Collaborative Creativity in TikTok Music Duets

Katherine O'Toole
kmarie.otoole@gmail.com
KatherineOaposToole2020@u.northwestern.edu
Northwestern University
Evanston, Illinois, USA

ABSTRACT

On the social media platform TikTok, users are able to engage with each other's content by using the Duet feature, which allows them to re-share another user's video while also layering on additional content to the original video. Through this, the affordances of the Duet feature facilitate a distributed and collaborative creative process, in which we can observe the evolution of cultural artifacts through the different versions that are produced from user contributions. As a result, the open-ended nature of these collaborations positions engagement as both a creative and social act. In this paper, we identify the ways in which the Duet feature supports decentralized co-creativity and engagement between users. We find that the cumulative nature of an artifact's creative evolution, along with the ability for multiple iterations of an artifact to develop in parallel, facilitates development of diverse creative artifacts.

CCS CONCEPTS

• **Human-centered computing → Collaborative and social computing**; **Collaborative interaction**; **Collaborative content creation**; **Collaborative and social computing systems and tools**.

KEYWORDS

Creativity, HCI, MIR, CSCW

ACM Reference Format:
Katherine O'Toole. 2023. Collaborative Creativity in TikTok Music Duets. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3544548.3581380

INTRODUCTION

Creativity is important for the generation of novel and innovative ideas and products, and collaboration can often support the development of creativity by increasing the diversity of the ideas and perspectives being considered [1, 10, 26, 57, 64, 74]. With digital technology, more and more collaborative work is being done online, meaning that these collaborative interactions are being mediated by digital platforms and tools [11, 61, 83]. This is particularly true in the case of large-scale online collaborations such as Wikipedia, which are developed through the contributions of many distributed individuals [47, 55].

Within the music industry, the development of digital platforms such as Spotify's Soundtrap [75] has allowed musical collaborations to occur entirely remotely. With interactions between collaborators being mediated by digital tools and platforms, we see the development of new collaborative practices emerge as a result [20, 54, 60, 61, 84]. In online music collaborations, the ease of duplicating and sharing digital artifacts has led to the rise of more open-ended approaches to collaboration. For example, on the music collaboration site BandLab, users are able to 'fork' a music track, allowing them to clone work by another user and build off of it without committing to a formal collaboration [79]. Projects created by layering together music from different sources can also be found on other online platforms, and serve as examples of musical artifactual creativity, in which existing ideas or objects are recombined or repurposed in order to generate new and innovative artifacts [36, 82]. For example, a YouTube video created by Andre Antunes features him playing a System of A Down song on electric bass over a simultaneous video of the musical group the Nooran Sisters, who are singing a traditional Sufi devotional song [5]. On the website Reddit, there is also an 'Improvised Harmony' community, where users take naturally occurring rhythmic sounds, such as the sound created by a running washing machine, and layer melodic and harmonic components in order to create a music artifact [38]. These differ from more 'traditional' creative collaborations, in that there is no defined group of collaborators, allowing for a collective and decentralized process of collaborative creativity to develop. One example of this is the social media platform TikTok, which has a feature specifically designed to enable these types of collaborations.

TikTok, a video sharing platform that allows users to create and share their own videos, currently has over 138 million active users in the US, and has risen to play a prominent role in modern popular culture as one of the fastest-growing social media platforms [16, 30, 51, 80]. On TikTok, users are able to use the 'Duet' feature, which allows a user to take an existing video and create a new version of it which includes both the original parent video content, alongside new content created by the user. In the new child video, this creates the effect of both the original and new content being played together at the same time, analogous to how two musicians would play a duet together, hence the name [44]. In this way, the virtual spaces offered by TikTok and other platforms serve as both a creative and a social outlet for many musicians [44, 49].

The Duet function also allows for longer Duet chains to form as new contributors sequentially add additional layers of content, as can be observed in the screenshot in Figure 1. It is also possible for an individual video to be Dueted by more than one user, meaning that a parent video might have multiple child videos, each

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04.
https://doi.org/10.1145/3544548.3581380

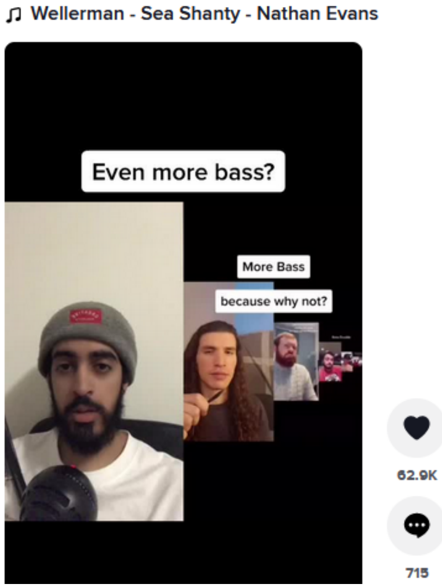

Figure 1: A Duet chain formed through the accumulation of multiple individual contributions.

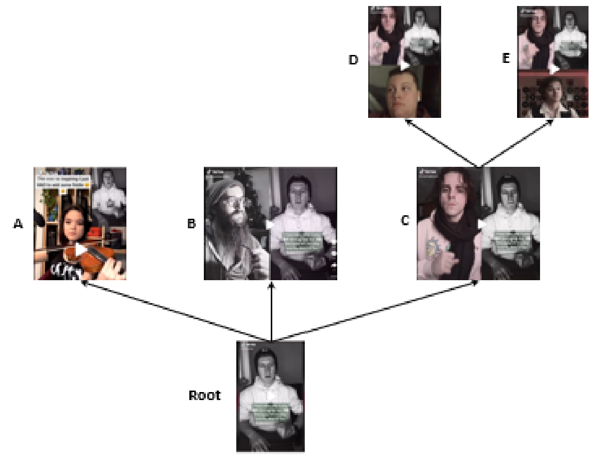

Figure 2: An example of the network structure formed by relationships between videos within Duet Chains. The initial video, referred to as the root video, is the parent video to videos A, B, and C, which were created by Dueting the root video. Video C is then the parent video to videos D and E, which were created when video C was Dueted. The screenshots illustrate how the content changes as additional contributions are added.

of which will be a different version of the original artifact. As a result, is it possible for an individual TikTok video to proliferate into numerous digital versions. Because we are able to capture the parent-child relationships between videos, it is possible to graph the relationships between these versions by representing individual videos as nodes, and the parent-child relationships between them as edges, as seen in the example in Figure 2. Here we see that each Duet chain can be traced back to one initial root video, which is the first video created in the chain. The root video itself is not a Duet, and therefore does not have a parent video, as is illustrated in Figure 2, where we can see that root video content contains a single video contribution. In this example, the root video was then Dueted 3 times, making it the parent video for videos A, B and C. Video C was then Dueted twice, making it the parent video for videos D and E. This open collaboration model allows anyone on the platform to contribute to the creative evolution of the artifact, facilitating a process which is distinct from coordinated online creative collaborations. In this paper, we examine how the technical affordances of the Duet feature structure these collaborative interactions and shape the development of the creative artifacts being produced.

Using digital trace data from TikTok Duet videos which featured musical content, we examined the interaction networks formed by individual videos' relationships to one another. This allowed us to consider the structural elements of the artifact networks created by users' collaborative interactions with TikTok Duets. Looking at three specific case studies, we then extracted quantitative data about the musical attributes of the videos using Music Information Retrieval (MIR) software in order to examine artifact development throughout the process of collaborative creativity. This allowed us to perform a computational analysis of the evolution of video content over the course of the Duet interactions by measuring how the quantitative musical features of the video artifacts changed over time. We then expanded our analysis to include over 2,000 parent-child video pairs, and collected data on the amount of time that passed between when a parent video was posted, and when the corresponding child video was posted. With this data we were able to model the probability distribution of time in between when videos were Dueted in order to identify patterns in user engagement with Duet videos.

We find that the open-ended nature of Duet collaborations and the social transparency of the TikTok platform encourages a diverse and collaborative exploration of the creative space by allowing users to directly engage with publicly accessible digital artifacts. Our work highlights the importance of digital tools for supporting collective ideation, and identifies key HCI affordances that can be leveraged to support collective creativity at scale. This research furthers our understanding of how the design of socio-technical tools and platforms can facilitate collective creativity, and how we can harness the innovation potential of these platforms.

RELATED WORK

Social Creativity

Research into creativity as a social phenomenona has shown that the larger socio-cultural context affects the ways in which creative ideas are developed and disseminated [4]. Social networks also play an important role in the process of creative development, particularly for their role in exposing creative individuals to a wider range of diverse ideas [6, 10, 22, 64]. For example, research has shown that musical artists with wider and more well-connected networks have been observed to create more novel and unique work [57]. In addition to affecting the creative process of an individual, social interactions can also lead to the emergence of a collective creative process within a group, a phenomenon known as distributed creativity [63, 71]. In this setting, the creative process is distributed across multiple actors, and their collaborative interactions of sharing and evaluating ideas, brainstorming, and providing feedback lead to the development of what Parolin and Pellegrinelli term a 'collective creative imagination' [35, 63, 71]. As a result of the web of interactions between actors, we can observe a collaborative emergence of creativity.

Because of this, distributed creativity cannot be understood without examining the interactions that make up these generative processes. This is done with interaction analysis, which involves examining the impact of each individual's contribution to the creative artifacts as the collaboration unfolds over time [63, 71]. When considering how distributed creativity may occur in digital spaces, we must consider the ways in which interactions are shaped by the technical affordances of the available tools, both for interactions between users, and for interactions between users and the creative artifacts being shared [20, 23, 50]. Currently, much of the research into distributed creativity has mainly looked at collocated, synchronous interactions, and the qualitative nature of interaction analysis makes it difficult to scale to larger interaction contexts such as online platforms [46, 63, 71]. As a result, we do not have a great deal of information on how the process of distributed creativity unfolds when these interactions are occurring in a digital space. Additionally, existing research into online music collaboration has mainly focused on cases where collaborators are able to directly communicate and coordinate with one other, typically through a shared digital work space, or user interfaces that encourage interaction [9, 13, 21, 43, 59, 90]. These studies have highlighted how shared representations and awareness of others' activities can help to encourage greater engagement between collaborators [21, 39, 59, 65, 78]. With this type of online collaboration, coordination tools are used to allow collaborators to all work on the same artifact throughout the project, or aggregate their work into one final product at the end [41, 76, 83]. However, this is not the case with the artifactual creativity we observe on TikTok and the other platforms mentioned above, where contributors do not directly interact with one another and do not share a common workspace. In order to understand how users are engaging with the collaboration under these circumstances, we examine the artifact network that is formed by representing each video as a node, with the edge between them representing the parent-child relationship. This allows us to model the structure of the interactions that take place between the artifact and the individual contributors, and identify the group collaboration patterns that emerge.

Computational Music Analysis

With TikTok Duets, the interactions occur solely between the user and an artifact, not between individual users. Since the contribution made by each individual user is captured when they interact with the artifact to create a new version, we can examine how the content of the artifact changes over time across these different versions. Here we see the video artifact serves the function of an 'intermediary object', acting as a shared representation within the collaborative process [18]. The content of the video provides the user with the information necessary to identify the available perceptual affordances, which are the ways a user perceives they can interact with and add to the artifact [27–29]. As a result, the artifact also serves a communication purpose and allows for the flow of information within the interaction network [34]. Through examining the artifact content and how it changes with each additional user contribution, we can therefore analyze the emergence of distributed creativity over the course of the artifact's development [2, 71].

