The Humming Box: AI-powered Tangible Music Toy for Children

***Chunhan Chen**
New York University
370 Jay Street, Brooklyn, NY
11201, USA
cc6268@nyu.edu

***Yihan Tang**
New York University
370 Jay Street, Brooklyn, NY
11201, USA
yt1559@nyu.edu

***Tianyi Xie**
New York University
370 Jay Street, Brooklyn, NY
11201, USA
tx520@nyu.edu

Stefania Druga
MIT Media Lab
75 Amherst St, Cambridge, MA
02139, USA
sdruga@mit.edu

* Joint first author

CHI PLAY EA '19, October 22 - 25, 2019, Barcelona, Spain

ACM ISBN 978-1-4503-6871-1/19/10.
https://doi.org/10.1145/3341215.3356990

Abstract
Music education benefits children's mental and intellectual development in early childhood [14]. Guided by child-centered design principles, we propose the Humming Box, an AI-powered tangible music toy for 4-6-year-old children. The main goal of our project is to blend traditional hand crank music boxes with software-based music manipulation and create new possibilities for children to create and remix music manually. Our paper presents findings from 3 design sprints; each demonstrating playtest results and new design iterations. The user evaluation of our final design suggests that multimodal musical creation leads to increased engagement. We also discuss how the modular design of our toy's functional components encourages creative expression and cooperative play.

Author Keywords
Multimodality; Music; Early childhood education; Tangible interaction; Modularity

CCS Concepts
•Human-centered computing → Usability testing; User centered design; Interface design prototyping; *Field studies; Sound-based input / output;* **Human computer interaction (HCI);** User studies; **•Social and professional topics → Children;**

	Multi-modality	Modularity	Tangible
Groove-Pizza	●	●	
Melody-Morph	●	●	
Mmmtss	○	●	
MaKey Makey		●	●
Magic Cube			●
Humming Box	●	●	●

● featured ○ partially featured

Figure 1: Comparison of Humming Box with other related projects

Figure 2: Traditional music box which inspired our toy. Photography by Huntmic

Introduction

Music education has been long proven [7] to benefit early childhood development [14]. However, music expression that requires professional instrument training or software assistance [5] presents a higher barrier of entry due to additional cost of time, money, and effort [26]. Moreover, preliminary approaches to express melodies like humming, singing, and whistling [25] set a high emotional barrier for most people.

In order to lower the barrier of entry for younger kids, we want to tap into natural and accessible ways that they already use to share songs. So we propose a new toy that transforms hummed melodies into MIDI signals by pitch detection. Our toy provides kids the option to modify the melody's speed, by mapping it to the physical speed of the cranking and to switching the melody's timbre, by inserting different cards into the box. Our planned interactions consist of two parts: generating melody with children's voice input and music remix interactions utilizing multimodality, such as visual, audio, and motion. Since children have an innate interest in playing music [15], we decide to prioritize the latter part and focus on improving its playfulness and intuitiveness.

Related Work

Based on our design goals, we explored related toys in multimodality, modularity and tangible interaction fields (as shown in Figure 1).

Well-designed tangible toys, such as Munchkin Mozart Magic Cube [27], Makey Makey [23], LittleBits [22], and Lego Education SPIKE Prime [21], prevent children from wasting time on other digital products [19] like YouTube. They help children to form healthy entertaining habits. Moreover, they are also beneficial to children's dexterity development and muscle growth [20]. Multimodality stimulates children's interests in multiple ways and extends their focus time. Technology-empowered toys such as Coding Critters [32], Sphero [34], Jibo [18], and Cosmo [4] were designed to attract children with multimodal interactions (e.g., motion, sound, light). Additionally, they are beneficial to children's personal development [14]. Throughout the multimodal learning process, children interact with new information in various ways, which can cultivate their curiosity [16]. Besides, different sensory modalities are effective as guides for their selective actions. Successful modular toys such as GroovePizza [10], Mmmtss [33] and Scratch [13] free children's creativity by encapsulating intricate technical details to accessible, functional modules for the younger audience. Additionally, constructive interactions with modules will benefit children-initiated activities in the integrated playful pedagogies [9].

