# From Music to Animacy:
# Causal Reasoning Links Animate Agents with Musical Sounds

**Minju Kim (minjukim@ucsd.edu)**
University of California, San Diego, Department of Psychology
9500 Gilman Drive, La Jolla, CA, 92092 USA

**Adena Schachner (schachner@ucsd.edu)**
University of California, San Diego, Department of Psychology
9500 Gilman Drive, La Jolla, CA, 92092 USA

## Abstract

Listening to music activates representations of movement and social agents. Why? We ask whether high-level causal reasoning about how music was generated can lead people to link musical sounds with animate agents. To test this, we asked whether people (N=60) make flexible inferences about whether an agent caused musical sounds, integrating information from the sounds' timing and from the visual context in which it was produced. Using a 2x2 within-subject design, we found evidence of causal reasoning: In a context where producing a musical sequence would require self-propelled movement, people inferred that an agent had been present causing the sounds. When the context provided an alternative possible explanation, this 'explained away' the agent, reducing the tendency to infer an agent was present for the same acoustic stimuli. People can use causal reasoning to infer whether an agent produced musical sounds, suggesting that high-level cognition can link music with social concepts.

**Keywords:** causal reasoning, inference, agency, animacy, music cognition, social cognition

## Introduction

The mere act of listening to music activates more than just auditory representations. Music also activates representations of movement, even when no movement is produced or viewed (Cannon & Patel, 2021; Chen, Penhune & Zatorre, 2008; Grahn & Rowe, 2009). Music appears to activate social representations as well, particularly for music that people believe was created by a person (Launay, 2015; Steinbeis & Koelsch, 2009).

Why would musical sounds make us think about movement and social agents? At least in part, musical sounds are linked with movement due to low-level, modality-specific connections: Features of rhythmic sound drive the motor system (and vice versa) via early-developing connections between auditory and motor systems (Cannon & Patel, 2021; Chen et al., 2008; Grahn & Rowe, 2009; Patel & Iversen, 2014). Beyond rhythm, other auditory features also drive the perception of animacy, movement, and emotion in music (Blust, Baker, Richard & Shanahan, 2015; Juslin & Laukka, 2003; Sievers, Polansky, Casey & Wheatley, 2013). The organized nature of musical sound may also lead listeners to think about agents: From early in life, people expect that orderly objects and sounds are associated with agents (Ma & Xu, 2013; Newman, Keil, Kuhlmeier, & Wynn, 2010; but see

Schachner & Kim, 2018). These accounts share a focus on the role of perceptual attributes of sound in driving links with movement and social concepts.

Here we propose that an additional cognitive process may also play a role. In particular, high-level causal reasoning about how the music was generated may lead people to link music with movement and agents, through a process of inference to the best explanation (Keil & Newman, 2015; Lipton, 2004; Ma & Xu, 2013; Tenenbaum, Griffiths & Kemp, 2006). Causal reasoning has the potential to explain why music is linked both to representations of movement as well as to animate agents: This may result from a joint inference (Baker, Jara-Ettinger, Saxe & Tenenbaum, 2017) about the type of movement required, and the type of force that could cause those movements. Under this account, to explain an observed sequence of musical pitches, people infer what types of movements would be required to produce them; and simultaneously infer the most likely cause of those movements, weighing multiple hypotheses (perhaps they were generated by a person; by an animal; by gravity; by wind; etc.). They then choose the hypothesis that is both plausible a priori and provides a good explanation of the observed events.

This type of inferential reasoning occurs in multiple related domains, including physical-mechanical reasoning, social reasoning, and mental state inference (Teglas et al., 2011; Baker et al., 2009; Battaglia, Hamrick & Tenenbaum, 2013). Here, we ask whether this domain-general causal reasoning process also occurs when listening to musical sounds, and explore the implications of this idea. To be clear, we view causal inference as a domain-general ability that is not unique to musical sounds, and likely occurs in a similar way for non-musical environmental sounds (Gerstenberg, Seigel & Tenenbaum, 2018; Traer, Norman-Haignere & McDermott, 2020). However, if people engage in causal reasoning about musical sounds in particular, this would provide a parsimonious explanation for several aspects of musical representations.

