Reshaping the songwriting process with AI

Removed for Blind Review

Abstract

This article is rooted in the professional practice of the lead author as a songwriter and producer of popular music and focuses on the application of artificial intelligence (AI)-based tools in the songwriting process. First, we look at the structure of the songwriting process and report on how various AI tools fit into this structure, from the standpoint of the assistance they supply to the songwriter at various stages of the process but also the need they require for human curation and correction. We also discuss a method to measure the proportion of AI in a song but refocus the understanding of AI contribution on the generation of variations curated by a human, rather than a perhaps misled understanding as a fully creative contribution. Then, we examine the value that AI-based tools contribute in terms of addressing a concrete range of songwriter's needs, then conclude on how AI tools are reshaping the songwriting process in ways which are preserving and supporting the agency of the songwriter.

Introduction and background

This article analyses how the usage of a variety of AI tools by a professional music songwriter and producer can reshape the songwriting artistic practices. Many market-leading AI music generators advertise quicker music creation for a variety of applications, straightforward and comprehensive user interface and mimicking older and modern composers alike (McFarland 2023). Whilst it is easy to take such software tools at face value and to incorporate them into a workflow, first it is useful

to understand why songwriters might need quicker music creation or might have the desire to mimic other composers. As AI-based tools are becoming increasingly implemented throughout the artistic fields (Sturm et. al 2019; Louie et. al 2020; Knotts and Collins 2020), there is also a need to define and analyse the use cases where these tools can provide amateur musicians with new and innovative ways to improve their processes and tell their stories (Deruty et. al 2022), thus delivering support rather than putting an end to human creativity or storytelling (Holzapfel, Sturm, and Coeckelbergh 2018).

Much of the recent institutional research available on music and AI is centred around the task of automatic music composition which has gradually received more and more attention over the last decades. The versatility of AI technology allows for a wide variety of musical applications to be addressed (Briot, Hadjeres, and Pachet 2020). Notable examples include the imitation of a specific style such as *DeepBach* (Hadjeres, Pachet, and Nielsen 2017) which can generate chorales in the style of Bach, and *folkRNN* (Sturm, Santos, and Korshunova 2015) which generates folk Irish music snippets from scratch. Other applications include interpolation between music content (Robert et al. 2018), multitrack generation (Dong et al. 2018), or experimental approaches such as the practice of Holly Herndon (2022) where she uses AI to transform her voice while performing. The work presented in our paper is focused on popular music songwriting practice, where the usage of AI is gradually growing (Avdeeff 2019; Huang et. al 2020; Micchi et. al 2021). Focusing on a popular genre and the detail of its production process is meant to help cultivating greater public understanding of AI technologies and their implications. From a commercial perspective, this is also meant to address the needs of most audio equipment consumers, who may have a greater need for empowerment through AI than the users who are expert producers in the above-cited genres.

One of the salient incentives to involve AI in the production of popular music is economic viability. Indeed, in 2018 it was shown that it takes an average of 5.34 songwriters to create the top 100 biggest hit singles (https://tinyurl.com/2facy6eu). In April 2020, it was calculated that it would take 357 streams to make £1, therefore an artist would need 3,114 streams to earn £8.72 which is the UK's minimum wage per hour. These figures assume that the artist owns one hundred percent of the rights, which is rarely the case (Murray 2020). Therefore, dividing what little royalties are paid from streaming services between multiple writers makes it difficult to create a financially viable career from songwriting alone. Such economic incentives may prompt music producers to look for ways to increase the efficiency and throughput of the songwriting process and to remove co-writers who would require credit and a cut of the royalties, but without hampering a natural sense of creativity or musical identity. One of the purposes of this article is thus to analyse the use cases of a variety of contemporary AI-based tools in the context of songwriting, to determine if and how AI tools may change the songwriting process in ways which increase or hamper efficiency.

Beyond efficiency, it is crucial to understand the validity that AI-based popular music production could have in a commercial music consumption environment. Such surveys have already been conducted, such as the study by TickPick (https://www.tickpick.com/ai-drops-an-album) that was "designed with the intent of testing human ability to correctly identify AI-generated lyrics, as well as their opinion on creativity, emotionality, and favourability of various lyrics presented to them". This study resorted to GPT-2 (Radford et. al 2019) to generate lyrics across four of the "world's most popular genres of music", namely rock, rap, country and pop. No modifications were applied to the songs generated by GPT-2, except to

censor the expletives often found in hip-hop/rap songs. Across all genres, the AI-generated lyrics were more often attributed to a known songwriter than to an AI, e.g., for the country genre almost one in four people (22.2%) thought that AI lyrics were attributed to American Country Singer-Songwriter Garth Brooks and only 12.3% of participants could tell that the country lyrics were AI-generated. A similar experiment, this time addressing music in addition to lyrics, took place with *folkRNN* (Sturm, Santos, and Korshunova 2015) which produced 100,000 new folk tunes causing a diverse range of reactions from music experts in folk music and the public. To test the quality of the generated tunes, the authors invited professional Irish traditional musicians to create new songs based on these tunes. The outcome is a full-length album (Sturm and Ben-Tal 2018), most of which coming from material generated by the model. Specifically, of the album's 31 tracks, 20 were created using AI-generated tunes with various degrees of alteration by the professional musicians. The proportion of human agency versus unaltered AI-generated contents can be evaluated from the technical report related to the album. This research focuses on audience perceptions of AI-based music, whereas our article is rather focused on the introduction of AI usage within human practice.

Inserting AI tools in the songwriting process appears as a natural extension of a lineage. Indeed, computer-generated music appeared in the 1950s, with an initial focus on algorithmic music creation via probabilistic or stochastic methods (Hiller 1981; Xenakis 1992). Indeed, algorithmic music diverted the definition of human creativity away from the notions of genius or sudden lightbulb moments, to focus it instead on (a) the notions of variation followed by curation, akin to Campbell's blind variation/selective retention model of creativity (Simonton 2011), and (b) the proportion of authoring which can be encoded in an algorithmic structure rather than in the resulting musical score (Jacob 1996). Sampling techniques as well as AI-

based music generation today are posing exactly the same questions about the definitions and boundaries of originality: (a) given a sample collection, or a varied set of musical patterns generated by an AI, how much human selection and correction are needed to compose a viable and original song and (b) to which extent are a collection of pre-authored samples, or an AI which was machine-learned from such music samples, inheriting or encoding some of the original collection of author's styles and creativity? Examining the proportion of human agency naturally deployed by the first author of this article when using AI tools in professional songwriting practice will shed some light on these questions.

