Does Background Music Matter in Data Videos? A Study of Music's Impact on Persuasion, Engagement, and Recall

Hessam Djavaherpour
Independent Researcher
Philadelphia, Pennsylvania, USA
djavaherpour@gmail.com

Leni Yang
Inria, CNRS, University of Bordeaux
Bordeaux, France
Inria, CNRS, University of Bordeaux
Bordeaux, France
leni.yang@inria.fr

Yvonne Jansen
CNRS, Inria, Univ. Bordeaux, LaBRI
Bordeaux, France
CNRS, Inria, Univ. Bordeaux, LaBRI
Bordeaux, France
yvonne.jansen@cnrs.fr

Pierre Dragicevic
Inria, CNRS, Univ. Bordeaux
Bordeaux, France
Inria, CNRS, Univ. Bordeaux
Bordeaux, France
pierre.dragice@gmail.com

Narges Mahyar
Computer Science
City St George's, University of London
London, United Kingdom
Computer Science
City St George's, University of London
London, United Kingdom
narges.mahyar@city.ac.uk

Mahmood Jasim
Louisiana State University
Baton Rouge, Louisiana, USA
Louisiana State University
Baton Rouge, Louisiana, USA
mjasim@lsu.edu

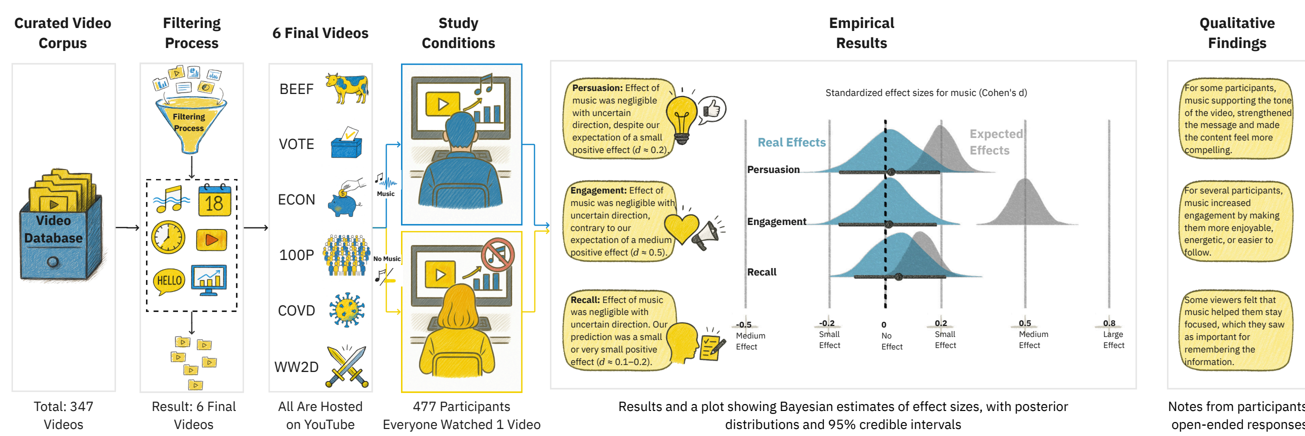

Figure 1: We assembled a corpus of 347 data videos, applied a multi-stage filtering pipeline to select 6 exemplars, and conducted a preregistered online experiment with 477 participants comparing versions with and without background music. Quantitatively, we did not observe measurable effects of music on persuasion, engagement, or recall. However, qualitative responses revealed nuanced perceptions; some participants reported that music enhanced enjoyment, focus, or emotional connection, indicating that music's impact is complex, subjective, and highly dependent on design and context.

Abstract

Data videos combine visualization, animation, narration, and often background music to tell stories with data. While music is widely believed to enhance emotion and persuasion, its impact in data videos remains unexplored. We conducted a preregistered between-subjects experiment comparing six widely-viewed data videos with or without background music. Using Bayesian modeling and thematic analysis, we did not observe consistent measurable effects of background music on persuasion, engagement, or information recall. Qualitative responses revealed a more nuanced picture: some participants described the music as distracting or mismatched, while others reported that it enhanced enjoyment, supported focus, or strengthened emotional resonance when well aligned with the video's tone. These findings suggest that the influence of background music in data videos is highly context-dependent, shaped

CHI '26, Barcelona, Spain

ACM ISBN 979-8-4007-2278-3/26/04
https://doi.org/10.1145/3772318.3791466

by genre, familiarity, and its alignment with visual–narrative structure. We discuss possible reasons for the limited measurable effects observed in real-world videos and outline opportunities for future work on purpose-designed, incidental, or adaptive music for data-driven storytelling.

CCS Concepts

• **Human-centered computing → Empirical studies in visualization**.

Keywords

Data videos, Music, Data Visualization, Persuasion, Recall, Engagement

ACM Reference Format:
Hessam Djavaherpour, Leni Yang, Yvonne Jansen, Pierre Dragicevic, Narges Mahyar, and Mahmood Jasim. 2026. Does Background Music Matter in Data Videos? A Study of Music's Impact on Persuasion, Engagement, and Recall. In *Proceedings of the 2026 CHI Conference on Human Factors in Computing Systems (CHI '26), April 13–17, 2026, Barcelona, Spain.* ACM, New York, NY, USA, 20 pages. https://doi.org/10.1145/3772318.3791466

1 Introduction

Data videos are *"custom motion graphics combining both visual and auditory stimuli to promote a data story"* [1]. They are thought to be an effective medium for educating the public, persuading, and raising awareness around global and social issues [1, 77]. For instance, there are videos with over tens of millions of views related to the COVID-19 pandemic [101], wars [32], and world development [8].

In order to inform, engage, or persuade viewers, data video designers have explored a plethora of techniques, including narrative structures [14], audio-visual components [83], and cinematic camera techniques [107]. Additionally, a substantial amount of data videos are accompanied by some form of background music [14]. Music has long been used to evoke emotions, inspire change, and deliver messages [58, Chap. XI], and it can have a profound impact on viewers' subjective interpretation of a scene [27]. Prior research has demonstrated mostly the impact of background music on consumer behavior [27], such as increasing time spent shopping [60] and customer loyalty [61] Music was also found to influence decision-making in other contexts, such as risk-taking in financial markets or compliance with public policy [62]. Yet, evidence of music's effect is mixed. In particular, some studies found no impact [22, 71] and others have even found detrimental effects of background music on behavioral and psychological outcomes [33, 38].

Although the effects of music have been widely studied in various domains, music has received little attention in information visualization. While it has been suggested that background music can make data videos more memorable, enjoyable, and immersive [81], these claims remain largely untested. To our knowledge, no prior studies have examined how background music shapes viewers' experiences of data videos. In this work, we address this gap by investigating three key ways in which music could impact the experience of data-video viewers:

- **RQ1**: How does the inclusion of background music affect the *persuasive power* of data videos?
- **RQ2**: How does the inclusion of background music affect *engagement* with data videos?
- **RQ3**: How does the inclusion of background music affect *information recall* with data videos?

To answer these questions, we conducted a preregistered, between-subjects online experiment with 477 participants. Participants were randomly assigned to watch one of six data videos in either their original form with background music or a version with the music removed. The six data videos were systematically selected from a large corpus of data videos based on corpora published in previous work [1, 105–109], partly based on their quality and popularity.

We employed Bayesian estimation [44] to evaluate the impact of background music on our three primary outcomes: persuasion, engagement, and information recall. We also analyzed participants' open-ended responses using thematic analysis [17]. Across all three outcomes, we did not observe a measurable overall influence of background music in this experimental context, as the differences between music and no-music conditions were close to zero. While music has been postulated to be an important design element in data videos [81] and has been shown to enhance emotional and persuasive impact in other contexts, our quantitative results indicate no consistent directional effect across persuasion, engagement, or information recall. However, qualitative responses revealed that some participants perceived benefits, such as increased enjoyment, improved focus, or stronger emotional resonance, suggesting that music's impact may be subjective and contingent on contextual or design-specific factors. Taken together, these findings suggest that, in our stimuli, measurable differences were not detectable; however, audience experience may still be shaped by music under certain conditions. This underscores the importance of considering other well-established factors, such as visual clarity, narrative structure, and informational content, while also acknowledging that auditory design may matter for some viewers and in certain contexts.

Our study provides the first empirical evidence on how background music influences viewers of data videos. It extends both prior work investigating the impact of music in various areas, and prior work in visualization looking at persuasion [77], viewer engagement [2, 107], and information recall [11, 12] in data visualization. Our contributions are threefold:

(1) *Empirical evidence.* A preregistered, between-subjects study in which we did not observe substantial or consistent measurable effects of background music on persuasion, engagement, or recall across six real-world data videos. Our results reflect what we were able to detect with these stimuli and this design, rather than a universal absence of influence.
(2) *Qualitative insights.* Thematic analysis of participant feedback revealed a more nuanced picture: while many participants described the background music in some videos as distracting, manipulative, or poorly matched, others reported that music enhanced enjoyment, supported focus, or strengthened emotional resonance when it fit the tone of the narrative. These responses also showed that some viewers believed music *would* improve certain videos, and others felt that the music they heard *did* enhance their experience.

(3) *Design and research implications.* A reframing of music as a context-dependent design element in data storytelling. Drawing on both quantitative and qualitative findings, we highlight how the influence of music may depend on genre, tone, familiarity, and integration with narrative structure. We discuss opportunities for aligning music with data-driven pacing, tailoring auditory choices to audience expectations, and exploring personalization. We also outline future research directions to examine subtle, delayed, or design-specific influences of music.

2 Related Work

In this section, we discuss prior works investigating the design elements of data video and their effects. Additionally, we highlight prior studies that evaluate the impacts of music in cases where music complements visual content.

2.1 Data Videos and their Components

In the last few decades, conveying data and disseminating knowledge through visual and audio narration has seen major successes in communicating scientific discovery [88], persuading informed decision-making [103], enabling sensemaking [111], and uncovering complex dynamics [37]. As a powerful yet complex multimedia form, data videos have attracted substantial research into their components and design strategies. Animations have been the dominant focus. Amini et al. [3] first summarized eight common types of chart animations in data videos, such as *creation and destruction* (animations that create or destroy a visualization) and *transitions* (animations that smooth the transition from one chart to another). Subsequent work introduced extensive animation designs for purposes such as conveying narrative intents [15, 86, 109], evoking emotions [46], and smoothing the transition between scenes [94, 96, 113]. In addition, much research provides strategies incorporating different components (e.g., narratives, visuals, sounds) to achieve high-level effects, such as realizing narrative structures [1, 14, 105, 109], building climaxes [104], and achieving cinematic styles [20, 106, 107]. In recent surveys by Shen et al. [81, 82] on empirical research and authoring tools of data videos, it is revealed that research in background music and sound effects constitutes only a small fraction of the literature. Past research recommended music as a tool for setting the tone and mood based on observed uses of music in real-world cases [95, 104–107]. However, none of them conducted experiments to evaluate the effects of music in data videos. To address this gap, we conducted comparative studies to provide empirical evidence to inform background-music design in data-video production.

2.2 Impact of Data Videos on Viewers' Experience

With the variety and complexity of data, there is a growing need to understand the real effects and impacts of data video designs. Prior studies have evaluated how the design of data videos can influence the viewer experience, knowledge acquisition, and decision-making. They reflected the common interests of this community. Our study design draws on their selections and designs of measurements.

In terms of viewer experience, enhancing viewer engagement in data videos has been a primary research focus. It has been found that increasing user engagement can benefit from pictographic and animations [2], visual foreshadowing [51], and immersive techniques [108]. To measure engagement, Amini et al. [2] developed a Likert scale by extracting an initial set of items that capture five key engagement attributes from questionnaires in related disciplines (e.g., game design, psychology, storytelling), and then refining these items through user studies. Our study adapts this highly used instrument for evaluating engagement.