### Animate Agents as Causes of Music

Causal inference should allow people to infer that many musical sounds were caused by animate agents. From early

in life, people have a mental model of the physical-mechanical world that allows them to predict and explain physical-mechanical events; and a distinct mental model of animate agents as differing from inanimate objects (Battaglia, Hamrick & Tenenbaum, 2013; Spelke & Kinzler, 2007). In particular, animate agents have unique capacities including self-propelled, goal-directed movement (Gelman, Durgin, & Kaufman, 1995; Gergely et al., 1995). Thus, when we see an event that would require self-propelled movement, we infer that an animate agent must be the cause (e.g. Saxe, Tenenbaum & Carey, 2005).

Similarly, when we hear sounds that would require self-propelled movement, we may infer that an animate agent was its cause. Existing work shows that we use rational causal inference to reason about the sources of non-musical sounds (Gerstenberg et al., 2018; Traer et al., 2020). For example, when faced with a pinball-machine-like box into which a ball could be dropped, participants were able to use the timing of impact sounds to infer the path the ball had taken (e.g. where it had entered the box). Participants were also able to use auditory information, together with visual context, to predict where the ball would likely land. In both cases, participants' judgements were well-predicted by a Bayesian model of rational causal inference (Gerstenberg et al., 2018). This work shows that people can use properties of sound to infer physical-mechanical information, such as an objects' path of movement. However, this work has not tested whether people can make the more abstract inference that certain sounds were produced by animate agents.

## Causal Reasoning and Explaining Away

Can people make inferences about whether an animate agent caused non-vocal, musical sounds they hear? We expect causal inference to occur in a similar way for musical sounds as for non-musical environmental sounds, and focus on musical sounds (short melodies played on a xylophone) for two reasons. First, causal reasoning about abstract, instrumental musical sounds would have implications for the nature of musical representations. Second, short instrumental sequences may be produced by people or by inanimate forces (e.g. a wind chime), leading to uncertainty that may be resolved with causal inference – allowing us to detect whether people can engage in this reasoning when hearing musical sounds.

One signature of causal inferential reasoning is 'explaining away' – the idea that a plausible alternative explanation should weaken evidence for other explanations. This phenomenon is a signature of causal inference, and serves as evidence that the underlying cognitive representation is rich and structured (Gopnik & Sobel, 2000; Pearl, 2000; Tenenbaum et al., 2006). If causal reasoning can link musical sounds with agents, then the presence of a plausible physical-mechanical cause should 'explain away' the agent, reducing the likelihood of inferring an agent as the cause, even for the very same musical stimulus.

In previous work, we explored a different topic – why orderly stimuli are associated with animate agents – by
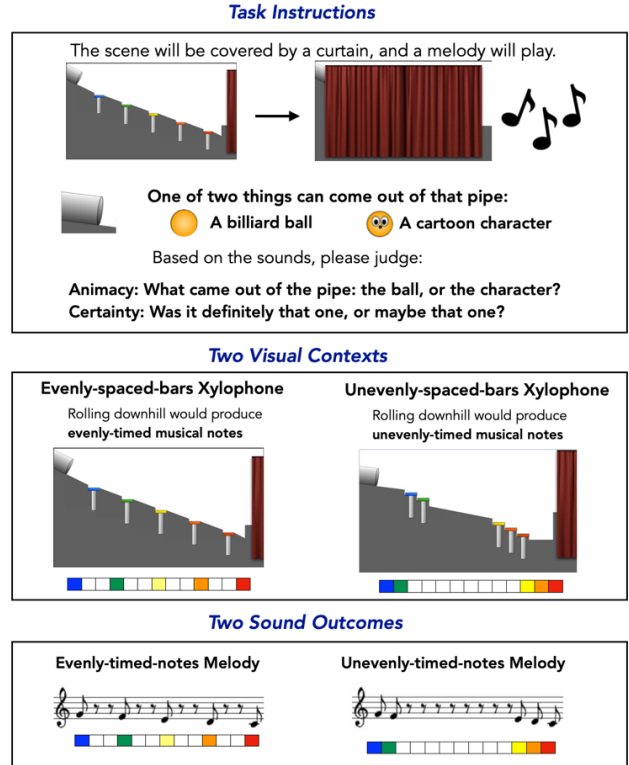


