Speculating on Reflection and People's Music Co-Creation with AI

COREY FORD, Queen Mary University Of London
NICK BRYAN-KINNS, Queen Mary University of London

Recent advances in generative Artificial Intelligence (AI) and Human-Computer Interaction (HCI) have led to a design trend of user interfaces supporting novices in co-creating with AI. Typically, these tools aim to extend novices' abilities and democratise the more difficult aspects of creative practice. We suggest that designing for reflection might be an interesting alternative avenue to explore in novice's creative co-creation with generative AI. We contribute a systematic literature review of 122 creative AI co-creators – mostly for music – with respect to HCI theories on reflection. We then explore findings from our literature review within the domain of music, suggesting an interface design for an interactive tool based on the current trend in music AI to expose the latent space of variational autoencoders.

CCS Concepts: • **Applied computing** → **Sound and music computing**; • **Human-centered computing** → **Interactive systems and tools**; • **Computing methodologies** → *Machine learning*; *Artificial intelligence*.

Additional Key Words and Phrases: machine learning, generative, creative AI, reflection, music, composition, co-creation, human-AI interaction, evaluation, novice, creativity, literature review, computational creativity

Reference:
Corey Ford and Nick Bryan-Kinns. 2022. Speculating on Reflection and People's Music Co-Creation with AI. In *Generative AI and HCI Workshop at CHI 2022, May 10.* Online, 6 pages.

1 INTRODUCTION

Advances in computing power have enabled the development of increasingly complex Artificial Intelligence (AI), especially with respect to machine learning models and deep neural networks [9, 13]. In particular, generative techniques such as generative adversarial networks [15] and variational autoencoders (VAEs) [22] have made it easier for people to incorporate generated outputs into their own artistic creations [13]. Indeed, it is becoming increasingly accessible for people to co-create creative artefacts with AI.

The evaluation of such tools for novice's co-creation with generative AI systems have generally been inspired by 'third-wave' HCI techniques [7], examining how tools support engagement as opposed to ease-of-use [6, 7]. For example, see [1, 10, 19, 23]. However, this does not necessarily support novices in exploring some of the holistic benefits of creative practice, such as growing themselves artistically (for example, a composer might consider how their music meets their artistic intentions [21, 42]) or complementing their understanding of other creative experiences (for example, composers might consider how performers might interpret their notations [21, 42]). In this paper, we take an alternative viewpoint which emphasises reflection in the co-creative process between people and generative AI systems. Indeed, we suggest that an alternative role for AI in a collaborative creative process could be to encourage reflection. We explore this through the lense of music AI co-creators, first i) systematically exploring reflection in *musical* co-creative AI literature, and ii) suggesting the design of a user interface to provide opportunities for reflection in novice's *music* co-creation with AI.

2 REFLECTION

Reflection is an ill-defined concept. We propose, based on related work [2, 5, 26, 35], that reflection is a process where different solutions are contemplated during or after moments of uncertainty. Key to this definition is the notion of

uncertainty — reflection is prompted at points where a person's actions lead to results that do not match their tactic know-how [2, 35]. It is also notable that reflection occurs both in-action and on-action, as discussed by Schön [35], who is frequently cited in HCI research on reflection [5]. Designing for reflection was a topic of concern in HCI since around the early 2000s [5] but saw a surge in interest following a CHI workshop [34] and a number of review papers mapping out the field [4, 5, 12] around the 2010s. Notably, Fleck and Fitzpatrick [12] reviewed interdisciplinary literature on reflection to produce a pragmatic framework for interaction designers, describing the ways in which computers could support different levels of reflection from R0 (no elaboration on a description of events) to R4 (considering wider issues). A key idea was to encourage people to reflect by displaying novel information — although others highlight that people also need a reason to reflect and encouragement to do so [37]. A different perspective is Gaver, Beaver and Benford's [14] work on designing intentionally suggestive artefacts to encourage reflection — for example, by implicating incompatible contexts into one design.