Ideation and Design Process

Inspiration

Humming is commonly used by children to create their music [28]. But this habit often gradually diminishes as they grow up. Even if music education has become commonplace, the established rules in the field set a barrier for children to express themselves with music [35] creatively. These frustrations may affect their interests and confidence in constructive music play [6]. Alex Ruthmann [2], from NYU MusEDLab, brought up frontier explorations of interactive digital music composition for kids, which inspired us to reconsider basic elements in traditional music education by mapping them in a more natural and accessible way for younger players. We aimed to offer children an innovative and intuitive physical experience of composing and playing with music without professional training. In this way, children could potentially gain confidence in making broader expressions and creations.

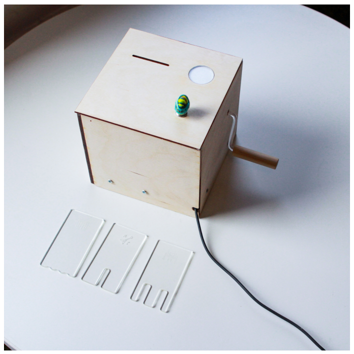

Figure 3: Set up of Humming Box (final design).

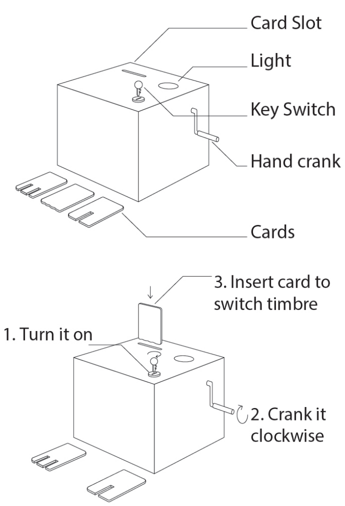

Figure 4: Components and key interactions with the Humming Box

Concept Validation
In order to assert children's interest in music, we surveyed our classmates about their childhood experiences. We invited a group of master students (9 women and 10 men) from various professions, including design, art, and technology within the ITP (Interactive Telecommunications Program) community at New York University. 18 out of 19 interviewees were interested in music as a child. Fifteen interviewees reflected they had more or less tried to create their melody. The validation from the ITP community supported our design motivation.

Minimum Viable Prototype
In order to further probe our initial interaction idea and determine the most appropriate form factor for our design, we made a minimum viable prototype out of clay and foam core. For the music functionality, we used a hidden smartphone with HumOn [1] app to demonstrate the key features of detecting a hummed melody and modifying its rhythm, pitch, and timbre.

Experts Review at FabLearn Conference
We presented the "dummy" prototype (conveys our core concept though not functioning) at FabLearn Conference 2019 [3], and held an expert review session with more than 20 UX designers and researchers. These experts highly acknowledged the concept of multimodal music education of our prototype. Besides, the group's review gave us more insights on children's behavior patterns and helped us to simplify interaction accordingly while maintaining their attention. For instance, one of the teachers suggested children might yell at the box or talk to the box instead of humming, which were proven right in the first round of playtesting. In other words, children may not interact with the toy expectedly. Incorporating children behavior patterns into consideration is critical to design playful and intuitive experiences.

Re-ideation
Based on the feedback received at FabLearn Conference, we decided to pivot the key interaction design direction towards a music box to utilize the existing conceptual model of children to simplify their cognitive burden of memorizing interactions while maintaining the valuable concept about multimodal composing. As a musical education and composing toy, the Humming Box exploits cranking from traditional music boxes (Figure 2) as the primary interaction of generating sound. According to Don Norman, establishing natural mappings can significantly reduce the cognitive burden [29] for children. By mounting the control of rhythm on the hand crank itself, we aim to use the cranking affordance to convey instructions about playing sound. In addition to generating sound, multiple modular designs have been built, allowing children to modify the timbre with natural physical gestures as a way of body-syntonic learning [30]. In the latest version, as shown in Figure 3, we designed and laser-cut three transparent acrylic cards with edged instrument icons on the surface and encoded curve edge at the bottom. When a child inserts a card into the slot on top of the box, the encoded curve edge will trigger a unique set of mechanical buttons inside, which eventually toggle the instrument on GarageBand application from the connected device (e.g., MacBook or iPad). In a nutshell, we hope to build on old patterns of play and create a vibrant and generative play experience by blending in new tangible ways to manipulate and remix songs.

Final Design
The Humming Box has three main components (Figure 4): a wooden box prototype with a USB 2.0 Micro cable, three transparent instrument cards and keys. It also requires a computer installed with GarageBand program.

