it. On the proximate level, several universal features of human music (Fitch, 2006; Stevens & Byron, 2009) — like its ritualized context, periodic pulse (beat), discrete pitches and a highly repetitive repertoire — may contribute to solving the proposed adaptive problem of maintaining group cohesion. Specifically, they all make music more predictable than, for example, language and thus enable coordination between multiple individuals at once via synchronization of body movements and blending of voices.

This hypothesis of music as a tool for supporting group cohesion predicts that joint music making ultimately increases prosocial commitment and fosters subsequent cooperation among the performers. Indeed, Anshel and Kipper (1988) found that adult Israeli males cooperate better in a prisoner's dilemma game and score higher on a questionnaire on trust after a group singing lesson, compared to passive music listening, active poetry reading or just watching a film together. Likewise, Wiltermuth and Heath (2009) showed that US students scored higher on a weak-link coordination-exercise and a public-goods game after joint singing along with a song played from headphones, compared to no singing or forced "asynchronous" singing (via playing the same stimulus at individual tempi). Adding synchronous movement (by moving plastic cups from side to side on a table) to the synchronous singing condition did not improve the scores in the subsequent economic games.

However, for the evolutionary argument, much stronger evidence would be provided if similar prosocial effects could be shown in young children. Kindergarten children are presumably not engaged in sexual advertising, nor do they have to form coalitions in fear of encountering rival neighboring groups. In terms of the group cohesion hypothesis, kindergarten children are a better test than adults because children this young, especially in Western cultures, have had few experiences of institutionalized music occurring for external pragmatic reasons. Therefore, we can probably neglect normative knowledge as a source for their interpretation of the manipulation phase and for their decision making during the dependent measures (Olson & Spelke, 2008). But since all human children have musical predispositions and skills (Trehub & Hannon, 2006; Zentner & Eerola, 2010), it would be very telling if involvement in joint music making and dancing somehow influences children's spontaneous altruistic and cooperative tendencies.

In the current study, therefore, we had pairs of 4-year-old children participate in a 3-min episode of interactive play. Using the same setup, procedure and cover story, children either interacted with one another (and an adult) in the context of traditional music — that is, with dancing, singing and playing percussion instruments to a novel, but easy-to-learn, children's song (Musical condition) — or they interacted with one another (and an adult) during basically the same joint activity but without singing, dancing or playing instruments (Non-musical condition). Immediately after this manipulation phase, each pair participated in two social interactions designed to test their willingness to (1) help their partner and (2) cooperate on a problem-solving task. We predicted that prior engagement in joint music making should make children behave more prosocial, i.e., spontaneously help each other more and solve a task rather jointly instead of alone.

2. Methods

2.1. Participants

A final sample of 96 four-year-old children were included in the study (48 males and 48 females, mean=4 years and 6 months, range=4.0 to 5.0 years). Children were recruited from 16 different German urban day-care centers and came from mixed socioeconomic backgrounds. The two children in a pair were always recruited from the same kindergarten group. This way we could assume that the peers knew each other from frequent previous interactions. In addition, both children were asked whether they knew the other peer, and they had to agree to play the new game together before the experimenter brought them to the testing room. Children in a pair were familiar with each other in order to create a situation analogous to those present in traditional small-scale societies. However, to ensure that each pair was randomly assigned to one of the two conditions, the order of conditions was defined prior to testing. Importantly, we ensured that the children in our study were not aware of participating in a scientific study; instead, they joined the experimenter in the belief they would play some novel games he had brought along. Another 15 pairs took part in the study but were not included in the final sample, either because they did not pass the warm-up task (three pairs), because they did not pass the cooperation test (two pairs) or they did not pass the helping test (10 pairs). The exclusion criteria for each test are detailed below.

2.2. General study design

In addition to the musical vs. non-musical manipulation, our study had gender as an independent variable. For simplification, we only paired children of the same gender. As a result, our study had a between-participants, 2×2 design, with two independent variables (condition and gender) and two dependent variables (voluntary helping and spontaneous cooperative problem solving). The final sample size was 48 pairs with 12 pairs for each condition–gender combination. The whole session lasted about 20 min and consisted of four main episodes: (1) experimental manipulation phase, (2) dependent measure one, (3) manipulation phase repeated and (4) dependent measure two. Our reason for running the manipulation phase again before the second dependent measure was to reinforce any prosocial effects of music in case they were transient and faded while the children were concerned with the tasks to follow. During data acquisition, we alternated the condition from session to session and counterbalanced the order of dependent measurements within sessions (either helping test first or cooperation test first).

2.3. Experimenter

All testing was done by the same experimenter (first author). He recruited the participants in the kindergarten groups and administered the whole session. A technical assistant was present in the testing room, but only spoke during a short introduction to the children and briefly as part of a demonstration of the cooperation test (see below). In order to avoid any third-party/authority biases or bystander effects, the experimenter and the technical assistant were not present during both dependent measures so that the two children had to solve both tasks on their own.

2.4. Manipulation phase

We designed both conditions such that the tasks were equally difficult, plausible and motivating for the children; that the children had to follow exactly the same sequence of actions in order to reach the same joint goal; and that the amount of movement, gestural and verbal interaction was leveled across conditions. To do so, the experimenter strictly followed the same script to demonstrate the consecutive actions and instruct the children, either embedded as dance and sung phrases during the Musical condition or as “non-dancing” movement and spoken phrases during the Non-musical condition.