2.1 Reflection in Human-AI Co-creation

As an initial investigation into the landscape of reflection in human-AI co-creators, a corpus of 122 papers from 2010-2020 were reviewed. 82 papers were gathered from the New Instruments for Musical Expression (NIME) conference series because it has a rich catalogue of systems designed to enhance novices creativity. 24 papers were also gathered from the International Conference on Computational Creativity (ICCC) because it is a relatively young conference (starting in 2010), distinctly exploring the role that computers play in creative practices. To identify papers, text from each conference paper's PDF was extracted and scanned for the keywords used in Pessoa *et al.* [30] to discover interfaces targeting novices. The abstract of these papers were then manually analysed before inclusion. 13 more papers were also added from Herreman *et al.*'s [16] review of generative music systems, selecting those which included a user interface, alongside three more papers recommended by colleagues [10, 23, 31], to capture a broader selection of AI systems.

Each paper was scored against a rubric describing design features which encourage reflection, listed below alongside examples of papers found in our literature review. This was based largely on Fleck and Fitzpatrick [12] because it provides a pragmatic operationalisation for determining the amount of reflection an interface might afford.

- **R1: The interface provides a log of knowledge and experience, including mechanisms for sharing designs (e.g. through online forums).** For example, MicroJam [24, 25] is an app where tiny musical performances can be created by drawing lines onto a screen. A log of these touch-screen interactions are then stored, which can be shared online to encourage reflection within a community.
- **R2: The interface offers reflective prompts, annotates people's creations, or encourages collaboration during the creative process.** For example, Huang and Chew's Palestrina Pal [18] analyses and annotates MIDI files provided by the user against a set of music theory rules, annotating melodic errors in their graphical user interface which people could reflect upon.
- **R3a: The interface presents more information than people are able to usually analyse.** For example, Holbrow, Jessop and Kleinberger's [17] ORB — a round ball held in the hand — produces "feedback about[...] the vibrations produced in a person's body" [17, pg. 433] when using their voice. People were thus more aware of the "vibration(s) caused by amplitude, frequency, and timbre" [17, pg. 433] beyond what they may usually notice when singing, providing opportunities for novel reflections on their creative practice.
- **R3b: The interface presents suggestive elements within the context of the design (cf. [14]).** For example, the Creative Sketching Partner [20] allows people to draw a sketch and then responds with an similar image,

but purposefully quantifies conceptual information to generate images with a high novelty, suggesting new ideas to people which might prompt reflection during their artistic process.

An overview of our survey results are visualised in Figure 1. 76/122 papers did not encourage reflection. Of the remaining 46 papers, 6 papers reached R1, 23 reached R2, 15 reached R3a, and 2 reached R3b. Indeed, at the highest levels of reflection, 15 papers provided extra information to people (R3a), whereas only 2 papers used 'suggestive' designs (R3b). No NIME papers in the corpus used 'suggestive' designs — the general trend was to use simplistic mappings, lowering the barrier of entry for novices, perhaps following advice from seminal works [36, 39, 41]. In summary, we found that using suggestive designs (R3b) to encourage reflection was largely unexplored in novice co-creators.

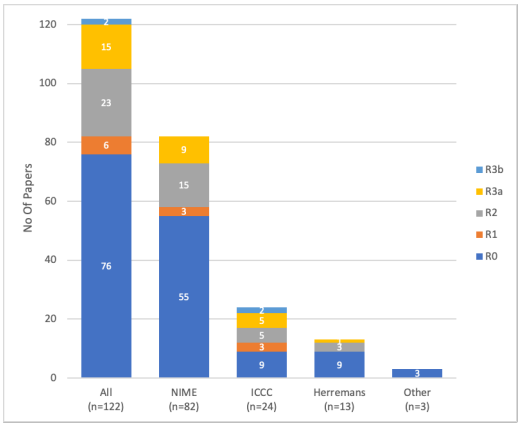

Fig. 1. A chart showing the number of papers for each category of reflection, split by publication venue.