The changes to the artifact's content will be informed by the perceptual affordances of the artifact, which are based on the user's understanding of the rules or conventions of the particular creative domain they are working in. This is also referred to as a 'conceptual space', meaning the structures of thought and generative rules which impact the ways in which we explore and develop new ideas [15]. Within the domain of music there are fairly universal principles of music composition which are grounded in how the human brain processes sound, and which help us to understand what affordances will be perceived by a listener. One example of this is the distinction between consonance and dissonance. While some musical styles will incorporate dissonant elements for color and decoration, it is a fairly universal principle that music is grounded in the construction of consonant sounds. Through research into Auditory Scene Analysis, we know that the human brain goes through the process of auditory grouping, which allows us to take the raw input of sound waves, and make sense of them based on features such as pitch, timbre, and amplitude [19]. This allows us to distinctly perceive the main components of musical organization, which are rhythm, melody, harmony, and affect, or is the emotional impact of music [42]. From these, we see the emergence of certain basic rules for musical organization, similar to how syntax structures language in a way that allows us to extract meaning from it [48]. This foundation of understanding informs the perceptual affordances that are available for a specific artifact. For example, if a Duet video features a song that is in the key of E minor, then that will inform the user of the kinds of chords and intervals that it would make sense for them to use in their contribution. Similarly, if the parent video has music with a Beats Per Minute (BPM) of 60, then the user would know that their musical contribution needs to be performed with a BPM of 30, 60, 120, or some other multiple that allows it to line up with the original.

We can therefore analyze the changes that occur within an artifact over time by using Music Information Retrieval (MIR) software to extract quantitative data about these perceptually meaningful music features [12, 32]. This data, referred to as audio features, or audio descriptors, is information that can be extracted from audio signals and has been demonstrated to provide accurate and robust data for comparisons of music similarity [12, 52] as well as for identifying trends in how music genres and styles evolve over time [17, 40, 56, 67, 73, 85]. Because we are able to capture this quantitative data, we can perform a computational analysis of the contributions made by individuals within the interaction network. As such, the methodology outlined in this paper contributes an analytical framework with which to examine distributed creativity in digital spaces.

Online Interactivity

Previous studies of human interactivity have shown that the temporal distribution of interaction events tends to be heavily right-tailed and 'bursty', meaning that there are a large number of events with short periods of time between them, and fewer events with longer periods of time in between [8, 14, 68]. Within online spaces like TikTok, the technical affordances of the platform are designed so that users are able, and encouraged, to engage and interact with existing content in ways that generate new content, facilitating what are known as 'participatory cultures' [44]. Often on these platforms, the content creation process incorporates a social element which helps to encourage user participation and engagement [25, 66]. This is what is known as an 'experiential-affection' type of participation, which emphasises collective experiences and interactions [45]. We would therefore expect to observe similar bursty behavior in the temporal dynamics of users sharing Duets. Additionally, bursty patterns of interactivity have also been observed during the process of information diffusion on online platforms as users share content with one another. During this process of information diffusion on online platforms, studies have shown that the majority of content propagation occurs shortly after the initial posting and quickly drops off [37, 53, 89]. Previous research has also found that when bursts of attention are recurring, for example when an article is re-shared on a social media platform, these later bursts tend to both decrease in volume and have increasingly larger periods of time between them [37]. The decrease of collective attention over time, seen in the diminishing returns of recurring bursts, has been attributed to a decrease in novelty due to previous audience exposure to the content [37, 53, 89]. In addition to being a form of interaction, Duets also serve as a mechanism of content diffusion, and for Duet chains, a new generation being added to an existing chain can be viewed as a recurring burst of activity. We examine whether, on average, the amount of time in between generations of a Duet chain also tends to increase, indicating decreasing collective attention.

METHODOLOGY

Data Collection

For this study, we collected digital trace data from publicly accessible TikTok Duet videos. We started by considering the case of the TikTok video of the song 'The Wellerman', sung by Nathan Evans, which became a viral sensation in early 2021 both on the TikTok platform and beyond. By searching the 'wellerman' hashtag on TikTok, we were able to manually collect an extensive list of the Duet videos which could be traced back to the original Wellerman video, providing us with a case study that allowed us to examine the structure and content of the video interaction network. In order to gather a larger data set that would also allow us to look at the temporal trends of interactivity, we additionally collected the first 130 Duet videos urls that were returned from a manual search run in mid March 2022 on the TikTok website for videos tagged with the 'duet' hashtag. From this initial data set, we took the list of TikTok users who had posted the Duet videos, and used a Selenium-based web scraper which checked the 10 most recent videos that had been posted to their account. We then collected data on those videos that were marked with the 'duet' hashtag. This data consisted of the video url, the TikTok user handle of the video poster, the date the video had been posted, and all hashtags that had been added to the video. The script also collected the same information for the original parent video that had been Dueted, meaning that thr account for the user who had posted the parent video would also be scraped for Duet videos the next time the script was run. This script was then run roughly 1-2 times a week for the months of May and June 2022. Additionally, the script checked all parent videos that were collected in order to determine if they had also been Dueted. If so, then the grandparent video data was also collected, and the parent-grandparent video pair would also be added to the data set. This was repeated until the original root video was returned. If we were unable to trace the chain back to the root video due to missing data, that chain was excluded from the dataset.

In order to filter our data to only include videos with musical content, we utilized the hashtags that accompanied each video. From our data set, we compiled a list of all unique hashtags, and manually coded them to identify those which indicated musical content. These included generic music terms (eg. music, harmony, song), instrumentation and voice parts (eg. violin, alto), genre and style descriptors (eg. rock, indiepop), and song information such as the song name, artist name, or lyrics. We then used this list to filter the Duet videos to only include those which had been tagged with one of the musical content hashtags, or which were in a Duet chain that included a video tagged with a music content hashtag. Additional manual review of the videos was also performed in order to remove Duet chains that contained musically relevant hashtags, but did not have music as the primary focus of their content, such as dance videos or videos where the music was used as background. This left us with 3,242 unique videos, 1,073 of which were root videos, created by 1,359 unique users and giving us 2,173 parent-child pairs in total.

Data Analysis

Tree Structure. Because our data collection process preserved the parent-child relationship between each child Duet video and the parent video it was Dueted from, we can capture the tree structures that are formed by longer Duet chains. We can see from the example in Figure 2 that by tracing the parent-child ties within a given tree, we can identify the unique paths of nodes between the root node and each of the terminal leaf nodes, which represent the final videos in each Duet chain which were not themselves Dueted. We can then identify which generation an individual video is based on its distance from the root node, with the root node indicating generation zero. With this, we can describe individual trees based on the total number of root-to leaf paths they contain, and the number of generations in their longest path. Within the trees that developed from these root nodes, the total number of paths per tree ranged from 1 to 189, with the total length of individual paths ranging from 1 to 18 generations.

Although the child video is linked to the parent video within our data set, this is a one directional tie, and the parent video does not contain metadata that links it to any of its child videos. As a result, it is not possible to determine what percentage of all child videos we have captured for each of the root videos in our data set. Because of this, we limited our tree structure analysis to the three largest trees in the dataset. For these, we performed additional manual searches on TikTok using relevant hashtags to ensure we had captured the majority of the videos that linked back to the root nodes, thereby capturing the majority of the tree structure. These three trees were the 'Scream It Out' tree, with 81 paths and a maximum path length of 3, the 'Misty Mountains' tree, with 14 paths and maximum path length of 18, and the 'Wellerman' tree, with 189 paths and a maximum path length of 16. The Scream It Out root video, which refers to one of its hashtags, is a short, humorous clip of a woman about to record herself singing, who accidentally falls, with the video catching the surprised yell she makes. The Misty Mountains root video features the song 'Misty Mountains' from The Hobbit movie being performed a cappella by 4 male bass vocalists. The Wellerman root video features an a cappella male tenor vocalist singing the sea shanty of the same name. One interesting note is that the root video for the Misty Mountains tree was actually created by the TikTok user account, The Wellermen, so named because it was created by a handful of male vocalists who all connected through contributing to the Wellerman Duet tree [86].

Video Content Analysis. In the case of Duet videos, each video is the result of a user interacting with an existing video by Dueting it. The methodology of interaction analysis involves examining the impact of each individual's contribution as the collaboration unfolds over time [71]. Since we are looking at this from the perspective of the creative development of the artifact, we want to focus on measuring novelty, as creativity is typically defined in terms of novelty or originality [69]. For Duet videos, this entails comparing the key perceptual features of the child video content to that of the parent video [7, 62, 70]. This can be done using computational analysis of Music Information Retrieval (MIR) data to map the development of the artifact as it undergoes each of these interactions. However, because Duet videos are created by layering additional audio over the parent video, many of the quantitative audio features that we could capture with MIR data would be effectively the same for both the parent and child video. These include things like tempo, musical key, whether the video is in a major or minor key, the duration of the audio, and the time signature. To account for this, we instead analyzed video content using automatic music tagging, a classification process which predicts the semantic tags most likely to be relevant for an audio sample. These semantic tags are descriptive keywords that contain high-level meta-data about the audio qualities of the music sample in questions, such as genre, mood, the time period the song is from, and the instruments being played [24, 31]. The process of automatic music tagging consists of first identifying which audio features extracted from the audio signal are most relevant for predicting the semantic tags, and then training a model on a multi-label classification task to assign the correct tags to the audio samples based on this feature data. For feature selection, deep convolutional neural networks are often used, as they are able to learn hierarchical features [24].

