

How musicality changes moral consideration: People judge musical entities as more wrong to harm

Psychology of Music
1–21

© The Author(s) 2022

Article reuse guidelines:

sagepub.com/journals-permissionsDOI: [10.1177/03057356221096507](https://doi.org/10.1177/03057356221096507)journals.sagepub.com/home/pom

Tanushree Agrawal¹ , Joshua Rottman²
and Adena Schachner¹

Abstract

A growing literature shows that music increases prosocial behavior. Why does this occur? We propose a novel hypothesis, informed by moral psychology: evidence of others' musicality may promote prosociality by leading us to judge musical individuals as having enhanced moral standing. This effect may be largely indirect, by increasing perceptions of how intelligent and emotionally sensitive musical individuals are. If so, simply *knowing* about others' musicality should affect moral evaluations, such as wrongness to harm. Across four experiments (total $N=550$), we found supportive evidence. Information that an animal or person had the capacity and motivation to engage with music led participants to judge these entities as more wrong to harm than matched neutral or non-musical counterparts. Similarly, knowing that a person was not musical made people judge them as less wrong to harm than neutral or musical counterparts. As predicted, musicality was positively associated with perceptions of capacities for emotionality and intelligence, and these broader factors partially mediated the relationship between musicality and wrongness to harm. These effects were not influenced by participants' own musicality. Thus, non-moral attributes like musicality can impact moral consideration, carrying implications for social behavior and for interventions to promote prosociality.

Keywords

moral standing, music, emotionality, agency, intelligence

Beyond its capacity to provide enjoyment, music can promote prosocial behavior. Engaging in musical activities makes people feel more connected, more helpful, and less prejudiced (Clarke et al., 2015; Savage et al., 2020; Tarr et al., 2015). These positive effects of music on social

¹Department of Psychology, University of California San Diego, La Jolla, CA, USA

²Department of Psychology, Franklin & Marshall College, Lancaster, PA, USA

Corresponding author:

Tanushree Agrawal, Department of Psychology, University of California San Diego, 9500 Gilman Drive, M/C 0109, La Jolla, CA 92093-0109, USA.

Email: t2agrawa@ucsd.edu

behavior appear across ages (e.g., Kirschner & Tomasello, 2010), across diverse populations (e.g., Neto et al., 2016, 2019), and across a variety of musical activities (Hove & Risen, 2009; Kreutz, 2014; Vuoskoski et al., 2017). Several theories have been proposed to explain why music increases prosocial behavior. Shared musical experiences promote coordination and synchrony of movement, which have been argued to increase prosocial behavior (e.g., Cirelli et al., 2018). Moral philosophers have also proposed that musical engagement enhances empathy in musicians, dancers, and listeners (Ansani et al., 2019; Bloom, 2010), and that musicality is a character-building virtue (Bicknell, 2001; Cox & Levine, 2016; Kivy, 2009). These accounts predict that musical engagement should increase prosocial behavior when individuals personally engage in musical activities. Here, we propose that others' musical capacities may also affect our judgments of their moral standing, or worthiness of moral consideration (Goodwin, 2015). Our account makes the novel prediction that simply knowing about others' musicality should be enough to affect our social decisions, even without directly producing or listening to music.

Music and moral consideration

Why would others' musicality have consequences for social judgments? We hypothesize that evidence of musicality may lead people to make broader inferences about others' mental capacities, such as their capacities to *experience* physical and emotional sensations (joy, sadness, pain); and their *intelligent* or *agentive* capacities to think complex thoughts, make decisions, and produce controlled behaviors (Gray et al., 2007; Knobe & Prinz, 2008; Robbins & Jack, 2006; Sytsma & Machery, 2010). These two factors appear to be the primary components of attributions of mental life, or mind perception (Epley & Waytz, 2010; Gray et al., 2007; Kozak et al., 2006; Sytsma & Machery, 2012; but see Weisman et al., 2017).

These two factors also drive people's judgments about moral standing. People represent moral worth on a continuum, with some entities considered more wrong to harm than others (Crimston et al., 2016). This is often measured by asking people to judge which entity would be more painful for them to harm, if they were forced to do so (Goodwin, 2015; Gray et al., 2007). On this measure, people reliably rate it least wrong to harm plants, slightly wrong to harm fish, more wrong to harm apes, and most wrong to harm a human infant (Crimston et al., 2016; Gray et al., 2007). Across many different entities, people's beliefs about entities' capacities for experience and intelligence/agency explain over 90% of the variance in their judgments of the wrongness to harm these entities (e.g., Goodwin, 2015).

We hypothesize that people can use evidence of a person's or animal's musicality to infer their broader mental and emotional abilities, and that these factors will mediate an effect of musicality on wrongness to harm. An entity's engagement with music may be taken to imply a capacity for emotional experience, more broadly. Musical engagement typically involves a high degree of emotionality, and most people are aware of this (Juslin, 2013; Zentner et al., 2008). Producing music also requires agency, in that it involves intentional, controlled decisions and actions (Sloboda, 1988). This evidence of heightened capacities for experience and intelligence may lead observers to infer higher moral worth as a result.

The role of musicality in moral judgments has specific potential consequences for behavior toward both non-human animals and humans, both of which vary in their levels of musicality (Ayotte et al., 2002; Honing, 2019; Loui et al., 2017; Schachner et al., 2009). We do not expect musicality to be the only behavior that provides evidence of others' mental and emotional abilities. However, we expect that musicality is one of a small set of behaviors that provide this evidence easily, quickly, and convincingly; potentially by demonstrating others' ability to value activities or experiences for their own intrinsic worth (e.g., appreciation of beauty in art, poetry,

nature; Fayn et al., 2015; McCrae & Sutin, 2009). We return to this issue in the “General discussion” section. Here we aim to test the impact of musicality on moral judgments about both animals and humans, with the aim of specifically informing how evidence of musicality impacts social attitudes.

The current studies

To test the impact of musicality on moral judgments, we conducted a series of four experiments in which we manipulated entities’ musicality, and measured people’s judgments of the wrongness to harm these entities. We first asked: does simply knowing that another entity is musical make participants judge the entity as more wrong to harm? We tested whether evidence of musicality made participants judge animals as more wrong to harm, both when considering individual animals (Exp. 1) and when comparing animal species, a situation relevant to conservation efforts (Exp. 2).

In addition, we asked whether musicality impacts our judgments about other humans. In the first two experiments, we compared judgments across matched human characters, one described as highly musical, and one for whom musicality was not mentioned (Exp. 1–2). While this manipulation matches the amount of information provided, it leaves open the possibility that differences are due to negative effects of attributes of the control character, rather than positive effects of musicality. To isolate the effect of musicality, in the final two experiments we compared a highly musical character to a neutral baseline character, as well as to a character who lacks musicality. This allowed us to ask both whether high musicality increases wrongness to harm above baseline, and whether low musicality decreases wrongness to harm (Exp. 3–4).

Finally, we teased apart two theories regarding how and why musicality impacts moral standing (Exp. 4). We hypothesize that experimental manipulations of musicality impact judgments of characters’ experience and/or intelligence, and that these factors mediate the relationship between musicality and wrongness to harm. Alternatively, musical people may prefer other musical entities, due to the attractiveness of similarity (Reis, 2007). In this case, we would expect that highly musical people would be most likely to select musical agents as more wrong to harm, rather than musicality being equally taken into account by all of our participants.

Stimuli, data, methods, and supplemental material for all experiments can be found on OSF (https://osf.io/e6gqc/?view_only=ae69977e977f43b6a52c2221221add8). We pre-registered Experiments 3 and 4 (<https://aspredicted.org/pd2yq.pdf>; <https://aspredicted.org/b4rm7.pdf>).

Experiment I

To test whether people judge musical entities as more worthy of moral concern, we introduced participants to a set of characters previously established to vary in moral standing (e.g., a robot, a frog, a baby; adapted from Gray et al., 2007, see Figure 1). Across multiple trials, participants viewed all possible pairings of these characters, and for each pair, they selected which character would be more painful for them to harm, if they were forced to do so. This measure is commonly used as an indicator of wrongness to harm (e.g., Goodwin, 2015; Gray et al., 2007).