In terms of knowledge acquisition, past studies have evaluated whether data videos enhance the understanding and retention of information [72, 76]. However, the results remain inconclusive regarding their effect on data comprehension or information retention. In these studies, both measures were assessed using the accuracy of participants' responses to quizzes about the video content. In our work, we focus specifically on information retention and therefore also employ a quiz-based assessment.

Finally, as for decision-making, past studies investigated the persuasive power of different data video designs in influencing people's attitudes. For instance, Choe et. al [16] found that incorporating four persuasive elements, including primary task, dialogue, system credibility, and social support, resulted in higher perceptions of persuasive potential of data videos by participants. Concannon et. al [19] found that participants had strong empathic responses after watching personalized data videos that explained young adults' financial pressure based on data from areas where viewers are. Sakamoto et al. [77] found that including affect-centered narratives in data videos increased participants' positive attitudes towards contact-tracing apps, especially when they hold a neutral to a positive pre-attitude towards the topic. Sallam et al. [78] emphasized the interaction between health-related data video content and viewers' affective states. They also found that viewers' personalities are interlinked with persuasion, showing that highly neurotic individuals are less likely to change their opinions than people with other personalities. Shifts in attitudes after watching data videos are commonly used to indicate their persuasive power. Our study followed this approach by asking participants about their attitudes toward each video's main message both before and after viewing. While prior research has explored data video designs across varied metrics, the impacts of music on data videos remain underexplored. Nonetheless, they informed the design of the evaluation metrics and questionnaires in our study.

2.3 Effects of Background Music on Viewers

Music is extensively studied as a complementary factor to the visual modality in supporting comprehension of visual representations [45]. Due to its pervasive nature, music has been demonstrated to influence human behavior, such as prosocial activities [41], social movements [91], decision-making tasks [70], and purchasing behaviors [4, 60]. For instance, strategic music in war-related documentaries might lead people to make humanitarian aid decisions or fair donations [80]. A selective music might reduce the fear of needles and promote blood donation campaigns (e.g., Keep the World Beating) [92]. A slower-tempo music has been shown to encourage shoppers to move more leisurely, leading to increased sales [60]. Additionally, music can complement narratives, such as dramatic sounds typically used in fiction films to form a sense of tension, create empathy with certain characters, or emphasize particular narratives [75]. For instance, in Jaws, John Williams' music defined the character of sharks as menacing and fearful [1].

However, the effects of background music are not always positive. A meta-analysis revealed that background music can have beneficial, detrimental, or no effect on adult listeners in different behavioral and psychological outcomes [38]. Although no existing work directly investigates the effects of background music in data videos, we reviewed studies on instructional and explanatory videos, as they similarly aim to present abstract or complex information clearly and engagingly. Contrary to the above literature's findings that highlight the influences of music on decision-making, its benefits for information retention, comprehension, and engagement have not been observed consistently. For example, in a study where students learn basic concepts with textual learning material, followed by a narrated instructional video, adding background music to the video does not affect students' information retention [71]. In another study about instructional videos, no significant differences were found in video watching time and frequencies and knowledge retention between students who watched with and without background music and sounds [22]. There are external factors that influence the effects of background music. A study found that students with prior knowledge performed better when the video had background music, while students without prior knowledge performed better without the background music [33].

Considering the mixed evidence on how background music influences viewers, it is important to assess its effects specifically within data videos. The major challenges for measuring the effectiveness of music in data videos are mainly from the ubiquity of highly tailored music in our everyday lives, increasing people's sonic awareness [73], and varied preferences of music and alignments on the visualization topic [110]. Prior to the present work, we conducted a preregistered study investigating the effects of background music on persuasion, engagement, and emotion in data videos using a single existing video presented without music, with default music, or with custom-composed music [29]. The study found increased perceived persuasiveness with default music, mixed effects for custom music, and no consistent increases in engagement; however, its focus on a single stimulus limited generalizability. In this work, we extend this line of inquiry by examining six widely viewed real-world data videos and the effects of pre-existing, production-authored music across contexts. We investigate three key dimensions related to the impact of music on data videos — persuasion, engagement, and information recall.

3 Methodology

We conducted a preregistered between-subjects online experiment on Prolific with 477 participants, using six data videos in their original form and the same six with background music removed. In this section, we describe the study design and methodological details of our study.

3.1 Data Video Selection

This section describes how we selected the data videos in our study.

3.1.1 Data Video Corpus Construction. We began by compiling a large corpus of existing videos from seven published corpora that focus specifically on the design, structure, or storytelling techniques of data videos. Next, focusing on alignment with our research objectives and content diversity, we applied a structured three-round filtering process to identify a final set of six videos. Each round introduced increasingly specific inclusion criteria, moving from general accessibility constraints to detailed qualitative assessments.

The seven published corpora used for our first round include: Wei et al.'s Hero's Journey framework for data storytelling [105], Xu et al.'s work on cinematic endings [106], Yang et al.'s exploration of 3D data videos [108], Xu et al.'s guidelines for compelling openings [107], Shi et al.'s design space for animated narratives [86], Yang et al.'s application of Freytag's pyramid to data stories [109], and Amini et al.'s work on narrative visualizations through a cinematographic lens [1]. We merged the video lists from these corpora using a custom Python script, which also removed duplicate entries based on video title and URL. The resulting collection comprised 347 unique videos in English and Mandarin. To ensure accessibility and usability, we manually reviewed all video links to confirm that they were functional and not behind a paywall. For videos with broken or outdated links, we attempted to locate alternate sources. This process resulted in a curated set of 320 videos with verified working links to target videos. The full list of videos and additional information about the filtering process is available in our OSF repository.[2]

This initial corpus of 320 videos served as the foundation for a multi-stage refinement process. Because our study required a small, well-defined set of videos that varied in topic, style, and visual density, but met both technical constraints and conceptual aims, a single round of filtering was insufficient. We therefore applied an iterative, three-round approach: 1) Apply broad, objective criteria for accessibility and baseline quality, 2) Assess visual and structural features of the videos, including the presence and type of visualizations and whether they employed sequential comparison formats, and 3) Apply qualitative coding for persuasive message presence and clarity. This stepwise narrowing allowed us to preserve thematic and stylistic diversity.

3.1.2 First Round of Filtering: Accessibility and Baseline Quality. The first round of filtering applied a set of baseline criteria to ensure accessibility, relevance, and quality. These criteria were determined iteratively: we initially experimented with restrictive thresholds (e.g., excluding voice narration, requiring ≥1M views, etc.), but relaxed them when they proved too limiting, resulting in fewer video candidates. The final criteria we used are as follows:

(1) Background music. Each video must include background music, as this was central to our experiment and created the basis for the music (♫) / no-music (🔇) comparison.
(2) Language. Videos must be in English, either through narration or on-screen text, to ensure consistency of comprehension across participants.
(3) Platform. Videos must be publicly available on YouTube with a working link. YouTube was chosen to ensure accessibility, stable hosting, and comparability of view counts across videos.

[1] https://www.soundstripe.com/blogs/the-psychology-of-music-for-film

[2] https://osf.io/htq82

(4) Publication date. Eligible videos must have been published in 2015 or later, so that the sample would reflect contemporary design and production practices in data videos.
(5) Duration. Videos must be between 30 and 300 seconds. Short videos (< 30 seconds) risk lacking sufficient narrative or persuasive development, while long videos (over 5 minutes) pose risks of participant fatigue and reduced comparability across conditions [112].
(6) Popularity. To ensure that the selected videos had reached a broad audience, we selected videos that had a minimum of 100,000 views. This threshold served as a proxy for visibility and relevance, reducing the likelihood of including obscure or low-quality videos that had little audience engagement.

Applying these criteria yielded 85 videos (∼ 25% of the original corpus), which provided a diverse pool of candidates for subsequent coding and experimental design.

3.1.3 Second Round of Filtering. In the second round of filtering, we introduced more nuanced criteria, including qualitative coding of 85 videos that passed the first round. While the first round relied on metadata and quantitative attributes, the second-round filters involved direct visual content analysis. The following four main dimensions guided this analysis:

(1) Visualization focus. We measured the proportion of a video's runtime during which a data visualization was shown (0%–100%). We defined *visualization* as a visual encoding of data; visuals that did not represent actual data (e.g., simulated or decorative graphics) were excluded. This criterion allowed us to emphasize actual data, rather than relying primarily on narration or illustrative imagery. To calculate visualization focus, we developed a custom pipeline that extracted frames from each video at 3-second intervals. The team members labeled each frame as containing a data visualization or not. The percentage of frames labeled as visualizations was used as the operational measure of visualization focus.
(2) Visualization type. We coded whether the video featured primarily 2D visualizations, 3D visualizations, or a combination of both. This criterion was included because 2D and 3D encodings may differ in accessibility, interpretability, and cognitive load. Retaining this distinction allowed us to preserve a mix of stylistic and structural approaches across the candidate pool.
(3) Sequential comparison. We coded whether a video presented different quantities in sequence, usually from smallest to largest. We defined *sequential comparison video* as videos that emphasize relative magnitudes by presenting values one after another, often using a horizontally scrolling 3D bar chart format. Prior work has shown that sequential comparison is a dominant but stylistically specific genre of data video [108], highlighting the importance of examining its presence in our candidate pool.
(4) Sequential size comparison. Finally, we identified whether a video used sequential comparison specifically to depict physical sizes. We defined *sequential size comparison video* as a subtype of sequential comparison focused on the physical dimensions of objects or phenomena. We separated this category because physical size comparisons may evoke different perceptual and rhetorical responses than other kinds of sequential comparisons, making it useful to distinguish in coding.

After applying these criteria, we tested the impact of excluding videos with less than 50% visualization focus. After excluding videos by applying these filters, the set of eligible videos was reduced to 17, which we designated as *Tentative Shortlist #1*. This collection served as an initial pool for final video selection. The full list is available as both a spreadsheet and a YouTube playlist.[3]

3.1.4 Third and Final Round of Filtering. In the third and final round of filtering, our focus shifted from structural and visual characteristics to the presence and clarity of a persuasive message. As part of this shift toward evaluating the videos' overall communicative design, the members of the research team also qualitatively examined the role and character of each video's background music. Specifically, we discussed whether the background music appeared to support, distract from, or be neutral with respect to the persuasive goals of the content. While these observations were not used as formal scoring criteria, they informed our broader discussions of each video's multimodal design. These qualitative observations were intended to contextualize the experimental stimuli and to support interpretation of the study results. They were not used to define experimental conditions, nor did they function as independent variables or formal classification schemes. With these contextual considerations in mind, each member of our research team independently applied the following four qualitative coding criteria (using a 1–5 scale where applicable) to the 17 candidate videos in *Tentative Shortlist #1*:
(1) Message presence. The extent to which a video conveyed an explicit or implicit persuasive message.
(2) Message vs. mere information. Whether the content reflected persuasive intent rather than neutral fact delivery.
(3) Connection to action. The degree to which the message conveyed in the video was linked to potential individual or societal actions.
(4) Message clarity. The consistency and salience of the persuasive message portrayed across the video.

In parallel, all members of the research team provided open-ended descriptive notes on the background music of each candidate video, including observations such as: "carefully designed selection of tracks," "generic or template-like," "engaging and well matched," "monotone or contradictory in tone," "distracting crescendos," "effective soundscape but overshadowed by narration," or "music not playing an important role." These comments offered additional context about how the soundscape functioned within each video's overall design. It is important to distinguish between the effects of background music as reported by study participants and the inferred design intentions derived from the research team's qualitative assessment of how the music appeared to function. The former captures audience response measured through experimental data, whereas the latter reflects the researcher's interpretation of communicative design. (Refer to our OSF epository[4], under Video Selection, for more details).