It is important to understand why AI by itself, versus used by a human author, cannot be creative by definition. Looking at AI based on contemporary machine learning which uses deep neural networks (DNNs): these require a large database to capture the concept, or perhaps the patterns, of what they are designed to generate (Najafabadi et al. 2015). The patterns found in the data after DNN optimisation are encoded as an immutable series of mathematical instructions, until a possible retraining of the model with new data – although DNNs are characterised by their high complexity, they still materialise as a series of programming instructions which cannot deviate from the programme. The development of AI-based music could therefore lead to the music creation industry being "locked in a loop of repetition, generating only minor incremental advancements" (Corpsey 2018). Developing one's own creative signature may rely on a proportion of prior references, or prior patterns, but is largely about breaking prior rules of taste, practice, technique, style etc. Along those lines, an effective definition of innovation is about entering a territory of "blindness" (Simonton 2011; Simonton 2012), or in other words finding ways to enter a domain where the rules are unknown and yet to be experimented with and discovered. If a computer was programmed to break the rules, then that

instruction would become a rule in itself: that leads to a paradox that machines cannot resolve but humans can. For example, “the only rule is that there is no rule” can be creatively resolved by a human, e.g., as a permission to cheat in a gaming context or as a permission to be ruthless in a battle context. A machine, on the other hand, would have no choice but to enter an infinite loop of breaking its own programme (Carter 2022).

From that standpoint, describing AI’s usage within popular music creation as “predominantly that of novelty, experimentation and largely as a tool for collaboration” (Avdeeff 2019) may imply a misuse of “collaboration” to describe the usage of a tool. Whilst there are a few instances of popular music produced with some involvement of AI tools, media stories might have entertained some confusion around the cases where the usage of AI was limited to generating backing tracks or automating production tasks. A prime example of this is Taryn Southern’s I AM AI (https://tarynsouthern.com/album), where she used tools such as IBM’s Watson Beat (https://www.ibm.com/watson/music/uk-en), AIVA (https://www.aiva.ai), Amper (https://www.ampermusic.com) and Google Magenta (https://magenta.tensorflow.org) to create only the instrumental aspects of her songs. Southern arranged the compositions, wrote vocal melodies and lyrics, while producer Ethan Carlson handled vocal production, mixing and mastering. Another example is SKYGGE’s Hello Word (Avdeeff 2019), which uses software Flow Machines (https://www.flow-machines.com), thus claiming AI usage, whereas Flow Machines is designed to “give music creators a plethora of options to play around with”. In both cases, the tools explicitly require control from the human operator to produce a structured and aesthetically pleasing result from partially structured data such as audio samples and style models, respectively. As such, the operation of those tools appears more akin to the control of a musical instrument or

to the operation of a sophisticated sampler than to the notion of collaboration with a colleague (Dennett and Lambert 2017). In other words, these cases are about someone working with something automatic rather than someone working with someone else with equal control capabilities. This questions the degree, or the definition of the so-called intelligence supplied by the machine (Dennett 1998).

Besides, without a human operator, another question is whether AI-generated pop music would be missing emotion or passion or might take away the human beauty and understanding behind music creation (Sturm et al. 2019). Thus, it might be that an AI won't create anything inherently creative nor emotional nor revolutionary, but a human creatively finding the right use cases for AI will. From that standpoint, it is useful to experiment with varying proportions of human agency versus AI-based automation at various stages of the songwriting process, in a scientific attempt to understand what may or may not make music creation inherently human.

The songwriting process

When choosing a set of tools to compose a song, it is important for a songwriter to determine whether they consider themselves a composer, a lyricist, or a songwriter who combines both. Furthermore, are they a single singer-songwriter or writing songs for a band which they are part of? Do they want to develop a career as a songwriter composing for other artists and individuals? Is the aspiration to co-write with other writers or producers or to offer services as a topliner writing melodies and lyrics for tracks? (Oliver, 2013). In the context of AI, this helps to identify which gaps in their songwriting skills could be filled by the usage of AI tools. For example, instrumentalists who are wishing to develop their toplining skills should seek out lyric and melody generators whilst lyricists may benefit from melody and accompaniment generators.

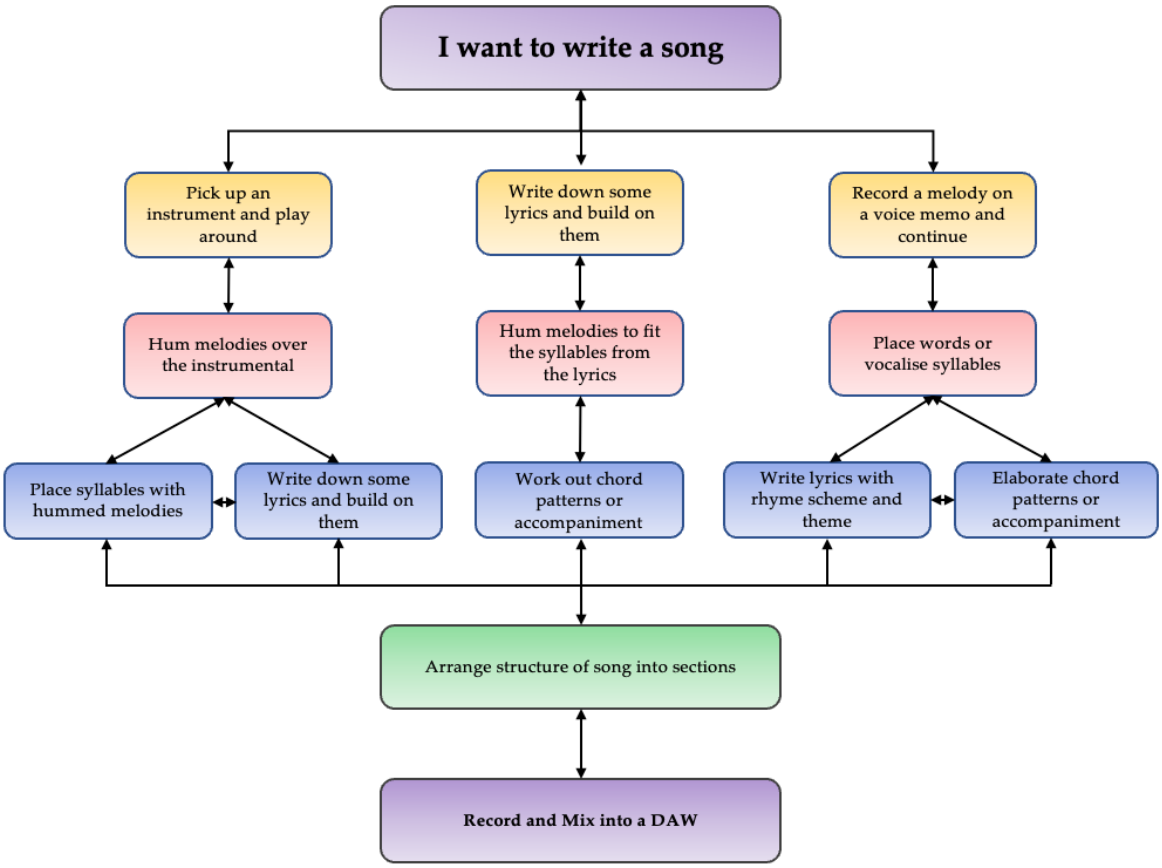

Figure 1: The songwriting process.