Within this set, we embedded our experimental stimuli: otherwise-matched pairs of musical and control humans and animals. This method was selected to minimize demand effects. We expected that if participants were able to guess the hypothesis or expected result, they might answer in the predicted manner to fulfill this perceived demand (Orne, 1962; Robson, 2011). By embedding our experimental manipulation in the context of this broader set of stimuli, we aimed to avoid revealing that musicality was the main factor of interest in the study (this was

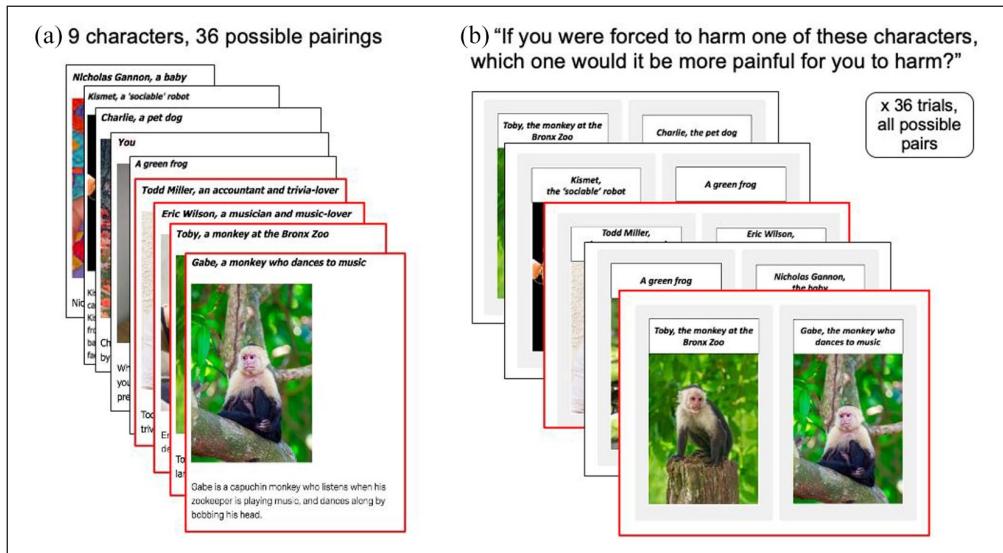


Figure 1. Method, Experiment 1.

Note: Participants were introduced to nine characters. Then, on each trial, participants saw two characters, and selected which would be more painful for them to harm (all possible pairings presented). The set of characters included two critical matched pairs (two monkeys and two humans), one of each described as having the capacity and motivation to engage with music, and the other described without providing information about musicality.

effective; see “Methods”). We predicted that participants would select both the musical animal, and musical human, as more painful to harm than their control counterparts.

We also measured participants’ own level of musical engagement, using the Active Engagement subscale of the GoldMSI (Müllensiefen et al., 2014). We asked whether highly musical participants were more likely to judge musical entities more painful to harm, as would be predicted if the effects were due to similarity attraction (Reis, 2007).

Method

Participants. Adults residing in the United States ($N = 100$) participated via Amazon Mechanical Turk (31% female; mean age = 34.79, $SD = 10.71$). All were over age 18, had 93% or higher approval rate of prior work on the MTurk platform, and received \$1.50 for participation. Five additional participants were tested but excluded based on two exclusion criteria determined *a priori*: choosing a pebble as more wrong to harm than an adult human ($n = 5$); or failing when asked to select the number “two” ($n = 0$).

Procedure and stimuli. Subjects were introduced to nine characters in random order, via a short vignette and a photographic image for each. Five characters were from Gray et al. (2007): a frog, pet dog, human baby, a sociable robot, and “You” (participants were shown an image of a mirror, and asked to consider themselves; images provided by K. Gray). The additional four characters, two monkeys, and two men, provided the manipulation of interest (musical vs control, see Table 1). One of each pair was described as having the capacity and motivation to engage with music, and the other was described without providing information about musicality. Descriptions were matched for length and wording, and information extraneous to our

Table I. Descriptions of Critical Characters, Exp. I.

	Musical character	Control character
Monkeys	"Gabe is a capuchin monkey who listens when his zookeeper is playing music, and dances along by bobbing his head."	"Toby, a six-year-old capuchin monkey, lives in a large enclosure at the Bronx Zoo in New York City."
Humans	"Eric Wilson is a thirty-year-old musician and devoted music-lover living in New York."	"Todd Miller, 32, is an accountant and a devoted trivia-player who lives in Chicago."

manipulation (which image, name, age, or location was assigned to the musical vs control character) was counterbalanced across participants. The musical animal was compared with a neutral control character. We aimed to include a similar amount of information for the matched control characters, to avoid possible effects of providing different levels of detail. Thus, for the human characters, both the musical and control individuals were described as having an age, profession, hobby, and city of residence. (For comparison to a neutral baseline human character, see Exps. 3 and 4.)

Each possible pairing of characters was presented separately (with images and short descriptive phrases), such that each participant made 36 randomly ordered, pairwise, forced-choice comparisons across all nine characters. Left-/right positions were randomized. For each pair of characters, participants were asked: "If you were forced to harm one of these characters, which one would it be more painful for you to harm?." Participants then completed the two attention checks, and were asked what they thought the study was about. Next, participants completed the Active Engagement subscale of the GoldMSI (Müllensiefen et al., 2014), nine questions regarding how frequently they listen to music and do music-related activities in their everyday lives (Cronbach's $\alpha = .87$). Last, participants filled out a series of standard demographic questions.

Participants remained unaware of hypotheses. Participants' answers regarding what the study was about were coded by two independent raters (unaware of condition and other participant responses) for mentions of music or the monkey/human character pairs. No participants mentioned any of these in their free responses (inter-rater reliability = 100%), suggesting that they did not guess that musicality was relevant to the purpose of the study.

Results & discussion

We first calculated the total number of times that each character was selected as more painful for the participant to harm, out of the eight comparisons per character. For characters from prior work (Gray et al., 2007), the average ranking of characters matched previously established judgments (see Figure 2[a]; note that as in Gray et al. (2007), the dog was described as the pet of a particular family; this factor, as well as cuteness (Sherman & Haidt, 2011) may partially explain the dog's high wrongness-to-harm ranking, as harming the dog would also violate human ownership rights). We then examined trials that pitted the musical versus control characters against one another (see Figure 2[b]). Seventy-five percent of participants selected the musical monkey as more wrong to harm than the control monkey (75/100, $p < .001$, binomial test). Sixty percent of participants selected the musical human as more wrong to harm than the control human; this proportion was not significantly different than chance (60/100, $p = .057$; binomial test). Post-hoc power analyses revealed that a larger

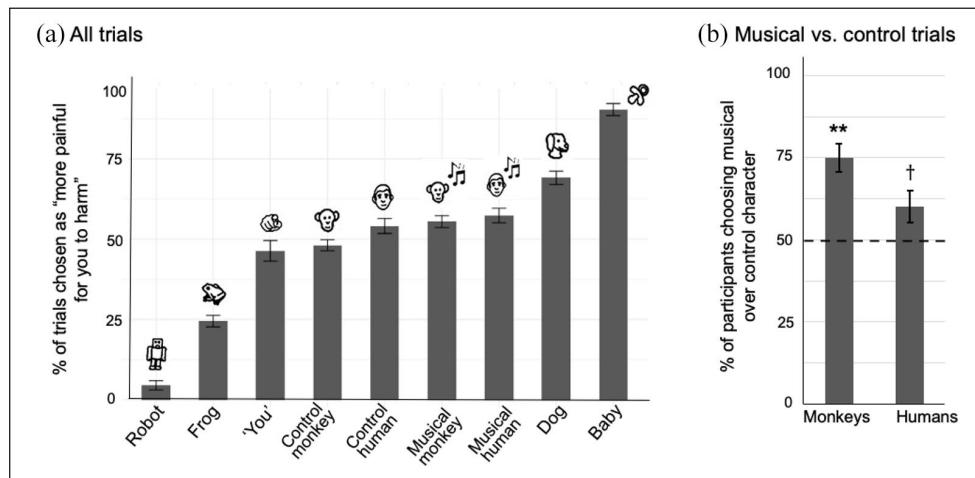


Figure 2. Results, Experiment 1.

Note: (a) Participants made nuanced judgments of wrongness to harm across the nine entities. The Y-axis shows the percentage of times that each character was selected as being more painful for the participant to harm, versus each of the other characters. Error bars are standard errors. Icons from Gray et al. (2007). (b) On trials that pitted the musical versus control characters against one another, participants chose the musical monkeys as more wrong to harm; this effect trended in the same direction for humans.

sample size was required to reliably detect a true effect of this size (96% power to detect the effect size seen in monkeys, only 29% power to detect the smaller effect size seen in humans). We thus sought to replicate this effect with a larger sample in Exp. 2.