Following this independent coding, we met as a team to review the ratings, discuss points of divergence, and collectively decide

[3] YouTube playlist link
[4] https://osf.io/htq82

which videos to retain or eliminate. Our decisions were guided by both the coded scores and broader design considerations. Specifically, we aimed to (a) balance thematic diversity across societal issues, (b) represent a range of visual and narrative styles, and (c) avoid including multiple videos from the same producer. The latter was not always possible: videos made by Vox Media [102] appeared frequently in the candidate pool, and we ultimately retained two of its videos because they scored highly across our criteria.

This rigorous filtration process resulted in the final six selected videos spanning climate change, democratic processes, economic forecasting, global demographics, pandemic data, and historical reflection (see Figure 2 for screenshots of these videos):

- BEEF: **Why beef is the worst food for the climate**: A data-rich and visually engaging explanation of beef production's impact on climate, combining charts, physicalization, and a clear call to action. The soundtrack consists of several short electronic–orchestral cues varying in intensity to accompany topic transitions. The titles for the credited music include "Out of Curiosity" (Pierre Henry Jacques Adenot, SACEM), "Cautious Optimism" (Matthew Sanchez and Sam Taylor, PRS), "Mischievous Pizzicato" (Paul Brian Hart and David John Arnold, PRS), and "Illustrious Prince" (Laurent Dury, SACEM).
- VOTE: **Why America needs automatic voter registration**: A persuasive breakdown of benefits and mechanics of automatic voter registration, framed with international comparisons to broaden relevance. The soundtrack consists of several energetic electronic and synth-driven cues, which maintain a steady rhythmic pulse and vary moderately in intensity. The titles for the credited music include "Cheeky Peek (B)" and "Wonky Step" (Rob Persaud, Matt Black, and Adrian Quick, ASCAP), "Illumination" (Benjamin MacDougall, ASCAP), and "Behind the Fence" (Michael Price, PRS).
- ECON: **The World's Largest 10 Economies in 2030**: A data-driven projection of the shifting global economic landscape, ranking emerging and established markets by their forecast GDP. The video uses a modern corporate–cinematic instrumental soundtrack that maintains a steady, driving rhythm throughout and varies modestly in intensity when new data segments are introduced. Specific music credits were not provided in the video.
- 100P: **If The World Were 100 People**: A vivid reimagining of global population statistics through unit-based visualizations, distilling large-scale demographics into a relatable 100-person world. The soundtrack consists of light, upbeat instrumental cues, using bell-like melodic instruments (such as xylophone or glockenspiel) supported by soft percussion and a steady rhythmic pulse, complementing the video's rapid-paced transitions. Although the video did not provide music credits, audio identification tools (e.g., Shazam[5]) indicate that the track is likely "A Flower" by Red Red Groovy.
- COVD: **Corona race meme (COVID-19 growth by country)**: A bar chart race tracking the rapid spread of COVID-19, juxtaposing somber data with intentionally incongruous meme music. The video uses the Eurobeat track "Gas Gas Gas" by Manuel, a song widely associated with online "race" memes and the *Initial D* franchise. The music features a very fast tempo, prominent electronic drums, and energetic synthesizer melodies. In this context, the background music is used primarily for comedic contrast, creating an exaggerated sense of acceleration as the bars depicting COVID-19 diagnosis cases update rapidly.
- WW2D: **Number of deaths in WW2 per country**: A sobering 3D visualization of wartime casualties, using physicalized coffin stacks to convey scale and human cost. The video uses the instrumental track "Space Coast" by Topher Mohr and Alex Elena. The soundtrack features cinematic-style orchestral elements, including strings and percussion, combined with a steady rhythmic progression.

After finalizing these six selected videos, we conducted an additional round of design-space coding [81] to characterize their multimodal properties. As part of this process, we rated the relative presence of background music in each video through assigning an ordinal tag of "Medium" or "High" based on the perceived loudness, continuity, and prominence of the soundtrack relative to the narration and/or data visuals. These tags are shown as the "Music Presence" column in Table 1.

Together, these six videos provided thematic breadth, stylistic variety, and consistent persuasive intent. They were published between 2016 and 2020, collected between 370,000 and 20 million YouTube views in 2025, and ranged from 45 seconds to 5 minutes in duration. Three videos included voice narration, while three did not. Their visualization focus ranged from 57% to 100%, and message presence scores ranged from 3.3 to 5.0. All details are in Table 1.

3.2 Study Design

We conducted a between-subjects online experiment to investigate how background music affects people's experience of data videos. Each participant was invited to watch a video with or without background music, after which they answered questions measuring persuasion, information recall, and engagement with the data video. Our study was approved by the institutional review board (IRB) of Louisiana State University.

3.2.1 Stimuli and Conditions. We created twelve experimental stimuli: six data videos in their original form with background music (see subsection 3.1 on how they were selected) and six matched versions with the music removed. For better experimental control, the "no music" versions retained all their original content (visuals and voiceover when present) except the music track, which was removed using the AI-based tool Audiostrip [5]. Videos were embedded in the Qualtrics survey and presented using Qualtrics's random assignment feature to balance conditions across participants.

3.2.2 Study Design. Participants were randomly assigned to watch one of twelve stimuli. We chose this between-subjects design *i)* to prevent fatigue, as watching six videos and completing questionnaires for each would have been time- and energy-intensive; and *ii)* to avoid carryover effects and reduce response bias due to demand characteristics [66]. While within-subject designs can increase statistical power, this benefit depends on correlated responses within participants [21, Chap. 8], which is unlikely in our case, given the large effect of video topic relative to music. Thus, although we initially planned for each participant to view two videos on different topics, one with music and one without, we abandoned this

[5]Shazam.com

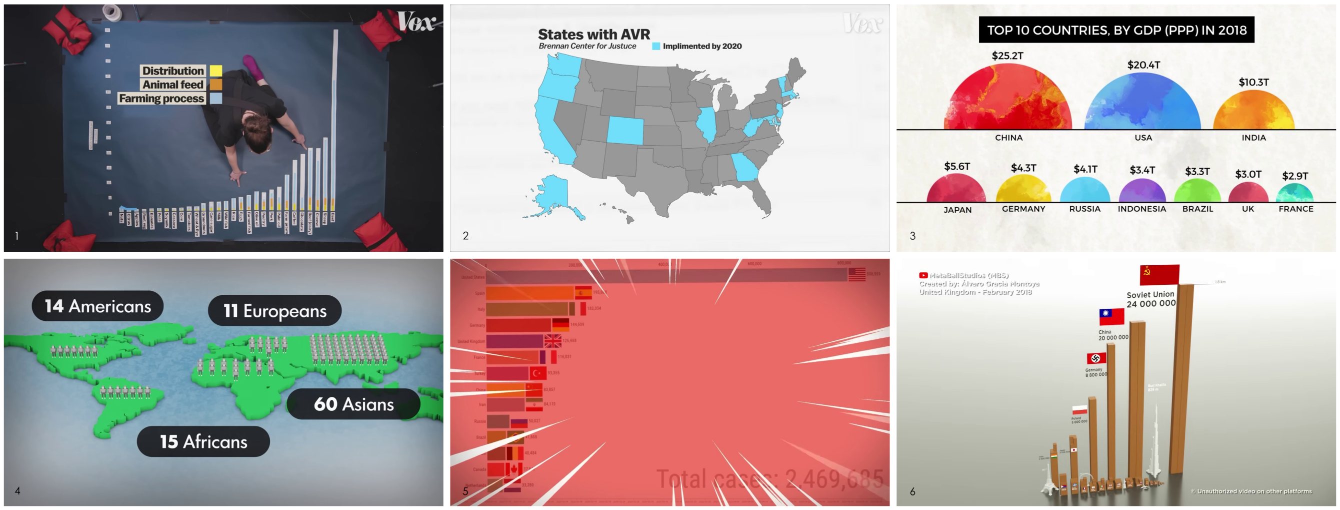

Figure 2: Data videos included in our study: (1) *Why beef is the worst food for the climate* (Vox Media, 2020) [100] ; (2) *Why America needs automatic voter registration* (Vox Media, 2018) [99] ; (3) *The World's Largest 10 Economies in 2030* (Visual Capitalist, 2019) [98] ; (4) *If The World Were 100 People* (GOOD Magazine, 2016) [30] ; (5) *Corona race meme (COVID-19 growth by country)* (Kori, 2020) [43] ; and (6) *Number of deaths in WW2 per country* (MetaBallStudios, 2018) [59] .

Table 1: Summary of videos used in the study

ID	Video title	Creator	Datasets	Year	Views	Length (min)	Narration	Music presence	Vis. focus	Message presence
BEEF	Why beef is the worst food for the climate [100]	Vox Media	[105–107]	2020	1.1 M[a]	4.5[b]	Yes	Medium	73%	5.0
VOTE	Why America needs automatic voter registration [99]	Vox Media	[106, 107, 109]	2018	370 K	4	Yes	Medium	57%	5.0
ECON	The World's Largest 10 Economies in 2030 [98]	Visual Capitalist	[86]	2019	480 K	3	Yes	Medium	68%	3.8
100P	If The World Were 100 People [30]	GOOD Magazine	[86]	2016	3.3 M	2.5	No	High	92%	3.5
COVD	[V3] Corona race meme (covid-19 growth by country) [43]	Kori	[109]	2020	20 M	1	No[c]	High	100%	3.3
WW2D	Number of deaths in the WW2 per country [59]	MetaBall-Studios	[86, 108]	2018	8.6 M	2.5	No	High	85%	4.3

[a] View counts are rounded to the nearest hundred thousand [23] (K = Thousand, M = Million).
[b] The time is rounded to the nearest half minute.
[c] There is a brief audio narration in some frames of the video.

approach due to concerns about carryover effects, response bias, and minimal expected benefit.

3.2.3 Procedure. The experiment session began with participants consenting to participate and answering a set of screening questions. These covered language fluency, auditory and visual impairments, and participants' ability to watch and listen to videos (including a short audio test). We also collected demographic information and viewing preferences (e.g., typical video consumption and music genres). The full question set is available as supplementary materials. Following the screening, participants answered a set of pre-study questions, watched a randomly-assigned data video, and answered a set of post-study questions. Participants completed the study in 13 minutes on average (SD = 8, range = 3–51). They were compensated $3.50 USD, corresponding to a rate of $14/hour on average.

We implemented multiple measures to identify bots and ensure data quality:

1. *Screening:* Participants were not included if they reported not being fluent in English, having significant visual or auditory impairments, or lacking the necessary technology to watch and listen to videos.
2. *Audio check:* Participants completed a brief audio test in which they typed a number spoken in an audio clip. Those who entered incorrect responses were not included.

3. *Attention check:* An attention-check question was used to confirm that participants were following instructions. Those who failed (about 14% of all submissions) were not included.
4. *CAPTCHA:* A bot detection question was included at the end of the survey. Responses failing this check were discarded.
5. *Duplicate or incomplete responses:* Any incomplete or multiple submissions from the same participant were removed.
6. *Completion code:* Participants who submitted invalid Prolific completion codes were not included.

Prolific replaced participants whose responses failed to meet criteria 1–5. This process yielded 480 responses, in line with our planned sample size (further discussed in subsection 3.4). Among these, three participants submitted invalid Prolific completion codes, resulting in 477 valid responses for analysis.

3.2.4 Measurements. Participants answered questions related to persuasion, engagement, and recall (our three primary outcomes), as well as music perception, along with demographic and screening items. All the following measurements, except information recall, use 7-point Likert-type scales. For instance, for questions that start with "To what extent would you likely...," participants selected from 1—extremely unlikely to 7—extremely likely. For questions that start with "To what extent do you agree...," the participants selected from 1—extremely disagree to 7—extremely agree. To measure information recall, participants were asked to answer multiple-choice questions about basic facts from the video.