Figure 1 shows the songwriting process experienced as a professional musician by the lead author. The first stage, shown in purple, indicates the songwriter wishing to write a song and being in that physical or mental place. This stage is followed by three possible starting points for creative action: instrumental practice, lyrics, and melodies. The yellow, red and blue blocks involve inspiration, the source of creative spark that will help the songwriter shape their work. In contrast, the following green and lowest purple blocks involve elaboration, i.e., applying music theory and production skills to build upon the original ideas and create a more definitively recorded and arranged work. Although this might not be indicative of every

songwriter's process, there is consensus around it (McCartney 2005; Barber 2017; Huang 2022) and it can be used to map out the areas of the songwriting process which can be improved or supplemented with AI technologies.

The first author of this article is a young professional musician. In 2014, she attended Newcastle-under-Lyme College and was awarded with a D*D*D* (Triple Distinction Star) in Music. Then, she studied bass guitar, songwriting, and production to obtain a BA (Hons) Music (1st) at The Liverpool Institute for Performing Arts in 2019. During this time, she was mentored one-to-one by songwriting greats such as Keith Mullin from The Farm, Eddie Lundon from China Crisis, and Sir Paul McCartney. She also started using commercially available AI songwriting tools to speed up her songwriting process and develop new and interesting ideas. In the Fall of 2021, a song that she co-wrote and co-produced with artist Celine Lyng achieved top one in the Norwegian charts. Following with her studies, she also completed a master's degree with Distinction in Popular Music Practice at the British and Irish Modern Music Institute in Birmingham, during which she furthered her research into AI and human collaboration in music. Throughout these periods, she has consistently performed in professional ensembles as a bassist, pianist and backing vocalist across the UK and Norway. The following sections are written at the first person as they reflect her first-hand experience as a professional songwriter.

Starting with Accompaniment

"Ordinarily, I would begin by picking up an instrument, such as guitar or piano, and playing around with various voices of chords or arpeggios until I find something memorable or significant. Then, I would improvise melodies over the devised accompaniment, through humming or vocalising, until I find a hook or suitable refrain. There is a constant evaluation throughout this process as I figure

out how I want the finished product to sound until I feel comfortable enough with the initial section to move onto the next. Once this is complete, I move onto recording the song into a DAW, while still evaluating and making alterations as I go. When working with AI, this process remains mostly unchanged."

Starting with Lyrics

"When starting a song by writing lyrics, I tend to begin with an idea I've been inspired by. This could be an object, a phrase that I've heard, a conversation or situation I've been involved in or a feeling I've been experiencing. Following this, I structure out the lyrics with different sections and a rhyme scheme to start humming a melody to it. Once the top-line is finished to an extent, I begin to find some chords or accompaniment through instrumental practice and arrange the song into various sections, expanding upon existing ideas as I go. Once the song has taken shape and has the various elements, I begin recording it into a DAW to finish the arrangement."

Starting with Melody

"Starting with a melody usually involves a hook coming to mind as the inspiration and recording a hummed or scatted version of this in a memo on my phone. This provides the intention that I can expand upon when I am available in my home set-up with the facilities needed. Once the melody has been sufficiently fleshed out, I place words or vocalise syllables on it, then write lyrics with a rhyme scheme and consistent theme or story. Typically, simultaneously to this, I elaborate chords or an accompaniment to guide the development of the song's structure. Once these elements feel finished, I arrange the overall structure of the instrumental aspects of the song and record the project into a DAW."

Songwriting Aspect	AI Usage	Benefits	Negatives	Example tools
Lyrics	Generate lyrical stimulus through either conditional or unconditional generation	• Alleviate writer's block • Broaden musical horizon • Speed up songwriting • Can be helpful to composers	• Often needs human curation, can become repetitive of the training data	"These lyrics do not exist", GPT1, 2 and 3, ChatGPT
Chords	Generate chord progressions through either conditional or unconditional generation	• Can inspire new ideas • alleviates the need for knowledge in music theory	• Can be random or unusable • Can be simple or not sound cohesive with existing material	Scaler 2, Chord Chord, Auto Chords, Melobytes, Magenta
Melodies	Generate melodic ideas or phrases through either conditional or unconditional generation	• Alleviate writer's block • Broaden musical horizons • Speed up songwriting • Can be helpful to lyricists	• Often needs human curation • Can be repetitive of training data • Often generates too many notes to correlate with syllables or note jumps that are near impossible to sing	Audoir, SoundDraw, Magenta, AI Music Pro
Rhythm and Percussion	Generate drum sequences either in audio or MIDI. Can be used for sections of songs, fills or full track.	• Useful for non-drummers or to change the genre or feel of a song	• If generating audio, drum sound could be inappropriate • If generating MIDI, it must be paired with appropriate samples	Logic Pro Drummer, Magenta Drum Bot, Emergent Drums, Musicaa, Infinite Drum Machine
String and Orchestral	Arrange pre-existing MIDI or Audio of strings.	• Useful for those who are not familiar or comfortable arranging these instruments but wish to include them in their music	• String samples can sound lifeless • Some arrangements could be impossible for string players to perform or be too classical for popular music scenarios	NSynth, MuseNet, Scaler 2
Accompani-ment	Generate a backing track either based on parameters or random generation.	• Allows vocalists or those without a formal music education to create a backing track without a need to record through a DAW or recording studio.	• These services will often have very limited customisation options.	Boomy, AIVA, Amper, Lalal, MuseNet
Accompaniment Arrangement	Often instrumental tracks, generated using a mood or descriptive word.	• A quick and easy way to acquire cheap or free music for advertisements or videos	• These services will often have very limited customisation options.	Boomy, AIVA, Amper, Lalal, MuseNet
Sourcing Samples	Sorting through samples and grouping them for ease of finding or generating for a specific sound desired.	• Can improve the time taken to find and sort samples when writing or producing music	• There are limited and specific ways these services will sort for you. • Software is expensive.	SoundDraw, Ecrett Music, NSynth
Mixing and Mastering	Mixing and mastering either full tracks, aspects of tracks or aspects of mixing.	• Allows vocalists or those without formal music education to produce quality material with little production knowledge.	• Mixing and Mastering engineers may feel like they're being replaced • Cannot apply creative mixing like desired	LANDR, AI Mastering, RoEx, Cryo Mix

Table 1: Categories of AI tools applicable to various parts of the songwriting process.