For the human characters, both were described as having a particular career (musician, accountant) and favorite activity (music, trivia games). Thus, participants' judgments of humans could indicate either a positive effect of musicality, or a negative effect of the characteristics of the non-musical character. We address this concern in Experiments 3 and 4. Lastly, we asked if only highly musical people chose musical entities as more wrong to harm. Participants' own musicality scores were not related to their judgments of whether musical characters were more wrong to harm, in either the case of monkeys, Wald Z(98) = 0.84, $p = .40$, or humans, Wald Z(98) = 1.59, $p = .11$ (nested logistic model comparisons).

Experiment 2

Exp. 1 suggests that musicality influences moral judgments for animal entities, and perhaps for humans as well. In a second experiment, we had two aims. First, we asked whether our finding would generalize to species-level comparisons. We were motivated by potential implications for animal conservation. When comparing the value of species, do people judge an animal from a musical species as more wrong to harm, even if others of that species share its musical qualities (making it less unique)? Thus, we introduce two animal species, one musical and one a neutral control. We then ask participants to judge wrongness to harm an individual from each species. We predicted that participants would judge individuals from a musical species of monkey more painful for them to harm than those from a neutral matched species.

In addition, with regard to the human entities, we asked whether participants' tendencies to select the musical human over the control human in Exp. 1 were reliable. In line with post-hoc

Table 2. Descriptions of Monkey Species, Exp. 2.

Musical monkey species	Control monkey species
"Calusan monkeys are native to Central and South America. Calusan monkeys are known to respond to music, by dancing along and bobbing their heads."	"Sequel monkeys live in tropical forests in the Amazon. Sequel monkeys are brown, grey, or white, with lighter fur around their face and eyes."

power analyses from Exp. 1, we increased the sample size to allow us to reliably detect true effects of this size; that is, musical alternatives chosen as more painful to harm 75% of the time for monkeys, and 60% of the time for humans, compared with a 50% probability by chance. Other aspects of the method remained the same as Exp. 1.

Method

Participants. Adults residing in the United States ($N = 150$) participated via Amazon Mechanical Turk (61 female; mean age = 35.32, $SD = 10.27$) using the same criteria and payment as in Exp. 1; none had participated in Exp. 1. Sample size was decided *a priori* based on power analyses from Exp. 1. Six additional participants were tested but excluded for failing one or more exclusion criteria determined *a priori* (Pebble vs man, $n = 4$; select "two," $n = 2$). Another 19 workers were excluded for being poor English speakers or bots (Dennis et al., 2018; TurkPrime, n.d.; see Procedure).

Procedure and stimuli. The procedure and stimuli were identical to those of Exp. 1, with two changes. First, participants were introduced to two novel species of monkey rather than individual monkeys (Table 2). At test, a single individual from each species was considered (Musical: "A calusan monkey, known to dance to music"; Control: "A sequel monkey, known to live in the Amazon"). Second, we added a free response question ("Please provide a summary of what you did in this study"), to detect poor English speakers or bots. Two independent coders (both unaware of participants' answers on the main dependent measures) analyzed descriptions from this question, as well as the question "Please tell us what you think this study was about," flagging any suspicious or nonsensical responses. Participants flagged by both coders were excluded. Coders agreed for 94.94% of participants. For cases of disagreement, the coders discussed and came to an agreement.

Participants remained unaware of hypotheses. Participants' answers regarding what the study was about were coded in the same way as Exp. 1. Again, no participants mentioned music or the monkey/human character pairs (IRR = 100%).

Results and discussion

As in Exp. 1, we replicated previously established rankings of wrongness-to-harm among the characters (see Figure 3[a]). On trials that pitted the musical vs control characters against one another (see Figure 3[b]), participants chose the monkey from a musical species as more painful to harm than the monkey from the control species (65%, or 98 of 150 participants; $p < .001$; binomial test). People also chose the musical human as more painful to harm than the control human (63%, or 94 of 150 participants; $p = .002$; binomial test). Participants' standardized

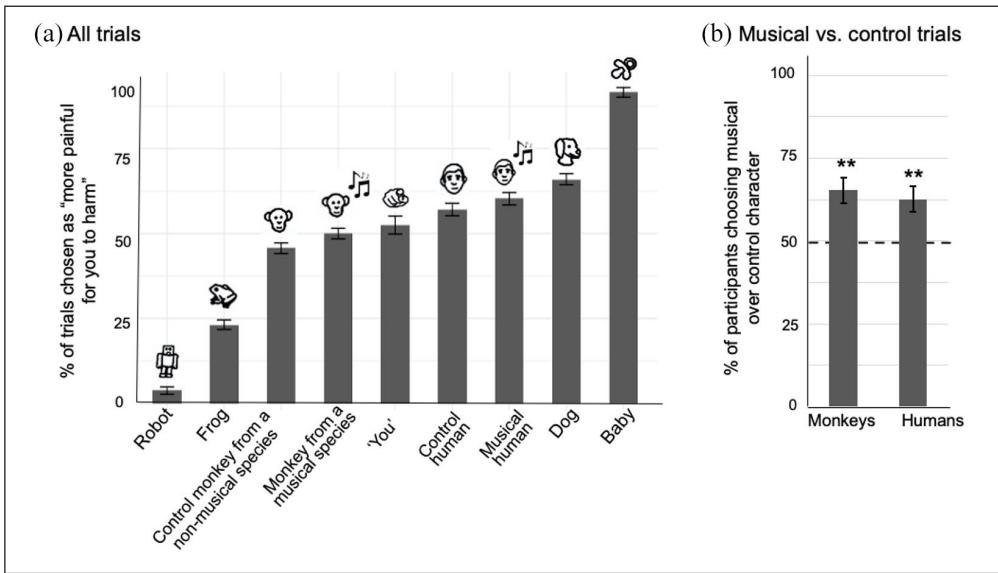


Figure 3. Results, Experiment 2.

Note: Participants chose the musical entities as more painful for them to harm. (a) The percentage of times each character was selected as being more painful for them to harm (vs each of the other characters). Error bars indicate standard errors. Icons from Gray et al. (2007). (b) On trials that directly pitted the musical versus control monkey species, or the musical versus control human individuals, participants chose the musical entities as more painful for them to harm.

musicality scores were again not related to whether they selected musical entities as being more wrong to harm, for monkeys, $\text{Wald } Z(148) = -0.41$, $p = .68$, or humans, $\text{Wald } Z(148) = -0.19$, $p = .85$ (nested logistic model comparisons).

Experiment 3

In both of the previous experiments, we found that musicality increased wrongness to harm. For animals (in particular, monkeys), we found that when compared with a neutral control character, musical individual monkeys were judged more painful to harm than matched control individuals (Exp. 1). When an entire animal species was described as musical, participants again chose an animal from the musical species as more painful for them to harm than a neutral control (Exp. 2). In both samples, we found that the tendency to prefer musical entities was not predicted by participants' own musicality.

The interpretation of human findings from Experiments 1 and 2 is more difficult. To match the information content of our two human characters, we described both as having a particular career (musician, accountant) and favorite activity (music, trivia games). Thus, participants' preferences on human comparisons could show a negative effect of characteristics of the non-musical character, rather than a positive effect of musicality.

Here we address this concern. We also ask a novel question: does lack of musicality decrease wrongness to harm, in addition to high musicality increasing it? We accomplish this by distinguishing a non-musical human individual and a neutral baseline human individual (for whom no information is provided about musicality), and thus comparing three levels of musicality: a highly musical, neutral, and explicitly non-musical human individual. If musicality impacts

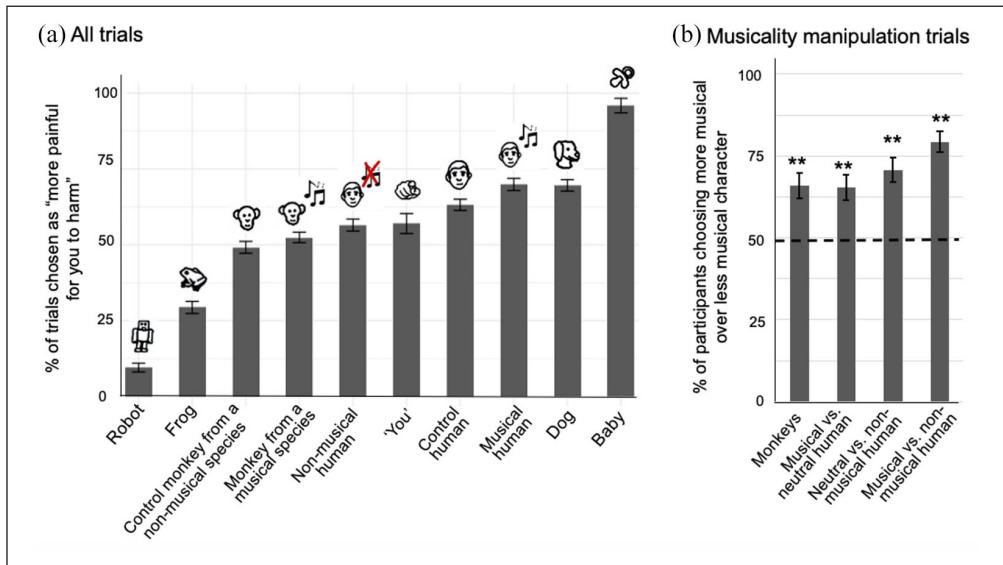


Figure 4. Results, Experiment 3.