Figure 1: Method. Participants were asked to judge whether an animate agent, or inanimate ball, was present behind a curtain when musical sounds had occurred. We manipulated the visual context and the timing of the musical sounds that occurred, to test whether people integrated auditory and visual information to infer whether an agent caused the sounds.

creating an experimental paradigm that may be useful for investigating the role of causal inference in music perception (Schachner & Kim, 2018). In particular, participants were shown a xylophone-like staircase, in which the bars of the xylophone were embedded in the steps. In one condition, the bars were arranged so that a ball rolling downhill by gravity would produce a descending scale (G-F-E-D-C). In another condition, the bars were placed in different positions, such that rolling downhill would play the same notes in a different order. The xylophone was occluded from view, and musical sounds played: Either a descending scale, or the same notes in a different order. Participants were asked to judge which of two things had been present when the sounds occurred: An inanimate ball, or an animate cartoon character. Participants reliably inferred that an animate character was present only when the sounds and placement of bars on the xylophone implied that the character would have had to change direction or jump (Schachner & Kim, 2018).

These data suggest that people may use pitch information (the order in which pitches were produced) to infer whether an agent caused musical sounds. However, the previous study allowed participants to solve the task without using auditory

information: Because we were not concerned with isolating the role of auditory information, the task provided a visual "legend" of numbers corresponding to each pitch as notes were played, to aid difficulties with pitch perception or memory (Schachner & Kim, 2018). Thus, participants could have solved this task without reasoning about the causes of musical sounds, but instead relying solely on visual cues.

## The Current Study

Here, we ask whether people can use information from musical sounds to infer whether the sounds were caused by an agent. We focus on the use of timing in musical sounds, as past work shows that people can make physical-mechanical inferences based on non-musical sounds' timing (Gerstenberg et al., 2018). Can people integrate auditory information about musical sounds with visual information about the environment, to rationally infer whether an agent or object was the most likely cause of a musical sequence? If people use their rich mental models of the physical and social world to reason about musical stimuli, they should be able to infer whether musical sounds were caused by an agent, by making a joint inference about the movements required to cause the sounds, and the type of agent or object required to cause the movements.

To test whether people integrate auditory and visual information to infer whether an agent was the cause of musical sounds, we adapted prior methods (Schachner & Kim, 2018), designing animated stimuli in which xylophone-like musical bars were embedded in a hill. In particular, we manipulated the *visual context* by changing the distance between the bars, either placing them at even distances from one another (evenly-spaced bars) or in unevenly-spaced groupings (unevenly-spaced bars; see Fig. 1).

On each trial, the visual context was occluded, a sound sequence played, and we asked participants to judge whether an animate character or inanimate object had been present when the sounds were played (ostensibly causing the sounds, although the idea of causality was not explicitly stated in the instructions). Across conditions, we manipulated the timing of the musical sounds, such that one of two *sound sequences* played: A descending scale with evenly-timed notes, or unevenly-timed notes. The two sound sequences were timed such that in the evenly-spaced visual context, a ball rolling downhill with gravity could be expected to play the evenly-timed sound sequence; a ball in the unevenly-spaced visual context would be expected to play the unevenly-timed notes.

If people can use causal reasoning to infer that animate agents cause musical sounds, then people should make different inferences about whether an agent was the cause of each sound sequence, depending on the visual context in which it was produced. That is, there should be an interaction between the nature of the sound, and its visual context. In a visual context where a sound sequence could be produced by gravity alone, this possible physical-mechanical explanation should "explain away" the agent, leading people to be less likely to infer that an agent caused the sounds. In contrast, in a context where the sound sequence would require self-propelled movement (e.g. jumping, changing speed or direction), people should infer that the sound was caused by an agent – thus changing their inferences about the cause of the sound sequences depending on their context.