To analyze the music content of the videos in our data set, we used a publicly available Pytorch implementation of a Harmonic Convolutional Neural Network for automatic music tagging [87, 88]. The model, developed by Minz Won, is trained on audio samples that have been notated with descriptive semantic feature tags containing information about the aural attributes of the music. When given

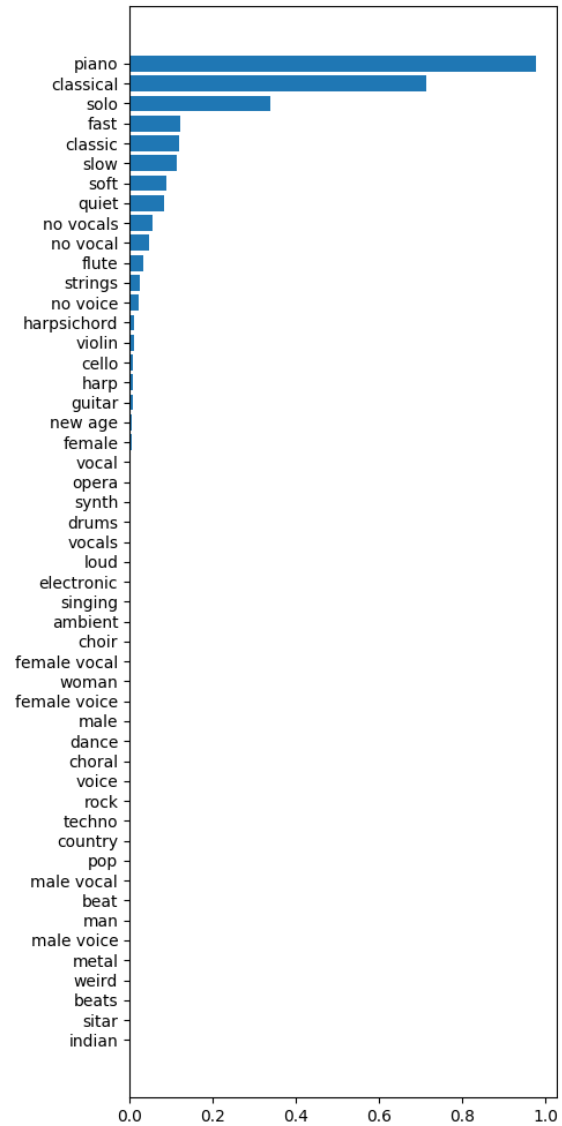

Figure 3: Output of a Harmonic CNN semantic tagging model, indicating the probability that each of the semantic features is accurate for the audio input being analyzed.

new audio samples, the network extracts the audio data and returns the probability distribution for the top 50 semantic features most frequently used in the training data, as demonstrated in Figure 3. This allowed us to generate a feature vector for each video from the values of the probability distribution across the same 50 semantic descriptors. With this data, we were able to compare the difference between each child video's feature vector and its parent video feature vector to quantify the change in aural features for each new version of the video artifact. For example, if a parent video consisted of an a cappella singer, and the child video added a trumpet player, the feature vectors for the two videos would be distinct, even if the trumpet player was doing nothing but doubling the singers part exactly. To calculate the aural novelty of each new video with respect to its parent video, we measured the Euclidean distance between the two feature vectors. We opted for a distance metric as opposed to cosine similarity, because the goal of our analysis was to not just compare similarity between parent-child pairs, but to also capture the magnitude of the aural feature change that occurred between the first and last video within a Duet chain.

Content Diversity Analysis. Using the individual feature vectors for the videos within a Duet tree, we are able to analyze how the artifact changes over the course of the interactions that it undergoes. For each of the paths within a given tree, we can look at the distance between each video's feature vector and the feature vectors of the videos that came before it in the path in order to understand how much overall change is being introduced to the artifact, and how the artifact is moving through the feature space. To examine the overall change within a path, we calculated the net distance between the root video's feature vector, and the feature vectors of each video in the path. This allows us to see the net distance from the root node that each path travels within the feature space for each successive generation. From this, we can see the total magnitude of change being made to the aural features of the artifact over time. Additionally, we calculated the distance between each child video's feature vector and its parent video feature vector, in order to see the relative magnitude of change between each generation of an individual path. We then also took the cumulative sum of these distances across the entire path in order to calculate the absolute distance within feature space that the path had traveled over each successive generation.

Temporal Metrics. In order to consider the temporal nature of user interactivity with Duet videos, we need to look at the distribution of inter-event times, which is the amount of time that passes between when the parent video is posted, and when the child video is posted. Within our data set, we were able to collect data about the posting date for 3,030 videos across 1,268 users, giving us 2,005 pairs of parent-child videos for which we could calculate the inter-event time. For our analysis, we considered both the amount of time that had passed between when the parent video was posted and the child video was posted, as well as which generation the child video was within its Duet chain. The available TikTok data for when each video was posted only included a date and not a time stamp, therefore the time in between video events for each child-parent pair was calculated by subtracting the parent video date from the child video date to yield the number of days that elapsed.

RESULTS

Tree structure

To examine the tree structures that were formed, we selected the three largest trees from the data as case studies. These were the 'Scream It Out' tree, with 81 paths and a maximum path length of 3, the 'Misty Mountains' tree, with 13 paths and maximum path length of 18, and the 'Wellerman' tree, with 189 paths and a maximum path length of 16. The tree structure for each of these can be seen in Figures 4, 5, and 6. These three examples illustrate that there can be a great deal of variety in the tree structures that are formed by Duet chaining. These differ from more traditional creative collaborations, in that there is no defined group of collaborators, and the collaborative process is both uncoordinated and decentralized. Most importantly, we see that there is no one final end product, rather the original artifact is able to branch off into multiple versions. This is distinct from the other types of online collaborations we have seen, both small and large-scale, where coordination tools are used to allow collaborators to all work on the same artifact throughout the project. Duet chains therefore contrast sharply with more traditional creative collaborations, in that there is no one defined group of collaborators, and the number of contributors is able to scale considerably.

Content Diversity

We then examined the content diversity of the videos contained within these three case studies. As seen in Figures 7, 8 and 9, the graphs on the left display the net distance within feature space traveled by each path, and the graphs on the right display the the relative magnitude of change between each generation of an individual path, and the cumulative amount of absolute distance each generation of the chain has traveled within the feature space.

We can observe how novelty can be introduced through the successive iterations that occur within an individual path, which captures the evolution of an individual instance of the artifact. With the Misty Mountains tree in Figure 8 and the Wellerman tree in Figure 9, we can see the aggregation process and its effect on the novelty of a path over many generations. From these examples we can observe that the net distance traveled within the feature space tends to increases over time. However, because the Duet feature preserves each iteration of the artifact along a path, the amount of novelty represented by a path is not solely limited to how far from the root node the leaf node of a path is, but also by the amount of exploration that is happening as it moves within the feature space. This is demonstrated by the fact that the cumulative distance a path travels through feature space is not necessarily the same as the net distance that it travels from the root node. We can also observe how different paths that develop from the same root node can end up moving into different areas of feature space. By applying Principal Component Analysis to the video feature vectors, we can map this movement within a 2-dimensional representation of the feature space, and look at the differences in the relative positioning of two different paths.

In Figure 10 we compare the movement within the feature space for two different paths from the Wellerman tree. Here we can see that the two paths overlap for the first few generations and then branch off. Following that, each successive generation within the individual paths brings them to a new location, and they ultimately end up in two different areas of the feature space. For the first few generations where the paths overlap, we saw a capella male bass harmonization added. Path A, which was 9 generations long, then diverges from the other path with the addition of an electronic bass line, created in the style of bump house music. Following that, additional acoustic violin voices are added. The leaf node video for Path A can be viewed at https://www.tiktok.com/@annaeyink/video/6921079536071347461. Path B, instead has only acoustic contributions. Following the point where the two

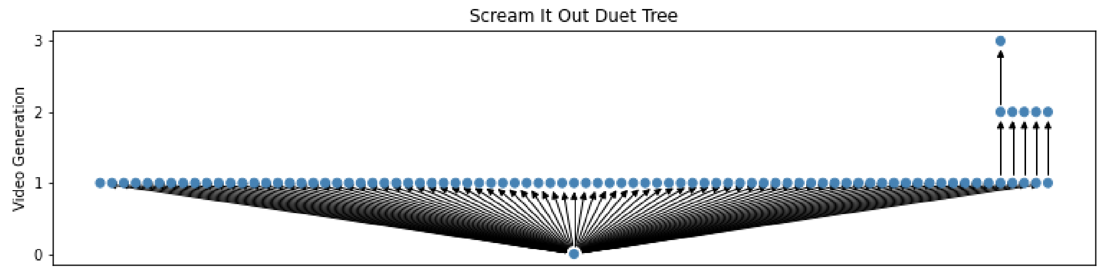

Figure 4: In the tree formed from Duet chains originating from the 'Scream It Out' root video, we see a large number of paths, but all of the paths are very short.

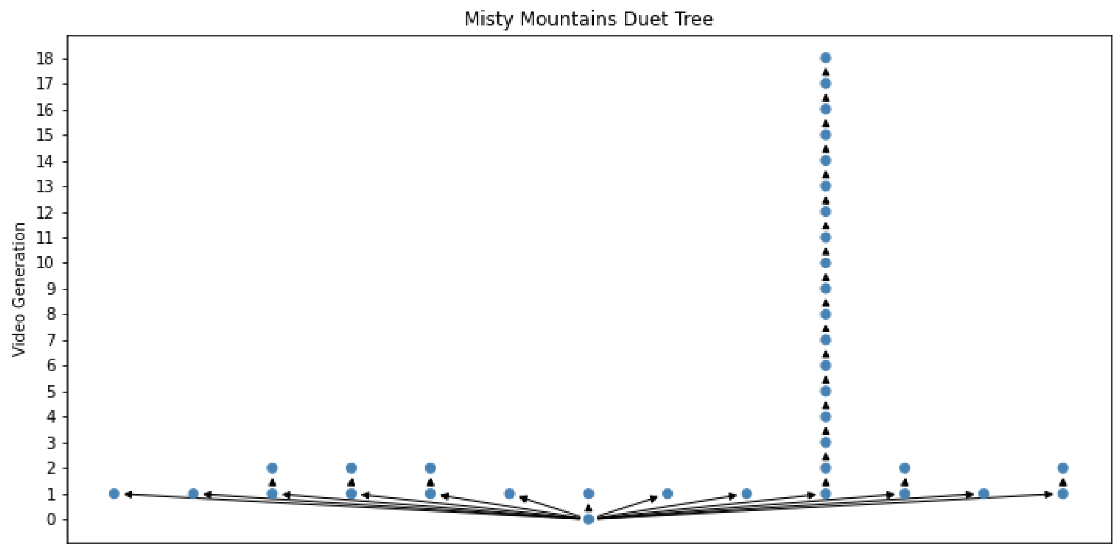

Figure 5: In the tree formed from Duet chains originating from the 'Misty Mountains' root video, we see fewer total paths, however this tree has a path which is 18 generations long, the longest in our data set.

paths diverge, Path B also sees the additional of acoustic violin, along with female soprano, flute, and harp parts. The leaf node for Path B can be viewed at https://www.tiktok.com/@harpistkt/video/6918390090163096838. We can see that despite the two paths sharing the same starting point, and having some commonalities in the types of instruments and voice parts they contain, the end results are aurally quite different, which is reflected by their diverging positions within the feature space. We can contrast this with the Misty Mountains tree, where we see the amount of aural change over time in the longest path is more gradual and consistent than those of the paths in the Wellerman tree. If we look at the content of this Misty Mountain path, which can be viewed at https://www.tiktok.com/@_luke.the.voice_/video/7072762537795931438, we can see that the majority of the contributions are additional male bass vocals, hence with each additional generation there is minimal change to the aural attributes, and therefore less distance between the feature vectors. Regardless of its starting point, where a path ends up traveling to in feature space is entirely dependent on the specific contributions made over the course of that path. This demonstrates that TikTok Duets are indeed examples of distributed creativity, as their creative development is an emergent property of the specific network of interactions that each individual path undergoes.