Note: Wrongness to harm tracked with experimentally manipulated levels of musicality. (a) The percentage of times each character was selected as being more painful for the participant to harm, versus each of the other characters. Error bars indicate standard errors. (b) On trials that pitted the musical versus control monkey species, or the musical, neutral, and non-musical human individuals, participants consistently chose the entity with higher musicality as more wrong to harm.

judgments of wrongness to harm, then we should see that wrongness to harm tracks with the amount of musicality. In contrast, if the human results from the previous experiments were driven primarily by attributes of the control character, then we should not find differences for these new stimuli.

The addition of a non-musical character also allows us to test an alternative account of our findings. Under this alternative account, rather than valuing musical individuals, participants simply value entities that are unique, unusual, or surprising. The inclusion of a non-musical human character allows us to test this idea, since people who lack enjoyment of music are relatively rare. Musical anhedonia, the inability to enjoy music, affects only 3–5% of the population (Stewart, 2014). If participants are simply selecting unique or unusual individuals as more wrong to harm, then they should select the non-musical human character as more wrong to harm than the neutral or musical characters. In contrast, if levels of musicality are driving the differences, then participants should judge the non-musical person as the least wrong to harm among the human individuals.

We preregistered all stimulus methods and analyses. Our preregistration is available here: <https://aspredicted.org/pd2yq.pdf>.

Method

Participants. As preregistered, 150 US adults participated via Amazon Mechanical Turk (62 female; mean age = 36.25, $SD = 11.88$) using the same criteria and payment as in Exps. 1 and 2; none of them had participated in any of the previous experiments. Eighty-three additional participants were tested but excluded for failing preregistered exclusion criteria (Pebble vs

Table 3. Descriptions of Human Characters, Exp. 3.

Musical human	Neutral human	Non-musical human
"Eric Wilson is a 34 year-old professional musician, who spends his free time listening to music and following his favorite artists."	"Todd Miller is a 35 year-old professional, who works at a full time job during the week, and lives in Chicago, Illinois."	"Matt Johnson, 36, has never enjoyed listening to music, and he doesn't understand why others seem to find music so appealing."

human, $n=21$; select "two," $n=2$), being poor English speakers or bots ($n=58$, same detection method as in previous experiments), or having technical difficulties ($n=2$; reported that images and vignettes did not load).

Procedure and stimuli. As preregistered, the procedure and stimuli were identical to those of prior experiments, with one change. Participants were introduced to three human individuals, instead of two (Table 3): one musical character, one neutral character, and one non-musical character. Images, names, and other non-musical aspects (including characters' ages) of the descriptions were counterbalanced across participants. The addition of a 10th character (the third human character) resulted in a total of 45 pairwise comparisons across all 10 characters.

Participants remained unaware of hypotheses. Participants' answers regarding what the study was about were coded in the same way as previous experiments. Three of the 150 participants mentioned music or the critical character pairs (IRR = 100%). All findings remain the same whether these participants' data are included or excluded, with no differences in conclusions or statistical significance. As preregistered, results reported include these data points.

Results

We again replicated the previously established rankings of wrongness to harm (see Figure 4[a]). Participants judged the human characters to be more wrong to harm as their musicality increased (see Figure 4[b]). They chose the musical over the neutral human (65%, 98/150 participants; $p<.001$; binomial test); the musical over the nonmusical human (79%, 119/150; $p<.001$); and the neutral over the non-musical human (71%; 106/150; $p<.001$). Participants also chose the animal from the musical species as more wrong to harm than the matched non-musical animal (66%, 99/150; $p<.001$). Participants' musicality scores were again not related to whether they selected musical entities as being more wrong to harm, for monkeys, Wald Z(148)=−1.12, $p=.26$, or humans, musical versus neutral: Wald Z(148)=1.48, $p=.14$; musical versus non-musical: Wald Z(148)=−0.37, $p=.71$; neutral versus non-musical: Wald Z(148)=−1.37, $p=.17$ (nested logistic model comparisons).

Discussion

Across three experiments, we find that people judge it more wrong to harm musical entities than less musical entities. Greater musicality led to participants to judge it more painful for them to harm individual animals (Exp. 1), individual humans (Exps. 1–3), and animals of species that differed in musicality (Exp. 2). For humans, we find both a positive effect of musicality, in increasing wrongness to harm above a neutral baseline, and a negative effect of lack of

musicality, in decreasing wrongness to harm below a neutral baseline (Exp. 3). These effects are not simply due to participants favoring characters with unique or surprising attributes; if this were the case, they should have chosen the low musicality person as more wrong to harm, as a lack of enjoyment of music is an unusual trait (Stewart, 2014). In addition, these effects are not simply due to high musicality participants favoring characters that were more similar to themselves; instead, we find that participants' level of musicality did not predict their tendency to choose musical entities as more wrong to harm.

Experiment 4

Why does musicality impact moral decisions? We hypothesize that people use evidence of a person's or animal's musicality to infer their broader mental and emotional abilities, and that these general psychological capacities mediate the effect of musicality on wrongness to harm. Evidence of musicality may lead people to infer others' capacities for rich physical and emotional *experiences* (joy, sadness, pain); and their capacities to think complex thoughts, make decisions, and produce controlled behavior, termed *intelligence* or *agency*. These two factors appear to be the primary components of attributions of mental life (e.g., Gray et al., 2007) and also drive people's judgments about moral standing (e.g., Goodwin, 2015).

To test our hypothesis, we asked participants to again compare each possible pair of entities, but to additionally judge characters' relative intelligence; capacity for emotional experience; and wrongness to harm (measured as in previous experiments), in a within-subject design. Participants also judged characters' relative musicality. This allowed us to examine the relationship of musicality judgments to harm, intelligence, and/or experience judgments across the entire set of 10 characters, not just the characters involved in our experimental manipulations. Second, measuring musicality allowed us to examine how participants make spontaneous inferences about the musicality of other entities, even when information about musicality is not explicitly provided (e.g., frogs, robots).

We measured only the capacity for emotional experience, and not the capacity for physical experience, because in a between-subjects version of this experiment (available in our OSF repository), we had previously found that these two factors were almost perfectly correlated (Pearson's $r = .96$). Capacity for physical experience also had a weaker relationship to musicality than capacity for emotional experience in this prior experiment. The regression results of the between-subject version of this experiment are in line with the findings of the within-subject version reported here (see OSF for more detail).

We preregistered all stimulus methods and analyses. Our preregistration is available here: <https://aspredicted.org/b4rm7.pdf>.

Method

Participants. Undergraduates at a large public university in Southern California ($N = 150$) participated in an online experiment in exchange for course credit (100 female; mean age = 20.61, $SD = 2.29$). Six additional participants were tested but excluded for failing one or more preregistered exclusion criteria (pebble vs human on any dependent measure, or select "two," $n = 5$), or technical issues ($n = 1$; participant noted that images did not load).