We developed a detailed background story to keep the children motivated throughout the whole session and to cover each condition as experimenter-guided pretend play. For a detailed protocol of the setup and procedure, please refer to the Supplementary Material online. The cover story of the manipulation phase included a “garden pond” (an oval blanket) inhabited by nine colored frogs, sitting in trios on three lily pads at the pond’s rim (Fig. 1). Each frog could be used — according to condition — either as a normal toy by letting it hop up and down the floor, or as a musical instrument by scraping its back with an additional stick. At the beginning of the first manipulation phase, the experimenter introduced one extra frog to the children and — as a warm-up task — asked each child to hold the frog by herself and copy the experimenter’s action according to condition. In order to pass the warm-up task, each child had to voluntarily pick up the frog at least once and imitate the experimenter’s action. After the warm-up, the experimenter pretended that the nine frogs in the pond were still asleep and needed to be woken up either by a “morning song” (Musical condition) or by some “morning exercise” (Non-musical condition). After one round of demonstration, where the children only had to watch the experimenter, he invited the children to pick up a frog by themselves in order to help

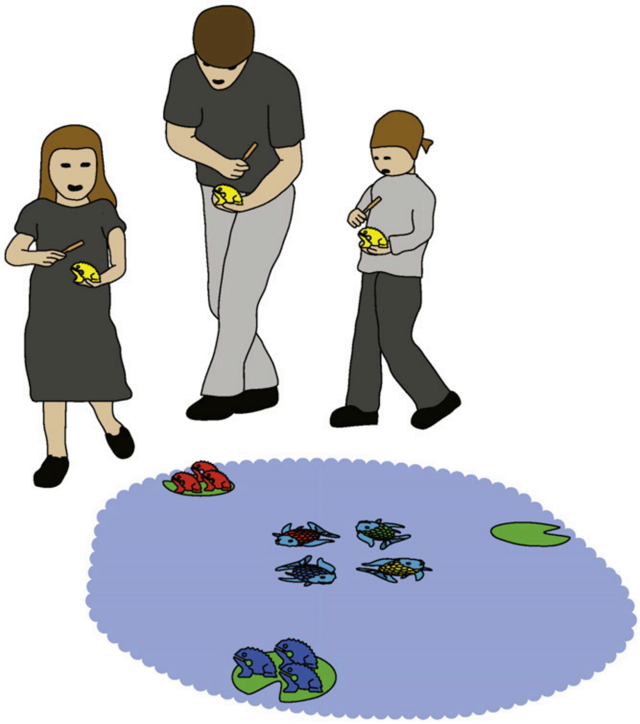

Fig. 1. *Manipulation phase*. During the Musical condition, children in a pair danced around a pond (blue blanket) together with the experimenter, while singing a novel children's song and playing percussion instruments (wooden bullfrogs) in time with their singing and additional background music. The Non-musical condition was based on exactly the same setup, procedure and cover story. However, we omitted any musical features while moving around the pond: the frogs were introduced as simple toys and only spoken language was used for communication.

waking them up. Then again, he demonstrated the task sequence according to condition, this time asking the children to "do as I do". During the next 3 min of semi-guided play, the children voluntarily imitated the experimenter's actions and copied his utterances.

In the Musical condition, children had the opportunity to follow the experimenter in walking around the pond while synchronizing their steps to the pulse of the music and singing a novel, but easy-to-learn, song (see Supplementary Material for details on the song) along with guitar chords (pre-recorded) using the frogs as instruments in synchrony with the song's lyrics. In the Non-musical condition, everybody walked and crawled around the pond while letting the frogs jump in non-synchronized intervals with accompanying utterances. The whole group song/exercise was performed three times during one manipulation phase until all frog trios were "awake".

The experimenter followed basically the same script as listed in the Supplementary Material during both conditions such that the order and timing of events were identical during the whole manipulation phase. Furthermore, in order to constantly control the level of joint activity during the play, we repetitively integrated short tasks that required the joint action of all three participants: (1) picking up the frogs from one lily pad at exactly the same moment, (2) putting them back simultaneously or (3) tapping the frogs together to pretend three "kisses" once every round (see Supplementary Material for details). All of these joint action tasks were triggered by certain key words in the experimenter's sung/spoken instructions and could only be accomplished in time if both children constantly paid attention to the play partner's behavior and coordinated their actions accordingly. Finally, the experimenter's verbal instructions during the Non-musical play had exactly the same content and were as frequent as those instructions embedded in the lyrics of the song during the Musical play. In order to sound natural during both conditions, the instructing phrases were highly repetitive and did rhyme (to be applied as lyrics in the song of the Musical condition), but at the same time had a straightforward grammatical structure (to be applied as "ordinary" verbal instructions in the Non-musical condition).

The only difference between the two conditions was that the musical manipulation phase included some distinct features of music which are — in this combination — not shared with speech and other forms of social interaction (Fitch, 2006; Nettl, 2005). First, a *periodic pulse* underlying the children's song functioned as a shared reference for the children to synchronize their body movements: i.e., (1) the scraping of the frogs with the sticks and (2) the footsteps while "dancing" around the pond. Therefore we played the song at a tempo of 115 beats per minute, which is close to the spontaneous tempo of human locomotion (MacDougall & Moore, 2005). Second, the use of *discrete pitches* and a highly *repetitive melodic structure* allowed the children to reproduce the song easily and sing in chorus with the experimenter. Third, the discretization of time and pitch in music made the children's actions and utterances more predictable and ritualized in the Musical condition, creating a *joint performative context*. Finally, the integration of music into the interactive game created an additional expressive mode that lies beyond the referential and propositional use of words in language. This a-referential *expressiveness* of music (Fitch, 2006) has been shown to effectively communicate mood, affect and distinct emotions between performer and listener (Juslin & Västfjäll, 2008; Sloboda & Juslin, 2001).