(1) Persuasion. We assessed persuasion by measuring attitudinal shift. It aligns with previous research [69] that suggests measuring the changes in participants' attitudes before and after treatment as an effect of a persuasive message to measure its persuasive power. Following the practice in [69], participants' attitudes are measured by single-item scale questions that plainly ask them to what extent they agree with the main messages from the videos. Given that the videos are not long and have a few key takeaways, we decided to tailor three questions for each video. Before designing questions, all members of the research team wrote and discussed what they believed were the key takeaways of the videos to try to avoid a wrong or incomplete understanding. Before viewing the data video, participants answered three questions gauging their initial level of agreement with the video's main messages. For example, participants who were assigned the BEEF video were asked *"To what extent do you agree that changing eating habits can help reduce greenhouse gas emissions?"*. We will refer to these questions as **pre-agreement** questions. After having watched the video, participants answered the same three questions (**post-agreement** questions). Persuasion was defined as the average difference between post- and pre-agreement, as will be further explained in section 4.

(2) Perceived video persuasiveness. As a complementary metric, we measured the video's perceived persuasiveness: after having watched the video, participants rated how persuasive and believable they found it. Items include: *"To what extent do you agree that the information in the video is reliable?"*; *"To what extent do you agree that the video has the potential to influence its viewers?"*; *"To what extent do you agree that the video gives viewers a new behavioral guideline?"*. The questions were the same for all videos.

(3) Behavioral intention. After having watched the video, participants indicated their intention to take action based on the video's message. The behavioral intention questions were tailored to each video. For example, for BEEF, participants were asked: *"To what extent would you likely take action to reduce greenhouse gas emissions?"* and for ECON, the question was: *"If you were planning a personal investment, how likely would you be to invest in Asian markets (e.g., by purchasing stocks or funds based in Asia)?"*

(4) Engagement. To assess viewer engagement, we adapted an 11-item scale developed by Amini et al. [2] for measuring viewer engagement for data videos. It covers five engagement attributes, including cognitive involvement, affective involvement, enjoyment, focused attention, and aesthetics. The original scale was developed by deriving an initial set of 53 items from surveys in psychology, HCI, and game design, followed by refining the list through a user experiment. Minor modifications were made to better align with our study goals, including rewording or removing entertainment-specific items and adding a question on interest in learning more about the topic. For example, among the four original items for measuring enjoyment, two focused on entertainment were removed (i.e., "This video was fun to watch" and "The video was entertaining") as entertainment is not the purpose of all videos. Moreover, some videos are about serious and heavy topics (e.g., WW2D), and the questions would be inappropriate. More details are provided in the supplementary materials. Example questions include *"I would like to learn more about the topic"* and *"The video triggered my emotions"*. The questions were the same for all videos.

(5) Information recall. At the end of the study, each participant answered multiple-choice questions assessing the extent to which they could recall basic facts from the video. Deciding on multiple-choice questions for measuring recall follows is inspired by previous work in visualizations [42]. We designed recall questions that reflect salient and important information in the videos. When the video had voiced narration, we segmented the video script into parts, such as the background, major arguments, or findings. From every part, we identified key facts and conclusions to inspire the recall question design. When the video did not have voiced narration, we selected salient data points (e.g., the largest value) as well as information emphasized through visual cues (e.g., prolonged camera focus or zoom-ins). We distributed questions across the video timeline as evenly as possible. An example includes: *"According to the video, what is the primary greenhouse gas emitted during the digestion process of cows and sheep? (a) Methane (b) Nitrous oxide (c) Ammonia (d) None of the above."* Compared to the attitudinal questions from before (i.e., pre- and post-agreement questions), these questions were more factual. There were 3 to 6 questions, depending on the video. Responses were scored for full or partial credit. All questions were piloted to ensure appropriate difficulty.

(6) Music perception. After watching the video, participants answered one music-related question based on their assigned condition. Participants who watched a data video with music were asked: *"To what extent do you agree that music enhances this video?"* Others who watched a data video without music were asked *"To what extent do you agree that adding music would enhance this video?"*

3.3 Research Questions

Our primary research questions were:

- **RQ1:** How does the inclusion of background music affect the **persuasive power** of data videos?
- **RQ2:** How does the inclusion of background music affect **engagement** with data videos?
- **RQ3:** How does the inclusion of background music affect **information recall** with data videos?

We sought to answer these research questions by estimating the overall effect of music on the relevant metric across all six videos. As secondary research questions, we also examined the effect of music within individual videos.

Expectations. Based on prior insights from visualization researchers [81] and our own intuitions, we expected data videos with background music to be more impactful than those without. At the same time, we anticipated the effect of music to be tenuous, since music is only one of many elements in a data video and is typically not the most salient. This aligns with findings from other domains, where some studies were unable to detect the effect of background music [38]. More specifically, we anticipated (and preregistered) a *small effect*[6] *of music on persuasion* overall (RQ1), with potentially medium effects for certain videos. We anticipated a *medium effect of music on engagement* (RQ2) and a *small to very small effect of music on recall* (RQ3). Our reasoning was that while background music can be reasonably expected to affect reported engagement, its causal link to persuasion seems less direct, and its causal link to information recall even less so. Furthermore, persuasion and information recall are likely dominated by other factors, such as the video's content and the participant's prior beliefs and knowledge, making any effect of music possibly hard to detect.

3.4 Participants

3.4.1 Sample Size Rationale. After coding our planned Bayesian analyses (further detailed in subsection 4.2), we conducted frequentist power analyses by simulating multiple datasets and calculating the proportion where the 95% credible interval excluded the null.[7] For our primary analyses—examining the overall effect of music across all videos—we estimated that $N = 480$ would provide approximately 0.75 power to detect a small standardized effect ($d \approx 0.2$) and >0.99 power to detect a medium effect ($d \approx 0.5$), under reasonable simulation assumptions (see code in supplementary material for details). For video-specific effects, power was roughly 0.15 for $d \approx 0.2$ and 0.6 for $d \approx 0.5$. Overall, given our effect size expectations (subsection 3.3), $N = 480$ should allow us to estimate the direction of most primary effects while offering at least limited sensitivity to subtler effects.

3.4.2 Recruitment and Inclusion/Exclusion Criteria. Following our sample plan, as detailed in our preregistration, we aimed to recruit 480 participants through four badges. We configured Prolific to propose our study only to participants corresponding to the following criteria:

- having a 95% or higher completion rate,
- having at least 10 prior studies approved, but no more than 10,000,
- being fluent in English (self-declared).

Additionally, we added our own screening procedure as described in subsubsection 3.2.3 to ensure that all participants were able to hear the music and or narration of the tested videos.

We included one attention check question in the form of an instructional manipulation check [65]: *"Please select the type of topic of the video that you just watched. When asked for the type of topic, you must select Video Game. What is the type of topic of the video?"* Initially, we provided the following four answer options: *(1) Politics, (2) Health (3), Climate Change, (4) Video Game.* However, while running the first batch of participants, we noticed that an unusually high number failed the attention check (47 out of 120, i.e., about 28%), and this specifically for videos whose topic was in the list of answer options. We therefore changed the options so that none of them matched the topic of any of the videos. The final total number of participants who failed the attention check is 79, which results in a failure rate of about 14% overall, and 8% when excluding the first batch from the calculation.

3.4.3 Sample Characteristics. Prolific provides basic demographic information for the participants recruited for a study. We summarize in Figure 3 information on age ranges, country of residence, sex (Prolific only provides data on binary sex not on self-reported gender), and the number of previously accepted contributions on the Prolific platform.

4 Quantitative Data Analysis

In this section, we outline the methodology we used to quantitatively evaluate the impact of background music on our outcomes of interest, and we present our results. We used Bayesian estimation as our statistical inference framework [44]. This approach bases conclusions on the interpretation of posterior distributions and credible intervals, viewing statistical evidence as a continuum rather than a binary outcome [9, 44]. All analyses were preregistered unless stated otherwise, and extensive supplementary materials are available online.[8]

4.1 Data Transformations

We aggregated and normalized the measurements presented in subsubsection 3.2.4 as follows:

- **Persuasion score** $\in [-1, 1]$: we computed pre- (resp. post-) agreement scores by summing up responses to the three pre- (resp. post-) agreement questions and normalizing between 0 and 1, with 0 meaning no agreement with the video's messages, and 1 meaning full agreement. The persuasion score is the difference between post- and pre- agreement scores. Therefore, 0 means no persuasion, i.e., no shift of the participant's agreement with the

[6] Following the convention proposed by Cohen [18, p.24], we use the term *medium effect* to refer to $d \approx 0.5$, *small effect* to refer to $d \approx 0.2$, and *very small effect* to refer to $d \approx 0.1$, with d being Cohen's standardized effect size d.

[7] We also tested frequentist counterparts of our Bayesian models (i.e., linear models and linear mixed models) and found highly consistent results. To save computation time, we used these frequentist models as approximations for our power estimations.

[8] Link to the preregistration on OSF: https://osf.io/fzde7/; link to the OSF project with supplementary material: https://osf.io/htq82

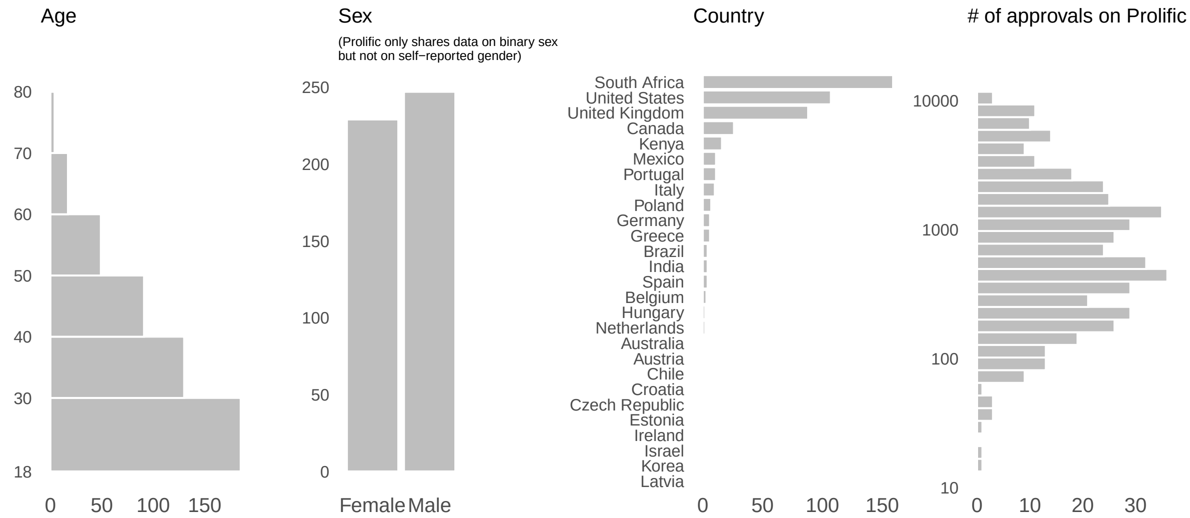

Figure 3: Basic demographics describing our participant sample.

video's messages, while 1 means maximum persuasion, and -1 means maximum negative persuasion.

- **Perceived video persuasiveness** $\in [0, 1]$: the normalized sum of the responses to the three questions capturing the extent to which participants found the video persuasive.
- **Behavioral intention** $\in [0, 1]$: normalized response to the question capturing the participant's intent to take action based on the video.
- **Engagement score** $\in [0, 1]$: composite of 11 items from an engagement questionnaire, with reverse-coding and normalization applied.
- **Recall score** $\in [0, 1]$: total score obtained to the video-specific questionnaire testing participant's recall of factual information in the video, with 0 being the worst possible score and 1 being the best possible score (see supplementary material for exact questions and scoring schemes).

4.2 Bayesian Models

We used Bayesian models to estimate the effect of music on our three primary outcomes of interest: persuasion score (addressing RQ1), engagement score (addressing RQ2), and recall score (addressing RQ3).