Procedure and stimuli. As preregistered, participants were first introduced to the same characters as in Exp. 3. They then answered two-alternative forced choice questions for all possible pairs of characters in four blocks, with each block involving a different question (e.g., wrongness to

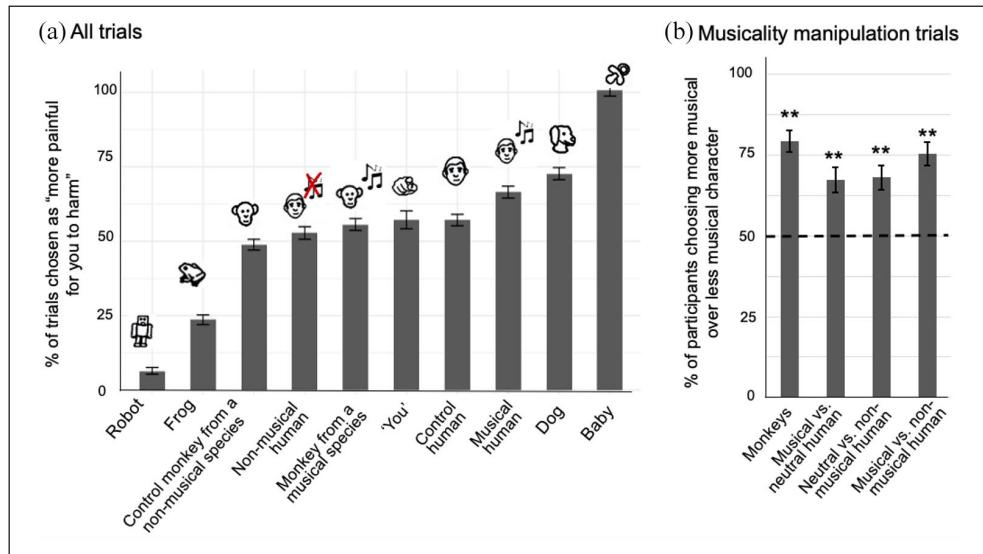


Figure 5. Wrongness to Harm Results, Experiment 4.

Note: (a) The percentage of times each character was selected as being more painful for the participant to harm, versus each of the other characters. Error bars indicate standard errors. (b) Wrongness to harm tracked with experimentally manipulated level of musicality. On trials that pitted the musical versus neutral monkey species, or the musical, neutral, or non-musical human individuals, participants chose the entity with more musicality as more painful for them to harm.

harm) and one complete set of pairwise comparisons. Participants always completed the wrongness to harm block first, and the remaining blocks in random order.

In the *wrongness to harm* block, participants answered the same harm question as in previous experiments. In the *musicality* block, participants were instructed to think of musicality as the extent to which the characters “possess the natural ability to perceive and enjoy music” (Honig, 2019), and were asked “Which one of these characters do you think is more musical?” In the *emotional experience* block, participants were asked “Which one of these characters is more emotionally sensitive?”; in the *intelligence* block, they were asked “Which one of these characters do you think is more intelligent?” (both measures were taken from Piazza et al., 2014).

Most participants remained unaware of hypotheses. A larger proportion of participants (23.3%, $n = 35$) mentioned music when asked to guess the aim of the experiment ($IRR = 100\%$). This was expected, since participants were asked to judge all characters’ musicality in this experiment. All analyses were conducted with and without these 35 data points, and findings were the same in both cases, with no differences in conclusions or statistical significance. As preregistered, results reported here include these data points.

Results

We again replicated previously established rankings of wrongness-to-harm among all characters (see Figure 5[a]). In addition, participants made judgments about rankings of intelligence and emotionality between the characters that replicated previous work (Gray et al., 2007; see Figures 7[a] and 8[a]). Participants again judged it to be more painful for them to harm an animal from a musical species (musical vs neutral monkeys: 119/150 or 79.3%, $p < .001$), and

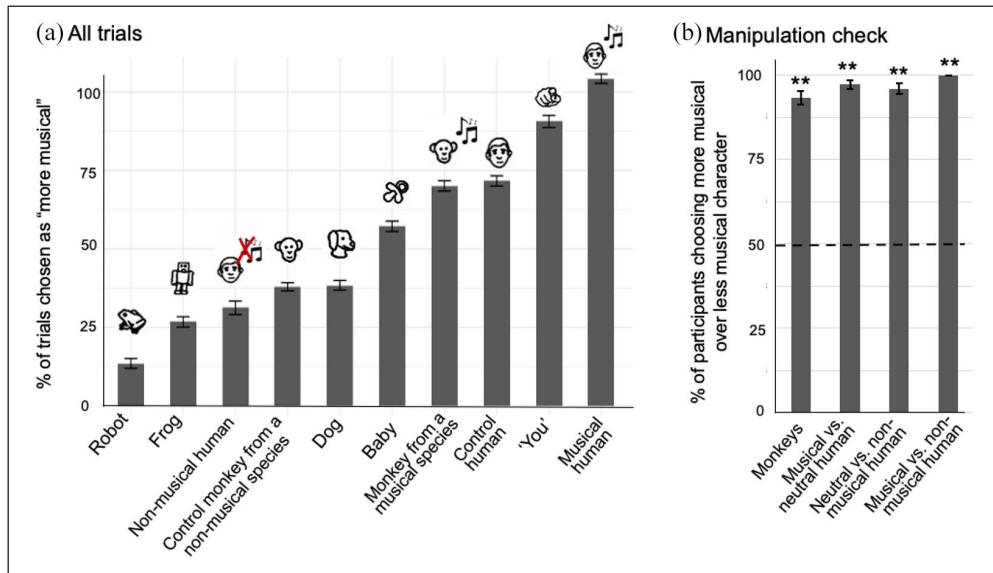


Figure 6. Musicality Results, Experiment 4.

Note: (a) People made sensible, consistent judgments about characters' levels of musicality, even when no explicit information was provided about musicality. The percentage of times each character was selected as being more musical, versus each of the other characters. Error bars indicate standard errors. (b) As expected, on trials where musicality was experimentally manipulated (musical vs control monkey species, or the musical, neutral, and, non-musical human individual characters), participants chose the more musical character as being more musical (providing a manipulation check).

again judged the human characters to be more wrong to harm as their musicality increased (musical over neutral human: 101/150, or 67.33%; $p < .001$; musical over non-musical human: 113/150, or 75.3%; $p < .001$; Neutral over Non-musical Human: 102/150, or 68%; $p < .001$; binomial tests, see Figure 5[b]).

Participants' own levels of musical engagement did not predict their tendencies to make most of these judgments (musical vs neutral monkeys: Wald Z(148) = -0.70, $p = .48$; musical vs non-musical man: Wald Z(148) = 0.96, $p = .34$; neutral vs non-musical man: Wald Z(148) = 0.02, $p = .98$; nested logistic model comparisons). When comparing whether it was more painful to harm the musical or the neutral man, participants' own levels of musical engagement were a significant predictor (musical vs neutral man: Wald Z(148) = 2.41, $p = .02$). We suspect that this effect is not robust, as this pattern did not appear in any of the previous studies.

Musicality ratings. Participants made sensible, consistent judgments about the musicality of other entities, even for entities with no explicitly stated musical information. Across the characters where musicality was not manipulated, the frog was ranked as least musical, followed by the robot, the dog, and then the baby (see Figure 6[a]).

As expected, for characters where musicality was manipulated, participants almost unanimously judged the more musical characters as being more musical than their less-musical counterparts (musical vs neutral monkeys: 140/150, or 93%; binomial test, $p < .001$; musical vs neutral human: 146/150, or 97.3%; $p < .001$; musical vs non-musical human: 150/150, or 100%, $p < .001$; neutral vs non-musical human: 144/150, or 96%, $p < .001$, see Figure 6[b]).

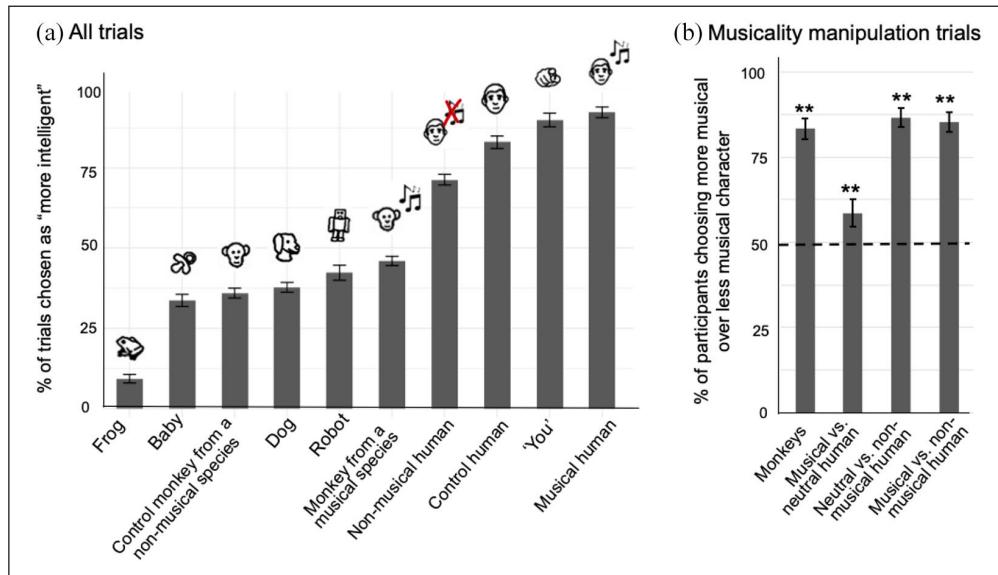


Figure 7. Intelligence Results, Experiment 4.