What is also interesting about TikTok Duet trees is that we can not only see the evolution of the artifact over time by following a unique path, but we can see all of the alternative branching paths that are created. As we can see in Figure 7, the Scream It Out tree is a good example of this, as it contains a number of unique paths that all branch out from the root node. From this, we are able to

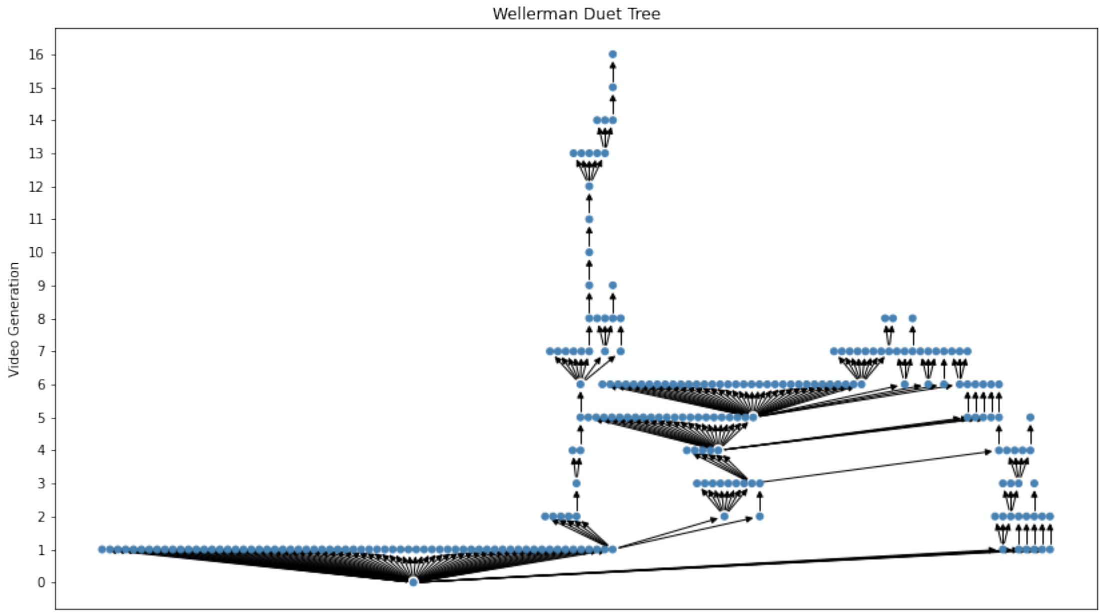

Figure 6: The tree formed by Duet chains originating from the 'Wellerman' root video is the largest in our data set, with both a large number of paths, and more paths containing a higher number of generations than the average for our data set.

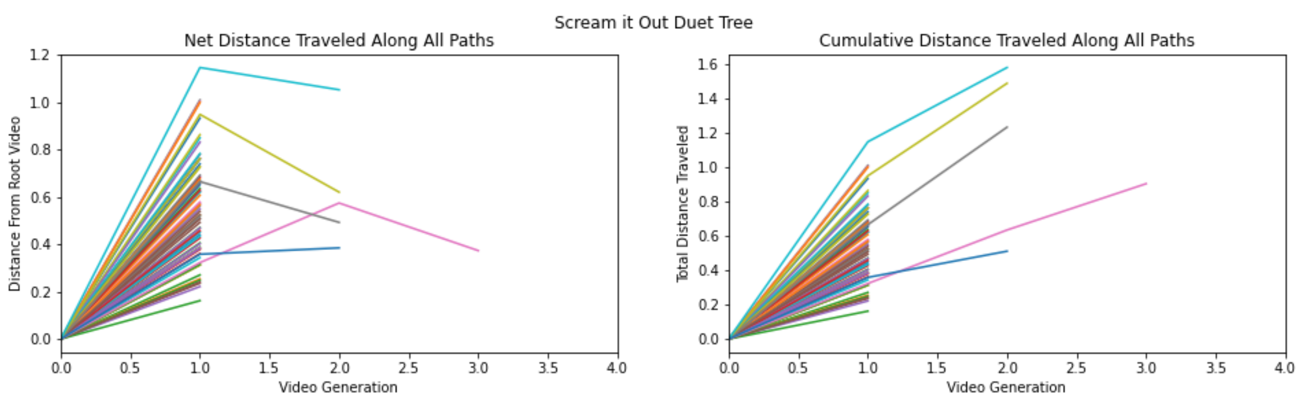

Figure 7: The figure on the left displays the net distance from the root node that each path travels within the feature space with each additional generation. The figure on the right displays the cumulative amount of absolute distance each generation of the chain has traveled within the feature space, allowing us to see the gradient of change between each successive generation. Because the Scream it Out tree contains a high number of short paths, we see lots of initial branching.

observe how multiple branching instances of an artifact allows for the parallel development of novelty. In addition to comparing the relative positioning of individual paths, we can also look at the total span of a Duet tree within feature space across all paths and all generations, as seen in Figures 11, 12, and 13. This makes TikTok Duets distinct, in that they are able to facilitate and capture the proliferation of diversity that results from this decentralized collaborative process.

Within these Duet trees, we can see how the amount of aural diversity increases both through the aggregation of additional content over successive generations within an individual path, and

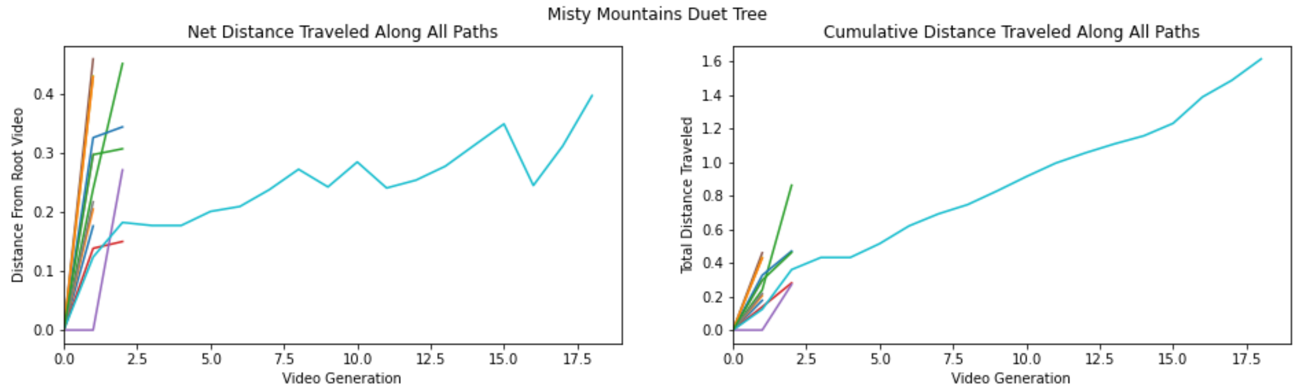

Figure 8: The figure on the left displays the net distance from the root node that each path travels within the feature space with each additional generation. The figure on the right displays the cumulative amount of absolute distance each generation of the chain has traveled within the feature space, allowing us to see the gradient of change between each successive generation. We see initial branching, with the longest path showing slight increases in total distance over successive generations.

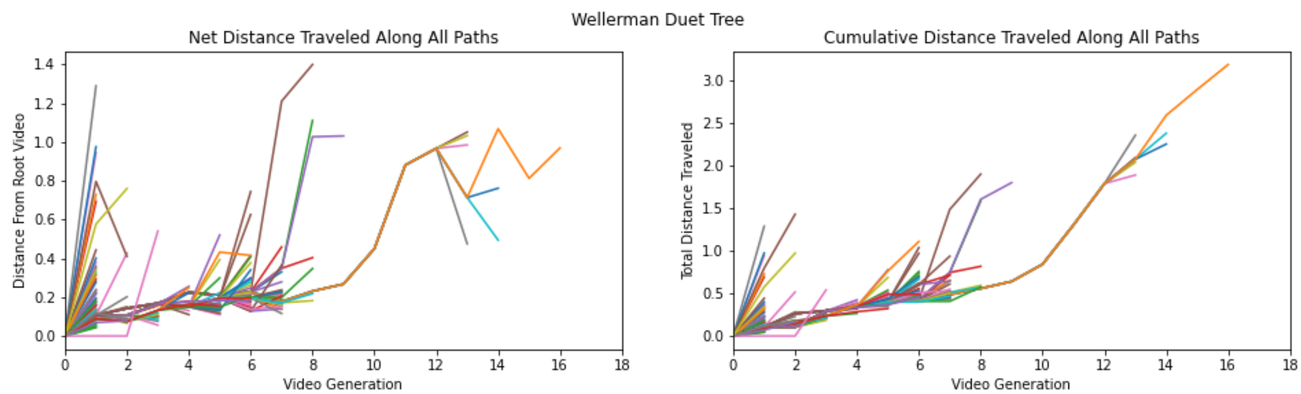

Figure 9: Figure on the left displays the net distance from the root node that each path travels within the feature space with each additional generation. The figure on the left displays the cumulative amount of absolute distance each generation of the chain has traveled within the feature space, allowing us to see the gradient of change between each successive generation. Here we see a great deal of variation in the distances traveled by the different paths, and can observe that some paths show sharp increases in net and cumulative distance at different generations.

through the proliferation of branching paths which can evolve independently. This distributed creative process is enabled by the Duet feature's ability to create multiple digital instances of an artifact which can each be interacted with and developed independently. We also observed examples of content diversity which were outside the scope of our music tagging data. For example, one of the videos within the Wellerman tree, which can be viewed at https://www.tiktok.com/@marypapageorge/video/6918860151374073093, features an Irish dancer, and is interesting because it introduces a new kind of interaction not previously seen in the chain. It is also interesting because while it is common for dance and music to be performed together, contextualizing the dancing within a musical Duet chain reframes it as a music generating activity through its addition of a percussive component. Additionally, another video in the Wellerman tree, seen at https://www.tiktok.com/@veepats/video/6918206108133870854 and in Figure 14, features a user performing a sign language translation of the song. Here we see the video modality creating new opportunities for expression and allowing a more diverse community to engage with and contribute to the content [58, 77].

Engagement

When considering the time between parent and child videos, we can see in Figure 15 that the distribution of time in between video events is very right tailed, suggesting that Duet interactivity exhibits bursty behavior [8]. Additionally, when we model the distribution of inter-event times with a probability density function as shown in Figure 16, we can see that the probability of a parent video being Dueted drops off very quickly as we get further from the time the video was originally posted.

To then test whether the time in between bursts also increases within Duet chains, we can compare the distributions of total time

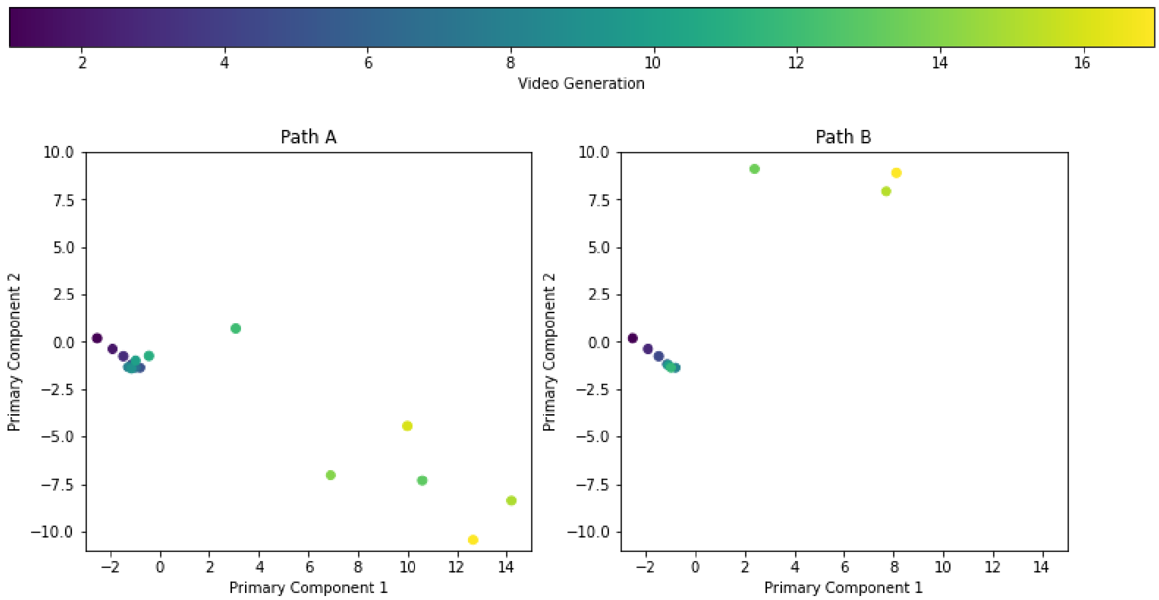

Figure 10: Using a 2 component PCA, we map two paths from the Wellerman tree to semantic feature space. We can observe that the paths branch off from one another and develop different aural characteristics, which is represented by their diverging positions within the feature space over successive generations.