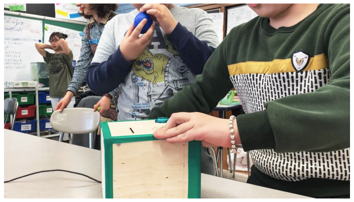

Figure 5: Children playing with our prototype in Playtest I (After-school session in Chinatown, New York)

Figure 6: Children interacting with Multimodal Prototype in Playtest II (Bring Your Child to Work Day at Tisch School of the Arts, NYU)

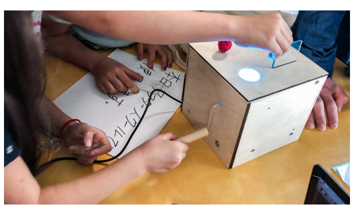

Figure 7: Children interacting with Final Prototype in Playtest III (a Saturday workshop at Brooklyn Nihonho Gakuen (BNG)

Set Up and Play
(1) Connect the Humming Box to GarageBand with a USB 2.0 Micro cable. (2) In GarageBand, set at least four instruments. Currently, we choose Guzheng, Electronic Piano, Guitar, and Funk. The order of instruments matters. (3) Use the key to unlock the box by spinning clockwise. A white animated LED ring will light up on top of the box. (4) Keep GarageBand active and start cranking the box. The "Baby Shark" melody will start to play. The speed of cranking would change the rhythm of the melody in real time.The LED ring will light up like a rainbow and gradually change color at the same pace as cranking. (5) Insert instrument card into the card slot on top of the box to switch the song's timbre. The different colors will flash on the LED ring as the box detects the different cards.

Iteration and Findings

We performed three playtests with approximately 50 children with age ranges 3-6 years old. We also gathered insights and feedback from more than 20 experts, and nearly 100 older children and adults from various backgrounds.

Version I: Key Interaction Prototype
We designed and built a wooden prototype box with a cranking handle connected to a rotary encoder, a button as the placeholder for recording and a bottomless slot. Although the prototype requires a laptop connection in order to output the MIDI signal to GarageBand, children are only interacting with the box. In this version, we used paper to represent instrument modules and manually toggle the timbre in GarageBand to mock-up the switching instrument effect.

Playtest I: After-school Session in Chinatown
Aiming to gather insights about how children would physically and emotionally interact with cranking, button pressing, and module inserting, we conducted the initial round of playtesting with 14 children (third-grades and fourth-grades) from an after-school session in Chinatown, New York. The children were grouped in two to interact (Figure 5) with the prototype for about 5 minutes, followed by a 2-minute interview session. We first observed their behavior, then raised questions to clarify children's choices among the interaction. Through observation, we found that nearly every child enjoyed pressing the button while cranking simultaneously. We were surprised that children got very involved in inserting almost everything they could find on the desk into the bottomless slot at once, which broke the prototype later. By inserting they expected timbre changes and were disappointed when the output failed to modify. We attributed this reaction to children's nature of exploring and that single sound output didn't meet their expectations. The impact of cranking speed variations on rhythm changes was not perceivable enough for most groups. Due to the lack of satisfying feedback, a reasonable focus time for children in this playtests ranged between 1 to 2 minutes. Therefore, a more precise mapping between cranking speed and rhythm output became critical, and more variations of feedback are needed to increase engagement and improve their playful experience. Besides, instrument module (card) inserting lacks useful feedback, in which case children tend to push cards to the bottom.

Version II: Multimodal Prototype
With lessons learned from the first playtest, we re-constructed the prototype, replacing the button placeholder and adding a NeoPixel ring as another visual feedback component. The NeoPixel ring was coded to perform light animation of spinning while changing in rainbow colors. We mapped the animation pace to the cranking speed. We designed, laser-cut, and laser-edged three rectangle acrylic cards as instrument switch modules. The card insertion mechanism was improved by attaching switches to a cardholder at the

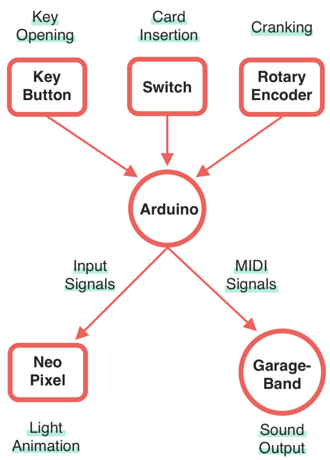

Figure 8: Technical structure of the latest version.