Note: (a) The percentage of times each character was selected as being more intelligent, versus each of the other characters. Error bars indicate standard errors. (b) Intelligence largely tracked with experimentally manipulated levels of musicality. On trials that pitted the musical versus neutral monkey species, or the musical, neutral, and non-musical human individual characters, participants chose the entity with more musicality as being more intelligent.

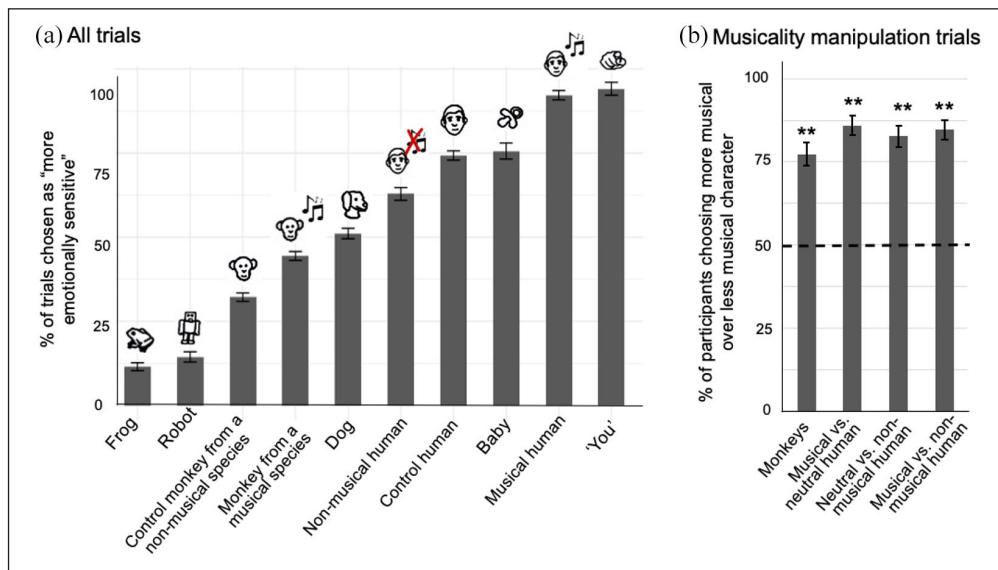


Figure 8. Capacity for Emotional Experience Results, Experiment 4.

Note: (a) The percentage of times each character was selected as being more emotionally sensitive, versus each of the other characters. Error bars indicate standard errors. (b) Emotional sensitivity tracked with experimentally manipulated levels of musicality. On trials that pitted the musical versus neutral monkey species, or the musical, neutral, and non-musical human individual characters, participants chose the entity with more musicality as being more emotionally sensitive.

Table 4. Mixed-Effects Logistic Regression Results, Experiment 4.

Fixed effects	Estimate	Std. error	Wald Z	p	df	AIC
<i>Simple logistic model of wrongness to harm</i>					5	5,655.9
(Intercept)	-1.75	0.31	-5.56	<.001		
Emo.	1.35	0.10	13.33	<.001		
experience						
Intelligence	0.65	0.11	6.21	<.001		
<i>Full logistic model (incl. musicality)</i>					6	5,636.2
(Intercept)	-2.04	0.31	-6.50	<.001		
Emo.	1.28	0.10	12.57	<.001		
experience						
Intelligence	0.58	0.11	5.43	<.001		
Musicality	0.61	0.11	5.69	<.001		

Effect of musicality manipulation on experience and intelligence ratings. On trials where musicality was experimentally manipulated, the monkey from the musical species was chosen as more emotionally sensitive than the control (116/150, or 76.7%; $p < .001$), and the human characters were judged more emotionally sensitive as their musicality increased (musical vs neutral human, 129/150, or 86%; $p < .001$; neutral vs non-musical human, 124/150, or 82.7%; $p < .001$; musical vs non-musical human, 127/150, or 84.7%; $p < .001$; binomial tests, see Figure 8[b]).

Similar judgments were found for intelligence. The monkey from the musical species was chosen as more intelligent than the control (125/150, or 83.3%, $p < .001$; binomial tests), and the human characters were judged to be more intelligent as their musicality increased (musical vs neutral human: 88/150, or 58.7%; $p = .04$; musical vs non-musical human, 128/150, or 85.3%; $p < .001$; neutral vs non-musical human, 130/151, or 86.7%, $p < .001$, see Figure 7[b]).

Participants' own levels of musicality did not predict their tendencies to select the more musical character as having higher emotionality or intelligence (more musical vs less musical matched character; emotional sensitivity: Wald Z(598) = -0.38, $p = .71$; Intelligence: Z(598) = 1.21, $p = .23$).

Relationship between wrongness to harm and other factors. Using a mixed-effects logistic regression, we asked whether entities' intelligence, capacity for emotional experience, and musicality predicted wrongness to harm judgments (with these factors as fixed effects; subject and trial number were entered as random effects). Intelligence and the capacity for emotional experience were both significant predictors (Table 4). Notably, musicality made an independent contribution to predicting wrongness to harm, beyond the other factors (model 1: harm predicted by intelligence, emotionality [fixed effects], subject, trial number [random effects]; model 2: the same model with musicality added [as a fixed effect; $\chi^2 = 31.8$, $p < .001$; nested logistic mixed-effect model comparison]).

To determine whether experience and intelligence mediate the effect of musicality on wrongness to harm, we conducted a multi-level mediation analysis (as in Bauer et al., 2006), the paths and coefficients of which are visualized in Figure 9. We found that intelligence and experience partially mediated the effect of musicality on wrongness to harm. The total effect of musicality on wrongness to harm was significant (direct path: $\beta = 0.86$, $p < .001$). The effect of musicality on capacity for experience ($\beta = 1.23$, $p < .001$) and intelligence ($\beta = 1.25$, $p < .001$) were also significant, as were the effects of experience and intelligence on wrongness to harm

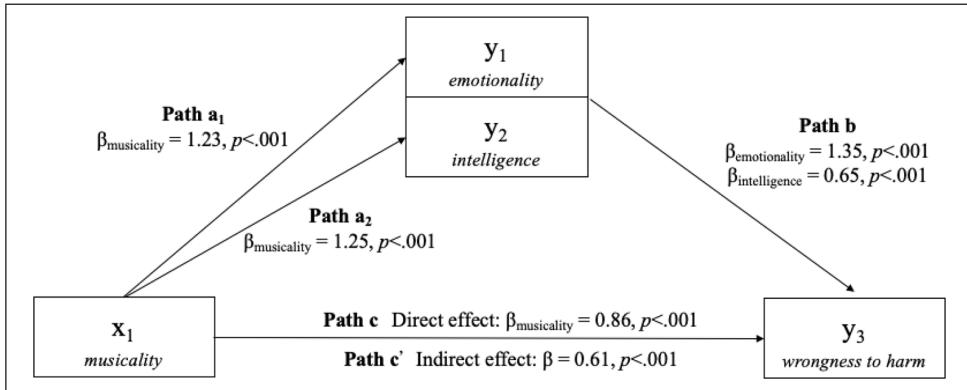


Figure 9. Mediation Analysis, Experiment 4.

Note: The effect of musicality on wrongness to harm is partially mediated by perceived traits such as emotionality and intelligence, as evidenced by the reduced (but statistically significant) coefficient for the indirect path versus the direct path between musicality and wrongness to harm.

($\beta_{\text{emotionality}} = 1.35, p < .001$; $\beta_{\text{intelligence}} = 0.65, p < .001$). When perceived intelligence and emotionality were controlled for, the direct effect of musicality on wrongness to harm remained significant, but at a smaller effect size (indirect path: $\beta = 0.61, p < .001$).

Discussion

Our findings here provide an explanation of why musicality shifts moral decisions. People use other entities' musicality to infer their broader underlying abilities, including capacity for experience and intelligence. Thus, in line with our hypothesis, musicality appears to provide evidence of mental life, in terms of intelligence and capacity for experience. These factors link seemingly amoral musical activities to the moral domain.