For **persuasion** (RQ1), we fit a Bayesian linear mixed-effects model [56] predicting persuasion score from the binary music condition, with a random intercept for video, and a Gaussian likelihood:

$$\texttt{persuasion\_score} \sim \text{Normal}(\mu, \sigma), \quad \mu = \texttt{music} + (1 \mid \texttt{video}) \tag{1}$$

This model accounts for some videos being naturally more persuasive, but assumes that the effect of music is the same across all videos. Note that since each participant only saw one video, we could not include participant effects explicitly in the model. This model was chosen over more complex models (e.g., the interaction model $\texttt{persuasion} \sim \texttt{music} \times \texttt{video}$ and the random-intercept model $\texttt{persuasion} \sim \texttt{music} + (1 + \texttt{music} \mid \texttt{video})$ based on power analyses and informal testing on simulated data, suggesting a better predictive power for the chosen model.

Although observations were theoretically bounded between -1 and 1, the Gaussian likelihood was chosen because the persuasion measure is a difference between pre- and post-agreement scores with a meaningful 0, and because the observations were expected to lie close to 0 and away from their theoretical bounds (confirmed by pilot data and then experimental data). We use weakly informative priors that do not favor any particular direction of effect. Full details about the analyses (including R code for the Bayesian analyses with comments explaining and justifying the choices made, R code for generating the simulated datasets, R code for running the power analyses, and a table with results from power analyses) were preregistered and available in the OSF storage.

For **engagement** (RQ2), the model was similar to the persuasion model (fixed effect of music and a random intercept for video), except that we used a beta likelihood with a logit link function [56] to account for the response variable being bounded between 0 and 1 and, in contrast to the persuasion measure, expected to produce values *near* the bounds:

$$\texttt{engagement\_score} \sim \text{Beta}(\mu, \phi), \quad \text{logit}(\mu) = \texttt{music} + (1 \mid \texttt{video}) \tag{2}$$

The model for **recall** (RQ3) was the same:

$$\texttt{recall\_score} \sim \text{Beta}(\mu, \phi), \quad \text{logit}(\mu) = \texttt{music} + (1 \mid \texttt{video}) \tag{3}$$

These three models were used to estimate the overall effect of music across all six videos. To examine how the effect of music varied across individual videos, we also fit separate Bayesian linear models for each video, predicting persuasion with only a fixed effect of music:

$$\texttt{persuasion\_score}_{video} \sim \text{Normal}(\mu, \sigma), \quad \mu = \texttt{music} \tag{4}$$

$$\texttt{engagement\_score}_{video} \sim \text{Beta}(\mu, \phi), \quad \text{logit}(\mu) = \texttt{music} \tag{5}$$

$$\texttt{recall\_score}_{video} \sim \text{Beta}(\mu, \phi), \quad \text{logit}(\mu) = \texttt{music} \tag{6}$$

Again, these models were chosen over alternative approaches (e.g., extracting video-specific effects from a general interaction model or random-intercept model) based on power analyses and informal testing on simulated data. All model choices were preregistered.

Plots. For all models, we extracted the posterior distribution of the estimated increase in the outcome due to music, on the original normalized scale (from -1 to 1, or 0 to 1). In all plots that follow, the blue density curve shows the posterior distribution, the dot shows its median, and the error bar indicates its 95% credible interval. Together, these elements convey the uncertainty around the effect: the point estimate is our "best bet" for the true increase, the interval very likely (with 95% probability) contains the true increase, while the density curve provides the full range of probabilities.

4.3 Results

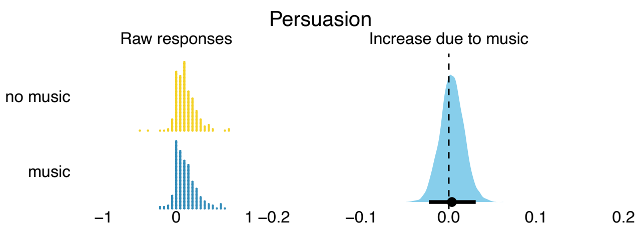

Figure 4: Effect of music on persuasion across all videos. *Left:* Raw responses for persuasion without music and with music; *Right:* Bayesian estimation of the effect of music on persuasion on the same -1 to 1 scale. The error bar is a 95% credible interval.

4.3.1 Persuasion (RQ1). Results for the persuasion metric are shown in Figure 4. The two gray histograms on the left display the distribution of all 477 persuasion scores. The top histogram shows scores from the 239 participants who saw a video *without* music, while the bottom one shows scores from the 238 participants who saw a video *with* music. These two histograms indicate that videos were generally persuasive, regardless of the inclusion of background music, which appears to have no discernible effect.

This is confirmed by the right plot showing our Bayesian estimate of the increase in persuasion score due to music (see end of subsection 4.2 on how to interpret the different elements of this plot). The overall effect of background music on persuasion, across all six videos, was 0.0036 (95% CI [-0.023, 0.031]). This effect is extremely small—likely less than 1% of the full range of possible persuasion scores (from -1 to 1)—and its direction is uncertain, as it could be slightly positive or slightly negative. Overall, the influence of background music on persuasion is negligible. In subsubsection 4.4.1, we will further discuss the magnitude of the effect and compare it with our prior expectations.

4.3.2 Engagement (RQ2). Results for the engagement metric are shown in Figure 5. Again, the two gray histograms show the raw distribution of responses for the no-music and the music condition—indicating rather positive engagement overall, while the blue plot shows our Bayesian estimate of the effect of music on persuasion score. The overall effect of music on engagement is 0.0033, 95% CI

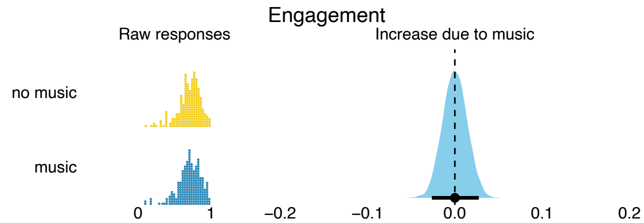

Figure 5: Effect of music on engagement across all videos. *Left:* Raw responses for engagement score without music and with music; *Right:* Bayesian estimation of the effect of music on engagement on the same 0–1 scale. The error bar is a 95% credible interval.

[-0.022, 0.030]. Once more, this effect is negligible, with no stronger evidence for a positive than for a negative effect. This finding is even less expected than the previous one, as we anticipated a substantial (medium) effect of music on engagement. We will again further interpret expected and measured effect sizes in subsubsection 4.4.1, and will discuss possible reasons for this surprising result in section 6.

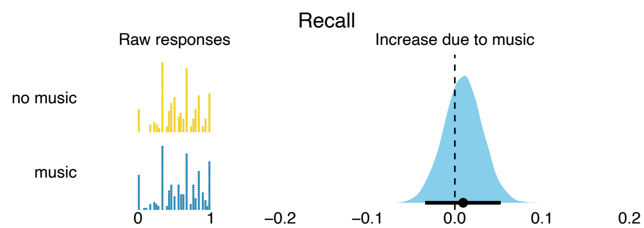

Figure 6: Effect of music on information recall across all videos. *Left:* Raw responses for recall scores without music and with music; *Right:* Bayesian estimation of the effect of music on recall on the same 0–1 scale. The error bar is a 95% credible interval.

4.3.3 Information Recall. Finally, results for the information recall metric are shown in Figure 5. The overall effect of music on recall was 0.01, 95% CI [-0.030, 0.049]. As before, the effect appears small, although there is slightly more uncertainty as to its size, likely due to a wider spread in raw responses. As for the direction of the effect, the results are again inconclusive. These results are, however, in line with our initial expectation of a small to a very small effect of music on information recall.

4.3.4 Per-Video Effects. As per our preregistration, we also examined the effect of music for each of the six videos separately. Although splitting the dataset in this way substantially reduces sample size ($N \approx 40$ per condition) and therefore the reliability of our inferences, we conducted this analysis to identify videos in which music may have an unusually large impact. All results are reported in Figure 7. There appears to be some variability across videos, especially for information recall, where the effect seems to be positive for some videos and negative for other videos. However, these patterns should be interpreted cautiously due to the small

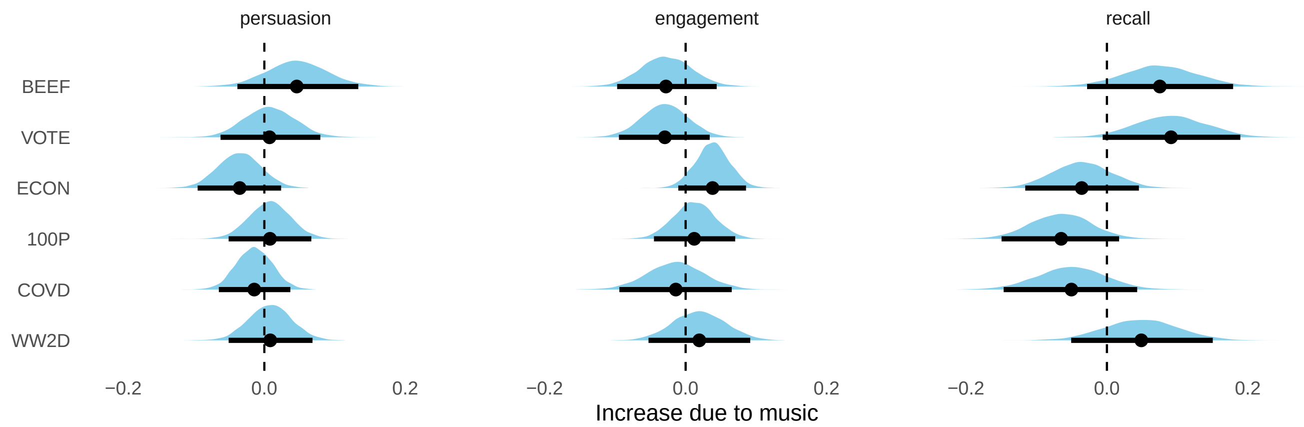

Figure 7: Effect of background music on our three key outcomes of interest (as columns) for each of the six videos (as rows). Positive values indicate higher scores in the *music* condition relative to the *no-music* condition, while negative values indicate lower scores. Error bars are 95% credible intervals.

sample sizes and the statistical multiplicity[9], and more focused studies are needed to confirm these effects. For the time being, the only conclusion we can reasonably draw is that there is no evidence of a very strong effect of background music for any of the videos.

4.4 Additional Analyses

The analyses in this section were not preregistered and were conducted post-hoc.

4.4.1 Standardized Effect Sizes. To help us interpret how small the effect of music is, we estimated Cohen's d [18] for our three primary outcomes. Cohen's d is a standardized, unitless measure of effect size, where the difference between two group means is expressed relative to the variability within the groups. We performed a Bayesian estimation of d by fitting Gaussian regression models and dividing the posterior draws of the fixed effect for music by the corresponding posterior draws of the residual standard deviation.

The resulting estimates are reported on the right of Figure 1 (blue density plots, black error bars and black dots).

All three point estimates (dots) are extremely close to zero, with the 95% credible intervals (error bars) suggesting that the maximum plausible effect size in either direction is less than $|d| = 0.2$, which is conventionally considered as a small effect size [18, p.24]. Overall, these results confirm that music likely has little to no practically meaningful impact on persuasion, engagement, or recall in this study.

Comparing these results with our prior expectations (gray plots in Figure 1, see subsection 3.3), we anticipated a small positive effect of music on persuasion ($d \approx 0.2$); while our results do not entirely rule out this possibility, the prediction was likely on the optimistic side. For engagement, we speculated a medium positive effect ($d \approx 0.5$), which is fully contradicted by the small and possibly negative effect we observed. Finally, for information recall, our predictions were already cautious, anticipating a small or very small positive effect ($d \approx 0.1$–0.2); our findings are consistent with these predictions, but we did not anticipate the possibility of a negative effect, which as likely as a positive one. Note that in order to get unbiased estimates, we did *not* use our effect size expectations as priors in our models.