Applying a selection of AI tools in the songwriting process

As seen in Table 1, a wide variety of commercial AI tools are available to address the various parts of the songwriting process: lyrics generation, melodic composition, arrangement etc. To focus our analysis, the tools selected and described below are the ones used in the preparation of an experiment originally aimed at measuring the correlation between the perception of song quality and the quantity of AI used in the songwriting process. For that experiment, the lead author has produced seven songs where the involvement of AI tools varies in 25% increments from 0% AI usage to 100% AI usage, which are available to listen on (https://tinyurl.com/2j6d3zh6) The present section describes the list of AI tools applied for that, as well as the methodology suggested to quantify the percentage of AI involvement in a song.

Lyrics

Bored Humans

"The lyric generator by Bored Humans (https://boredhumans.com/lyrics_generator.php) is based on OpenAI's GPT-2 (Radford et. al 2019), trained on 7,000 song lyrics and 13,000 poems and generated in a song-like structure. This web service only generates lyrics and does not allow for a seed input, meaning that the outcome is completely random. The lyrics tend to be cohesive, not changing topic midway through the song, but require countless generations to achieve the theme desired in the track. I chose to limit myself to three generations from which to select a subject which I thought was closest to my original idea. Bored Humans can occasionally get stuck in a loop: although repetition of choruses is commonplace in songwriting, I frequently found repeated verses instead within the generated song. As a result, I had to use material which I either composed myself or obtained from a previous generation to have enough original and distinctive sections for a complete set of lyrics."

Audoir

"Audoir (https://www.audoir.com) includes two separate apps for lyrics generation versus melody and harmony generation. I used Audoir's lyrics generator into the inspiration phase by inputting the seed phrase which I wanted to include in the lyrics and then choosing which sections I wanted to focus on in the potential song. Audoir generates lyrical phrases in the form of verses with either three or four lines, typically with an AAAA or ABAB rhyme scheme. This implies that there is less input required from myself to shape the generated material into a typical song format, which results in speeding up the lyrics writing process. There again, I only allowed three generations per seed sentence, followed by lyric curation, to limit the time spent on potentially infinite repetitions of the lyrics generation. The only alterations I made to the lyrics from Audoir were to correct grammar or tense errors and to move different generated sections to different areas of the song, such as swapping verses and choruses."

GPT-2

"To improve over Bored Humans and Audoir, I decided to train my own personalised lyrics generator. For that, started from OpenAI's GPT-2, which is a generative pre-trained language model that predicts the next word in a text sequence, and retrained that model using TensorFlow (https://www.tensorflow.org) and 18,000 tokens of my own lyrics as the retraining data. By tuning the parameters of length, temperature, and batch size, I obtained a model able to achieve sentences better adapted to songwriting than the original model, and able to emulate my personal songwriting style. At the lyrics generation stage, I was able to use a sentence seed to shape the subject of the lyrics generated. This resulted in less repeated generations than Audoir or Bored Humans to get some

material which I could use in a song. As such, GPT-2 was the most favourable way to generate lyrics. However, the lack of a consistent rhyme scheme and of clear sections required curation. The material suggested by GPT-2 was most like my own work without being derivative. I still restricted myself to three repeated generations with the same parameters and to avoiding lyrics alterations other than to fix grammar errors or to move existing lyrical sections to other parts of the song - the same rules that were applied to all lyrics generators, to keep their results comparable and avoid bias related to different usage rules."

Intriguing feature of the lyrics generated from GPT-2 are the grammatical mistakes and unusual phrases produced by the algorithm, which might affect the correctness of subject and verb agreement, pronoun use, sentence fragments, modifiers target and general grammar. Looking at the following example:

> Acre my feet, tired and weak
> On my knees think logically,
> I watched your every move
> I watched your every move
>
> Sitting on my hands and knees
> The outside mired in mystery
> I watched your every move
> I watched your every move

'Acre my feet' does not make grammatical sense, with the common phrase being 'anchor my feet'. Then, the word 'mired, referring to being stuck, entangled or hindered, may feel unusual in the context of 'outside mired in mystery'. Finally, the repetition of a segment that is not a chorus does not make common structural sense in a pop song. These suggest issues of coherence in the generated lyrics. However, it belongs to the artist to decide if these should be corrected or not. If completely corrected by the author, then the AI-generated results stop being specific to AI: human corrections would automatically get them closer to passing a Turing test. On

the other hand, the author can choose to leave the errors uncorrected and acting as features of the song, if that is deemed artistically interesting: for the above example, the lead author chose to record the song without correcting the mistakes. This could lead to aesthetics specific to AI, akin to overlaying audio defects to create a particular sonic atmosphere, but in the lyrical space. However, this artistic choice belongs to the artist, not to the AI per se.

These Lyrics Do Not Exist

"Compared to other AI tools tried, *These Lyrics Do Not Exist* (https://theselyricsdonotexist.com) appeared friendlier and able to deliver suitable lyrics in a Verse/Pre-Chorus/Chorus/Bridge format; however, the control options were more limited, as they consisted in choosing a single word for the sentence seed, one emotion from five (very sad - very happy) and one genre from six options. The genre options felt like they were heavily relying on tropes and cliches and did not feel unique nor likely to contribute anything new. It all felt quite derivative of some else's material. Furthermore, the delivered lyrics lacked a rhyme scheme and included too many repeated lines in the verses and chorus. This tool therefore appeared as the least favourable one for lyrics, because it required too many corrections, lacked control options, and lacked a general sense of supplying originality or surprise in what it generated."

Melody

Audoir

"Regarding the usage of Audoir for melody and harmony creation, I found that the melodies generated by this part of the tool contained frequently repeated notes, large jumps between notes and a lack of pauses typically used for breathing space in vocal lines. This meant that I typically used it to find melodic ideas or phrases rather

than taking it exactly as it was generated. Fixing Audoir's outcomes did not take long as I would sing through the melodies and adjust the syllables of the lyrics where it felt natural to pause, take a breath or reduce the size of any note intervals that felt too large.