Alternatively, people may believe that all musical sounds were created by an agent (Launay, 2015). In this case, participants should answer that all of the sound sequences are more likely to be produced by the agent, regardless of visual context. Similarly, participants may expect that musical sounds with certain features (e.g. more orderly sequences, Newman et al., 2010) are always associated with agents. In this case, participants should infer that the evenly-timed notes were produced by an animate agent equally in both visual contexts. These patterns of findings would suggest failure to engage in causal reasoning, and failure to consider alternative physical-mechanical explanations for the sounds that are suggested by the visual context. The study design, sample size, and analyses were pre-registered before data collection (https://aspredicted.org/rf47e.pdf).

## Methods

### Participants

As pre-registered, sixty undergraduate students were recruited from an online study pool at a large public university in Southern California (Age: M = 20.97 years, Range = 18 to 31; 19 males), and were tested online. All participants gave informed consent, and received credit as compensation. 17 additional participants were tested but excluded due to pre-registered exclusion criteria: failing one or more attention check questions (13); technical issues (4).

### Stimuli

Animated videos and accompanying sounds were constructed using Apple Keynote, Apple GarageBand, and Apple QuickTime software. Two familiarization videos demonstrated how the xylophone-like bars produced sounds when struck (one video for each visual context). Both videos showed a hill with colored bars and a pipe at the top of the hill (see Fig. 1). A mallet entered the scene and struck a sequence of bars before exiting the scene. The sequences played were designed to differ equally from both sound sequences played during the test trials, and thus consisted of the ascending scale played twice, once with even timing and once with uneven timing. Two additional familiarization videos provided evidence of the ball's inanimacy and the character's animacy, by showing the character moving in a self-propelled way, and the ball rolling with gravity.

During each trial, a test video was shown in which the visual context was covered by an animated curtain, and one of two sound outcomes played: a descending scale with evenly-timed notes (G, F, E, D, C, with fundamental frequencies of 778, 702, 647, 569, and 511 Hz), or a descending scale with unevenly-timed notes (the same pitches, played G-F-E-pause-pause-D-C). Both sound sequences were of overall equal duration (4 seconds).

## Design

The experiment used a 2 (visual context) x 2 (sound outcome) within-subject design. Each participant completed four trials, composed of two blocks (by visual context) of two trials each (one for each sound outcome). The order of the blocks, and the order of test trials within a block, were counterbalanced between subjects.

## Procedure

Participants first viewed a familiarization video which demonstrated how the bars produced sounds when struck, were asked to describe what happened, and answered two attention check questions (whether each bar played the same note, different notes, animal sounds, or no sounds; and whether the bars on the xylophone were evenly-spaced, or unevenly-spaced).

On the next page, participants were asked to notice the pipe on the left-hand side of the scene, and were told that one of two things could come out of the pipe: An inanimate ball, or a cartoon character. To provide evidence of the ball's inanimacy and the character's animacy, brief videos were shown (of the character moving in a self-propelled way, or the ball rolling with gravity), and they were identified in ways consistent with inanimacy and animacy: As a billiard ball, or as cartoon character with a name (Fred).

Participants then completed test trials. On each test trial, participants watched a video where the visual context was occluded, and a sound sequence played (either the evenly-timed or unevenly-timed descending scale). Participants were then asked to judge whether an agent or inanimate object had been present in the scene when the sounds occurred, using 2-alternative forced choice ("What came out of the pipe? Was it the ball? Or was it Fred?"). Participants then reported their certainty about this answer, again using a 2-alternative forced choice ("Was it definitely that one, or maybe that one?"). This response format was used in order to allow for developmental comparison to children in future studies. The two scales were then combined, to create a single four-point scale (Definitely the ball, Maybe the ball, Maybe the agent, Definitely the agent).

After completing two test trials, participants received similar instructions for the second block (the other visual context condition), with a new visual context introduced. They then completed two additional test trials, as in the first block, but with the videos showing the new visual context. Participants lastly were asked if they had experienced any technical issues, and to guess what the experiment was about, before submitting their answers.