4.4.2 Additional Measures. We additionally measured the degree to which participants perceived the video as wanting to persuade them. Responses were overall high irrespective of the presence of music (see Figure 8–left).

We also measured quite a high behavioral intent, operationalized as willingness to take action after having watched the video. However, again, we were not able to find evidence for an effect of music on this measure either (see Figure 8–right).

Finally, we had asked participants if they thought that the *music enhances the video* (for those who were in the music condition) or if they thought that the *addition of music would enhance the video*. Raw responses to these two questions are summarized as histograms in Figure 9. The patterns in participants' responses reveal meaningful differences across video types. For videos with voiced narration viewed in the no-music condition (9, top row, leftmost three columns), opinions were mixed: some participants felt that adding music would improve the video, while others were unsure or believed it would add little. In contrast, for videos without voiced narration, participants more consistently indicated that music would enhance the experience, suggesting that when narration is absent, viewers may expect or value a musical layer to support pacing or emotional tone. In the music condition (9, bottom row), many participants reported that the background music enhanced the video they watched, though with notable variability for COVD (), whose intentionally incongruous meme music was polarizing, being appreciated by some participants who were familiar with the reference but considered distracting by others.

[9] Since we fit a total of $3 \times 6 = 18$ models, each on a different dataset, the probability of observing at least one 95% credible interval excluding zero by pure chance is $1 - (1 - 0.05)^{18} \approx 0.6$, assuming there is no effect in any of the datasets [50].

Figure 8: Effect of music on perceived persuasion (left) and behavioral intent (right). The two histograms on the respective left sides show raw responses for the two measures while the respective right sides show the Bayesian estimation of the effect of music on the two measures. The error bar is a 95% credible interval.

5 Qualitative Data Analysis

To complement our quantitative results, we conducted a qualitative analysis of participants' open-ended feedback to better understand their experiences during the study. Of the 477 participants, 318 either left the comment section empty or provided only brief, non-substantive remarks (e.g., "no," "thank you," "n/a," etc.). This resulted in 159 substantive comments, which we coded inductively using an open coding approach. Two coders independently coded the data, resulting in 198 individual codes. Then, across two sessions, the codes were discussed, conflicts resolved, and consolidated in groups following the thematic analysis approach [17]. The groups were then iteratively discussed and finalized into themes through discussions within the research team.

5.1 Results

17 participants commented explicitly on the audio design, with 16 remarks focusing on the background music and one on the narration. These responses reflect a wide range of opinions, highlighting the challenge of creating soundtracks that are both effective and unobtrusive in data videos.

Participant comments were unevenly distributed across the six videos (see again Table 1 for details about the videos). The largest number concerned COVD (♫), which accounted for eight comments. Three comments focused on WW2D (🔇), two on 100P (🔇), and one each on VOTE (♫), ECON (♫), and BEEF (🔇). The sole remark about narration occurred in the BEEF (♫). As a reminder, BEEF, VOTE, and ECON featured *voice narration* and were rated by coders as having a *medium presence* of music (in the music condition, where the music track was intact). Meanwhile, 100P, COVD, and WW2D had *no voice narration* and were rated as having a *high presence* of music.

5.1.1 Background music influenced focus, sometimes in distracting ways.

5.1.1 Background music influenced focus, sometimes in distracting ways. Out of 16 comments on the music, 14 participants expressed concerns about the music's impact on their ability to focus or trust the message. For instance, *P003* stated, the music used for COVD (♫) *"was actually very distracting."* Some other responses were more severe in tone: *P067* described the soundtrack of the same video as *"some cheap ... theme song,"* and *P081* offered a particularly strong critique on VOTE (♫): *"The music is standard, cookie cutter, commercialized, soulless and boring ... I'd prefer no music at all if the outcome ... comes off as just insulting to intelligence."* This participant (*P081*) also framed the video itself as part of a broader *"carefully constructed exercise of manipulation"*; reflecting deeper skepticism toward persuasive media. Similarly, *P138* warned that music might make WW2D (♫) feel *"emotionally manipulative."* Some participants explicitly preferred the absence of music due to the clarity it offered. *P383* remarked, *"I was glad that the video wasn't accompanied by music because I could concentrate better without it"*; referring to 100P (🔇). In contrast, for the same video (100P (🔇)), *P144* mentioned how music could have served as a focusing aid: *"Without music it was very hard to stay focused. I had to actively retain my focus."* This is consistent with prior research suggesting that music can support attentional engagement if appropriately matched to content and tone [40, 52, 67].

One participant directly addressed the narration, noting that it made concentration difficult. *P059* wrote, *"... the voice over [on BEEF (♫)] was a little bit boring and made it hard to concentrate well to the video."* While not a major focus in participant feedback, this comment suggests that narration quality, similar to music, can modulate engagement and comprehension, and indicates the importance of narration design in data videos.

5.1.2 Background music depends on selection and execution. While most of our participants who commented on music (14/16) were negative about background music in the data video they viewed, some of them emphasized that the effect of music depends heavily on its specific selection and execution. As *P129* pointed out, *"It would really depend on the music—this can be a challenge to find the right one"* ; referring to ECON (🔇). In addition, two participants mentioned being familiar with the music being played in the COVD (♫) video they watched. The video follows a template for a popular internet meme, amassing over 20 million views. However, the participants reacted differently to the familiar music. *P274* mentioned *"I knew the song used in the video"*, but *P275* said, *"Cringe 2020 meme."*

5.1.3 Background music shaped emotional responses. 9 participants described strong emotional responses to the videos, underscoring how data-driven narratives can resonate on a personal and affective level. These reactions ranged from sadness and empathy to shock and surprise. Many of these comments reflected how participants connected with human suffering or broader social realities. For example, *P124* found the content of WW2D (♫) *"very moving and shocking"*; and *P330* described it as *"heartbreaking to think most of the deaths were young men with so much life ahead of them."* Other participants expressed astonishment or disbelief toward specific

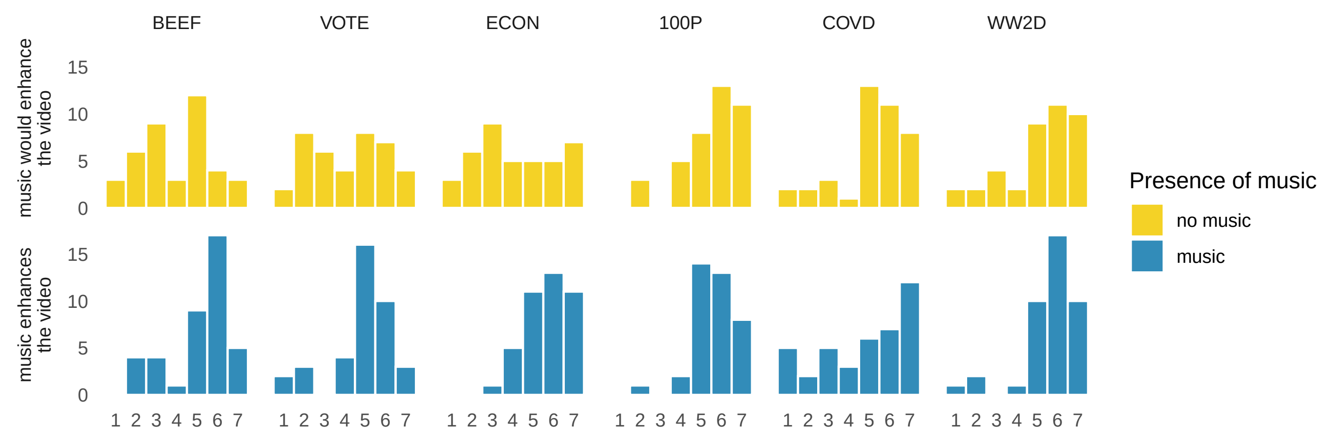

Figure 9: Top row: Responses to the question if *music would enhance the video* (only asked in the *no music* condition. Bottom row: Responses to the question if music was experienced as enhancing the video watched (only asked in the *music* condition).

statistics. For instance, *p325* remarked, *"If these numbers are correct, it is surprising"*; suggesting that the magnitude or nature of the figures in 100P (♫) challenged their prior expectations. Another participant (*P266*) found the COVD (♫) video amusing despite the seriousness of the topic: *"Although about a serious subject, I thought the video was hilarious."*

Notably, every participant who reported an emotional reaction to the WW2D video viewed the music condition (WW2D (♫)). While such anecdotes may not establish causality, this pattern suggests that the somber background music may have heightened the emotional salience of the already intense visual content.

The diversity of these participant responses points to an important nuance in our findings: many critical comments may reflect not general dislike or indifference to background music, but dissatisfaction with a specific music track used, including its genre, tone, or mixing. Themes of commercialism, manipulation, and tonal mismatch suggest that music's role in data videos may not be absolute, but contingent on aesthetic, emotional, and technical factors. Moreover, the comments on music, together with those specifically tied to individual videos, indicate that familiarity may play a role in shaping audience reactions. When viewers recognize a theme or background track, they tend to respond more positively and enjoy the video more, such as in COVD (♫), rather than feeling distracted or offended.

6 Discussions and Future Work

While prior work suggests that background music can enhance emotional impact or persuasiveness in videos, our findings suggest that its effect may be limited, subjective, or context-dependent when applied to data videos. Surprisingly, we found no measurable effect of background music on persuasion, engagement, or recall, suggesting that the communicative power of data videos may lie more in factors such as visual and narrative structure than in the presence of background music. In this section, we speculate on possible reasons for this limited impact, drawing on prior research and our own findings, and we discuss implications for both future studies and design practices to further tease apart when and how music could have a potential impact on data videos.

6.1 The Delicate Interplay Between Auditory and Visual Narrative Elements

While prior works [25, 26] on integrating music with visualization suggest that music can enhance engagement and emotional resonance, our quantitative results did not reveal measurable effects in our stimuli, yet the qualitative responses indicate that the influence of music may depend on how it interacts with the visual and narrative layers of the video. One possible reason for the limited effect of background music is that viewers of data videos may focus mainly on elements most critical for comprehension. Research in multimedia learning suggests that people give most of their attention to information that directly helps them understand the material, such as graphics and spoken explanations [55]. Similarly, in data videos, the visualization is the central source of meaning, while the narration provides context and guides interpretation [34]. At the same time, background music occupies a more delicate position: it does not provide factual content, but it can shape pacing, tone, and emotional resonance when it aligns with the narrative. This interplay may help explain why some participants perceived music as enhancing focus or emotional connection, even though these effects did not emerge in our quantitative measures. Taken together, this highlights a distinction between the inferred communicative role of background music in the original video design and the effects of music as perceived and reported by viewers, which did not always align. Rather than being purely peripheral, music's influence appears contingent on how it complements the visual storyline and the viewer's interpretive needs.

This idea is supported by studies on selective attention: when people face several types of information at once, they tend to ignore elements that do not seem essential to the task [48]. For data video viewers, the main task is to understand the patterns and arguments shown in the visuals and narration. This suggests that in data storytelling, visual and narrative design carry most of the communicative weight, while music plays a supporting role. This emphasis on narration is also reflected in our findings. Even though Figure 7 shows no measurable difference in persuasion, engagement, or recall across conditions, participants often indicated a preference for having some form of audio rather than complete silence. As illustrated in Figure 9, many in the no-music condition who had watched the videos without narration (100P, COVD, and WW2D) noted that the videos felt less engaging without background sound. This suggests that when narration is present, background music may become less critical, but in its absence, silence can make the viewing experience feel incomplete or awkward.