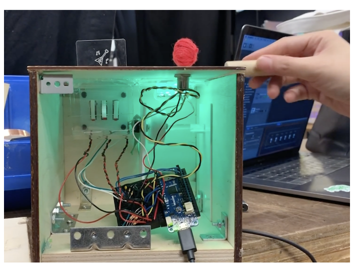

Figure 9: Mechanism anatomy.

bottom of the slot aiming to provide firmer support and the click feedback, which can indicate that the insertion had been detected and prevent users from further pushing the card through. At this stage, however, we were still manually switching the instruments in GarageBand. Throughout the process, we realized that the system was still in lack of clear state signal when it was initialized and ready for a new interaction session. Hence, we added a key switch for the “power control” effect. The keys also created a belonging relationship between children and the toy [8].

Playtest II: "Bring Your Child to Work" Day at Tisch

During the "Bring Your Child to Work" Day at Tisch School of the Arts, New York University, we invited around 10-15 children to play with the new prototype (Figure 6). The age of this group varies from 3 to 6 years old. Observations took the most part during this playtest. All the children responded actively to the light animation of the NeoPixel ring. By the first playtest, they also expressed strong interests when the timbre got toggled. Nevertheless, most failed to pick up the melody. Some younger children (approximately 3 years old) struggled a little bit with the first key. Overall, children reacted positively to the Humming Box. A girl even came back to play with it for a second time. As for card detection mechanism, the clicking feedback successfully kept children from inserting too far, but two children were confused about the orientation of cards at their first trials. Besides, we noticed that card inserting was not smooth enough that sometimes, cards would be stuck.

Version III: Final Prototype

Based on observations during the second playtest, we improved the rotary encoder reading algorithm such that it would stabilize the cranking input. We observed that children found the previous melody less attractive. As a result, we changed the MIDI notes to a more popular song, Baby Shark, in this version. Additionally, in order to provide appropriate feedback for card insertion detection, we added instant color changes to the NeoPixel ring. Each instrument card would trigger different color wipe out. Meanwhile, it would also send a switching command to GarageBand by triggering keyboard inputs (Figure 9). In this way, we achieved automatic timbre manipulation. Finally, we reorganized and refactored our code.

During the fabrication, the most challenging part was to re-design and build the card insertion mechanism. By bending the acrylic board, we built a cardholder with calibrated angles and slots for switches (Figure 8). Each card would trigger different switch combinations. Our observation about children’s confusion of the orientation of the cards guided us through adding a unique, symmetrical, encoded curve to the bottom of each card. Additionally, we wrapped the keys with wool to not only give it a more attractive look and more delicate texture but also make it easier for children to grab. We used transparent acrylic for cards to make them identical for both sides and also to allow the NeoPixel light to become visible. As a result, no matter how children inserted, they would always behave in the right way. We also stabilized the cranking mechanism and calibrated the rotation angles to ensure a smoother movement.

Playtest III: Japanese Elementary School and ITP Spring Show

We evaluated the latest prototype at the Brooklyn Nihongo Gakuen Saturday workshop (Figure 7) and the ITP Spring Show [17] with 10 boys and 12 girls. The age of the attendees varies between 3 to 6. Children expressed a strong willingness to play with multiple interactions, figure out connections between multimodal feedback, and share their findings with others. The productive interactions and improved melody significantly increased their focus time (around five minutes) and drove their curiosity to the next level. Moreover, some children invited friends to explore new possibilities cooperatively. They even primarily distributed responsibilities by assigning a representative to each interaction. The audience was also highly engaged by contributing thoughts and suggestions. When there are guardians, the Humming Box acted as a linkage of the cooperation between younger children and parents. Children usually initiated the interaction. Parents would guide children to finish more challenging tasks such as card insertion and key switching. They would also inspire children to experiment with different instrument cards, which lead to a structured play activity that benefit children from an integrated pedagogical approach [9]. During the ITP Spring Show [17], older children and adults from various professional fields expressed strong interests in the prototype. A 9-year-old boy immediately mastered basic interactions and established an original way to mix with the instrument choices.