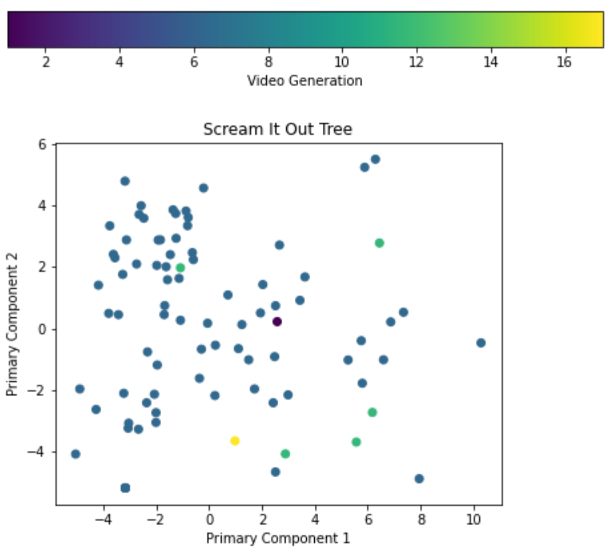

Figure 11: A 2 component PCA analysis of the Scream It Out tree videos, mapping each video to its relative position within the semantic tagging feature space.

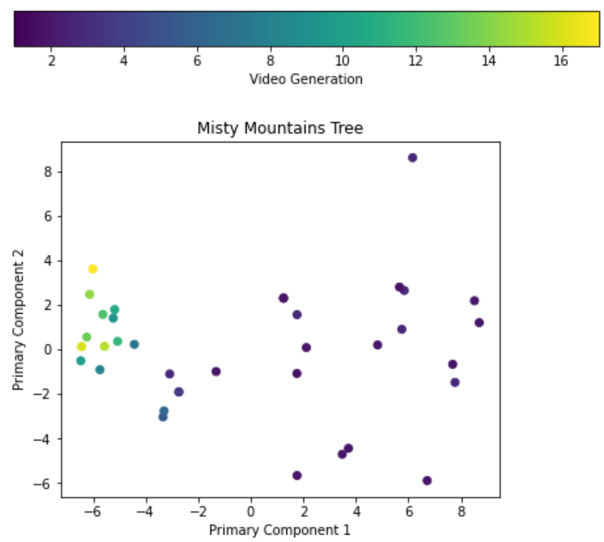

Figure 12: A 2 component PCA analysis of the Misty Mountains tree videos, mapping each video to its relative position within the semantic tagging feature space.

between parent-child video pairs for successive generations of Duets. Because these distributions are heavy-tailed, we can model each distribution using a power law, which takes the form:

$$p(x) \propto x^{-\alpha} \tag{1}$$

Since power laws are scale invariant, this allows us to compare the different generations' distributions to one another, even though we do not have the same amount of data for each generation. This can be done by comparing the alpha values that are returned when

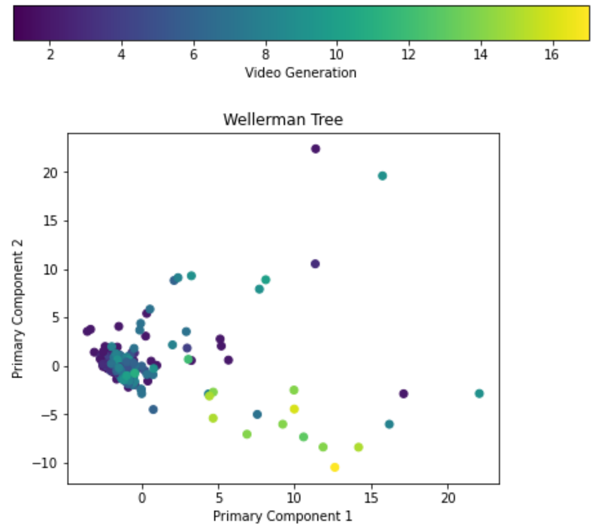

Figure 13: A 2 component PCA analysis of the Wellerman tree videos, mapping each video to its relative position within the semantic tagging feature space.

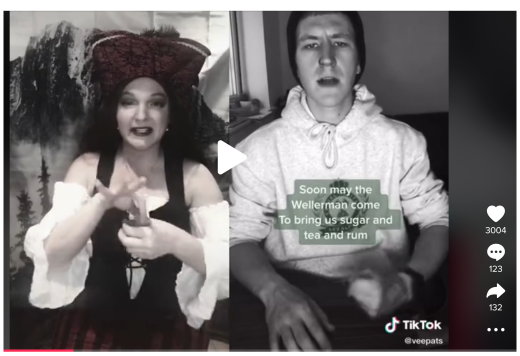

Figure 14: A sign language interpretation of 'The Wellerman' is added to the original song via the Duet feature.

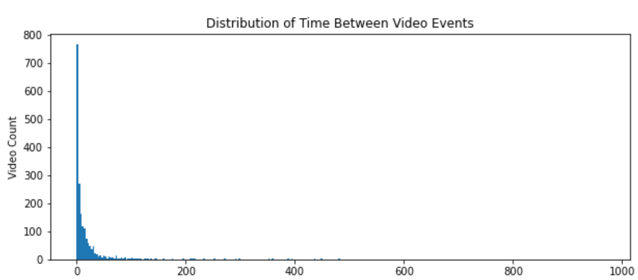

Figure 15: Distribution of total time in days between when parent video is posted and child video is posted.

each generation's distribution is fit to a power law. For power law distributions, a higher alpha value indicates a higher proportional

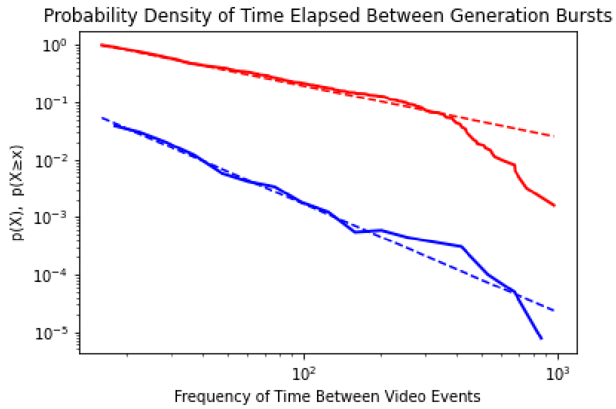

Figure 16: Probability density function (p(X), blue) and complementary cumulative distribution function (p(X §x), red) of frequencies of times in between parent-child video pairs.

concentration of smaller values within the data. If on average, time in between generations increases for Duet Chains, like we've seen with other examples of online diffusion, then we would expect to see the alpha value decrease for each successive generation [37]. Using the Python powerlaw package, we fit the different generations to a truncated power-law distribution to find the alpha value and standard error, as seen in Table 1 [3]. We tested each distribution for goodness of fit using a one-way chi-squared test between the observed data and simulated data generated from the fitted distribution, and found no statistically significant differences within the bounds of the observed data.

Table 1: Power-law distribution fit for inter-event times of different generations of Duet chains.

Generation	Alpha	Standard Error
1	1.2	0.005
2	1.36	0.02
3	1.46	0.05
4	1.27	0.05
>=5	1.26	0.02

Interestingly, what we find is that the alpha values increase slightly for the first three generations before then decreasing. When considering why this might occur, it is important to note that for TikTok Duets, interacting with a Duet chain doesn't just involve re-sharing the video, but also adding to the video content in some way. With TikTok, the defining feature of the Duet function is that it allows a user to interact with existing videos by adding their own material to create and share a new version of the artifact [44]. Therefore, unlike in previous studies of online information dissemination [37, 53, 89], where content remains static during the process of propagation, the act of sharing content via the Duet function also introduces a change to the content. It is possible that the dynamic nature of the Duet video content means that novelty does not immediately begin to decay after the first burst due the creation of new iterations of the original artifact.

Another consideration is that while the Duet chain videos from later generations might help bring attention to the Duet tree that is developing, users viewing them might opt to create a new chain by Dueting the root video, rather than adding to an already existing chain. This could potentially explain why we see longer inter-event times for first generation videos. In Figure 17 we can see from the timeline of when videos in the Scream It Out tree were posted that there is initially a delay before activity peaks. It is possible that users saw the child videos, rather than the root video, and this is what inspired them to then Duet the root video. If we look at the timeline for the Wellerman tree in Figure 18, we can also see a similar pattern on 2021-01-15, where a peak in posts of first generation videos occurs soon after we observe a larger peak in posts of later generation videos. This suggests that the social visibility of the platform also plays a role in motivating user engagement with Duets.

DISCUSSION

Our results highlight the ways in which the technical affordances of the TikTok Duet feature shape the creative process of its users, and the ways that creative artifacts can evolve on the platform. We see the occurrence of distributed creativity at scale, which is facilitated by the social nature of the platform, the ability of individuals to easily view and contribute to creative projects, and the ability for digital artifacts to proliferate into multiple versions.

Collaborative Exploration of Conceptual Space

The Duet feature allows for an iterative, chaining process of creative development to occur, which has been shown to lead to greater diversity in cultural artifacts over time [33]. Additionally, due to the digital nature of the artifacts, the Duet feature is able to preserve all versions of an artifact that have been created, and can 'clone' any version of the artifact. Because Dueting a video creates a new instance of the artifact, a Duet chain is able to branch into multiple paths at any given point, which allows it to diverge and proliferate into multiple distinct versions. Because the development of one path doesn't necessitate pruning of the other paths, we see a form of decentralized collaboration emerge, where multiple individuals are able to engage in parallel development of creative ideas that spring from the same initial seed idea. As a result, we see that users are able to engage in a collective exploration of the conceptual space that the artifacts exist in [15]. Because there is no structure put in place to enforce a convergence to a single final product, such as we see with Wikipedia where there can only be one version of the article at a time, the TikTok Duet function is able to support collaborative ideation at scale by enabling the development of diverse artifacts [1, 74, 92]. With the example of Paths A and B from the Wellerman tree in Figure 10, we can see that this can enable greater exploration within the conceptual space [72, 91]. Allowing for both chained and parallel development of artifacts also means that the scale and diversity of the platform's user base can be leveraged to bring in a wide range of ideas and perspectives. This can lead to unexpected creative developments, as seen with the the sign language version of the Wellerman video, in which we see new ways for users to engage with the content. By introducing new paradigms for how individuals can experience and engage with music, these artifacts have the potential to impact creative culture beyond the TikTok platform. The methodological approach we have introduced in this paper allows us to analyze this process of distributed creativity in terms of the exploration of the overall conceptual space and the degree of content diversity present within the collaboration. Additionally, this approach allows us to explore how a single seed idea can proliferate into a range of artifacts through different combinations of ideas, which could provide valuable insight into the process of creativity more generally.