We also found that musicality made an independent contribution to judgments of wrongness to harm, such that adding musicality as a predictor explained participants' harm judgments better than attributed capacities for experience and intelligence alone. Thus, mediation may not fully explain the effects of musicality on moral judgments. People appear to intrinsically value musicality, in a way not explained by experience and intelligence alone. In prior work, while experience and intelligence explain most of the variance in moral harm judgments, additional factors such as harmfulness have been found to make independent contributions (Piazza et al., 2014). These findings are consistent with the broader idea that moral standing is not driven by only two factors (as in the two-source model; Sytsma & Machery, 2012); instead, judgments also appear to depend on other factors, such as harmfulness and musicality.

However, there is reason to believe that the unique predictive power of musicality in our dataset may be due to limitations of the measurement instruments. This is due to a general statistical truth: noisy measurements cannot fully control for the impact of the constructs they represent (Westfall & Yarkoni, 2016). Because, like all subjective measures, measurements of experience and intelligence are noisy, it may be possible for another variable to improve prediction above these measurements, even if the latent construct of musicality does not actually add unique predictive value above experience and agency, if all constructs were measured perfectly (Westfall & Yarkoni, 2016). As such we interpret findings of musicality's unique predictive

power with caution. Primarily, the current data show that musicality impacts judgments of capacity for experience and agency, which mediate impacts on wrongness to harm.

General discussion

Across four experiments, we found that simply knowing that certain entities are capable of engaging with music leads participants to judge them as being more wrong to harm. When musicality was experimentally manipulated, participants judged it to be more painful for them to harm musical rather than control animal individuals (Exp. 1) and members of a musical versus control animal species (Exps. 2–4). This effect extended to judgments about humans. Participants judged it to be more painful for them to harm a highly musical individual, over a neutral baseline or control individual; and less painful to harm an individual with low musicality, below a neutral baseline (Exps. 2–4).

A major question in music psychology is why music promotes prosocial behavior, and when it can be expected to do so. Existing theoretical frameworks across both music psychology and moral psychology predict that first-person engagement with music should increase prosociality, through personally listening, playing, or dancing to music (e.g., Clarke et al., 2015). We find that third-person observations, even written evidence of others' musicality, affect evaluations about wrongness to harm. These findings broaden the range of contexts in which music may be expected to impact social and moral behaviors.

Our findings further provide an explanation of why musicality shifts moral decisions: musicality provides broader evidence of inner mental life, including the capacity for emotional experience and intelligence (Exp. 4). Inferences about experience and intelligence act as mediators, linking evidence of musicality to increased wrongness to harm. Our data also speak against two alternative accounts. We find no evidence that similarity attraction explains these effects (Reis, 2007). Participants' own level of musicality did not predict their tendencies to choose musical entities as more wrong to harm. Our data also cannot be parsimoniously explained by the idea that participants simply choose the more unusual, unique or surprising character as more wrong to harm. Non-musical people are relatively rare (Stewart, 2014), yet we find they are judged least wrong to harm among human characters (Exps. 3 and 4).

These findings suggest that beliefs in animal musicality may aid animal conservation efforts. Anecdotally, the movement to save the whales is believed to have gained momentum from whalesong research, with scientists taking evidence of musicality as evidence of wrongness to harm (e.g., biologist Roger Payne: "Do you make cat food out of composer-poets? I think that's a crime"; May, 2014). This sentiment was echoed by laypeople and taken up as a banner issue by Greenpeace (May, 2014). The current findings provide a framework for explaining this phenomenon in terms of moral psychology.

Within the human population, people differ in their levels of musicality, with some people congenitally lacking the ability to perceive music (congenital amusia; Ayotte et al., 2002) or the ability to enjoy music (musical anhedonia; Loui et al., 2017; Mas-Herrero et al., 2014). In contrast, other clinical conditions like Williams syndrome are linked with high musicality (Levitin et al., 2004; Ng et al., 2013). Our findings suggest that people may judge it less wrong to harm individuals with amusia or musical anhedonia, and may be particularly compassionate toward individuals with conditions like Williams. Understanding these biases has the potential to inform our understanding of attitudes toward members of these clinical populations.

Is musicality unique in its impact on social and moral judgments? We do not expect musicality to be the only behavior that provides evidence of others' mental and emotional abilities,

or moral worth. However, we theorize that musicality is one of a small set of behaviors that provide this evidence easily, quickly, and convincingly. In ongoing work, we are testing the hypothesis that behaviors that demonstrate others' ability to value activities or experiences for their own intrinsic worth (such as aesthetic judgments, e.g., appreciation of beauty in art, poetry, or nature; Fayn et al., 2015; McCrae & Sutin, 2009) provide particularly strong evidence of others' emotional sensitivity, and thus their wrongness to harm. We contrast this with behaviors that are instrumental or extrinsically motivated (e.g., reaching a certain object or location), which we hypothesize may provide weaker evidence of emotional sensitivity. Aesthetic and moral judgments may be more deeply related than previously believed. Recent neuroimaging work shows similar neural activity and representations of aesthetic and moral judgments, suggesting similar cognitive processes in these domains (Heinzelmann et al., 2020; Tsukiura & Cabeza, 2011; Watson, 2013). We are testing our broader hypothesis regarding the role of aesthetic appreciation and intrinsic versus instrumental value in ongoing work.

The implications of human musicality for social attitudes are also likely to be more complex than is evident from the present studies. For example, the extent to which musicality is seen as a sign of intelligence may differ based on the genre of music (e.g., pop vs classical); or the extent to which the music is seen as sophisticated (e.g., Loomba, 2015). In the current experiments, we intentionally designed our vignettes to remain vague regarding the characters' preferred genre of music and the mode of their engagement with music. Future work may explore the impact of various forms of engagement with different types of music on intergroup attitudes and related social judgments. Overall, we find that reasoning about others' musicality is deeply interwoven with moral thought, providing a framework for understanding why music and dance, seemingly amoral aesthetic behaviors, have consequences for social and moral behaviors.

Acknowledgements

The authors thank Pascal Gagneux, David Huron, and David Barner for helpful feedback on this project.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by funding from the Grammy Foundation's Research Grant to AS, as well as a fellowship to TA from the Center for Academic Research and Training in Anthropogeny (CARTA) at the University of California, San Diego.

ORCID iD

Tanushree Agrawal  <https://orcid.org/0000-0001-5814-3109>

References

- Ansani, A., D'Errico, F., & Poggi, I. (2019). 'You will be judged by the music I hear': A study on the influence of music on moral judgment. *Web Intelligence*, 17(1), 53–62. <https://doi.org/10.3233/WEB-190400>
- Ayotte, J., Peretz, I., & Hyde, K. (2002). Congenitalamusia: Group study of adults afflicted with a music-specific disorder. *Brain*, 125(2), 238–251. <https://doi.org/10.1093/brain/awf028>