## Results

To test whether participants integrated auditory information with visual information to infer the causes of musical sounds, we used an ordinal logistic regression model to predict participants' judgements of what had caused the sound ('definitely the ball', 1, to 'definitely the agent', 4; ordinal variable), from the predictors of the visual context (evenly-
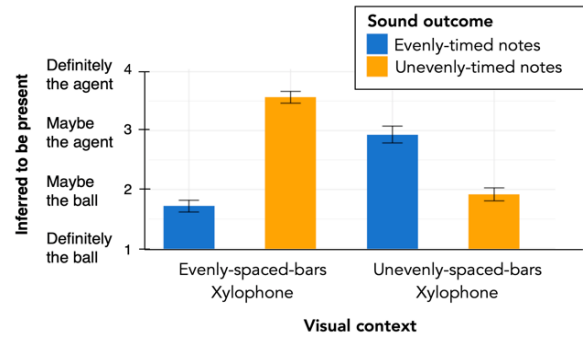


Figure 2: Results. Participants' judgments of whether an animate agent (the character) or an inanimate ball had been present when the sounds played, for each visual context (evenly-spaced-bars xylophone, unevenly-spaced-bars xylophone) and sound outcome (evenly-timed-note melody, unevenly-timed-note melody).

spaced bars, unevenly-spaced bars), the sound outcome (evenly-timed notes, unevenly-timed notes), the interaction of these two factors, and subject (as a random effect). As predicted, there was a significant interaction between visual context and sound outcome ($\chi^2(1) = 112.21$, $p < 0.001$, nested model comparison; $Z = -9.51$, $p < 0.001$, coefficient test; see Fig. 2). In the context of the evenly-spaced-bars xylophone, participants judged it more likely that the agent had been present after hearing the unevenly-timed notes ($M = 3.57$, $SE = 0.10$) than the evenly-timed-notes ($M = 1.72$, $SE = 0.10$; Wilcoxon Test, $Z = -6.13$, $p < 0.001$). This pattern was reversed in the other visual context: In the context of the unevenly-spaced-bars xylophone, participants judged it more likely that the agent had been present after hearing the evenly-timed notes ($M = 2.94$, $SE = 0.14$) than the unevenly-timed notes ($M = 1.92$, $SE = 0.11$; Wilcoxon Test, $Z = -3.98$, $p < 0.001$). There were also significant main effects of the visual context ($Z = 6.44$, $p < 0.001$), and of the sound outcome ($Z = 8.90$, $p < 0.001$).

## Discussion

Why is music linked with movement and social concepts? Many influential accounts focus on the role of particular perceptual attributes of sound, or on low-level perceptual-motor links, which have been well-established (Cannon & Patel, 2021; Juslin & Laukka, 2003). The current findings suggest that high-level cognition may also play a role. We found evidence that people can use causal reasoning to infer that an animate agent was the cause of musical sounds: Participants made different inferences about whether an agent was present causing each sound sequence, depending on whether self-propelled movement would be required to cause the sounds. When the visual context implied that the notes could be produced in an alternative way, this 'explained away' the agent, such that people were less likely to infer that an agent was present, causing the sounds.

These findings support the idea that people can reason about the kind of force – animate, or inanimate – that caused musical sounds they hear. This kind of causal inference is a domain-general ability that is not unique to musical sounds, and likely occurs in a similar way for non-musical environmental sounds (Gerstenberg et al., 2018; Traer et al., 2020). However, the finding that people can reason about agents as causes of musical sounds in particular helps to explain why music listening activates concepts of movement and animate agency: These abstract ideas may be inferred through causal reasoning. In particular, people may engage in a process of joint inference about the causes of sounds (including music), jointly inferring the type of movement required to cause the sounds, and the type of force that could cause those movements (e.g. a person, the wind). To make and test nuanced quantitative predictions about people's reasoning, future work may model this hypothesized reasoning as joint Bayesian inference (e.g. Baker et al., 2017).

The current work focused on the use of timing cues in causal reasoning, rather than other features of sound (pitch, timbre). For non-musical environmental sounds, these other features also constrain inferences about possible causes (e.g. Gaver, 1993), and people use multiple acoustic features to make causal inferences about physical-mechanical sound sources (e.g. intensity and reverberation, Traer et al., 2020; see also Agrawal et al., 2020; Cusimano, Hewitt, Tenenbaum & McDermott, 2019). Future work should investigate the extent to which people integrate and flexibly use multiple cues to make causal inferences about musical sounds, and whether causal reasoning about music resembles that regarding other environmental sounds.