Using Audoir to generate melodies raised another issue of coherence, that of creating motifs and phrases which don't combine well with the lyrics. indeed, combining melody and lyrics plays a defining role in the songwriting process: without melody one has a poem, without lyrics one has an instrumental track. Audio will generate melodies based on key, tempo, and temperature, with no indication of lyric syllable consideration. This leads to either an excess or an insufficient number of syllables for lyrical ideas to fit with or be written to. Shown in Figure 2 is a melody generated from Audoir, where it is easy to see the repetition of notes in a rapid rhythm and large jumps between notes: that would be atypical for songwriting, especially in the pop genre.

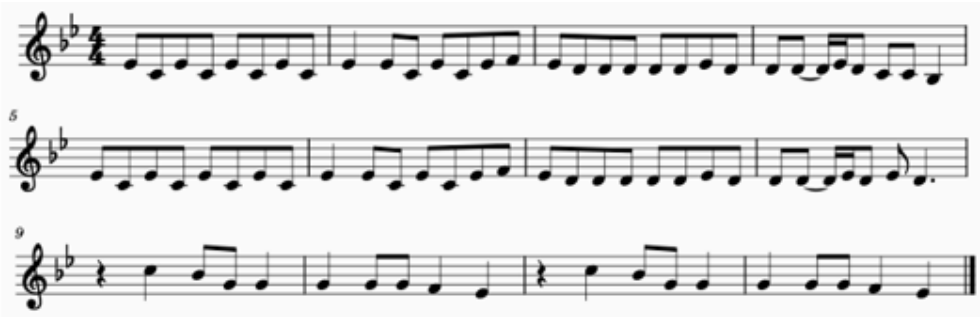

Figure 2: A melody generated using Audoir.

In that case, the lead author corrected the melody, as shown in Figure 3, to achieve less static repeated notes, e.g., in bar 3, and to transform the octave jump between bars 8 and 9 into a fifth. The melody is still recognisable as the AI-generated one, but

there is more movement throughout, and a vocalist can fit lyrics more easily. In particular, AI will not have any need to leave space for breathing in a generated melody, so breathing spaces have to be manually included to accommodate the physical constraints of performance – such corrections might also be mandatory for other instruments than voice, where the physical constraints may impact the songwriting process."

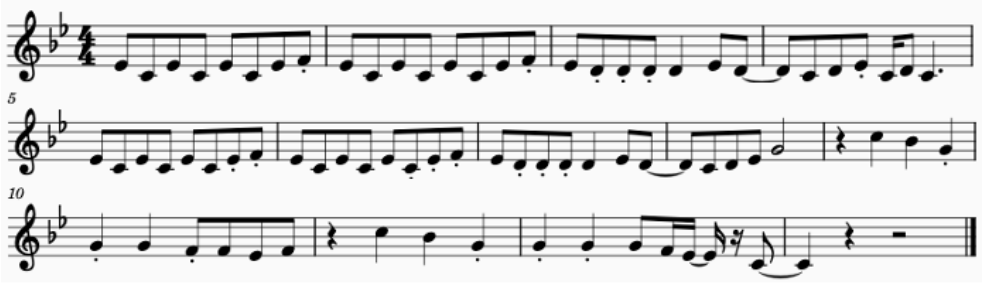

Figure 3: The melody after correction by the artist.

Arrangement

Amper

"Advertised as 'music for your defining moment', Amper is used for creating royalty free music often used in videos and podcasts. I have used this software mostly to create backing tracks for a separately written or generated top-line, although it can also be used to generate melodies. There is not much musical control when using Amper, as it is designed for non-musicians. There are only parameters such as genre, length, or emotion to choose from before the music is generated. There is an unlimited amount of tweaking and editing that can be performed, consisting of section lengths, samples sounds and additional instruments. My biggest issue with this program arose when recording its result in a DAW: I only had the audio samples and not the MIDI, so I was unable to change the instrument

sounds or to conduct any mixing or mastering as there was already reverb and EQ on all the tracks."

Google Magenta

"Described as an open-source research project that explores the role of AI and machines as tools in the creative process, I used Google Magenta to devise drumbeats, bass lines and melodies using the tools *Drumify*, which generates grooves based on the rhythm of the MIDI input, *Generate*, which generates a 4 bar phrase without requiring any input, and *Continue*, which is similar to *Generate* but requires an input to generate notes that are likely to follow an existing drum beat or melody. These had less customisable parameters than other tools but allowed me to provide an existing MIDI file for Magenta to continue writing or to rework specific MIDI sections. I mostly used this tool after writing lyrics and melody, either with other AI tools or by itself, then would incorporate Magenta into the arrangement and recording phases."

Logic Pro Drummer

"Logic Pro (https://www.apple.com/uk/logic-pro), a Digital Audio Workstation (DAW) by Apple, comes with the Drummer tool. This tool allows users to choose from a variety of styles and genres with a sliding parameter to alter a synthetic drum track's MIDI performance live whilst working on the arrangement. I found this drummer tool the most useful as it allowed changes and modifications, some of them automatic, throughout the development of the piece. I never used it during the inspiration stage, only during arrangement and recording. As a non-drummer, I felt like it wasn't limited to basic beats. This tool created distinct and appropriate sections for verses and the chorus, with fills to blend these sections seamlessly. This tool saved me a lot of time: typically, without the Drummer tool, I would play the

MIDI drum pattern I wanted on a keyboard, then quantise, then find the samples I wanted, whereas using the Drummer I only needed to add the track and adjust the slider accordingly."

Scaler 2

"Scaler 2 (https://www.scalerplugin.com) works mostly as an arrangement or accompaniment tool, being able to detect the key from MIDI or audio data and to suggest chord progressions to expand upon the existing ideas. I used Scaler 2 as the base for many songs, by asking it for suggestions of chord progressions based only on a genre or feeling. Using the performance tool, I arranged string sections and piano performances to reduce the time it would have taken to do this by hand for so many instruments. This kind of application may be incredibly beneficial to those unfamiliar with arrangement techniques for string or piano. Furthermore, as an individual who happens to be comfortable with arrangement techniques, this tool suggested ways that I may not have considered otherwise. It also enabled me to explore those options without wasting hours of labour on experimenting, only to find out that I would have preferred the first arrangement which I tried."