2.5. Dependent measure one: spontaneous helping

In the helping test, one child had a sudden accident which presented the partner with the choice to help or to continue pursuing her own play activity (van Baaren, Holland, Kawakami, & Knippenberg, 2004). We designed this test to elicit spontaneous behavior in the children's familiar environment and to create actual costs for helping. The cover story of the helping test was to "prepare some food in order to feed the fish in the pond". For this task, each child had to carry six out of 24 marbles at a time in a tube towards a novel apparatus, where they could be used in a fun game ("grind

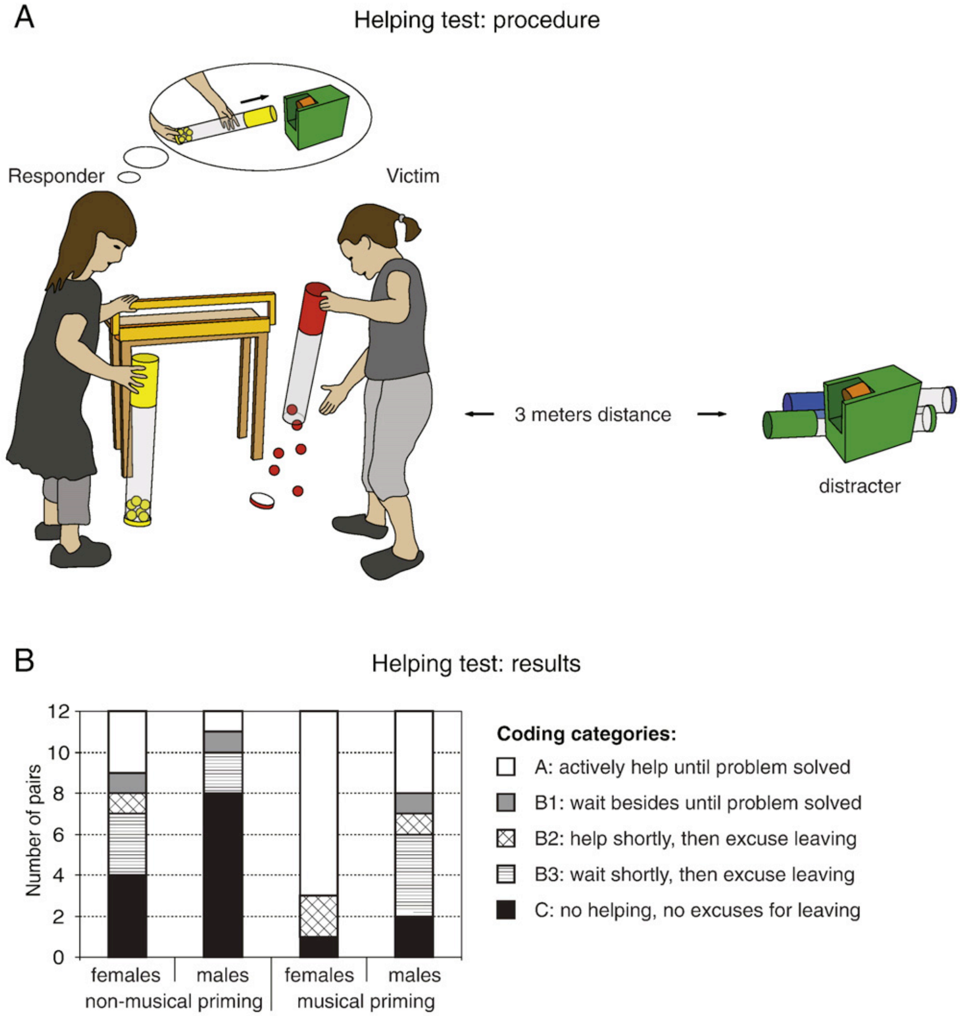

Fig. 2. *Dependent measure one: spontaneous helping*. We created a situation where one child (victim) had a sudden accident, which presented the other child (responder) with the free choice to either help actively, wait or continue pursuing her own play activity. Both children were instructed to carry tubes filled with colored marbles from a rack towards a novel apparatus, where each tube could be applied in an attractive game (Panel A). During the actual test, this apparatus served as a distracter, creating costs for the responder, because she had to postpone her own goals in order to wait for and/or help the victim. The accident happened when both children lifted their last tubes: the victim's tube's bottom lid fell off such that all marbles spilled over the floor. We only coded the responder's behavior between the moment she noticed the accident and the moment she continued using her tube with the distracter. We categorized her response according to the effort she made to fix the victim's tube and/or collect the marbles (Panel B).

the fish food") independent of the other peer (Fig. 2A). After the experimenter gave a demonstration of the task to each child by following a fixed script (Supplementary Material), he (and the assistant) left the room for 2 min, allowing both children to "go and grind" their remaining marbles on their own. However, the tube of one child ("victim" in Fig. 2A) was prepared so that the bottom fell out at the moment both children lifted the tubes and all six marbles spilled out onto the floor. At that stage, the novel apparatus served as a distracter to create explicit costs for the responding child ("responder" in Fig. 2A) because she had to postpone her private goals in order to wait for or help the victim. Importantly, we only assessed the responder's behavior between the moment she observed the accident and the moment she ground her own yellow marbles (38 out of 48 trials), or — in case she did not grind her own marbles within 2 min — the moment the experimenter returned and solved the victim's problem (10 trials).