Causal reasoning is likely one process among several which links music with movement and agency. For example, learned associations between musical sounds and people may also play a role (e.g. Launay, 2015). Based on experience, once we categorize sounds as 'music', we may simply assume that the sounds were caused by people. However, our data show that causal reasoning contributes above and beyond other known processes: In our study, people did not judge that all musical sequences were produced by agents (as hypothesized by Launay, 2015); and did not expect that more orderly sequences were always associated with agents (as seen e.g. by Newman et al., 2010). Overall, we suggest that causal reasoning is likely sufficient, but not the only way, that listeners link music with agency.

**Can We Generalize to Everyday Music Listening?**

Do people engage in causal reasoning during everyday music listening? In our study, participants were explicitly asked to reason about whether an animate agent or inanimate object was present. Our data therefore serve as a proof of concept that when prompted, people can engage in causal reasoning about musical sounds. It remains an open question the extent to which this process occurs implicitly and spontaneously during everyday music listening. We hypothesize that causal reasoning occurs during real-world music listening, though perhaps not automatically, and not in all cases. It may

particularly occur when deeply engaging with music, as in active or creative music listening (Dunn, 1997; Kratus, 2017; Peterson, 2006); or during mental imagery or mental narrative related to musical content (Margulis et al., 2019).

In our study, we also specified the context in which the sound was produced (the scene with the xylophone) for participants. This allowed us to manipulate the visual context, and test its effects on people's judgements. However, people often hear music without seeing the context in which it was played. Do people reason about the role of animate agents when hearing music without seeing the physical context?

We believe that people likely do engage in causal reasoning in this case (though again, perhaps not automatically or at all times). Music listening without visual access can be thought of as introducing a greater amount of uncertainty to the causal reasoning process, or starting with a larger hypothesis space. Even with more uncertainty, people could make an inference about the events needed to cause musical sounds, guided by their prior knowledge of how environmental sounds are caused, more generally; and of musical instruments, in particular. People could weigh the likelihood of the musical sequence under a variety of hypothesized causal events, and then marginalize over those possibilities (Griffiths & Tenenbaum, 2009). The framework of probabilistic causal reasoning thus applies equally well to the case where visual context is not provided. Future work should test the extent to which people engage in causal reasoning spontaneously during everyday listening to music.

One last way that naturalistic music differs from our stimuli is in its length. Our musical stimuli consisted of sequences of five pitches, whereas a typical musical sequence would contain a greater number of notes. The causal reasoning process shown here generalizes easily to longer musical sequences: Indeed, longer musical sequences likely provide even stronger evidence that an agent must be the cause. Producing long, complex musical sequences almost always requires repeatedly changing direction and speed in a self-propelled and intentional way – and this is the case in almost any plausible physical environment. That is, as songs get longer, one would have to posit the existence of increasingly implausible and complex Rube-Goldberg machines in order to think that the notes could be produced without an agent (e.g. by an inanimate object rolling downhill). Thus, particularly for longer musical sequences, causal reasoning should allow the listener to infer that an agent was the most likely cause of musical sounds, even when they are uncertain of the exact nature of the musical instrument or physical context of production.

**Developmental Implications**

Our findings raise additional questions about the development of music perception. Do infants and children spontaneously reason about the causes of musical sounds? If so, this may provide one explanation for infants' early interest in instrumental music: They may infer the presence of an agent, causing the musical sounds they hear. However, children's early experience with recorded music may provide

counter-evidence for this idea: In modern times, young children often have more experience with recorded music – for which no agent is present causing the sound – than with live instrumental performance (Mehr, 2014).