Overall, the lead author reports that she spent more focus on AI tools for lyrics generation because "I was struggling with lyrics, but also the lyrics from the commercial tools were more noticeably AI, with their strange grammatical and wording errors, whereas generated melodies and arrangements required less curation, so for melody and arrangements I generally felt more content with the results of the tools as they were."

Quantifying the amount of AI usage in a song

The seven songs composed with AI were written with 25% incremental increase

of AI involvement. There were no strict parameters for how these songs were to be written: either the AI was the prompt and starting point of the creative process that helped the author to start a song, or the author provided an initial idea and used AI to extend it towards a finished track. For example, with lyrics the author could either extend a fully AI-generated paragraph or provide words from her own inspiration as the seed for AI to predict additional lyrics. Similarly, for melodies and arrangements, the author could either continue an AI-generated stem or ask the AI to provide an extension conditioned on a stem that she had imagined without the support of an AI tool. The proposed calculation of the percentage of AI involvement is therefore based on the resulting song, across word units for the lyrics and bar units for the musical aspects of melody, accompaniment, and arrangement. The lyrical and musical components of a song are treated with equal weight, while the musical part is divided into the three aspects of melody, accompaniment, and arrangement, themselves treated with equal weights within the musical part. The following percentage formulae are therefore used to yield the final percentage:

$$\textit{AI music percentage: } \%_{AI}^{music} = \left(\frac{B_{AI}^{melody}}{B_{AI}^{melody} + B_{human}^{melody}} + \frac{B_{AI}^{accomp}}{B_{AI}^{accomp} + B_{human}^{accomp}} + \frac{B_{AI}^{arrgt}}{B_{AI}^{arrgt} + B_{human}^{arrgt}} \right) / \ 3$$

where B_x^y are counts of bars affected or unaffected by AI, respectively B_{AI}^y and B_{human}^y, pertaining to melody generation, accompaniment or arrangement, respectively B_x^{melody}, B_x^{accomp} and B_x^{arrgt}, and where $B_{AI}^y + B_{human}^y$ sum up to the total number of bars in the song.

$$\textit{AI lyrics percentage: } \%_{AI}^{lyrics} = \frac{W_{AI}}{W_{AI} + W_{human}}$$

where W_x refers to counts of words generated either by AI or by the author, respectively W_{AI} and W_{human}, and where $W_{AI} + W_{human}$ sum up to the total number

of words in the song. The word counts include articles such as “a”, “the” etc., but exclude the parts repeated for structural purposes, e.g. several repetitions of the chorus.

AI percentage for the whole song: $\%_{AI}^{song} = \left(\%_{AI}^{music} + \%_{AI}^{lyrics}\right) / 2$

To demonstrate how these formulas can be used within the music, looking at the lyrics for one of the songs of our experiment, titled Silhouette (https://on.soundcloud.com/vfXuR) we can identify the number of words in distinct sections. The human input data is the pre chorus and chorus, which contain a combined 57 words:

I feel like a dying light
Am I losing my mind?
Searching for the marks you left behind?

I was living in the dark
Wishing I’d never known a spark
To watch it burn out like a star

Lately I’ve been thinking I should turn off the light
But I can’t let myself fall asleep at night

I still see your silhouette on the walls
I can feel the marks you left through the scars
I still leave the lights on so it’s never dark
So I can watch the shadows dance around till dawn

I think I’m still dreaming
Can you feel me suffering?
I’m out in the cold with a broken wing

The verses and pre-chorus are italicised and were solely AI generated therefore, we know that 57 out the 114 lyrics are AI generated. Using the formula above, this translates to $\%_{AI}^{lyrics} = \frac{57}{114} = 50\%$, meaning that about half of all lyrics are AI generated. In the instance of this song, there was no involvement of AI with regards

to the musical elements of melody, chord progression or arrangement, so $\%_{AI}^{music}$ is zero. Using the final formula to calculate the overall AI percentage in the musical work concludes that the song is 25% AI generated.

Using this method as a consistent measure of AI involvement in the composed songs allows to test which amount of AI involvement works best. However, measuring the exact amount of human curation needed to fix grammatical errors or to help shaping a coherent story from the AI-generated material is more difficult, if feasible at all. Thus, this percentage is an indication of the quantity of AI-generated content but conditioned on the understanding that there is an unmeasured degree of human intervention that the author may choose to apply to the AI-generated sequences to achieve a good song.

Assets and shortcomings of AI tools usage in songwriting

We found that incorporating elements of AI-based generation of lyrics, melody, arrangement, and accompaniment can help the inspiration stage of the creative process and thus decrease the time taken to at the initial stages of composing a song. Two stages can be distinguished: inspiration and continuation. At the inspiration stage, using AI to broaden musical horizons allows a songwriter to "get out of their box" and to view their work in a different light by putting it in the perspective of potentially unexpected AI stimuli. At the extreme, its best use is to assist human songwriters to overcome writer's block, which can stem from "self-criticism or perfectionism as a source of block" (Flaherty 2004, p. 82) and lead to a "strangled feeling of inarticulateness, of ideas coming faster than words, of not being able to express what's inside". In contrast, AI will generate something regardless of self-criticism, and the generated contents appears useful to refocus the mind and assist

in restructuring the author's thinking. Models such as OpenAI's GPT-3 (Brown et. al 2020) are large 'text in, text out' models which can generate 100 pages of content from a trained model, or which only require a single word to produce additional text. In that sense, songwriters can use this software to generate wording variations at the same rate as their ideas, if not quicker. This method becomes quantity over quality, wherein the software will consistently yield words and results at the expense of some lyrics becoming derivative or meaningless. It then becomes the responsibility of the songwriter to turn the generated material into something that is not devoid of narrative nor impersonal.

At the continuation stage, AI based on conditional generation, i.e., which allows an input of key and time signatures, existing melody lines or chord sequences, or any other musical idea which was itself either automatically generated or conceived by the author at the inspiration stage, for example Google Magenta's *Continue*, is designed to supply extensions of the initial idea. There again, the helping mechanism is about stirring the author's mind away from some form of mental block, e.g., when the author might see no obvious continuation to the original idea. This was experienced in the lead author's work when using a chorus sentence to seed an AI tool with the intention of continuing the chorus's narrative, in the song titled Silhouette (https://tinyurl.com/2j6d3zh6).

When using AI generated material in conjunction with pre-existing written elements, we have found that it can provide a new meaning when recorded or performed. An AI will suggest the next logical word or note, which may influence the songwriter away from the direction to which they had initially chosen to take the song. However, the author is still in control: several variations can be generated until one is deemed suitable by the author to complete a song.