2.6. Dependent measure two: spontaneous cooperative problem solving

In the cooperation test, children from a pair played a novel game comprising a task that could be solved either individually or cooperatively (Azmitia, 1988). The cover story of this test was to "feed the baby fishes living in two small aquariums". The task comprised the use of 12 marbles (six red and six yellow) in two connected "slides" that were functionally identical, but differed in color (again, one red and one yellow; Fig. 3A). At each slide, one could drop a marble into one end (action X in Fig. 3A) and then, from the other end, pull a string to bring it out into a container (the

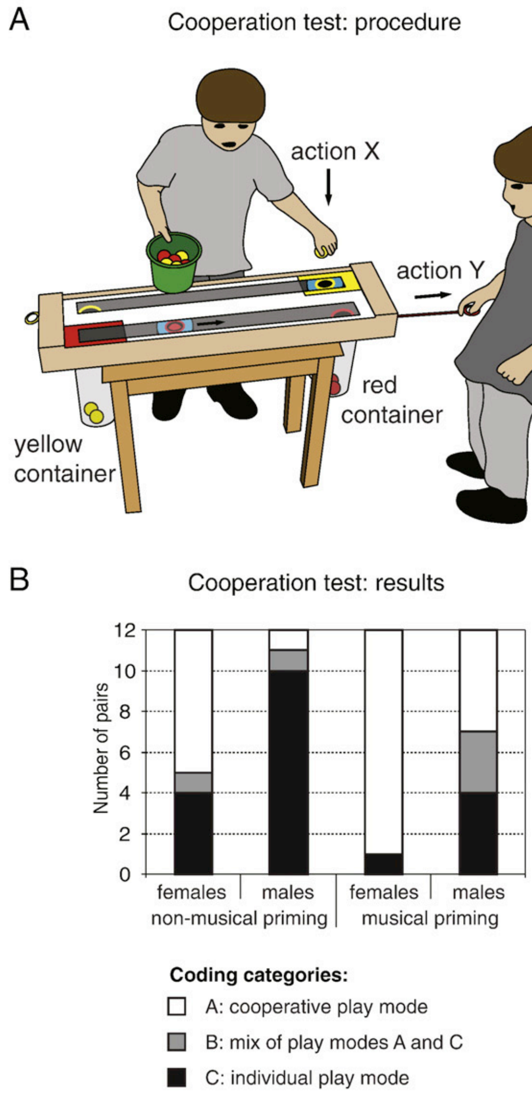

Fig. 3. *Dependent measure two: spontaneous cooperative problem solving*. To test the children's willingness to cooperate with each other spontaneously, we designed a game that could be played either individually or cooperatively (Panel A). The task was to transport red and yellow marbles into their appropriate containers via two functionally independent slides. Both slides were installed in the same apparatus, but working in opposite directions. The marbles needed to be placed one by one into the proper slide (action X), before pulling a string on the opposite end of the apparatus (action Y). The individual solution to this task was that both children only used the marbles of one color, performing both actions X and Y on one slide only, while walking around the apparatus. However, during cooperative play mode each child could stay at one end of the apparatus, easily pull "her" string and also load marbles for the partner. After the experimenter demonstrated both play versions, he let the children decide to solve the task either cooperatively or individually. During five trials, children from a pair switched between play modes (Panel B).

"aquarium") of appropriate color (action Y in Fig. 3A). The two slides were installed in parallel fashion into a single apparatus but in opposite orientations. The individual solution to this task was to "split" the apparatus, claim one color side and only feed those six marbles of appropriate color by performing both actions alone by walking around to each end in turn. However, the overall task could also be solved cooperatively if the children stayed at either end of the apparatus and successfully coordinated their actions. This way, one child could easily pull "her" string and also load marbles for the other child. After the experimenter demonstrated both individual use of the apparatus (again, by following a fixed script, as detailed in the Supplementary Material) and the joint use (by asking the assistant for help), he and the assistant left the room and let the children decide how to solve the task.

2.7. Coding and reliability

All coding was done from short video clips that had been extracted from the original videotapes. To ensure condition-blind coding, all clips were labeled with an arbitrary index number to conceal the condition from the coders. To assess inter-rater agreement, 25% of the trials were independently coded by a hypothesis-blind research assistant. For the helping test, we reached an inter-observer agreement of κ=0.99 and for the cooperation test the agreement was κ=0.85.

In the helping test, we coded the responder's reaction to the accident according to the effort she made to fix the peer's tube and/or collect the marbles. The highest rank of helpfulness was ascribed to children who actively contributed to the solution of the peer's problem plus waited until it was finally solved (Category A in Fig. 2B). Such "active helping" could have been picking up marbles, picking up the lid, trying to reattach the lid to the tube, suggesting strategies to fix the tube or holding the tube while the victim recollected the marbles. We defined the problem as "being solved", as soon as all red marbles were back in the red tube and the victim could proceed with grinding. The lowest rank of helpfulness was ascribed to those children who saw the peer's accident, but instantly turned away without any verbal leave-taking or excuses. Instead, they walked over to the distracter and immediately poured all of their own marbles into the apparatus (Category C in Fig. 2B).

Intermediate ranks were coded for those trials where the responder (1) did not actively contribute to the solution of the victim's problem, but withheld from grinding until the problem was solved (Category B1); (2) started to actively help, but then decided to abort helping although the problem was not solved yet and proceeded grinding (Category B2); or (3) withheld from grinding without getting involved actively, but then also decided to leave the peer before the problem was solved and start grinding alone (Category B3). Importantly, in all trials coded as Category B2 or B3, the responder used verbal leave-taking or excuses for refusing to help or wait any longer (all utterances are listed in Supplementary Table 1). For the rank-based ANOVA, we grouped Categories B1, B2 and B3 together, resulting in three ranks of helpfulness: high (A), intermediate (B1–B3) and low (C).

Another 10 trials had to be excluded from the final sample because the responder did not notice the accident and, thus, was not aware of the victim's need for help. In seven cases, this happened because the responder disobeyed the experimenter's order to wait until he gives the signal to lift the tubes simultaneously (see Supplementary Material), instead immediately ran towards the distracter before the experimenter had left the room. In another three cases, the responder picked up her tube much faster than the victim and ran off. As a consequence of both circumstances, the victim's tube broke when the responder was already half way at the distracter, with her back turned towards the accident scene. Because the application of the distracter only took about 2 sec, all of the 10 excluded children had finished their task before they even noticed that their partner was still at the rack trying to repair her own tube.