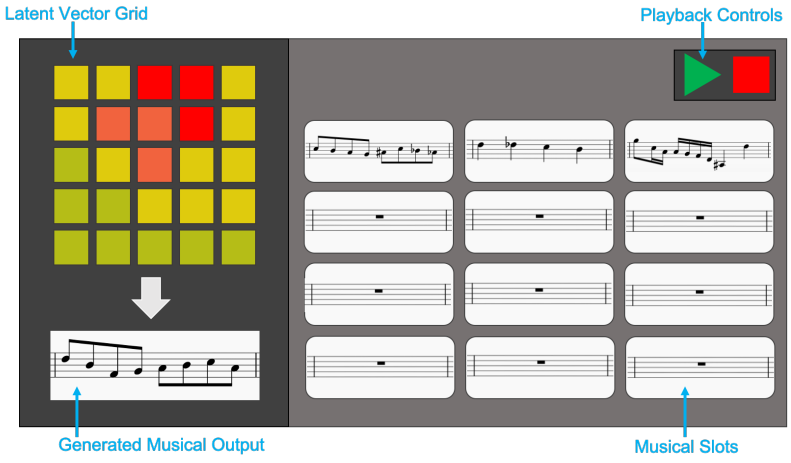

Fig. 2. A low fidelity mock-up of a latent space based interface, designed for exploring reflection in novice's music co-creation with AI.

3 USER INTERFACE DESIGN PROPOSAL

We take generative music AI systems as a key form of creative practice in which to explore how reflection could be used in the design of co-creative systems. Generative models have made an impact in music, where researchers have leveraged generative AI to scaffold novice's music composition process [1, 10, 11, 19, 23, 27], making it "conceivable for [people] with no composition training to create music" [23]. Indeed, generative AI has been used to support people's music composition process in multiple ways, such as by evaluating their creations against music theory [18], acting as an improvisation partner [27], or composing harmonies [19]. Specifically, we decided to focus on a current trend in human-AI music co-creation to expose the latent space of generative music models [23, 29, 38], allowing people to create their own music by sampling from a compressed possibility space — for example, latent space models have been used for generating [40, 43], exploring [8] and in-painting [29] music.

A low fidelity mock-up of our proposed interface is shown in Figure 2. On the left hand side is a grid mapping to a VAEs latent vector. People can select points from this grid to generate short musical fragments, displayed as music notation below the latent grid. These fragments can then be dragged into the slots on the right to create musical compositions, and played back (from left to right) using the playback controls. In our design, we aimed to incorporate features encouraging the higher levels of reflection, R3a and R3b, as described above. Firstly, we were inspired by related work [3, 28] to implement a heat map within the latent vector visualisation (see top left of Figure 2). Hot areas thus represent locations in the latent space that people have frequently explored, making it possible for them to analyse how they interacted with the latent space. By identifying areas in the latent space that they have not recently used, people are presented with more information than they would otherwise have been able to analyse, satisfying R3a. Furthermore, we suggest that the underlying generative model could sample from two VAEs trained on contrasting musical datasets — such as Bach pieces [19] and rock guitar pieces [33]. During the composition process, we could then use these outputs (displayed in the bottom left of Figure 2) to interpolate within the latent space of a larger, pre-trained music generation model (such as [32]), outputting fragments of generated music which are in some way 'between genres'. This would suggest new elements which could provide more opportunities for people to reflect on their music whilst composing, satisfying R3b. This might also cause uncertainty as outputs are blending between incompatible contexts, perhaps encouraging reflection in a different way.

4 CONCLUSION

By reviewing 122 papers on novice's co-creation with generative AI, we found that the majority of creative AI tools for novices in our sample did not offer opportunities for reflection. As first steps towards exploring this further, we then proposed a user interface design, motivating our work within the domain of music. In future work, we plan to make our proposed interface a reality, exploring how novices might engage in reflection during interaction. We are curious as to whether i) novices will engage in reflection more so when areas of the latent space they frequently explore are visualised, and ii) understand how people might react when interacting with seemingly incompatible latent space representations from different music genres. Overall, we hope to better understand how reflection occurs in human-AI interactions, and strive to demonstrate the validity of our approach of designing for reflection.

ACKNOWLEDGMENTS

The first author is a research student at the UKRI Centre for Doctoral Training in Artificial Intelligence and Music, supported by UK Research and Innovation [grant number EP/S022694/1].