Social Drivers of Creative Participatory Cultures

The Duet feature creates an environment which supports the generation of diverse ideas and artifacts through user interaction. The interaction networks we can observe serve to illustrate how the visibility and accessibility of shared representations on the platform lead to collaborative engagement [21]. This is enabled by the high degree of social transparency and visibility of users activities and content which is available on the TikTok platform. The scale of the platform therefore provides users with a wide range of opportunities for participation, while its high degree of social transparency allows users to make inferences about the potential creative and social value of engagement with a given artifact [25]. The visibility of the platform can also positively impact the creativity of individual contributors by exposing them to a wider range of diverse ideas [68]. As a result, we see that when distributed creativity occurs in digital spaces, the boundaries of what defines a collaborative group become more permeable. As a result of both the platform's social transparency and the flexibility of how and when new artifact versions can be created, the porous nature of collaborations on TikTok allows for ideas to be shared across a larger network of users, increasing the potential diversity of ideas that can be achieved.

Additionally, when examining creative systems, it is important to recognize that there is a conceptual distinction between the act of creativity, and a creative product [81]. On the TikTok platform, participation in the creative process is framed as a way of connecting and engaging with the wider community. This is evidenced in our analysis of the temporal dynamics of Dueting, where we observe the burstiness characteristic of human interactivity, but not an immediate drop off in engagement with later bursts that has been observed in cases of static content sharing. Due to the accessible and dynamic nature of Duet artifacts, the act of Dueting allows users to engage with one another through a shared creative experience [44, 49]. The case of the Wellerman Duet chain sparking the formation of the collaborative Wellermen group account, is one example of how Duets allow users to connect and engage with the wider community through participation in a shared creative process. Social interaction is therefore a driver for Duet engagement, and is able to be leveraged into distributed creativity due to the transparency of the TikTok platform, and the open-ended nature of Duet collaborations.

Implications

Our results highlight the importance of digital tools for collective ideation, as well as the necessity of examining how interactions

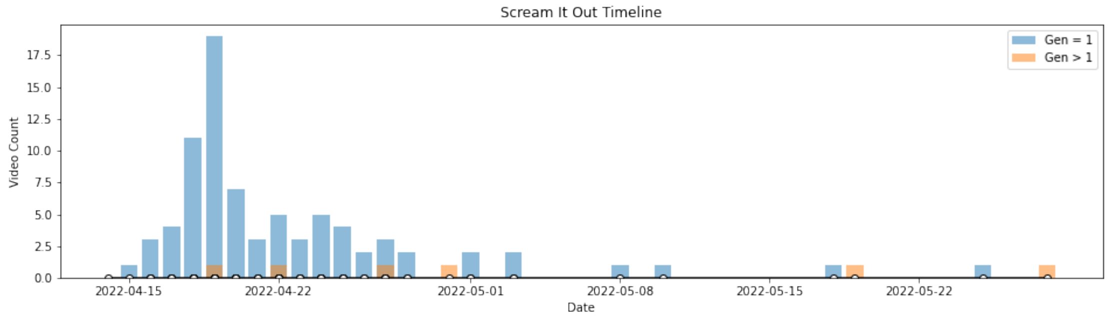

Figure 17: Timeline depicting the total number of times per day a video within the Scream It Out tree was Dueted. The blue bar represents that number of first generation videos, while the orange bar represents that number of videos that are second generation or later.

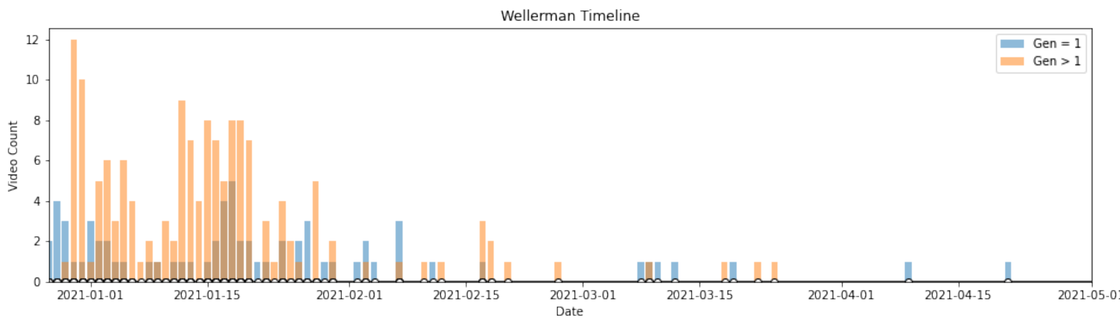

Figure 18: Timeline depicting the total number of times per day a video within the Wellerman tree was Dueted. The blue bar represents that number of first generation videos, while the orange bar represents that number of videos that are second generation or later.

between social and technical forces affect the development of participatory cultures online. From this we can identify key HCI affordances that can be leveraged to support distributed creativity at scale. We see that artifacts are easily accessible, and serve as shared representations encourage engagement across larger collaboration networks. Additionally, it is important that these digital artifacts are able to support dynamic versioning. In addition to preserving the full history of the artifacts development, this also allows for digital artifacts to proliferate into multiple versions, and allows for parallel development across multiple collaborators. This flexibility in copying and preserving multiple versions of digital artifacts allows for a collective exploration of the conceptual space, which leads to increased idea diversity. Finally, transparency and visibility into the activities of other users on the platform is needed in order to tap into social motivators that drive creative engagement and the development of larger collaborative networks. These are considerations which can be used to design platforms that support large scale distributed creativity and facilitate diverse creative ideation in digital environments [1, 74].

Limitations

This study analyzed TikTok Duet videos, which means the data was limited to the TikTok platform, and to the Duet feature, which is currently a functionality specific to TikTok. Although similar examples of musical collaboration have been observed on other platforms, additional research comparing these patterns in different contexts and with different kinds of content would help to further examine the influence of specific technical affordances on the creative process. Additionally, this analysis did not include social network information or success metrics such as views, likes, or the potential effect of Tiktok recommendation algorithms. Since these are factors that can influence the visibility of shared content and the motivation for users to engage with it, future work should investigate their impact on user interactivity patterns.

Additionally, this study looked solely at music collaboration. To understand digital distributed creativity more broadly, future work should also examine digital collaborations within other creative domains, such as visual art, literature, etc. Although the HCI affordances outlined above are applicable to different platforms and different creative domains, it is possible that we might see the emergence of different collaboration practices and interactions patterns due to variations in the available perceptual affordances of different systems and creative domains.

CONCLUSION

In this paper, we analyze TikTok video content to identify the ways in which the technical affordances of the Duet feature support co-creativity and engagement between its users. We examine how the affordances of the Duet feature facilitate a distributed and collaborative creative process, in which we can observe the evolution of the cultural artifacts through the different versions that are produced through user contributions. By analyzing the audio content of these videos with an automated semantic music tagging neural network, we are able to quantify information about how video artifacts evolve over time. Additionally, we find that the open-ended nature of collaboration with the Duet function prioritizes engagement as both a creative and social act.

REFERENCES

[1] Piotr D. Adamczyk, Kevin Hamilton, Michael B. Twidale, and Brian P. Bailey. 2007. Tools in support of creative collaboration. In *Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition (C&C '07)*. Association for Computing Machinery, New York, NY, USA, 273–274. https://doi.org/10.1145/1254960.1255010

[2] Ali Algarni. 2019. Exploring Neuroscience of Creativity Theories in designing Co-creative Agents. In *Proceedings of the 2019 on Creativity and Cognition (C&C '19)*. Association for Computing Machinery, New York, NY, USA, 686–690. https://doi.org/10.1145/3325480.3326559

[3] Jeff Alstott, Ed Bullmore, and Dietmar Plenz. 2014. powerlaw: A Python Package for Analysis of Heavy-Tailed Distributions. *PLOS ONE* 9, 1 (Jan. 2014), e85777. https://doi.org/10.1371/journal.pone.0085777 Publisher: Public Library of Science.

[4] Teresa M. Amabile and Julianna Pillemer. 2012. Perspectives on the Social Psychology of Creativity. *The Journal of Creative Behavior* 46, 1 (2012), 3–15. https://doi.org/10.1002/jocb.001 _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jocb.001.

[5] Andre Antunes. 2022. If System of a Down were from India. https://www.youtube.com/watch?v=mi106DZJhuQ

[6] Berkeley Andrus and Nancy Fulda. 2022. A Data-Driven Architecture for Social Behavior in Creator Networks. In *Proceedings of the 13th International Conference on Computational Creativity (ICCC)*. pages 339–348.

[7] Noah Askin and Michael Mauskapf. 2014. Cultural Attributes and their Influence on Consumption Patterns in Popular Music. In *(eds) Social Informatics. SocInfo 2014*, Luca Maria Aiello and Daniel McFarland (Eds.). Lecture Notes in Computer Science, Vol. 8851. Springer International Publishing, Cham, 508–530. https://doi.org/10.1007/978-3-319-13734-6$_$36

[8] Albert-László Barabási. 2005. The origin of bursts and heavy tails in human dynamics. *Nature* 435, 7039 (May 2005), 207–211. https://doi.org/10.1038/nature03459 arXiv: cond-mat/0505371.

[9] Álvaro Barbosa. 2003. Displaced Soundscapes: A Survey of Network Systems for Music and Sonic Art Creation. *Leonardo Music Journal* 13 (Dec. 2003), 53–59. https://doi.org/10.1162/096112104322750791

[10] Raiyan Abdul Baten, Daryl Bagley, Ashely Tenesaca, Famous Clark, James P Bagrow, Gourab Ghoshal, and Ehsan Hoque. 2020. Creativity in temporal social networks: how divergent thinking is impacted by one's choice of peers. *J. R. Soc. Interface* 17, 171 (Oct. 2020), 11. https://doi.org/10.1098/rsif.2020.0667

[11] Yochai Benkler. 2006. *The wealth of networks: how social production transforms markets and freedom.* Yale University Press, New Haven [Conn.]. OCLC: ocm61881089.

[12] Thierry Bertin-Mahieux. 2013. *Large-Scale Pattern Discovery in Music.* Ph. D. Dissertation. Columbia University. https://doi.org/10.7916/D8NC67CT

[13] Tina Blaine and Sidney Fels. 2017. 2003: Contexts of Collaborative Musical Experiences. In *A NIME Reader*, Alexander Refsum Jensenius and Michael J. Lyons (Eds.). Vol. 3. Springer International Publishing, Cham, 71–87. https://doi.org/10.1007/978-3-319-47214-0_6 Series Title: Current Research in Systematic Musicology.