There are several reasons to expect that infants and young children may be able to infer that agents are the causes of musical sounds they hear. From infancy, children differentiate between animate agents and inanimate objects, and know that agents have the unique capacity for self-propelled, goal-directed movement (Spelke & Kinzler, 2007; Gelman et al., 1995; Gergely et al., 1995). When infants see an event that would require self-propelled movement, they infer that an animate agent must be the cause (Saxe et al., 2005). In the first months of life, infants link sounds with specific events, expecting to hear a sound when they see an impact between two objects, or between an object and a surface (Bahrick, 1983, 1988; Kopp, 2014; Lewkowicz, 1992, 1994; Spelke, 1979). Thus, even infants may engage in causal reasoning about sounds, and realize that certain sounds could not be produced without an agents' intervention. In ongoing work, we are investigating the development of the ability and tendency to infer social causes of musical sounds in childhood.

Overall, the current work shows that adults can use causal reasoning to infer whether an agent or inanimate object produced musical sounds, integrating information from the visual context and the timing of musical sounds. This provides an additional means by which musical sounds are linked with movement and social concepts: Through high-level cognition, as a form of rational causal inference about the movements and agents required to cause the sounds.

## Acknowledgments

## References

Agrawal, T., Lee, M., Calcetas, A., Clarke, D., Lin, N., Schachner, A. (2020). Hearing water temperature: Characterizing the development of nuanced perception of auditory events. In S. Denison., M. Mack, Y. Xu, & B.C. Armstrong (Eds.), *Proceedings of the 42nd Annual Conference of the Cognitive Science Society* (pp. 494-499). Cognitive Science Society.

Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, *1*(4), 1-10.

Baker, C. L., Saxe, R. R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition, 113*(3), 329–349.

Bahrick, L. E. (1983). Infants' perception of substance and temporal synchrony in multimodal events. *Infant Behavior and Development, 6*(4), 429–451.

Bahrick, L. E. (1988). Intermodal learning in infancy: Learning on the basis of two kinds of invariant relations in audible and visible events. *Child Development*, *59*(1),197–209.

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327–18332.

Blust, A. S., Baker, D. J., Richard, K., & Shanahan, D. Music, animacy, and rubato: what makes music sound human?. In *Proceedings of the 14th international Conference on Music Perception and Cognition* (pp.186-190). CA: San Francisco.

Cannon, J. J., & Patel, A. D. (2021). How Beat Perception Co-opts Motor Neurophysiology. *Trends in Cognitive Sciences, 25*(2), 137–150.

Chen, J. L., Penhune, V. B., & Zatorre, R. J. (2008). Listening to musical rhythms recruits motor regions of the brain. *Cerebral Cortex, 18*(12), 2844–2854.

Cusimano, M., Hewitt, L. B., Tenenbaum, J., McDermott, J. H. (2018). Auditory scene analysis as Bayesian inference in sound source models. In C. Kalisch, M. Rau, T. Rogers, & J. Zhu (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp. 510-516). Madison, WI: Cognitive Science Society.

Dunn, R.E. (1997). Creative Thinking and Music Listening. *Research Studies in Music Education* 8, 42–45.

Gaver, W. W. (1993). What in the World Do We Hear?: An Ecological Approach to Auditory Event Perception. *Ecological Psychology, 5*(1), 1–29.

Gelman, R., Durgin, F., & Kaufman, L. (1995). Distinguishing between animates and inanimates: Not by motion alone. In D. Sperber, D. Premack, & A. J. Premack (Eds.), Symposia of the Fyssen Foundation. *Causal cognition: A multidisciplinary debate* (p. 150–184). Clarendon Press/Oxford University Press.

Gergely, G., Nádasdy, Z., Csibra, G., & Bíró, S. (1995). Taking the intentional stance at 12 months of age. *Cognition*, *56*, 165–193.

Gerstenberg, T., Siegel, M. H., Tenenbaum, J. B. (2018) What happened? Reconstructing the Past From Vision and Sound. In C. Kalisch, M. Rau, T. Rogers, & J. Zhu (Eds.), In *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (p.409), Madison, WI.

Gopnik, A., & Sobel, D.M. (2000). Detecting Blickets: How young children use information about novel causal powers in categorization and induction. *Child Development*, *71*(5), 1205-1222.

Grahn, J. A., & Rowe, J. B. (2009). Feeling the beat: premotor and striatal interactions in musicians and nonmusicians during beat perception. *Journal of Neuroscience*, *29*(23), 7540-7548.