REFERENCES

[1] Anna Rita Addessi, Laura Ferrari, and Felice Carugati. 2015. The Flow Grid: A Technique for Observing and Measuring Emotional State in Children Interacting with a Flow Machine. *Journal of New Music Research* 44, 2 (April 2015), 129–144. https://doi.org/10.1080/09298215.2014.991738

[2] Pengcheng An, Saskia Bakker, Sara Ordanovski, Ruurd Taconis, Chris L.E. Paffen, and Berry Eggen. 2019. Unobtrusively Enhancing Reflection-in-Action of Teachers through Spatially Distributed Ambient Information. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3290605.3300321

[3] Jeremy N. Bailenson, Nick Yee, Jim Blascovich, Andrew C. Beall, Nicole Lundblad, and Michael Jin. 2008. The Use of Immersive Virtual Reality in the Learning Sciences: Digital Transformations of Teachers, Students, and Social Context. *Journal of the Learning Sciences* 17, 1 (2008), 102–141. https://doi.org/10.1080/10508400701793141

[4] Eric P.S. Baumer. 2015. Reflective Informatics: Conceptual Dimensions for Designing Technologies of Reflection. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 585–594. https://doi.org/10.1145/2702123.2702234

[5] Eric P.S. Baumer, Vera Khovanskaya, Mark Matthews, Lindsay Reynolds, Victoria Schwanda Sosik, and Geri Gay. 2014. Reviewing Reflection: On the Use of Reflection in Interactive System Design. In *Proceedings of the 2014 Conference on Designing Interactive Systems* (Vancouver, BC, Canada) *(DIS '14)*. Association for Computing Machinery, New York, NY, USA, 93–102. https://doi.org/10.1145/2598510.2598598

[6] M.A. Blythe, K. Overbeeke, A.F. Monk, and P.C. Wright. 2004. *Funology: From Usability to Enjoyment*. Springer, Dordrecht. https://doi.org/10.1007/1-4020-2967-5

[7] Susanne Bødker. 2015. Third-wave HCI, 10 years later — participation and sharing. *Interactions* 22, 5 (aug 2015), 24–31. http://dl.acm.org/citation.cfm?doid=2818696.2804405

[8] Nick Bryan-Kinns, Berker Banar, Corey Ford, Courtney N. Reed, Yixiao Zhang, Simon Colton, and Jack Armitage. 2021. Exploring XAI for the Arts: Explaining Latent Space in Generative Music. In *Proceedings of the 1st Workshop on eXplainable AI Approaches for Debugging and Diagnosis (XAI4Debugging@NeurIPS2021)*. https://xai4debugging.github.io/files/papers/exploring_xai_for_the_arts_exp.pdf

[9] Filippo Carnovalini and Antonio Rodà. 2020. Computational Creativity and Music Generation Systems: An Introduction to the State of the Art. 3 (2020). https://www.frontiersin.org/article/10.3389/frai.2020.00014

[10] Chris Donahue, Ian Simon, and Sander Dieleman. 2019. Piano Genie. In *Proceedings of the 24th International Conference on Intelligent User Interfaces* (Marina del Ray, California) *(IUI '19)*. Association for Computing Machinery, New York, NY, USA, 160–164. https://doi.org/10.1145/3301275.3302288

[11] Morwaread Farbood, Henry Kaufman, and Kevin Jennings. 2007. Composing with Hyperscore: An Intuitive Interface for Visualising Musical Structure. In *Proceedings of the 2007 International Computer Music Conference (ICMC '07)*. 111–117.

[12] Rowanne Fleck and Geraldine Fitzpatrick. 2010. Reflecting on Reflection: Framing a Design Landscape. In *Proceedings of the 22nd Conference of the Computer-Human Interaction Special Interest Group of Australia on Computer-Human Interaction* (Brisbane, Australia) *(OZCHI '10)*. Association for Computing Machinery, New York, NY, USA, 216–223. https://doi.org/10.1145/1952222.1952269

[13] David Foster. 2019. *Generative Deep Learning: Teaching Machines to Paint, Write, Compose, and Play*. O'Reilly Media.

[14] William W. Gaver, Jacob Beaver, and Steve Benford. 2003. Ambiguity as a Resource for Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. Association for Computing Machinery, New York, NY, USA, 233–240. https://doi.org/10.1145/642611.642653

[15] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. Generative Adversarial Nets. In *Proceedings of Advances in Neural Information Processing Systems*.