Another potential reason why music may have shown little effect in our study is that the data videos used in our experiment might already have been sufficiently informative and visually engaging. For example, *P337* who watched the BEEF video without music commented that the video is *"very informative and well made."* When the content itself is rich in information and visually well designed, additional auditory cues may not add much to the viewer's experience. Studies in multimedia learning show that extra elements often help only when the main material is difficult to process or lacks clarity [55]. Research on cognitive load also suggests that adding non-essential elements can sometimes have neutral or even negative effects, especially if the core material is already demanding [90]. Data videos often rely on dynamic graphics, comparisons, and explanatory narration, which already provide multiple cues for engagement [7]. In such contexts, music may not significantly increase attention or motivation where the data videos are designed well enough to demand viewers' full attention. Therefore, the effectiveness of music may depend on situations where the content alone is less capable of sustaining attention, or where emotional framing is needed to guide interpretation.

6.2 Tailoring Music for Effective Data Storytelling

In some cases, the background music tracks used in our study may not have been well aligned with the content of the videos. Consistent with this are participants' comments on COVD (♬) and VOTE (♬) finding the background music to be "generic" (see Section 5.1 for more details). Prior research in film and media studies suggests that music is most effective when carefully synchronized with narrative shifts or emotionally important scenes [10]. Generic or template-based tracks may not be strong enough to shape perception, whereas *incidental* and *ambient* [47] music (broadly, background music that supports mood or pacing without dominating attention) have the potential to more effectively guide emphasis or transitions when selected or edited to align with narrative structure. In data videos, this alignment does not require bespoke composition; rather, even modest timing or stylistic adjustments can help the soundtrack better reinforce key data-driven moments. For example, musical cues can emphasize a sudden change in scale, a surprising comparison, or a narrative turning point. Music may even be directly driven by the data visualized [89]. Research has suggested that music carefully synchronized with a video's emotional and narrative content can enhance audience engagement and the perceived quality of the storytelling [87]. Prior studies integrating visualization, sonification, and music also indicate that when sound or music is designed deliberately, it can set tempo, direct attention, and integrate with graphics to create richer meanings [74].

A related consideration is that data videos vary in how their temporal structure is established. In some cases, creators begin with a preselected music track and edit the visuals to match its tempo and intensity; the COVD video in our corpus is an example of this approach, where the rapid acceleration of the COVID-19 diagnosis bars is deliberately aligned with a high-energy Eurobeat song. Participants who recognized the meme associated with this track often reported that the match between the music and the visual pacing enhanced their experience. Such workflows are analogous to *pre-lay* or voice-over editing in animations [28, 53], where the audio is recorded first, and then the visuals are adjusted to it. Both approaches represent meaningful forms of audio–visual alignment.

At the same time, our results leave open the possibility that music could make a difference under the right conditions. As shown in Figure 7, the estimated effects of music were small and uncertain, but Figure 9 indicates that many participants nonetheless perceived music as enhancing the video compared to silence. This contrast suggests that while generic tracks may fade into the background, music chosen or adapted to better fit the intended narrative rhythm may more effectively support attention or perceived engagement. Future research could therefore examine whether incidental or ambient tracks curated for narrative fit, or simple timing adjustments at key transitions can translate perceived improvements into measurable effects. Building on this possibility, an important next step is to consider practical strategies for aligning music with data-driven storytelling without requiring large-scale compositional effort.

Personalization could allow viewers to select from multiple musical layers or alternative tracks, adapting the tone of the video to their own preferences, cultural background, or emotional state. Additionally, our qualitative findings highlight how familiarity and preference strongly shapes reactions to background music (e.g., in the COVD (♬) condition). Considering these, personalized music systems could help accommodate such divergent reactions by aligning the soundtrack with individual expectations and prior associations. Adaptive systems could take this further by tailoring music automatically based on contextual or user-specific factors, such as prior viewing behavior, demographic profile, or even real-time physiological responses. Beyond personalization, music could also be leveraged as a supplementary tool for comprehension, such as using auditory icons or melodic transitions to provide content-dependent cues [45, 97] or using music as a complementary data encoding channel [89]. Advancing these approaches will require closer collaboration between designers and musicians, while grappling with the persistent challenge of musical subjectivity. From a research perspective, the central question is whether personalization and adaptivity enhance persuasion, engagement, and recall by aligning with user expectations, or whether they risk diluting the intended message. Addressing these trade-offs will be key to establishing design guidelines for more interactive and adaptive uses of music in data videos.

6.3 Reassessing Measures and Methods for Evaluating Music's Impact

Integrating music into visualization exposes unresolved design and evaluation challenges across visual and auditory modalities [25]. Traditional measures and methods used to evaluate persuasion, engagement, and recall may not fully capture the effects of music. While these metrics are commonly used in visualization research [2, 12, 69], applying them to music is complicated. Music often influences viewers indirectly, by shaping mood, framing attention, or guiding interpretation, rather than by directly improving comprehension or recall [13, 49, 57, 93]. If the measures are not sensitive to these more subtle effects, the role of music could be underestimated. Moreover, single post-study measures may miss longer-term outcomes. Persuasion or recall tests given immediately after viewing cannot reveal whether music has lasting impacts on memory, attitudes, or behavior. Insights from multimedia learning theory suggest that evaluation should extend beyond factual recall to include transfer of learning, motivational outcomes, and emotional responses [54].

As a richer evaluation of user experience is essential for determining whether music serves more than an enhancer and clarifying its role in communicative visualization [24, 36, 74], future work should therefore adopt a broader methodological toolkit. Such toolkits may combine traditional tests with continuous engagement tracking, affective self-reports, or physiological measures.

Psychometric instruments such as standardized scales of cognitive load or emotional response [64, 68, 79] could be used in future work to capture additional dimensions of the viewing experience, such as mental effort or affective intensity, that were beyond the scope of our outcome measures. Likewise, implicit measures such as the Implicit Association Test (IAT) [31, 63] provide a complementary way to examine subconscious associations that are not easily accessed through self-report. In persuasion research, the IAT can reveal whether exposure to a message shifts automatic associations with a target concept, providing evidence of implicit attitude change. For engagement, it may help uncover audiences' positive or negative associations with content or themes, possibly indicating levels of involvement not captured in self-reports. Prior research suggests that music can influence comprehension, distraction, and implicit persuasion in ways invisible to explicit ratings [85]. Systematically integrating these measures would provide a more complete picture of how music operates in data videos and help establish evaluation frameworks that balance overt outcomes with implicit dimensions of user experience.

6.4 Possible Negative Effects When Music Is Mismatched

While our quantitative results did not show consistent effects of background music, our qualitative findings suggest that music can introduce challenges when it is not well aligned with the video's tone, message, or intended audience. In our qualitative findings, the role of music emerged as contingent rather than absolute: music was appreciated when it aligned with the video's tone, message, and viewer expectations, but it was reported as distracting or alienating when mismatched in genre, tone, or cultural familiarity. Recognition of familiar tracks, such as the background music in COVD (♫), sometimes led viewers to focus more on the music than on the data or argument being presented. This possibility is supported by research in advertising, which suggests that incongruent or overly salient music can compete for cognitive resources, reduce message comprehension, or divert attention away from intended content [6, 35, 39, 84]. Multimedia learning studies similarly warn that elements not essential to understanding can impose unnecessary cognitive load and impair learning [90]. Importantly, these observations must be interpreted within the scope of our stimuli: the background music in the selected videos was chosen by their original creators but was not purpose-composed to optimize persuasion, engagement, or comprehension. Our findings, therefore, speak to how background music in *existing*, widely circulated data-videos function in practice, rather than to the effects of custom-designed or intentionally crafted music. Purposeful audio design, such as composing music to match narrative arcs, timing cues to key transitions, or using incidental or ambient music to support pacing, may lead to different outcomes and remains an open avenue for future work. In data videos, mismatched or overly familiar music may therefore interfere with rather than enhance communication, underscoring the importance of careful alignment between auditory and visual design.

Future research should investigate these potential negative impacts and how to mitigate them to ensure that the integration of music in data videos supports communication rather than undermining it. Taken together, our results suggest that mismatched or overly familiar music can interfere with communication in some circumstances, but they do not preclude the possibility that well-designed audio could enhance comprehension or engagement. Future research should explicitly compare off-the-shelf soundtracks with custom-designed musical treatments to further determine when music supports—or undermines—the communicative goals of data videos.

6.5 Limitations

Similar to other studies, our study should also be interpreted in light of limitations. Despite our attempt to balance thematic breadth, stylistic variety, and consistent persuasive intent, our selected videos do not exhaustively capture or represent the full breadth of visual, topical, and stylistic practices across data videos available across all domains. We included only six videos to maintain a minimum level of experimental control, verifying that all videos met rigorous selection criteria and were able to detect any substantial video-specific effects. Future studies could improve generalizability by including a larger sample of videos, though assessing the contribution of individual videos would likely become challenging: in our study, dividing $N = 480$ across six videos reduced the group-wise sample size to just $N = 40$ per condition (with vs. without music), providing sufficient statistical power only to detect large video-specific effects (subsubsection 3.4.1).

Moreover, we decided to conduct our studies with videos with their original background music, which might have impacted viewers' perceptions of said music. Different results may emerge if music is composed or adapted specifically to accentuate narrative beats or highlight key data moments. Furthermore, we measured persuasion, engagement, and recall immediately after participants were exposed to the stimulus, which might have limited their response to the viewed data videos. We did not measure if other factors, such as the emotions of participants, their perceived credibility of the video, and their aesthetic satisfaction, may also interact with the existence of background music and contribute to participants’ responses. Our work does not account for long-term or subtle effects such as trust, enjoyment, or aesthetic resonance, which create avenues for future investigations.

Moreover, participants watched videos chosen by us, which may not reflect ecological situations where people freely choose data videos to watch online or follow links shared with them. Finally, our participants were recruited from an online crowdsourcing platform, and despite the screening procedure and participation across all continents, they may not represent the relevant cultural musical backgrounds and diversity for a broad range of international audiences.

These limitations can serve as a scaffolding for potential future research where the data videos for investigation could be sourced from an expanded, language-agnostic, and culturally-sensitive corpus. Furthermore, bespoke or adaptive background music could be composed by engaging with expert musicians or generative AI to accentuate the narrative beats of critical data elements that could impact persuasion, engagement, or recall. Or perhaps longitudinal and socially embedded effects beyond these measures could be investigated to deepen our understanding of the interplay between background music and visual narratives in data videos.

7 Conclusion

We set out to test a widely held assumption that background music enhances the communicative power of data videos. In our preregistered study with 477 participants, we did not observe consistent measurable effects of background music on persuasion, engagement, or information recall across six widely viewed real-world data videos. At the same time, our qualitative findings revealed that viewers’ responses to background music were varied: some described certain tracks as distracting or mismatched, while others reported that music enhanced enjoyment, focus, or emotional resonance when well aligned with the video’s tone or familiar to them. These findings challenge designers to rethink the role of music in data storytelling. Rather than treating music as a default layer, we suggest exploring custom—possibly data-driven—scoring, soundtrack customization, and evaluation frameworks that capture both subtle and long-term effects. These results suggest that the influence of background music in data storytelling is contingent on selection, genre, familiarity, and narrative alignment. Rather than treating music as a default design element, future work may benefit from exploring intentional musical choices (e.g., incidental, ambient, purpose-designed, or data-driven) and from studying how different forms of alignment can shape perception and engagement. By reframing background music as a context-dependent component of data videos rather than a universal enhancer, we provide a foundation for future research on when, how, and for whom music may meaningfully contribute to data-driven communication. Advancing the role of music in data videos requires both cross-disciplinary collaboration and more nuanced evaluation methods. On the design side, collaboration among visualization researchers, musicians, cognitive scientists, and communication scholars may help translate the diverse roles background music can play into practical strategies for data-video production. On the evaluation side, interdisciplinary methods (e.g., psychometric instruments, implicit measures, and longitudinal designs) offer ways to capture subtle, delayed, or user-specific effects that may not appear in immediate self-report outcomes.