Future Work

We aim to integrate multiple AI-powered pitch detection algorithms (e.g. Google Duet [24], HumOn [1]) to allow children sing directly to the box then play their own melody in future iterations of our Humming Box design. Moreover, we would also provide various pre-set melody selections as an alternative solution for children who are not interested in writing melody but like to explore music composing elements such as determining pitch, rhythm or even remixing. Additional modifications would be made to provide personal features for children, such as allowing customized instrument modules by RFID detection and encouraging customized painting to the box. Besides, we are looking into executive functions and uses of the instrument, trying to test with children on the spectrum based on dexterity. Children, without experiential knowledge, can be easily affected by the overwhelming technology explosion. Thus, we designers are responsible for shaping the next generation's perceptions and habits of new technologies. We hope to continue exploring the horizon of music composing for younger creators by offering a smoother form of tangible interaction.

Conclusion

We proposed a new music creation interaction for children 4-6 years old by leveraging new AI powered technologies to manipulate songs. The toy's modular design expands children's ability in exploration [36]. Likewise, the combination of sound and visual interactions would provide advantages for children's neurocognitive development [12]. Moreover, input such as grabbing, inserting and twisting the key switch, slotting instrument cards, cranking the candle can be beneficial to their finger development [31, 11] and hand-eye coordination ability. These interactions also decreased training barrier for music performing, yet, maintained the aesthetic sense to adjust tempo, pitch, etc. As we discovered in the playtests, multimodal modularized interactions increased the attractiveness and focus time for children. We will continue to explore advanced pitch detection and modification algorithms, hoping this pilot toy would inspire many new forms of music creation for kids that can successfully blend old toys with new AI-powered technologies.

Acknowledgements

We would also like to express our gratitude to the ISS afterschool program at P.S.126/ Manhattan Academy of Technology for the opportunity to test our prototype with their students. We thank our professors: Stefania Druga, Daniel Shiffman, Tom Igoe and the ITP residents who provided insight and expertise that greatly assisted the our research and technical issues.

REFERENCES

1. 2017. HumOn - The Simplest Music Creation App. (2017). Retrieved June 24, 2019 from http://hum-on.com/
2. 2017. NYU Music Experience Design Lab. (2017). Retrieved June 24,2019 from https://musedlab.org/
3. 2019. FabLearn 2019 8211; 8th Annual Conference on Maker Education. (Jun 2019). Retrieved June 23, 2019 from https://nyc2019.fablearn.org/
4. ANKI. 2019. Cozmo | Meet Cozmo. (2019). Retrieved June 30, 2019 from https://anki.com/en-us/cozmo.html
5. Robert Anthony. 2013. Beginner's Introduction to Composing on the Computer - Software. (Sep 2013). Retrieved July 3, 2019 from https://music.tutsplus.com/tutorials/beginners-introduction-to-composing-on-the-computer-software--audio-19607
6. Dembo T. Lewin K. Barker, R. 1941. Frustration and regression: an experiment with young children. *University of Iowa Studies: Child Welfare* 18, 1 (1941), xv + 314. https://psycnet.apa.org/record/1941-04439-001
7. Terry D Bilhartz, Rick A Bruhn, and Judith E Olson. 1999. The Effect of Early Music Training on Child Cognitive Development. *Journal of Applied Developmental Psychology* 20, 4 (Dec. 1999), 615–636. DOI: http://dx.doi.org/10.1016/S0193-3973(99)00033-7
8. Peter R. Blake and Paul L. Harrisn. 2009. Children's understanding of ownership transfers. *Cognitive Development* 24, 2 (2009), 133–145. DOI: http://dx.doi.org/10.1016/j.cogdev.2009.01.002
9. P. Broadhead, J. Howard, and E. Wood. 2010. *Play and Learning in the Early Years: From Research to Practice*. SAGE Publications, Thousand Oaks, CA, USA.
10. Adam November Ethan Hein and Alex Ruthmann. 2017. NYU Music Experience Design Lab. (2017). Retrieved June 30, 2019 from https://musedlab.org/groovepizza
11. Centers for Disease Control and Prevention. 2019. Preschooler (3-5 years old) | CDC. (Feb 2019). Retrieved June 24,2019 from https://www.cdc.gov/ncbddd/childdevelopment/positiveparenting/preschoolers.html
12. Irini Giannopulu. 2013. Multimodal interactions in typically and atypically developing children: natural versus artificial environments. *Cognitive Processing* 14, 4 (Nov. 2013), 323–331. DOI: http://dx.doi.org/10.1007/s10339-013-0566-0
13. Lifelong Kindergarten Group. 2003-2019. Scratch - Imagine, Program, Share. (2003-2019). Retrieved June 24,2019 from https://scratch.mit.edu/
14. Susan Hallam. 2010. The power of music: Its impact on the intellectual, social and personal development of children and young people. *International Journal of Music Education* 28, 3 (Aug. 2010), 269–289. DOI: http://dx.doi.org/10.1177/0255761410370658
15. Carolyn Hildebrandt. 1998. Creativity in Music and Early Childhood. *ERIC - Young Children* 53, 6 (Nov. 1998), 68–74. https://eric.ed.gov/?id=EJ576005
16. Corinne Hutt. 1970. Curiosity in young children. *Science Journal* 6, 2 (1970), 68–71. https://psycnet.apa.org/record/1970-20592-001