[14] Norbert Blenn and Piet Van Mieghem. 2016. Are human interactivity times lognormal? *arXiv:1607.02952 [physics]* (July 2016). https://doi.org/10.48550/arXiv.1607.02952 arXiv: 1607.02952.

[15] Margaret A. Boden. 2004. *The Creative Mind: Myths and Mechanisms.* Routledge.

[16] Maximilian Boeker and Aleksandra Urman. 2022. An Empirical Investigation of Personalization Factors on TikTok. In *Proceedings of the ACM Web Conference 2022 (WWW '22)*. Association for Computing Machinery, New York, NY, USA, 2298–2309. https://doi.org/10.1145/3485447.3512102

[17] Sylvie Le Bomin, Guillaume Lecointre, and Evelyne Heyer. 2016. The Evolution of Musical Diversity: The Key Role of Vertical Transmission. *PLOS ONE* 11, 3 (March 2016), e0151570. https://doi.org/10.1371/journal.pone.0151570 Publisher: Public Library of Science.

[18] Jean-François Boujut and Eric Blanco. 2003. Intermediary Objects as a Means to Foster Co-operation in Engineering Design. *Computer Supported Cooperative Work (CSCW)* 12, 2 (June 2003), 205–219. https://doi.org/10.1023/A:1023980212097

[19] Albert S. Bregman. 1994. *Auditory Scene Analysis: The Perceptual Organization of Sound.* MIT Press. Google-Books-ID: jI8muSpAC5AC.

[20] Melanie S. Brucks and Jonathan Levav. 2022. Virtual communication curbs creative idea generation. *Nature* (April 2022), 1–5. https://doi.org/10.1038/s41586-022-04643-y Publisher: Nature Publishing Group.

[21] Nick Bryan-Kinns. 2013. Mutual engagement and collocation with shared representations. *International Journal of Human-Computer Studies* 71, 1 (Jan. 2013), 76–90. https://doi.org/10.1016/j.ijhcs.2012.02.004

[22] Ronald S Burt. 2021. SOCIAL NETWORK AND CREATIVITY. In *Handbook of Research on Creativity and Innovation.* Edward Elgar Publishing, pp82–104. https://www.e-elgar.com/shop/usd/handbook-of-research-on-creativity-and-innovation-9781788977265.html

[23] Ana Caraban, Evangelos Karapanos, Daniel Gonçalves, and Pedro Campos. 2019. 23 Ways to Nudge: A Review of Technology-Mediated Nudging in Human-Computer Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3290605.3300733

[24] Keunwoo Choi, George Fazekas, and Mark Sandler. 2016. Automatic tagging using deep convolutional neural networks. https://doi.org/10.48550/arXiv.1606.00298 arXiv:1606.00298 [cs].

[25] Laura Dabbish, Colleen Stuart, Jason Tsay, and Jim Herbsleb. 2012. Social coding in GitHub: transparency and collaboration in an open software repository. In *Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work (CSCW '12)*. Association for Computing Machinery, New York, NY, USA, 1277–1286. https://doi.org/10.1145/2145204.2145396

[26] Marija Mitrović Dankulov, Roderick Melnik, and Bosiljka Tadić. 2015. The dynamics of meaningful social interactions and the emergence of collective knowledge. *Scientific Reports* 5, 1 (July 2015), 12197. https://doi.org/10.1038/srep12197 Number: 1 Publisher: Nature Publishing Group.

[27] Nicholas Davis, Ellen Yi-Luen Do, Pramod Gupta, and Shruti Gupta. 2011. Computing harmony with PerLogicArt: perceptual logic inspired collaborative art. In *Proceedings of the 8th ACM conference on Creativity and cognition (C&C '11)*. Association for Computing Machinery, New York, NY, USA, 185–194. https://doi.org/10.1145/2069618.2069650

[28] N. Davis, C. Hsiao, K. Y. Singh, B. Lin, and B. Magerko. 2017. Quantifying Collaboration with a Co-Creative Drawing Agent. *ACM Transactions on Interactive Intelligent Systems* 7, 4 (Dec. 2017), 19:1–19:25. https://doi.org/10.1145/3009981

[29] Nicholas Davis, Chih-Pin Hsiao, Kunwar Yashraj Singh, Brenda Lin, and Brian Magerko. 2017. Creative Sense-Making: Quantifying Interaction Dynamics in Co-Creation. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition (C&C '17)*. Association for Computing Machinery, New York, NY, USA, 356–366. https://doi.org/10.1145/3059454.3059478

[30] Brandon Doyle. 2022. TikTok Statistics - Everything You Need to Know [Aug 2022 Update]. https://wallaroomedia.com/blog/social-media/tiktok-statistics/

[31] Shufei Duan, Jinglan Zhang, Paul Roe, and Michael Towsey. 2014. A survey of tagging techniques for music, speech and environmental sound. *Artificial Intelligence Review* 42, 4 (Dec. 2014), 637–661. https://doi.org/10.1007/s10462-012-9362-y

[32] Anders Friberg, Erwin Schoonderwaldt, Anton Hedblad, Marco Fabiani, and Anders Elowsson. 2014. Using perceptually defined music features in music information retrieval. *arXiv:1403.7923 [cs]* (March 2014). https://doi.org/10.48550/arXiv.1403.7923

[33] Liane Gabora and Maryam Saberi. 2011. An agent-based model of the cognitive mechanisms underlying the origins of creative cultural evolution. In *Proceedings of the 8th ACM conference on Creativity and cognition (C&C '11)*. Association for Computing Machinery, New York, NY, USA, 299–306. https://doi.org/10.1145/2069618.2069667

[34] Andrea Gaggioli, Elvis Mazzoni, Martina Benvenuti, Carlo Galimberti, Antonio Bova, Eleonora Brivio, Pietro Cipresso, Giuseppe Riva, and Alice Chirico. 2020. Networked Flow in Creative Collaboration: A Mixed Method Study. *Creativity Research Journal* 32, 1 (Jan. 2020), 41–54. https://doi.org/10.1080/10400419.2020.1712160 Publisher: Routledge _eprint: https://doi.org/10.1080/10400419.2020.1712160.

[35] Vlad Petre Glăveanu. 2014. Distributed Creativity: What Is It? In *Distributed Creativity: Thinking Outside the Box of the Creative Individual*, Vlad Petre Glăveanu (Ed.). Springer International Publishing, Cham, 1–13. https://doi.org/10.1007/978-3-319-05434-6_1

[36] Dejan Grba. 2020. Alpha Version, Delta Signature: Cognitive Aspects of Artefactual Creativity. *Journal of Science and Technology of the Arts* 12, 3 (Dec. 2020), 63–83. https://doi.org/10.34632/jsta.2020.9491 Number: 3.

[37] Igor Zakhlebin and Emoke-Agnes Horvat. 2020. Diffusion of Scientific Articles across Online Platforms. In *Proceedings of the Fourteenth International AAAI Conference on Web and Social Media*, Vol. 14. 762–773. https://doi.org/10.1609/