[16] Dorien Herremans, Ching-Hua Chuan, and Elaine Chew. 2017. A Functional Taxonomy of Music Generation Systems. *Comput. Surveys* 50, 5, Article 69 (Sept 2017). https://doi.org/10.1145/3108242

[17] Charles Holbrow, Elena Jessop, and Rebecca Kleinberger. 2014. Vocal Vibrations: A Multisensory Experience of the Voice. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Goldsmiths, University of London, London, United Kingdom, 431–434. http://www.nime.org/proceedings/2014/nime2014_378.pdf

[18] Cheng Zhi Anna Huang and Elaine Chew. 2005. Palestrina Pal: A Grammar Checker for Music Compositions in the Style of Palestrina. In *Proceedings of the 5th Conference on Understanding and Creating Music*.

[19] Cheng-Zhi Anna Huang, Curtis Hawthorne, Adam Roberts, Monica Dinculescu, James Wexler, Leon Hong, and Jacob Howcroft. 2019. The Bach Doodle: Approachable music composition with machine learning at scale. In *International Society for Music Information Retrieval (ISMIR)*. https://goo.gl/magenta/bach-doodle-paper

[20] Pegah Karimi, Mary Lou Maher, Nicholas Davis, and Kazjon Grace. 2019. Deep Learning in a Computational Model for Conceptual Shifts in a Co-Creative Design System. In *Proceedings of the Tenth International Conference on Computational Creativity* (Charlotte, North Carolina, USA) *(ICCC)*, Kazjon Grace, Michael Cook, Dan Ventura, and Mary Lou Maher (Eds.). Association for Computational Creativity, 17–24. http://computationalcreativity.net/iccc2019/papers/iccc19-paper-40.pdf

[21] Michele Kaschub and Janice Smith. 2009. *Minds on Music: Composition for Creative and Critical Thinking*. R&L Education.

[22] Diederik P Kingma and Max Welling. 2013. Auto-Encoding Variational Bayes. https://arxiv.org/abs/1312.6114

[23] Ryan Louie, Andy Coenen, Cheng Zhi Huang, Michael Terry, and Carrie J. Cai. 2020. Novice-AI Music Co-Creation via AI-Steering Tools for Deep Generative Models. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for

Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376739

[24] Charles Martin and Jim Torresen. 2017. MicroJam: An App for Sharing Tiny Touch-Screen Performances. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Aalborg University Copenhagen, Copenhagen, Denmark, 495–496. http://www.nime.org/proceedings/2017/nime2017_paper0096.pdf

[25] Charles Patrick Martin and Jim Torresen. 2019. Data-Driven Analysis of Tiny Touchscreen Performance with MicroJam. *Computer Music Journal* 43, 4 (12 2019), 41–57. https://doi.org/10.1162/comj_a_00536 arXiv:https://direct.mit.edu/comj/article-pdf/43/4/41/2005769/comj_a_00536.pdf

[26] Jennifer A Moon. 2013. *Reflection in Learning and Professional Development: Theory and Practice*. Routledge.

[27] François Pachet. 2003. The Continuator: Musical Interaction With Style. *Journal of New Music Research* 32, 3 (2003), 333–341. https://doi.org/10.1076/jnmr.32.3.333.16861

[28] Olena Pastushenko, Wilk Oliveira, Seiji Isotani, and Tomáš Hruška. 2020. A Methodology for Multimodal Learning Analytics and Flow Experience Identification within Gamified Assignments. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3334480.3383060

[29] Ashis Pati, Alexander Lerch, and Gaëtan Hadjeres. 2019. Learning To Traverse Latent Spaces For Musical Score Inpainting. In *Proceedings of the 20th International Society for Music Information Retrieval Conference*. https://archives.ismir.net/ismir2019/paper/000040.pdf

[30] Margarida Pessoa, Cláudio Parauta, Pedro Luís, Isabela Corintha, and Gilberto Bernardes. 2020. Examining Temporal Trends and Design Goals of Digital Music Instruments for Education in NIME: A Proposed Taxonomy. In *Proceedings of the International Conference on New Interfaces for Musical Expression*, Romain Michon and Franziska Schroeder (Eds.). Birmingham City University, Birmingham, UK, 591–595. https://www.nime.org/proceedings/2020/nime2020_paper115.pdf