Instead of using one generation tool to compose an entire song, we found it more tractable to combine several tools for several jobs. There are countless available tools, and it is not a one-size fits all scenario, with some tools focusing on different genres or styles that might not be suitable for every songwriter. For example, the lead author has used *These Lyrics Do Not Exist* to combine AI-generated words with a melody that she has written first-hand. Another example at this phase of the songwriting process is the use of *Scaler 2* to suggest chord progressions for a top-line that she has already composed but did not yet have an accompaniment for.

Software like Amper, Boomy (https://boomy.com/) and AIVA (https://www.aiva.ai/) can compose entire backing tracks for compositions with few requirements from the songwriter. Although these generated instrumental compositions offer less tailored options, compared to incorporating AI earlier in the process, they are very easy to use and are intended to democratise accessibility to music creation, as they alleviate the need to know about music theory, arrangement techniques or how to play an instrument. Rather, they provide a performance or MIDI arrangements for those who aren't familiar with these techniques, and suggest how they might be used in the track. This is at the risk of producing unrealistic performance scores, as observed with Audoir's lack of breathing spaces, if the author doesn't have the knowledge to validate the results.

Ordinarily, when writing with another songwriter or producer, one would need to explain ideas clearly about what the finished product should sound like. Conversely, AI tools can generate an entire backing track in minutes, thus allowing the workshopping or arranging to be iterated quickly, letting the music speak for itself and the rapid stream of musical ideas be experienced first-hand, rather than

being interfered with or slowed down by the constraints of verbal (mis)communication. Besides, AI is non-judgemental: in an ordinary co-writing scenario, it could be difficult to communicate a distaste in the material written together without offending or upsetting fellow writers, whereas an AI software will continually generate material without being dependent on maintaining politeness or pushing a prior musical agenda. In that sense, the author retains more control over the generated material, at the risk of missing out on the innovation opportunities that come with interaction with co-authors, but where the innovative variation is instead supplied via the quantity of innovations generated by the AI tool.

Resorting to AI-generated stimuli at the inspiration or initiation stage can pose questions about the authenticity, personalisation, and originality of the musical content, particularly the emotional character, which is integral to the process of creating music, whether that be as part of creation, performance, or interpretation (Juslin and Sloboda 2011). Along those lines, a significant aspect of a song is the context in which the artist decided to create their music and the personal relation of the artist to their own creation (Kania 2017). This, in turn, influences the consumer's emotional experience, for example driving them to find and connect with their own humanity in the work (https://www.youtube.com/watch?v=q5yxIzs5Wug). The narrative and lyrics themselves can be reinterpreted, yet an AI would never be able to experience these emotions. It may be able to replicate or mimic loss, grief, happiness, and various human-like emotions but would never have its own personal history in relation to these feelings. However, we have observed that AI-generated lyrics could lend themselves to more open-ended interpretations, which might be an asset for songwriting. For example, the lyrics from song Streetlamp Spotlights,

which were entirely human composed, lend themselves to a clear and concise narrative demonstrated in the chorus section:

Streetlamp Spotlights illuminate the way we walk home
You're by my side, makes me feel like I'm invincible
Something's implied by the way you're holding me tonight
I can see the sunlight creeping up behind us
I guess in hindsight we made to many detours, but they were justified
So that we could get a little more time.

These lyrics clearly depict two people walking home together at night, not wanting the moment to end. In contrast, the lyrics of song Homesick, half of which were AI-generated without human curation, lend themselves to a more open-ended interpretation:

I hide no bones about it, I'm a pragmatist never weakened by steel
I can't take it back, like a hole in a punch card make it disappear

In that case, there is attractiveness in that the lyrics boggle the audience's mind into finding some meaning.

One final aspect of this discussion is whether AI-based compositions are enjoyable and viable as commercially released music. This is probably more dependent on the author and the audience than the AI tools themselves. In the lead author's personal experience, the involvement of AI tools in her songwriting process has led to more experimentation and enjoyment overall and has delivered on the expectations of supplement of efficiency and reduction of dependency on co-producers, to the extent that the usage of AI can legitimately be claimed as an integral part of her songwriting process, without removing, threatening or even be comparable with the

possibility to collaborate with other songwriters or producers.

Conclusions and future work

This article reports the insights obtained from the introduction of AI tools in a professional musician's songwriting process. These indicate that the usage of AI tools achieves the desired goals of added efficiency and limitation of the need to resort to co-workers. It also underlines the role that human curation and correction, applied on top of the AI-generated elements, retain to achieve the musician's artistic goals. Indeed, throughout the use cases it appears that the main role of AI tools is to supply variation, in the sense applied by Simonton in his theory of creativity (Simonton 2011; Simonton 2012), or divergence, in the sense of the "Double Diamond" definition of the design process (Ball 2020): in both definitions, a phase of variation or divergence is followed by a phase of selection or convergence, where the amount of surprise at the divergent stage may vary depending on the generation process, but where the artist or the designer remains in charge of the corrections or judgement calls which are leading to the final production (reference removed for anonymity 2020). In other words, songwriting, like other forms of creation or design, is about generating choices prior to making choices (Brown 2009, p.73). In that context, most AI tools generate something which is neither completely random nor completely unsurprising, thus potentially hitting a sweet spot to aid in the creative process of songwriting. In all the observed cases, however, the musician remains in control of the creation process through selection and correction, so there may not be a reason to think that AI tools will replace the role that musicians hold in the creative process of songwriting.

Future Work

We are currently distributing a listening test aiming to rate the public's appreciation

of the seven songs composed for this experiment, in relation to their proportion of AI tools usage. The results are not yet available at the time of submission of this article, but will be reported in future publications.

Bibliography

Avdeeff, M. 2019. "Artificial intelligence & popular music: SKYGGE, flow machines, and the audio uncanny valley." In *Arts* 8(4):130.

Ball, J. 2019. "The Double Diamond: A universally accepted depiction of the design process." Blog post. Available online at www.designcouncil.org.uk/news-opinion/double-diamond-universally-accepted-depiction-design-process. Accessed January 2023.

Barber, S. 2017. "Professional Songwriting Techniques: A Range of Views Summarised from the Sodajerker Interviews (2011-2015)." in *The Singer-Songwriter Handbook*. Bloomsbury.

Braga, M. 2016. "The Verbasizer was David Bowie's 1995 Lyric-Writing Mac App." Available online at www.vice.com/en/article/xygxpn/the-verbasizer-was-david-bowies-1995-lyric-writing-mac-app. Accessed January 2023.

Briot, J. P., G. Hadjeres, and F. D. Pachet. 2020. *Deep learning techniques for music generation*. Heidelberg: Springer.