- Bauer, D. J., Preacher, K. J., & Gil, K. M. (2006). Conceptualizing and testing random indirect effects and moderated mediation in multilevel models: New procedures and recommendations. *Psychological Methods*, 11(2), 142–163.
- Bicknell, J. (2001). Music, listeners, and moral awareness. *Philosophy Today*, 45(3), 266–274.
- Bloom, P. (2010). How do morals change?. *Nature*, 464(7288), Article 490.
- Cirelli, L. K., Trehub, S. E., & Trainor, L. J. (2018). Rhythm and melody as social signals for infants: Rhythm and melody as social signals for infants. *Annals of the New York Academy of Sciences*, 1423(1), 66–72. <https://doi.org/10.1111/nyas.13580>
- Clarke, E., DeNora, T., & Vuoskoski, J. (2015). Music, empathy and cultural understanding. *Physics of Life Reviews*, 15, 61–88. <https://doi.org/10.1016/j.plrev.2015.09.001>
- Cox,D.,&Levine,M.(2016).*Musicandethics*.<https://doi.org/10.1093/oxfordhb/9780199935321.013.145>
- Crimston, D., Bain, P. G., Hornsey, M. J., & Bastian, B. (2016). Moral expansiveness: Examining variability in the extension of the moral world. *Journal of Personality and Social Psychology*, 111(4), 636–653. <https://doi.org/10.1037/pspp0000086>
- Dennis, S. A., Goodson, B. M., & Pearson, C. (2018). MTurk workers' use of low-cost "virtual private servers" to circumvent screening methods: A research note. <https://papers.ssrn.com/abstract=3233954>
- Epley, N., & Waytz, A. (2010). Mind perception. In S. T. Fiske, D. T. Gilbert, & G. Lindzey (Eds.), *Handbook of social psychology* (pp. 498–541). Wiley.
- Fayn, K., MacCann, C., Tiliopoulos, N., & Silvia, P. J. (2015). Aesthetic emotions and aesthetic people: Openness predicts sensitivity to novelty in the experiences of interest and pleasure. *Frontiers in Psychology*, 6, Article 1877. <https://doi.org/10.3389/fpsyg.2015.01877>
- Goodwin, G. P. (2015). Experimental approaches to moral standing: Moral standing. *Philosophy Compass*, 10(12), 914–926. <https://doi.org/10.1111/phc3.12266>
- Gray, H. M., Gray, K., & Wegner, D. M. (2007). Dimensions of mind perception. *Science*, 315(5812), 619–619. <https://doi.org/10.1126/science.1134475>
- Heinzelmann, N. C., Weber, S. C., & Tobler, P. N. (2020). Aesthetics and morality judgments share cortical neuroarchitecture. *Cortex*, 129, 484–495. <https://doi.org/10.1016/j.cortex.2020.04.018>
- Honing, H. (2019). *The evolving animal orchestra: In search of what makes us musical*. The MIT Press.
- Hove, M. J., & Risen, J. L. (2009). It's all in the timing: Interpersonal synchrony increases affiliation. *Social Cognition*, 27(6), 949–960. <https://doi.org/10.1521/soco.2009.27.6.949>
- Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unified theory of musical emotions. *Physics of Life Reviews*, 10(3), 235–266. <https://doi.org/10.1016/j.plrev.2013.05.008>
- Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children ☆, ☆☆. *Evolution and Human Behavior*, 31(5), 354–364. <https://doi.org/10.1016/j.evolhumbehav.2010.04.004>
- Kivy, P. (2009). *Antithetical arts: On the ancient quarrel between literature and music*. Oxford University Press.
- Knobe, J., & Prinz, J. J. (2008). Intuitions about consciousness: Experimental studies. *Phenomenology and the Cognitive Sciences*, 7(1), 67–83.
- Kozak, M. N., Marsh, A. A., & Wegner, D. M. (2006). What do I think you're doing? Action identification and mind attribution. *Journal of Personality and Social Psychology*, 90(4), 543–555.
- Kreutz, G. (2014). Does singing facilitate social bonding?. *Music and Medicine*, 6(2), 51–60.
- Levitin, D. J., Cole, K., Chiles, M., Lai, Z., Lincoln, A., & Bellugi, U. (2004). Characterizing the musical phenotype in individuals with Williams syndrome. *Child Neuropsychology*, 10(4), 223–247.
- Loomba, A. (2015). *Colonialism/postcolonialism*. Routledge.
- Loui, P., Patterson, S., Sachs, M. E., Leung, Y., Zeng, T., & Przysinda, E. (2017). White matter correlates of musical anhedonia: Implications for evolution of music. *Frontiers in Psychology*, 8, Article 1664. <https://doi.org/10.3389/fpsyg.2017.01664>
- Mas-Herrero, E., Zatorre, R. J., Rodriguez-Fornells, A., & Marco-Pallarés, J. (2014). Dissociation between musical and monetary reward responses in specific musical anhedonia. *Current Biology*, 24(6), 699–704. <https://doi.org/10.1016/j.cub.2014.01.068>
- May, M. (2014, December 26). *Recordings that made waves: The songs that saved the whales*. <https://www.npr.org/2014/12/26/373303726/recordings-that-made-waves-the-songs-that-saved-the-whales>

- McCrae, R. R., & Sutin, A. R. (2009). Openness to experience. In M. R. Leary & R. H. Hoyle (Eds.), *Handbook of individual differences in social behavior* (Vol. 15, pp. 257–273). Guilford.
- Müllensiefen, D., Gingras, B., Musil, J., & Stewart, L. (2014). The musicality of non-musicians: An index for assessing musical sophistication in the general population. *PLOS ONE*, 9(2), Article e89642. <https://doi.org/10.1371/journal.pone.0089642>
- Neto, F., da Conceição Pinto, M., & Mullet, E. (2016). Can music reduce anti-dark-skin prejudice? A test of a cross-cultural musical education programme. *Psychology of Music*, 44(3), 388–398. <https://doi.org/10.1177/0305735614568882>
- Neto, F., Pinto, M. da C., & Mullet, E. (2019). Can music reduce national prejudice? A test of a cross-cultural musical education programme. *Psychology of Music*, 47, 747–756. <https://doi.org/10.1177/0305735618774867>
- Ng, R., Lai, P., Levitin, D. J., & Bellugi, U. (2013). Musicality correlates with sociability and emotionality in Williams syndrome. *Journal of Mental Health Research in Intellectual Disabilities*, 6(4), 268–279.
- Orne, M. T. (1962). On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist*, 17(11), 776–783.
- Piazza, J., Landy, J. F., & Goodwin, G. P. (2014). Cruel nature: Harmfulness as an important, overlooked dimension in judgments of moral standing. *Cognition*, 131(1), 108–124. <https://doi.org/10.1016/j.cognition.2013.12.013>
- Reis, H. T. (2007). Similarity-attraction effect. In R. Baumeister & K. Vohs (Eds.), *Encyclopedia of social psychology* (pp. 876–876). SAGE. <https://doi.org/10.4135/9781412956253.n517>
- Robbins, P., & Jack, A. I. (2006). The phenomenal stance. *Philosophical Studies*, 127(1), 59–85.
- Robson, C. (2011). *Real world research: A resource for users of social research methods in applied settings* (3rd ed.). Wiley.
- Savage, P. E., Loui, P., Tarr, B., Schachner, A., Glowacki, L., Mithen, S., & Fitch, W. T. (2020). Music as a coevolved system for social bonding. *Behavioral and Brain Sciences*, 44, Article e59.
- Schachner, A., Brady, T. F., Pepperberg, I. M., & Hauser, M. D. (2009). Spontaneous motor entrainment to music in multiple vocal mimicking species. *Current Biology*, 19(10), 831–836. <https://doi.org/10.1016/j.cub.2009.03.061>
- Sherman, G. D., & Haidt, J. (2011). Cuteness and disgust: The humanizing and dehumanizing effects of emotion. *Emotion Review*, 3(3), 245–251.
- Sloboda, J. A. (1988). *Generative processes in music: The psychology of performance, improvisation, and composition*. Clarendon Press/Oxford University Press.
- Stewart, R. (2014, March 10). Musical anhedonia: Why some people just don't like music. *The Telegraph*. <https://www.telegraph.co.uk/culture/music/music-news/10687919/Dont-like-music-Scientists-explain-why.html>
- Sytsma, J., & Machery, E. (2010). Two conceptions of subjective experience. *Philosophical Studies*, 151(2), 299–327. <https://doi.org/10.1007/s11098-009-9439-x>
- Sytsma, J., & Machery, E. (2012). The two sources of moral standing. *Review of Philosophy and Psychology*, 3(3), 303–324.
- Tarr, B., Launay, J., Cohen, E., & Dunbar, R. (2015). Synchrony and exertion during dance independently raise pain threshold and encourage social bonding. *Biology Letters*, 11(10), Article 20150767. <https://doi.org/10.1098/rsbl.2015.0767>
- Tsukiura, T., & Cabeza, R. (2011). Shared brain activity for aesthetic and moral judgments: Implications for the Beauty-is-Good stereotype. *Social Cognitive and Affective Neuroscience*, 6(1), 138–148. <https://doi.org/10.1093/scan/nsq025>
- TurkPrime. (n.d.). *After the bot scare: Understanding what's been happening with data collection on MTurk and how to stop it*. <https://blog.turkprime.com/after-the-bot-scare-understanding-whats-been-happening-with-data-collection-on-mturk-and-how-to-stop-it>
- Vuoskoski, J. K., Clarke, E. F., & DeNora, T. (2017). Music listening evokes implicit affiliation. *Psychology of Music*, 45(4), 584–599. <https://doi.org/10.1177/0305735616680289>
- Watson, G. (2013). XIV—Psychopathic agency and prudential deficits. *Proceedings of the Aristotelian Society*, 113(3_pt_3), 269–292. <https://doi.org/10.1111/j.1467-9264.2013.00355.x>

- Weisman, K., Dweck, C. S., & Markman, E. M. (2017). Rethinking people's conceptions of mental life. *Proceedings of the National Academy of Sciences*, 114(43), 11374–11379. <https://doi.org/10.1073/pnas.1704347114>
- Westfall, J., & Yarkoni, T. (2016). Statistically controlling for confounding constructs is harder than you think. *PLOS ONE*, 11(3), Article e0152719. <https://doi.org/10.1371/journal.pone.0152719>
- Zentner, M., Grandjean, D., & Scherer, K. R. (2008). Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion*, 8(4), 494–521. <https://doi.org/10.1037/1528-3542.8.4.494>