icwsm.v14i1.7341
[38] ImprovisedHarmony. 2022. r/ImprovisedHarmony. https://www.reddit.com/r/ImprovisedHarmony/
[39] Nanna Inie and Peter Dalsgaard. 2020. How Interaction Designers Use Tools to Manage Ideas. *ACM Transactions on Computer-Human Interaction* 27, 2 (March 2020), 7:1–7:26. https://doi.org/10.1145/3365104
[40] Myra Interiano, Kamyar Kazemi, Lijia Wang, Jienian Yang, Zhaoxia Yu, and Natalia L. Komarova. 2018. Musical trends and predictability of success in contemporary songs in and out of the top charts. *Royal Society Open Science* 5, 5 (May 2018), 171274. https://doi.org/10.1098/rsos.171274 Publisher: Royal Society.
[41] Gerardo Iñiguez, János Török, Taha Yasseri, Kimmo Kaski, and János Kertész. 2014. Modeling social dynamics in a collaborative environment. *EPJ Data Science* 3, 1 (Dec. 2014), 1–20. https://doi.org/10.1140/epjds/s13688-014-0007-z Number: 1 Publisher: SpringerOpen.
[42] Ray Jackendoff and Fred Lerdahl. 2006. The capacity for music: What is it, and what's special about it? *Cognition* 100, 1 (May 2006), 33–72. https://doi.org/10.1016/j.cognition.2005.11.005
[43] Sergi Jordà. 1999. Faust Music On Line: An Approach to Real-Time Collective Composition on the Internet. *Leonardo Music Journal* 9 (Dec. 1999), 5–12. https://doi.org/10.1162/096112199750316730
[44] D. Bondy Valdovinos Kaye. 2022. Please Duet This: Collaborative Music Making in Lockdown on TikTok. *Networking Knowledge: Journal of the MeCCSA Postgraduate Network* 15, 1 (Feb. 2022). https://ojs.meccsa.org.uk/index.php/netknow/article/view/654 Number: 1.
[45] Christopher Kelty and Seth Erickson. 2018. Two modes of participation: A conceptual analysis of 102 cases of Internet and social media participation from 2005–2015. *The Information Society* 34, 2 (March 2018), 71–87. https://doi.org/10.1080/01972243.2017.1414092 Publisher: Routledge _eprint: https://doi.org/10.1080/01972243.2017.1414092.
[46] Ailbhe Kenny. 2014. 'Collaborative creativity' within a jazz ensemble as a musical and social practice. *Thinking Skills and Creativity* 13 (Sept. 2014), 1–8. https://doi.org/10.1016/j.tsc.2014.02.002
[47] Aniket Kittur, Jeffrey V. Nickerson, Michael Bernstein, Elizabeth Gerber, Aaron Shaw, John Zimmerman, Matt Lease, and John Horton. 2013. The future of crowd work. In *Proceedings of the 2013 conference on Computer supported cooperative work (CSCW '13)*. Association for Computing Machinery, New York, NY, USA, 1301–1318. https://doi.org/10.1145/2441776.2441923
[48] Daniel J. Levitin. 2006. *This is Your Brain on Music : The Science of a Human Obsession*. Dutton/Penguin Books.
[49] Maruša Levstek, Rubie Mai Barnby, Katherine L. Pocock, and Robin Banerjee. 2021. "It All Makes Us Feel Together": Young People's Experiences of Virtual Group Music-Making During the COVID-19 Pandemic. *Frontiers in Psychology* 12 (2021). https://www.frontiersin.org/articles/10.3389/fpsyg.2021.703892
[50] J C R Licklider and Robert W Taylor. 1986. The Computer as a Communication Device. *Science and Technology* 76 (1986), 21–31. http://memex.org/licklider.pdf
[51] Chen Ling, Jeremy Blackburn, Emiliano De Cristofaro, and Gianluca Stringhini. 2022. Slapping Cats, Bopping Heads, and Oreo Shakes: Understanding Indicators of Virality in TikTok Short Videos. In *14th ACM Web Science Conference 2022 (WebSci '22)*. Association for Computing Machinery, New York, NY, USA, 164–173. https://doi.org/10.1145/3501247.3531551
[52] S. Lippens, J.P. Martens, and T. De Mulder. 2004. A comparison of human and automatic musical genre classification. In *2004 IEEE International Conference on Acoustics, Speech, and Signal Processing*. Montreal, QC, Canada, iv–iv. https://doi.org/10.1109/ICASSP.2004.1326806 ISSN: 1520-6149.
[53] Philipp Lorenz-Spreen, Bjarke Mørch Mønsted, Philipp Hövel, and Sune Lehmann. 2019. Accelerating dynamics of collective attention. *Nature Communications* 10, 1 (April 2019), 1759. https://doi.org/10.1038/s41467-019-09311-w
[54] Kurt Luther, Kelly Caine, Kevin Ziegler, and Amy Bruckman. 2010. Why it works (when it works): success factors in online creative collaboration. In *Proceedings of the 16th ACM international conference on Supporting group work (GROUP '10)*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/1880071.1880073
[55] Thomas W. Malone, Robert Laubacher, and Chrysanthos Dellarocas. 2009. *Harnessing Crowds: Mapping the Genome of Collective Intelligence*. SSRN Scholarly Paper ID 1381502. Social Science Research Network, Rochester, NY. https://doi.org/10.2139/ssrn.1381502
[56] Matthias Mauch, Robert M. MacCallum, Mark Levy, and Armand M. Leroi. 2015. The evolution of popular music: USA 1960–2010. *Royal Society Open Science* 2, 5 (May 2015), 150081. https://doi.org/10.1098/rsos.150081 Publisher: Royal Society.
[57] Michael Mauskapf, Eric Quintane, Noah Askin, and Joeri M. Mol. 2017. Embeddedness and the Production of Novelty in Music: A Multi-Dimensional Perspective. *Academy of Management Proceedings* 2017, 1 (Aug. 2017), 16678. https://doi.org/10.5465/AMBPP.2017.2
[58] Thomas Barlow McHugh, Abir Saha, David Bar-El, Marcelo Worsley, and Anne Marie Piper. 2021. Towards Inclusive Streaming: Building Multimodal Music Experiences for the Deaf and Hard of Hearing. In *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems (CHI EA '21)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3411763.3451690
[59] Liang Men and Nick Bryan-Kinns. 2019. LeMo: Exploring Virtual Space for Collaborative Creativity. In *Proceedings of the 2019 on Creativity and Cognition*. ACM, San Diego CA USA, 71–82. https://doi.org/10.1145/3325480.3325495
[60] Gary M. Olson and Judith S. Olson. 2000. Distance Matters. *Human–Computer Interaction* 15, 2-3 (Sept. 2000), 139–178. https://doi.org/10.1207/S15327051HCI1523_4 Publisher: Taylor & Francis _eprint: https://doi.org/10.1207/S15327051HCI1523_4.
[61] Judith S. Olson, Dakuo Wang, Gary M. Olson, and Jingwen Zhang. 2017. How People Write Together Now: Beginning the Investigation with Advanced Undergraduates in a Project Course. *ACM Transactions on Computer-Human Interaction* 24, 1 (March 2017), 4:1–4:40. https://doi.org/10.1145/3038919
[62] Doheum Park, Juhan Nam, and Juyong Park. 2020. Novelty and influence of creative works, and quantifying patterns of advances based on probabilistic references networks. *EPJ Data Science* 9, 1 (Dec. 2020), 1–15. https://doi.org/10.1140/epjds/s13688-019-0214-8 Number: 1 Publisher: SpringerOpen.
[63] Laura Lucia Parolin and Carmen Pellegrinelli. 2020. Where does it come from? Collaborative emergence in creative work practices. *New Ideas in Psychology* 59 (Dec. 2020), 100800. https://doi.org/10.1016/j.newideapsych.2020.100800
[64] Jill E. Perry-Smith and Christina E. Shalley. 2003. The Social Side of Creativity: A Static and Dynamic Social Network Perspective. *The Academy of Management Review* 28, 1 (2003), 89–106. https://doi.org/10.2307/30040691 Publisher: Academy of Management.
[65] David Pinelle and Carl Gutwin. 2005. A groupware design framework for loosely coupled workgroups. In *Proceedings of the ninth conference on European Conference on Computer Supported Cooperative Work (ECSCW'05)*. Springer-Verlag, Berlin, Heidelberg, 65–82.
[66] Jennifer Preece and Ben Shneiderman. 2009. The Reader-to-Leader Framework: Motivating Technology-Mediated Social Participation. *AIS Transactions on Human-Computer Interaction* 1, 1 (March 2009), 13–32. https://doi.org/10.17705/1thci.00005
[67] Matthew Prockup, Andreas F Ehmann, Fabien Gouyon, Erik M Schmidt, Oscar Celma, and Youngmoo E Kim. 2015. Modeling Genre with the Music Genome Project: Comparing Human-Labeled Attributes and Audio Features.. In *Proceedings of the 16th ISMIR Conference*. Malaga, Spain, 7.
[68] Christoph Riedl and Anita Williams Woolley. 2017. Teams vs. Crowds: A Field Test of the Relative Contribution of Incentives, Member Ability, and Emergent Collaboration to Crowd-Based Problem Solving Performance. *Academy of Management Discoveries* 3, 4 (Dec. 2017), 382–403. https://doi.org/10.5465/amd.2015.0097
[69] Mark A. Runco and Garrett J. Jaeger. 2012. The Standard Definition of Creativity. *Creativity Research Journal* 24, 1 (Jan. 2012), 92–96. https://doi.org/10.1080/10400419.2012.650092 Publisher: Routledge _eprint: https://doi.org/10.1080/10400419.2012.650092.
[70] Babak Saleh, Kanako Abe, Ravneet Singh Arora, and Ahmed Elgammal. 2016. Toward automated discovery of artistic influence. *Multimedia Tools and Applications* 75, 7 (April 2016), 3565–3591. https://doi.org/10.1007/s11042-014-2193-x
[71] R. Sawyer and Stacy Lee Dezutter. 2009. Distributed creativity: How collective creations emerge from collaboration. *undefined* (2009). /paper/Distributed-creativity%3A-How-collective-creations-Sawyer-Dezutter/bd5e034dfae4054552bce7f0c43b60514370a750
[72] Jimmy Secretan, Nicholas Beato, David B. D Ambrosio, Adelein Rodriguez, Adam Campbell, and Kenneth O. Stanley. 2008. Picbreeder: evolving pictures collaboratively online. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '08)*. Association for Computing Machinery, New York, NY, USA, 1759–1768. https://doi.org/10.1145/1357054.1357328
[73] Joan Serrà, Alvaro Corral, Marián Boguñá, Martín Haro, and Josep Ll Arcos. 2012. Measuring the Evolution of Contemporary Western Popular Music. *Scientific Reports* 2, 1 (July 2012), 521. https://doi.org/10.1038/srep00521 Number: 1 Publisher: Nature Publishing Group.
[74] Pao Siangliulue, Kenneth C. Arnold, Krzysztof Z. Gajos, and Steven P. Dow. 2015. Toward Collaborative Ideation at Scale: Leveraging Ideas from Others to Generate More Creative and Diverse Ideas. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing (CSCW '15)*. Association for Computing Machinery, New York, NY, USA, 937–945. https://doi.org/10.1145/2675133.2675239
[75] Spotify. 2022. Soundtrap - Make music online. https://www.soundtrap.com/
[76] Shweta Suran, Vishwajeet Pattanaik, and Dirk Draheim. 2020. Frameworks for Collective Intelligence: A Systematic Literature Review. *Comput. Surveys* 53, 1 (Feb. 2020), 14:1–14:36. https://doi.org/10.1145/3368986
[77] Ana Tamayo. 2022. Sign Languages, Translation, and Interpreting: Creative Practices in Audiovisual Content. *Sign Language Studies* 22, 3 (2022), 484–519. https://doi.org/10.1353/sls.2022.0003 Publisher: Gallaudet University Press.
[78] Anthony Tang, Melanie Tory, Barry Po, Petra Neumann, and Sheelagh Carpendale. 2006. Collaborative coupling over tabletop displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Montréal Québec Canada, 1181–1190. https://doi.org/10.1145/1124772.1124950
[79] The BandLab Team. 2015. Forks – An exciting way to inspire and be inspired - From the Team. https://blog.bandlab.com/forks-an-exciting-way-to-inspire-

and-be-inspired/

[80] TikTok. 2022. TikTok - Make Your Day. https://www.tiktok.com/

[81] Chetan Walia. 2019. A Dynamic Definition of Creativity. *Creativity Research Journal* 31, 3 (July 2019), 237–247. https://doi.org/10.1080/10400419.2019.1641787 Publisher: Routledge _eprint: https://doi.org/10.1080/10400419.2019.1641787.

[82] Ian Wallace. 2014. The History of the Found Object in Art. http://www.artspace.com/magazine/art_101/the-history-of-the-found-object-in-art

[83] Dakuo Wang, Judith S. Olson, Jingwen Zhang, Trung Nguyen, and Gary M. Olson. 2015. DocuViz: Visualizing Collaborative Writing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, Seoul Republic of Korea, 1865–1874. https://doi.org/10.1145/2702123.2702517

[84] Dakuo Wang, Haodan Tan, and Tun Lu. 2017. Why Users Do Not Want to Write Together When They Are Writing Together: Users' Rationales for Today's Collaborative Writing Practices. *Proceedings of the ACM on Human-Computer Interaction* 1, CSCW (Dec. 2017), 107:1–107:18. https://doi.org/10.1145/3134742

[85] Christof Weiß, Matthias Mauch, Simon Dixon, and Meinard Müller. 2019. Investigating style evolution of Western classical music: A computational approach. *Musicae Scientiae* 23, 4 (Dec. 2019), 486–507. https://doi.org/10.1177/1029864918757595

[86] The Wellermen. 2022. The Wellermen on TikTok. https://www.tiktok.com/@thewellermen/video/7112521194460482822

[87] Minz Won, Sanghyuk Chun, Oriol Nieto, and Xavier Serrc. 2020. Data-Driven Harmonic Filters for Audio Representation Learning. In *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, Barcelona, Spain, 536–540. https://doi.org/10.1109/ICASSP40776.2020.9053669

[88] Minz Won, Andres Ferraro, Dmitry Bogdanov, and Xavier Serra. 2020. Evaluation of CNN-based Automatic Music Tagging Models. https://doi.org/10.48550/arXiv.2006.00751 arXiv:2006.00751 [cs, eess].

[89] Fang Wu and Bernardo A. Huberman. 2007. Novelty and collective attention. *Proceedings of the National Academy of Sciences* 104, 45 (Nov. 2007), 17599–17601. https://doi.org/10.1073/pnas.0704916104 Publisher: Proceedings of the National Academy of Sciences.

[90] Anna Xambó, Eva Hornecker, Paul Marshall, Sergi Jordà, Chris Dobbyn, and Robin Laney. 2013. Let's jam the reactable: Peer learning during musical improvisation with a tabletop tangible interface. *ACM Transactions on Computer-Human Interaction* 20, 6 (Dec. 2013), 1–34. https://doi.org/10.1145/2530541

[91] Lixiu Yu and Jeffrey V. Nickerson. 2011. Combining concept maps to catalyze creativity. In *Proceedings of the 8th ACM conference on Creativity and cognition (C&C '11)*. Association for Computing Machinery, New York, NY, USA, 13–20. https://doi.org/10.1145/2069618.2069623

[92] Lixiu Yu and Jeffrey V. Nickerson. 2011. Cooks or cobblers? crowd creativity through combination. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1393–1402. https://doi.org/10.1145/1978942.1979147