In the cooperation test, we only coded the children's behavior between the moment the experimenter handed the bucket of marbles over to the children and the moment all 12 marbles were used. We categorized the task as being solved either through cooperative play (Category A in Fig. 3B) or individual play (Category C) based on the frequency of slide switches, shared actions and communicative utterances. A "slide switch" occurred when one child or both children switched from using the red to using the yellow slide or vice versa. We coded as "shared actions" when both children intentionally contributed to the processing of one marble on the same slide, i.e., one child refilled the slide and the other child pulled it towards the appropriate container. Finally, we recorded the frequency of those utterances that communicated actions, intentions and strategies that were directly relevant to the joint solution of the task (examples are listed in Supplementary Table 2).

For a trial to be coded as cooperative play mode, both children had to switch slides at least once and also share actions X and Y on the same slide at least once. In addition, at least two communicative utterances had to accompany the joint solution of the task. During those trials coded as individual play mode, such slide switches or shared actions never occurred because each child stuck to one slide only — nor did the children use any communicative utterances to assist a joint task solution. However, five out of 48 trials could not be clearly categorized as either cooperative or individual play because the children switched between play modes at a later point of the game. Therefore we put those trials into a third, intermediate category. For the ANOVA, we ranked all three categories according to the level of cooperativeness: high (A), intermediate (B) and low (C).

Another two trials had to be excluded from the final sample because a child used the apparatus improperly, either attempting to put two marbles at a time into one slide and blocking the apparatus, or pulling the string only halfway such that the marble never fell into the corresponding container.

2.8. Statistical methods

We ranked all coding categories from the helping and cooperation test according to three levels of prosociality: high (A), intermediate (B) and low (C). Therefore, we applied a non-parametric, rank-based two-way ANOVA (Meddis, 1984) with two crossed factors: condition and gender. With increasing sample size the distribution of the test statistic of this test ('H') asymptotically approaches a chi-square distribution and can be approximated using the latter. Since our response variables represent ranked scores rather than interval scale measures, we refrained from indicating effect sizes, which would be hard to interpret and potentially unreliable. Instead, we present each individual response from all 48 trials in Figs. 2B and 3B. For all tests, we considered p values $\leq .05$ as significant.

3. Results

3.1. Spontaneous helping

Results from the helping test are shown in Fig. 2B. Children of both genders helped one another more after joint music making [χ^2 (1)=9.92; $p<.01$]. Girls were more helpful than boys overall [χ^2 (1)=4.46; $p=.03$], but the magnitude of the difference between conditions was similar for both genders [test of the interaction: χ^2 (1)< 0.01; $p=.97$]. Interestingly, looking more closely at those trials in which children decided not to help or stay until the peer's problem was solved, the children in the Musical condition more frequently offered verbal excuses (70% of trials) than children in the Non-musical condition (33%). These excuses suggest that, even when they did not help, children in the Musical condition still showed greater empathy or greater social commitment than children in the Non-musical condition (all utterances are listed in Supplementary Table 1).

3.2. Spontaneous cooperative problem solving

Results from the cooperation test are presented in Fig. 3B. Consistent with the results from the helping test, children of both genders chose the cooperative solution to the task more often after joint music making [χ^2 (1)=6.04; $p=.01$]. Again, girls were more cooperative than boys [χ^2 (1)=9.87; $p<.01$], but the magnitude of the difference between conditions was similar for both genders [test of the interaction: χ^2 (1)=0.07; $p=.78$]. During cooperative play, various types of verbal utterances were used to communicate strategies ("Let's try to pull together!"), to coordinate alternating actions ("Now it's your turn!") or to announce the joint completing of the task ("We have finished!") (more examples are listed in Supplementary Table 2) — whereas during individual play such communicative utterances never occurred. Consequently, the mean frequency of communicative utterances per trial was higher in the Musical condition (mean=6.38, S.D.=5.11) than in the Non-musical condition (mean=3.38, S.D.=5.27; Mann–Whitney U test, U=178.5, $p=.02$).

4. Discussion

4.1. Summary of results

Our results show that, when performed in a manner akin to that typical in traditional small-scale societies, joint music making enhances prosocial behavior in 4-year-old children. Since children in the Non-musical condition interacted with one another in the same way as those in the Musical condition — shared the same goals, coordinated their actions in time, even imitated each other's movements and verbal comments, only without music and dance — our study isolates the specific effect of music over and above social and linguistic interaction in general (van Baaren et al., 2004).

The fact that music making has a positive effect on prosociality in addition to its entertaining effects can be interpreted as evidence against the hypothesis that today's music and dance are completely non-adaptive byproducts of the human mind. However, the byproduct hypothesis may provide a plausible starting point for the evolution of musical behaviors which only later became exapted to serve a range of adaptive functions. One of these functions could have been the maintenance of social bonds and prosocial commitment among the members of individual social groups, ultimately increasing cooperation and prosocial in-group behavior (Huron, 2001; McNeill, 1995; Roederer, 1984). This theory treats music and dance as effective tools for creating a positive collective experience among many people at the same time, thereby generating an intuitive feeling of community and bonding among the performers.

In contrast, theories that see the adaptive function of music and dance as being display signals rather than as social tools (Hagen & Bryant, 2003; Merker, 2000) are not as well supported by our data. For example, regarding the helping test, it is unlikely that children at the age of 4 made any rational choice like "because we just played music together, I will help you now". More likely, the children in our study made an intuitive decision to help the other child because they felt immediate empathic concern with the peer's misfortune the moment they saw the accident happening and felt committed to give support somehow. This interpretation of increased commitment after joint music making is supported by the finding that the amount of verbal excuses among those children who decided not to help or wait any longer was higher in the Musical condition.