[31] Adam Roberts, Jesse Engel, Yotam Mann, Jon Gillick, Claire Kayacik, Signe Nørly, Monica Dinculescu, Carey Radebaugh, Curtis Hawthorne, and Douglas Eck. 2019. Magenta Studio: Augmenting Creativity with Deep Learning in Ableton Live. In *Proceedings of the International Workshop on Musical Metacreation (MUME)*. http://musicalmetacreation.org/buddydrive/file/mume_2019_paper_2/

[32] Adam Roberts, Jesse Engel, Colin Raffel, Curtis Hawthorne, and Douglas Eck. 2018. A Hierarchical Latent Vector Model for Learning Long-Term Structure in Music. In *Proceedings of the 35th International Conference on Machine Learning* (Stockholm, Sweden). https://goo.gl/magenta/musicvae-code

[33] Pedro Sarmento, Adarsh Kumar, CJ Carr, Zack Zukowski, Mathieu Barthet, and Yi-Hsuan Yang. 2021. DadaGP: A Dataset of Tokenized GuitarPro Songs for Sequence Models. In *Proceedings of the 20th International Society for Music Information Retrieval Conference*. https://arxiv.org/abs/2107.14653

[34] Corina Sas and Alan Dix. 2009. Designing for Reflection on Experience. In *CHI '09 Extended Abstracts on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI EA '09)*. Association for Computing Machinery, New York, NY, USA, 4741–4744. https://doi.org/10.1145/1520340.1520730

[35] Donald A. Schön. 1983. *The Reflective Practitioner. How Professionals Think in Action*. Basic Books Inc., London.

[36] Ben Shneiderman, Gerhard Fischer, Mary Czerwinski, Mitch Resnick, Brad Myers, Linda Candy, Ernest Edmonds, Mike Eisenberg, Elisa Giaccardi, Thomas T. Hewett, Pamela Jennings, Bill Kules, Kumiyo Nakakoji, Jay Nunamaker, Randy Pausch, Ted Selker, Elisabeth Sylvan, and Michael Terry. 2006. Creativity Support Tools: Report From a U.S. National Science Foundation Sponsored Workshop. *International Journal of Human-Computer Interaction* 20, 2 (2006), 61–77. https://doi.org/10.1207/s15327590ijhc2002_1

[37] Petr Slovák, Christopher Frauenberger, and Geraldine Fitzpatrick. 2017. Reflective Practicum: A Framework of Sensitising Concepts to Design for Transformative Reflection. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 2696–2707. https://doi.org/10.1145/3025453.3025516

[38] Notto J. W. Thelle and Philippe Pasquier. 2021. Spire Muse: A Virtual Musical Partner for Creative Brainstorming. In *Proceedings of the International Conference on New Interfaces for Musical Expression*. Shanghai, China. https://nime.pubpub.org/pub/wcj8sjee

[39] Marcelo Mortensen Wanderley and Nicola Orio. 2002. Evaluation of Input Devices for Musical Expression: Borrowing Tools from HCI. *Computer Music Journal* 26, 3 (2002), 62–76. http://www.jstor.org/stable/3681979

[40] Ziyu Wang, Yiyi Zhang, Yixiao Zhang, Junyan Jiang, Ruihan Yang, Junbo Zhao, and Gus Xia. 2020. Pianotree VAE: Structured Representation Learning for Polyphonic Music. In *Proceedings of the International Society for Music Information Retrieval Conference* (Montréal, Canada). https://arxiv.org/abs/2008.07118

[41] David Wessel and Matthew Wright. 2001. Problems and Prospects for Intimate Musical Control of Computers. In *Proceedings of the International Conference on New Interfaces for Musical Expression* (1-2 April, 2001). Seattle, WA, 11–14. http://www.nime.org/proceedings/2001/nime2001_011.pdf

[42] Arnold Whittall. 2002. Composition. In *The Oxford Companion to Music*, Alison Latham (Ed.). Oxford University Press, 279–283.

[43] Ruihan Yang, Dingsu Wang, Ziyu Wang, Tianyao Chen, Junyan Jiang, and Gus Xia. 2019. Deep Music Analogy via Latent Representation Disentanglement. In *Proceedings of the International Society for Music Information Retrieval Conference* (Delft, The Netherlands). https://archives.ismir.net/ismir2019/paper/000072.pdf