17. ITP. 2019. ITP/IMA Spring Show 2019. (May 2019). Retrieved June 25, 2019 from https://tisch.nyu.edu/itp/events/spring-2019/save-the-date--itp-spring-show-2019-

18. Inc. Jibo. 2019. Jibo Robot - He can't wait to meet you. (2019). Retrieved June 30, 2019 from https://www.jibo.com/

19. Davis Katie, Weigel Margaret, James Carrie, and Gardner Howard. 2009. *Social Development in the Era of New Digital Media*. Technical Report. Cambridge, MA, USA.

20. Arthur Lefford, Herbert G. Birch, and George Green. 1974. The Perceptual and Cognitive Bases for Finger Localization and Selective Finger Movement in Preschool Children. *Child Development* 45, 2 (June 1974), 335–343. DOI: http://dx.doi.org/10.2307/1127953

21. LEGO. 2019. Build Students' Confidence in STEAM and STEM Learning with LEGO Education. (2019).

22. littleBits Electronics Inc. 2011-2019. littleBits | Electronic Building Blocks for the 21st Century. (2011-2019). Retrieved June 24,2019 from https://littlebits.com/

23. Makey Makey LLC. 2018. Makey Makey – Makey Shop. (2018). Retrieved July 3, 2019 from https://makeymakey.com/

24. Yotam Mann. 2017. AI Duet by Yotam Mann | Experiments with Google. (May 2017). Retrieved June 24,2019 from https://experiments.withgoogle.com/ai-duet

25. Jr. Marcian E. Hoff. 2019. Entertainment and creative expression device for easily playing along to background music. (Jun 2019). Retrieved Jun 26 2019 from https://patents.google.com/patent/US4771671A/en#citedBy Patent No. US4771671A, Filed Jan. 8th.,1987, Issued Sep. 20th., 1988.

26. G. Mazzola. 2018. *The Topos of Music II: Performance: Theory, Software, and Case Studies*. Springer International Publishing, New York, NY, USA.

27. Inc. Munchkin. 2019. Mozart Magic Cube | Musical Toys for Babies | Munchkin. (2019). Retrieved July 3, 2019 from https://www.munchkin.com/mozart-magic-cube-new.html

28. Amanda Niland. 2009. The Power of Musical Play: The Value of Play-Based, Child-Centered Curriculum in Early Childhood Music Education. *General Music Today* 23, 1 (April 2009), 17–21. DOI: http://dx.doi.org/10.1177/1048371309335625

29. Donald A. Norman. 2002. *The Design of Everyday Things*. Basic Books, Inc., New York, NY, USA.

30. Seymour Papert. 1980. *Mindstorms: children, computers and powerful ideas*. Basic Books, Inc., New York, NY, USA.

31. MD Renee A. Alli. 2018. 4-to-5 Year Old Child Developmental Milestones. (Oct 2018). Retrieved June 24,2019 from https://www.webmd.com/parenting/4-to-5-year-old-milestones#2

32. Learning Resources. 2019. Coding Critters™ Rumble & Bumble. (2019). Retrieved June 30, 2019 from https://www.learningresources.com/coding-critters-dino

33. Eric Rosenbaum. 2008. MmmTsss Playful Looper. (Jul 2008). Retrieved June 30, 2019 from http://web.mit.edu/~eric_r/Public/mmmtsss/

34. Sphero. 2018. Sphero. (2018). Retrieved June 30, 2019 from https://www.sphero.com/
35. Susan. Young. 2008. *Music 3-5*. Routledge, Abingdon-on-Thames, UK.
36. Jiangmei Yuan and Robin T Bowen. 2018. Lifelong Kindergarten: Cultivating Creativity Through Projects, Passion, Peers, and Play. *Interdisciplinary Journal of Problem-Based Learning* 12, 2 (Aug. 2018). DOI: http://dx.doi.org/10.7771/1541-5015.1797