Likewise, the cooperation test in our study was not designed to measure any kind of rational choice of whether collaborating with someone will or will not create higher personal profits as in a classic public goods game for example (as used by Wiltermuth & Heath, 2009). Instead, this test confronted children with a task that could be solved efficiently either individually or jointly. Right after the demonstration, the children had to decide quickly how to play the game, so this test measured the children's intuitive feeling of connectedness and affiliation which influenced their spontaneous choice of playing it either alone or together. Consequently, the fact that in the Musical condition more children solved the task as a team, instead of everyone on her own, provides clear evidence for music and dance functioning as behavioral tools for mutual social bonding.

Interestingly, we found robust gender differences in prosocial behavior: girls helped and cooperated more than boys. However, this effect was independent of condition but replicated similar findings from other studies (Maccoby, 2002). By 3 to 4 years of age, the majority of both boys' and girls' social interactions in preschool settings are with members of the same sex (Fabes, Shepard, Guthrie, & Martin, 1997). Consequently, girls and boys grow up in two adjacent subcultures with different play types and communication forms. Girls' play generally reinforces interpersonal closeness, nurturance, talking and social sensitivity, whereas boys' play emphasizes independence, task orientation, dominance and competition. It seems likely that the gender effects we found in our helping and cooperation test are a reflection of such gender differences in play behavior.

Our findings do not rule out display-signal theories and other hypotheses for the adaptive function of music since they may all account for different aspects of human music and musicality today which were shaped by different selective pressures during different time periods in biological and cultural evolution. They do, however, present serious problems for any theory claiming that music today has no adaptive function. Although Western cultures use music for social bonding much less frequently than some other cultures, it is still used as a matter of course at many public occasions, such as sporting events, worship and weddings, presumably at least partly to enhance the group experience.

4.2. Proximate mechanisms

It may be possible that the enhanced prosocial behavior in the Musical condition was due to some music-specific evolved psychological mechanism. However, it may also be possible that the observed effects were caused by one or a mix of psychological mechanisms that all evolved for other, non-musical purposes and are just triggered very effectively by the evolutionary more recent cultural product called music. The cheerful song we used in our Musical condition, for example, may have, among other effects, simply induced a positive mood in each listener (Thompson, Schellenberg, & Husain, 2001), by means of psychological mechanisms that originally evolved for the vocal communication of emotions in general (Juslin & Laukka, 2003). Long before our ancestors invented any form of peer group music, kin selection may have played a large part in creating our strong and unreflective emotional reaction (Juslin & Västfjäll, 2008) to humanly organized sound (Blacking, 1976), since emotional contagion mediated by vocalizations could have been highly adaptive for hominid mothers and their infants in the scenarios described by Dissanayake (2000) and Falk (2004). And so in our study, if the children retained a music-induced positive mood throughout the dependent measures, it could indeed have caused additional prosocial behaviors as it has been previously shown in adults (Fried & Berkowitz, 1979; North, Tarrant, & Hargreaves, 2004). Likewise, using resonating objects as instruments to create sounds — like the toy frogs in our study — could have created additional “fun”, thereby inducing an even more positive mood in each individual. The human disposition to create sounds by slapping and stomping on resonating objects is apparently much older than music itself, as it is shared by many non-human primates, including gorillas and chimpanzees (Arcadi, Robert, & Boesch, 1998; Pika, Liebal, & Tomasello, 2003).

Another domain-general psychological mechanism that might be responsible for the social bonding effect of joint music making could be the so-called chameleon effect (Chartrand & Bargh, 1999) which refers to a type of non-conscious behavioral mimicry often observed among communicating human partners. It has been shown in experimental settings that such mimicking behavior can increase affiliation and induce later prosocial behavior (van Baaren et al., 2004), suggesting an underlying adaptive function in fostering relationships among interacting partners (Lakin, Jefferis, Cheng, & Chartrand, 2003). In human music, the ritualized context and highly repetitive repertoire may create particular condensed sequences of matched behavior, thereby effectively triggering this bonding mechanism.

While strong repetition of gestures and utterances is a general feature of human rituals, the integration of a periodic pulse and discrete pitches are unique features of human music. As observed in our study, the use of discrete pitches in culture-specific scales helped to harmonically blend the children’s voices along with the experimenter’s singing and the playback music. This joint singing experience alone might have created the stronger commitment and cooperative attitude we observed during the dependent measures as shown in preceding studies with adults (Anshel & Kipper, 1988; Wiltermuth & Heath, 2009). Likewise, the periodic pulse in music is essential to allow synchronization of the performer’s body movements during dancing and instrument play. The resulting interpersonal synchrony alone has been argued to produce positive emotions that weaken the boundaries between the self and the group, leading to feelings of collective delight that enable groups to remain cohesive (Hove & Risen, 2009; Wiltermuth & Heath, 2009).

Our own favorite candidate for an underlying proximate mechanism (not exclusive of the other hypotheses) is shared intentionality. Thus, we hypothesize that joint music making, as in our study, not only creates a higher level of coordination compared to other non-musical group activities, but also encourages the participants to maintain a constant audiovisual representation of the collective intention and shared goal of vocalizing and moving together in time, thereby strengthening their sense of acting together as a unit. The intrinsic desire to share emotions, experiences and activities with others clearly separates humans from other primates. It appears very early in human ontogeny and enables children to participate in cooperative activities involving joint intentions and joint attention from early infancy onwards (Tomasello, Carpenter, Call, Behne, & Moll, 2005). Understanding music as a collectively intended activity — with specially designed features that satisfy this human desire to share emotions, experiences and activities with others — might explain why the children in our study felt a stronger commitment after joint music making and so spontaneously helped or cooperated with each other. Compared to language which is most effective in mobilizing joint intentionality for goal-directed behavior, music might be more efficacious in mobilizing joint intentionality per se — in the sense of feeling a “we” unit, thus getting people to experience each other as co-active, similar and cooperative members of a group (cf. Cross, 2007). The next step will be to design follow-up studies with additional conditions that disentangle the components of traditional music making — as they were integrated into the current manipulation phase — in order to test precisely which of the proximate mechanisms proposed above might cause the prosocial effects and in which particular ways.