Brown, T. 2009. *Change by design*. Harper Business. (2019 edition revised and edited.)

Brown, T., et al. 2020. "Language models are few-shot learners." *Advances in neural information processing systems*, pp. 1877-1901.

Carter, A. 2022. "The Paradox of Spontaneity and the 'Genius for Ambiguity': on the (im)possibility of creative AI." *The 2022 International and Interdisciplinary Conference on Creativity, Innovation, and Entrepreneurship (IICCIE22)*. Talk also available to watch on: https://www.youtube.com/watch?v=FFWSMYW7pzQ&t=5824s

Copsey, R. 2018. "It Now Takes an Average of 4.84 Songwriters to Pen a Hit Single." Available online at www.officialcharts.com/chart-news/it-now-takes-an-average-of-484-songwriters-to-pen-a-hit-single__23039. Accessed January 2023.

Dennett, D.C. 1998. *Brainchildren: Essays on designing minds*. The MIT Press.

Dennett D.C., and E. Lambert, 2017. "Thinking like animals or thinking like colleagues?" *Behavioural and Brain Sciences* 40:34–35.

Deruty, E., et. a 2022. "On the development and practice of ai technology for contemporary popular music production." *Transactions of the International Society for Music Information Retrieval* 5(1).

Dong, H. W., W.Y. Hsiao, L. C. Yang, and Y. H. Yang. 2018. "Musegan: Multi-track sequential generative adversarial networks for symbolic music generation and accompaniment." In *Proceedings of the AAAI Conference on Artificial Intelligence*.

Flaherty, A. (2004). "The Midnight Disease: The Drive To Write, Writer's Block and the Creative Brain." New York, Houghton Mifflin.

Hadjeres, G., F. Pachet, and F. Nielsen. 2017. "Deepbach: A steerable model for Bach Chorales generation." In *International Conference on Machine Learning,* pp. 1362-1371.

Herndon, H. 2022. "Holly Herndon: How Ai Can Transform Your Voice." NPR. Available online at www.npr.org/2022/08/26/1119220726/holly-herndon-how-ai-can-transform-your-voice. Accessed January 2023.

Hiller, L. 1981. "Composing with computers: A progress report." *Computer Music Journal*, 5(4):7-21.

Holzapfel, A., B. Sturm, and M. Coeckelbergh. 2018. "Ethical dimensions of music information retrieval technology." *Transactions of the International Society for Music Information Retrieval* 1(1):44-55.

Huang, C.-Z. A., et. al 2020. "AI Song Contest: Human-AI co-creation in songwriting." In *Proceedings of the 21st International Society for Music Information Retrieval Conference*, pp. 708–716.

Huang, A. 2022. "The 5 best ways to start a song." Available online at www.youtube.com/watch?v=m6OgNpTjuBs. Accessed January 2023.

Iser, W. 1972. "The Reading Process: A Phenomenological Approach." *New Literary History*, 3(2):279–299.

Jacob, B. L. 1996. "Algorithmic composition as a model of creativity." *Organised Sound*, 1(03):157-165.

Juslin, P. N., and J. Sloboda. 2011. *Handbook of music and emotion: Theory, research, applications.* Oxford University Press.

Kania, A. 2017. "The Philosophy of Music. Stanford Encyclopedia of Philosophy." Available online at plato.stanford.edu/entries/music. Accessed January 2023.

Knotts, S., and N. Collins. 2020. "A survey on the uptake of Music AI Software." In *Proceedings of the International Conference on New Interfaces for Musical Expression* pp. 499-504.

Levitin, D. 2019. *This is Your Brain on Music: Understanding a Human Obsession.*

Penguin.

Louie, R., et. al 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* pp. 1-13.

McCartney, P. 2005. “Paul McCartney Reveals his Songwriting Secrets - BBC Radio ‘Sold on Song’ - 17 September 2005.” Available online at www.youtube.com/watch?v=3x-T3QKRtxA. Accessed January 2023.

McFarland, A. 2023. “7 Best AI Music Generators.” Unite AI. Available online at www.unite.ai/best-ai-music-generators. Accessed January 2023.

Micchi, G., et al. 2021. "I Keep Counting: An experiment in human/AI co-creative songwriting." *Transactions of the International Society for Music Information Retrieval* 4(1):263-275.

Murray, E. 2020. “New Figures Show How Many Track Streams Needed for Artists to make Minimum Wage”. DJ Mag. Available online at djmag.com/news/new-figures-show-how-many-track-streams-needed-artists-make-minimum-wage. Accessed January 2023.

Najafabadi, M.M., et al. 2015. "Deep learning applications and challenges in big data analytics." 2(1):1-21.

Oliver, B. 2013. *How [Not] to Write a Hit Song! 101 Common Mistakes to Avoid If You Want Songwriting Success*. United States: Createspace Independent Pub.

Pitoniak, A. 2017. *The Futures*. New York, Lee Boudreaux Books.

Radford, A., et al. 2019. “Language models are unsupervised multitask learners.” *OpenAI blog*, 1(8):9.

Roberts, A., et al. 2018. "A hierarchical latent vector model for learning long-term structure in music." *International conference on machine learning*.

Simonton, D.K. 2011. “Creativity and Discovery as Blind Variation: Campbell’s (1960) BVSR Model After the Half-Century Mark”. *Review of General Psychology*, 15(2): 158–174.

Simonton, D.K. 2012. “Taking the U.S. Patent Office Criteria Seriously: A Quantitative Three-Criterion Creativity Definition and Its Implications”. *Creativity Research Journal*, 24(2-3): 97-106.

SongFancy 2022. “The 5 in 5 Song Challenge.” Available online at songfancy.com/5-in-5-songwriting-challenge. Accessed January 2023.

Sturm, B., J.F. Santos, and I. Korshunova. 2015. "Folk music style modelling by recurrent neural networks with long short term memory units." In *16th international*

society for music information retrieval conference.

Sturm, B. and Ben-Tal, O. 2018. "Let's Have Another Gan Ainm - An experimental album of Irish traditional music and computer-generated tunes." *KTH Technical report.* Available online at kth.diva-portal.org/smash/get/diva2:1248565/FULLTEXT02.pdf. Accessed January 2023.

Sturm, B., et al. 2019. "Machine Learning Research that Matters for Music Creation: A Case Study." *Journal of New Music Research* 48(1):36–55. Wallas, G. 1926. *The art of thought*. Harcourt, Brace.

Xenakis, I. 1992. *Formalized music: thought and mathematics in composition.* Pendragon Press.