Griffiths, T. L., & Tenenbaum, J. B. (2009). Theory-based causal induction. *Psychological Review, 116*(4), 661–716. doi:10.1037/a0017201

Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance:

Different channels, same code? *Psychological Bulletin, 129*(5), 770–814.

Keil, F. C., & Newman, G. E. (2015). Order, order everywhere, and only an agent to think: The cognitive compulsion to infer intentional agents. *Mind and Language, 30*(2), 117–139.

Kopp, F. (2014). Audiovisual temporal fusion in 6-month-old infants. *Developmental Cognitive Neuroscience, 9*, 56–67.

Kratus, J. (2017). Music listening is creative. *Music Educators Journal*, *103*(3), 46-51.

Launay, J. (2015). Musical sounds, motor resonance and agency detection. *Empirical Musicology Review, 10*(1), 30–40.

Lewkowicz, D. J. (1992). Infants' responsiveness to the auditory and visual attributes of a sounding/moving stimulus. *Perception & Psychophysics, 52*(5), 519–528.

Lewkowicz, D. J. (1994). Limitations on infants' response to rate-based auditory-visual relations. *Developmental Psychology, 30*(6), 880.

Lipton, P. (2003). *Inference to the best explanation.* Routledge.

Ma, L., & Xu, F. (2013). Preverbal infants infer intentional agents from the perception of regularity. *Developmental Psychology*, *49*(7), 1330–1337.

Margulis, E. H., Wong, P. C., Simchy-Gross, R., & McAuley, J. D. (2019). What the music said: Narrative listening across cultures. *Palgrave Communications, 5*(1), 1-8.

Mehr, S. A. (2014). Music in the home: New evidence for an intergenerational link. *Journal of Research in Music Education, 62*(1), 78–88

Newman, G. E., Keil, F. C., Kuhlmeier, V. A., & Wynn, K. (2010). Early understandings of the link between agents and order. *Proceedings of the National Academy of Sciences of the United States of America*, 107(40), 17140–17145.

Patel, A. D., & Iversen, J. R. (2014). The evolutionary neuroscience of musical beat perception: the Action Simulation for Auditory Prediction (ASAP) hypothesis. *Frontiers in Systems Neuroscience, 8*, 57.

Pearl, J. (2000.) *Causality: Models, Reasoning, and Inference.* Cambridge, UK: Cambridge Univ. Press

Peterson, E.M. (2006). Creativity in Music Listening. *Arts Education Policy Review*, 107 (3), 15–21.

Saxe. R., Tenenbaum. J. B., Carey, S. (2005). Secret Agents: Inferences about hidden causes by 10- and 12-month-old infants. *Psychological Science*, *16*, 995–1001.

Schachner, A., & Kim, M. (2018). Entropy, order and agency: The cognitive basis of the link between agents and order. In C. Kalisch, M. Rau, T. Rogers, & J. Zhu (Eds.), *Proceedings of the 40th Annual Conference of the Cognitive Science Society* (pp.2430-2435). Madison, WI.

Sievers, B., Polansky, L., Casey, M., & Wheatley, T. (2013). Music and movement share a dynamic structure that supports universal expressions of emotion. *Proceedings of the National Academy of Sciences of the United States of America, 110*(1), 70–75.

Spelke, E. S. (1979). Perceiving bimodally specified events in infancy. *Developmental Psychology, 15*(6), 626–636.

Spelke, E. S., & Kinzler, K. D. (2007). Core knowledge. *Developmental Science, 10*(1), 89–96.

Steinbeis, N., & Koelsch, S. (2011). Affective priming effects of musical sounds on the processing of word meaning. *Journal of Cognitive Neuroscience, 23*(3), 604–621.

Teglas, E., Vul, E., Girotto, V., Gonzalez, M., Tenenbaum, J. B., & Bonatti, L. L. (2011). Pure Reasoning in 12- Month-Old Infants as Probabilistic Inference. *Science, 332*(6033), 1054–1059.

Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, 10, 309–318.

Traer, J., Norman-Haignere, S. V., & McDermott, J. H. (2020). Causal inference in environmental sound recognition. *BioRxiv.*