4.3. Music as a biocultural phenomenon

Again granting that most, if not all, “technical” components of human musicality first arose as byproducts of other adaptations (like human speech, bipedalism or social cognition), we think that to explain the manifestation of music and dance as promoters for group cohesion one must posit some later stage of evolution involving *cultural group selection* (Brown, 2000a). Here, distinct cultural groups, rather than individuals, are the unit of selection, with normative behaviors, rather than genes, changing during the process of adaptation (Wilson & Sober, 1994). In the case of music, cultural groups that sang and danced together during certain rituals would have experienced a stronger group cohesion which could have led to many different types of group beneficial behaviors in competition with other groups. However, selection of behaviors at the level of the cultural group can, over many generations, lead to some innate predispositions that shape behaviors which may be individually costly — and so disadvantageous for within-group competition among individuals — but that are at the same time beneficial to the group in its competition with other groups (Gintis, Smith, & Bowles, 2001). Such *gene-culture co-evolution* (Richerson & Boyd, 2005) might have led not only to musical behaviors that are especially effective in creating the cohesive effect but also to an innate proclivity for exactly these kinds of behaviors (cf. Trehub & Hannon, 2006).

Contemporary humans are thus born with this innate proclivity for music, with the desire to learn the musical practices of the surrounding culture (Trehub, 2001) even if the latter largely abandoned traditional forms of music application. For example, already 5- to 24-month-old Swiss and Finnish infants move and adjust their body rhythmically to playback music and other periodic sound patterns (Zentner & Eerola, 2010). The inherently social nature of the human sense of rhythm is suggested by the finding that young German children's motivation to spontaneously synchronize to an external beat increases when that beat is created by another human in a social context (Kirschner & Tomasello, 2009). In the modern Western world, such an innate proclivity for music may be no longer as adaptive as in small-scale societies, but would explain the perpetual existence of song, music and dance — notably in social settings — even if practicing and specializing in certain techniques consume a lot of time and energy (e.g., learn a dance choreography, study an instrument, sing in a choir, play in an orchestra, go clubbing). Likewise, playing or listening to music on one's own may provide a kind of substitute feeling of sociality (Cross, 2007).

In conclusion, the best explanation for our results is that human children today have an innate proclivity to produce and to enjoy musical behaviors like the ones used in our study. This proclivity together with music's efficiency in coordinating voice and action — thereby highlighting the shared intention of acting together as a "we" unit — encouraged the children in our study to behave more cooperatively and pro-socially towards each other.

Acknowledgments

We thank G. Große, D. Mietchen, A.D. Patel, A. Schachner, F. Warneken, two anonymous reviewers and the editor D.M.T. Fessler for discussion and comments. We acknowledge the statistical advice of R. Mundry and the technical assistance from K. Greve and M. Sera. We thank the children who participated in the study and their families.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.evolhumbehav.2010.04.004.

References

Anshel, A., & Kipper, D. A. (1988). The influence of group singing on trust and cooperation. *Journal of Music Therapy, 3*, 145−155.

Arcadi, A., Robert, D., & Boesch, C. (1998). Buttress drumming by wild chimpanzees: Temporal patterning, phrase integration into loud calls, and preliminary evidence for individual distinctiveness. *Primates, 39*, 505−518.

Azmitia, M. (1988). Peer interaction and problem solving: When are two heads better than one? *Child Development, 59*, 87−96.

Blacking, J. (1976). *How Musical is Man?* London: Faber.

Bohlman, P. (2000). Ethnomusicology and music sociology. In D. Greer (Ed.), *Musicology and sister disciplines* (pp. 288−298). Oxford: Oxford University Press.

Brown, S. (2000a). Evolutionary models of music: From sexual selection to group selection. In F. Tonneau, & N. S. Thompson (Eds.), *Perspectives in ethology: 13. Behavior, evolution and culture* (pp. 231−281). New York: Plenum.

Brown, S. (2000b). The "musilanguage" model of music evolution. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 271−300). Cambridge, MA: The MIT Press.

Chartrand, T. L., & Bargh, J. A. (1999). The chameleon effect: The perception–behavior link and social interaction. *Journal of Personality and Social Psychology, 76*, 893−910, doi:10.1037/0022-3514.76.6.893.

Clayton, M. (2009). The social and personal functions of music in cross-cultural perspective. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 35−44). Oxford: Oxford University Press.

Conard, N. J., Malina, M., & Munzel, S. C. (2009). New flutes document the earliest musical tradition in southwestern Germany. *Nature, 460*, 737−740.

Cross, I. (2007). Music and cognitive evolution. In R. Dunbar, & L. Barrett (Eds.), *Handbook of evolutionary psychology* (pp. 649−667). Oxford: Oxford University Press.

Darwin, C. (1871). *The descent of man and selection in relation to sex.* London: John Murray.

Dissanayake, E. (2000). Antecedents of the temporal arts in early mother–infant interaction. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 389−410). Cambridge, MA: The MIT Press.

Dissanayake, E. (2006). Ritual and ritualization: Musical means of conveying and shaping emotion in humans and other animals. In S. Brown, & U. Voglsten (Eds.), *Music and manipulation: On the social uses and social control of music* (pp. 31−56). New York: Berghahn Books.

Fabes, R. A., Shepard, S., Guthrie, I., & Martin, C. L. (1997). The roles of temperamental arousal and same-sex play in children's social adjustment. *Developmental Psychology, 33*, 693−702.

Falk, D. (2004). Prelinguistic evolution in early hominins: Whence motherese? *Behavioral and Brain Sciences, 27*, 491−503.

Fernald, A. (1989). Intonation and communicative intent in mothers' speech to infants: Is the melody the message? *Child Development, 60*, 1497−1510.

Fitch, W. T. (2006). The biology and evolution of music: A comparative perspective. *Cognition, 100*, 173−215.

Fried, R., & Berkowitz, L. (1979). Music hath charms... and can influence helpfulness. *Journal of Applied Social Psychology, 9*, 199−208.

Gintis, H., Smith, E. A., & Bowles, S. (2001). Costly signaling and cooperation. *Journal of Theoretical Biology, 213*, 103−119.

Hagen, E., & Bryant, G. (2003). Music and dance as a coalition signaling system. *Human Nature, 14*, 21−51.

Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition, 27*, 949−961.

Huron, D. (2001). Is music an evolutionary adaptation? *Annals of the New York Academy of Sciences, 930*, 43−61.

James, W. (1890). *The principles of psychology*. New York.

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological Bulletin, 129*, 770−814.

Juslin, P. N., & Västfjäll, D. (2008). Emotional responses to music: The need to consider underlying mechanisms. *Behavioral and Brain Sciences, 31*, 559−575.

Kirschner, S., & Tomasello, M. (2009). Joint drumming: Social context facilitates synchronization in preschool children. *Journal of Experimental Child Psychology, 102*, 299−314.

Lakin, J. L., Jefferis, V. E., Cheng, C. M., & Chartrand, T. L. (2003). The chameleon effect as social glue: Evidence for the evolutionary significance of nonconscious mimicry. *Journal of Nonverbal Behavior, 27*, 145−162.

Maccoby, E. E. (2002). Gender and group process: A developmental perspective. *Current Directions in Psychological Science, 11*, 54−58.

MacDougall, H. G., & Moore, S. T. (2005). Marching to the beat of the same drummer: The spontaneous tempo of human locomotion. *Journal of Applied Physiology, 99*, 1164−1173.

McNeill, W. (1995). *Keeping together in time: Dance and drill in human history.* Cambridge, MA: Harvard University Press.

Meddis, R. (1984). *Statistics using ranks: A unified approach.* Oxford, UK: Blackwell.

Merker, B. (2000). Synchronous chorusing and human origins. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 315−327). Cambridge, MA: The MIT Press.

Merriam, A. (1964). *The anthropology of music.* Evanston: Northwestern University Press.

Miller, G. (2000). Evolution of human music through sexual selection. In N. L. Wallin, B. Merker, & S. Brown (Eds.), *The origins of music* (pp. 329−360). Cambridge, MA: The MIT Press.

Mithen, S. (2005). *The singing neanderthals: The origins of music, language, mind, and body.* London: Weidenfeld Nicolson.

Nettl, B. (2005). *The study of ethnomusicology: Thirty-one issues and concepts* (2nd ed.). Urbana-Champaign: University of Illinois Press.

North, A. C., Tarrant, M., & Hargreaves, D. J. (2004). The effects of music on helping behavior: A field study. *Environment and Behavior, 36*, 266−275.

Olson, K. R., & Spelke, E. S. (2008). Foundations of cooperation in young children. *Cognition, 108*, 222−231.

Pika, S., Liebal, K., & Tomasello, M. (2003). Gestural communication in young gorillas (Gorilla gorilla): Gestural repertoire, learning, and use. *American Journal of Primatology, 60*, 95−111.

Pinker, S. (1997). *How the mind works.* New York: Norton.

Richerson, P., & Boyd, R. (2005). *Not by genes alone: How culture transformed human evolution.* Chicago: University of Chicago Press.

Richman, B. (1993). On the evolution of speech: Singing as the middle term. *Current Anthropology, 34*, 721.

Roederer, J. G. (1984). The search for a survival value for music. *Music Perception, 1*, 350−356.

Sloboda, J. A., & Juslin, P. N. (2001). *Music and emotion: Theory and research.* Oxford, UK: Oxford University Press.

Stevens, C., & Byron, T. (2009). Universals in music processing. In S. Hallam, I. Cross, & M. Thaut (Eds.), *Oxford handbook of music psychology* (pp. 14−23). Oxford: Oxford University Press.

Thompson, W. F., Schellenberg, E. G., & Husain, G. (2001). Arousal, mood, and the Mozart effect. *Psychological Science, 12*, 248−251.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences, 28*, 675−691.

Trehub, S. E. (2001). Musical predispositions in infancy. *Annals of the New York Academy of Sciences, 930*, 1−16.

Trehub, S. E., & Hannon, E. E. (2006). Infant music perception: Domain-general or domain-specific mechanisms? *Cognition, 100*, 73−99.

van Baaren, R. B., Holland, R. W., Kawakami, K., & Knippenberg, A. v. (2004). Mimicry and prosocial behavior. *Psychological Science, 15*, 71−74.

Wallin, N. L., Merker, B., & Brown, S. E. (2000). *The origins of music.* Cambridge, MA: The MIT Press.

Wilson, D. S., & Sober, E. (1994). Re-introducing group selection to the human behavioral sciences. *Behavioral and Brain Sciences, 17*, 585−654.

Wiltermuth, S. S., & Heath, C. (2009). Synchrony and cooperation. *Psychological Science, 20*, 1−5.

Zentner, M., & Eerola, T. (2010). Rhythmic engagement with music in infancy. *Proceedings of the National Academy of Sciences, 107*, 5768−5773.