Acknowledgments

We thank Eugénie Brasier for ideating on the teaser figure. Individual elements of our teaser (Figure 1) were created using ChatGPT’s text-to-image generative AI, based on researcher-provided ideas and sketches; the final composition was manually assembled using non-AI tools. We also thank Léana Petiot, Eugénie Brasier, Aymeric Ferron, and Martin Hachet at Inria for their valuable feedback and comments on the introduction.

References

[1] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Christophe Hurter, and Pourang Irani. 2015. Understanding data videos: Looking at narrative visualization through the cinematography lens. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1459–1468. doi:10.1145/2702123.2702431

[2] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Jason Leboe-McGowan, and Pourang Irani. 2018. Hooked on data videos: assessing the effect of animation and pictographs on viewer engagement. In *Proceedings of the International Conference on Advanced Visual Interfaces*. Association for Computing Machinery, New York, NY, USA, 1–9. doi:10.1145/3206505.3206552

[3] Fereshteh Amini, Nathalie Henry Riche, Bongshin Lee, Andres Monroy-Hernandez, and Pourang Irani. 2016. Authoring data-driven videos with dataclips. *IEEE Transactions on Visualization and Computer Graphics* 23, 1 (2016), 501–510. doi:10.1109/TVCG.2016.2598647

[4] Charles S Areni and David Kim. 1993. The influence of background music on shopping behavior: classical versus top-forty music in a wine store. *Advances in Consumer Research* 20 (1993), 336–340.

[5] AudioStrip. 2025. Audiostrip: Remove Vocals from Music Tracks Using AI. https://audiostrip.com/ Accessed: 2025-07-04.

[6] Jose M Ausin, Enrique Bigne, Javier Marín, Jaime Guixeres, and Mariano Alcaniz. 2021. The background music-content congruence of TV advertisements: A neurophysiological study. *European Research on Management and Business Economics* 27, 2 (2021), 100154. doi:10.1016/j.iedeen.2021.100154

[7] Scott Bateman, Regan L Mandryk, Carl Gutwin, Aaron Genest, David McDine, and Christopher Brooks. 2010. Useful junk? The effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 2573–2582. doi:10.1145/1753326.1753716

[8] BBC. 2011. Hans Rosling’s 200 Countries, 200 Years, 4 Minutes. https://www.youtube.com/watch?v=jbkSRLYSojo. Accessed: 2025-08-31.

[9] Lonni Besançon and Pierre Dragicevic. 2019. The continued prevalence of dichotomous inferences at CHI. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–11. doi:10.1145/3290607.3310432

[10] Valerie J Bolivar, Annabel J Cohen, and John C Fentress. 1994. Semantic and formal congruency in music and motion pictures: Effects on the interpretation of visual action. *Psychomusicology: A Journal of Research in Music Cognition* 13, 1-2 (1994), 28. doi:10.1037/h0094102

[11] Michelle A Borkin, Zoya Bylinskii, Nam Wook Kim, Constance May Bainbridge, Chelsea S Yeh, Daniel Borkin, Hanspeter Pfister, and Aude Oliva. 2015. Beyond memorability: Visualization recognition and recall. *IEEE Transactions on Visualization and Computer Graphics* 22, 1 (2015), 519–528. doi:10.1109/TVCG.2015.2467732

[12] Michelle A Borkin, Azalea A Vo, Zoya Bylinskii, Phillip Isola, Shashank Sunkavalli, Aude Oliva, and Hanspeter Pfister. 2013. What makes a visualization memorable? *IEEE Transactions on Visualization and Computer Graphics* 19, 12 (2013), 2306–2315. doi:10.1109/TVCG.2013.234

[13] Elvira Brattico and Thomas Jacobsen. 2009. Subjective appraisal of music: Neuroimaging evidence. *Annals of the New York Academy of Sciences* 1169, 1 (2009), 308–317. doi:10.1111/j.1749-6632.2009.04843.x

[14] Ruochen Cao, Subrata Dey, Andrew Cunningham, James Walsh, Ross T Smith, Joanne E Zucco, and Bruce H Thomas. 2020. Examining the use of narrative

constructs in data videos. *Visual Informatics* 4, 1 (2020), 8–22. doi:10.1016/j.visinf.2019.12.002
[15] Hao Cheng, Junhong Wang, Yun Wang, Bongshin Lee, Haidong Zhang, and Dongmei Zhang. 2022. Investigating the role and interplay of narrations and animations in data videos. *Computer Graphics Forum* 41, 3 (2022), 527–539. doi:10.1111/cgf.14560
[16] Eun Kyoung Choe, Yumiko Sakamoto, Yanis Fatmi, Bongshin Lee, Christophe Hurter, Ashkan Haghshenas, and Pourang Irani. 2019. Persuasive data videos: Investigating persuasive self-tracking feedback with augmented data videos. In *Proceedings of the AMIA Annual Symposium*. American Medical Informatics Association, Bethesda, MD, USA, 295–304. doi:10.1037/h0037050
[17] Victoria Clarke and Virginia Braun. 2014. Thematic analysis. In *Encyclopedia of Critical Psychology*. Springer, New York, NY, USA, 1947–1952.
[18] Jacob Cohen. 1988. *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates, Hillsdale, NJ, USA. doi:10.4324/9780203771587
[19] Shauna Concannon, Natasha Rajan, Parthiv Shah, Davy Smith, Marian Ursu, and Jonathan Hook. 2020. Brooke leave home: Designing a personalized film to support public engagement with open data. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3313831.3376462
[20] Matthew Conlen, Jeffrey Heer, Hillary Mushkin, and Scott Davidoff. 2023. Cinematic techniques in narrative visualization. arXiv:2301.03109 [cs.HC] doi:10.48550/arXiv.2301.03109 Preprint submitted to arXiv.
[21] Geoff Cumming and Robert Calin-Jageman. 2024. *Introduction to the new statistics: Estimation, open science, and beyond*. Routledge, New York, NY, USA. doi:10.4324/9781032689470
[22] Efren de la Mora Velasco, Atsusi Hirumi, and Baiyun Chen. 2021. Improving instructional videos with background music and sound effects: A design-based research approach. *Journal of Formative Design in Learning* 5, 1 (2021), 1–15. doi:10.1007/s41686-020-00052-4
[23] Andrew SC Ehrenberg. 1977. Rudiments of numeracy. *Journal of the Royal Statistical Society Series A: Statistics in Society* 140, 3 (1977), 277–297. doi:10.2307/2344922
[24] Kajetan Enge, Elias Elmquist, Valentina Caiola, Niklas Rönnberg, Alexander Rind, Michael Iber, Sara Lenzi, Fangfei Lan, Robert Höldrich, and Wolfgang Aigner. 2024. Open Your Ears and Take a Look: A State-of-the-Art Report on the Integration of Sonification and Visualization. *Computer Graphics Forum* 43, 3 (2024), e15114. doi:10.1111/cgf.15114
[25] Kajetan Enge, Alexander Rind, Michael Iber, Robert Höldrich, and Wolfgang Aigner. 2021. It's about time: adopting theoretical constructs from visualization for sonification. In *Proceedings of the International Audio Mostly Conference*. Association for Computing Machinery, New York, NY, USA, 64–71. doi:10.1145/3478384.3478415
[26] Kajetan Enge, Alexander Rind, Michael Iber, Robert Höldrich, and Wolfgang Aigner. 2023. Towards a unified terminology for sonification and visualization. *Personal and Ubiquitous Computing* 27, 5 (2023), 1949–1963. doi:10.1007/s00779-023-01720-5
[27] Jean-Philippe Galan. 2009. Music and responses to advertising: The effects of musical characteristics, likeability and congruency. *Recherche et Applications en Marketing (English Edition)* 24, 4 (2009), 3–22. doi:10.1177/205157070902400401
[28] Ian Garwood. 2012. Roto-Synchresis: Relationships between Body and Voice in Rotoshop Animation. *Animation* 7, 1 (2012), 39–57. doi:10.1177/1746847711428850
[29] Sarmistha Sarna Gomasta, Mahmood Jasim, Hossein Hadisi, Yvonne Jansen, Pierre Dragicevic, Narges Mahyar, and Ali Sarvghad. 2026. Investigating How Music Affects Persuasion, Engagement, and Emotion in Data Videos. arXiv:2601.17893 [cs.HC] https://arxiv.org/abs/2601.17893
[30] GOOD Magazine. 2016. If The World Were 100 People. https://www.youtube.com/watch?v=QFrqTFRy-LU&ab_channel=GOODMagazine. Accessed: 2025-04-09.
[31] Anthony G Greenwald, Debbie E McGhee, and Jordan LK Schwartz. 1998. Measuring individual differences in implicit cognition: the implicit association test. *Journal of Personality and Social Psychology* 74, 6 (1998), 1464. doi:10.1037//0022-3514.74.6.1464
[32] Neil Halloran. 2015. The fallen of world war ii. https://www.youtube.com/watch?v=DwKPFT-RioU. Accessed: 2025-08-31.
[33] Thomas Huk, Silvia Bieger, Silvia Ohrmann, and Beate Weigel. 2004. Computer Animations in Science Education: Is Background Music Beneficial Or Detrimental?. In *Proceedings of the World Conference on Educational, Multimedia, Hypermedia & Telecommunications*. AACE, Chesapeake, VA, USA, 4227–4234.
[34] Jessica Hullman and Nick Diakopoulos. 2011. Visualization rhetoric: Framing effects in narrative visualization. *IEEE Transactions on Visualization and Computer Graphics* 17, 12 (2011), 2231–2240. doi:10.1109/TVCG.2011.255
[35] Kineta Hung. 2000. Narrative music in congruent and incongruent TV advertising. *Journal of Advertising* 29, 1 (2000), 25–34. doi:10.1080/00913367.2000.10673601
[36] Michael Iber, Kajetan Enge, Niklas Rönnberg, Annika Neidhardt, Norbert Schnell, Katharina Pollack, Maria kallionpää, and Alan Chamberlain. 2024. What you hear is what you see? Perspectives on modalities in sound and music interaction. *Personal and Ubiquitous Computing* 28, 5 (2024), 655–656. doi:10.1007/s00779-024-01835-3
[37] Mohaiminul Islam and Shangzhu Jin. 2019. An overview of data visualization. In *Proceedings of the International Conference on Information Science and Communications Technologies*. IEEE, Piscataway, NJ, USA, 1–7. doi:10.1109/ICISCT47635.2019.9012031
[38] Juliane Kämpfe, Peter Sedlmeier, and Frank Renkewitz. 2011. The impact of background music on adult listeners: A meta-analysis. *Psychology of Music* 39, 4 (2011), 424–448. doi:10.1177/0305735610376261
[39] James J Kellaris, Anthony D Cox, and Dena Cox. 1993. The effect of background music on ad processing: A contingency explanation. *Journal of Marketing* 57, 4 (1993), 114–125. doi:10.2307/1252223
[40] Luca Kiss and Karina J Linnell. 2024. The role of mood and arousal in the effect of background music on attentional state and performance during a sustained attention task. *Scientific Reports* 14, 1 (2024), 9485. doi:10.1038/s41598-024-60218-z
[41] Kevin M Kniffin, Jubo Yan, Brian Wansink, and William D Schulze. 2017. The sound of cooperation: Musical influences on cooperative behavior. *Journal of Organizational Behavior* 38, 3 (2017), 372–390. doi:10.1002/job.2128
[42] Ha-Kyung Kong, Wenjie Zhu, Zhicheng Liu, and Karrie Karahalios. 2019. Understanding visual cues in visualizations accompanied by audio narrations. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3290605.3300280
[43] Kori. 2020. Corona race meme (COVID-19 growth by country). https://www.youtube.com/watch?v=4f6-MDhygWc. Accessed: 2025-04-09.
[44] John K Kruschke and Torrin M Liddell. 2018. The Bayesian New Statistics: Hypothesis testing, estimation, meta-analysis, and power analysis from a Bayesian perspective. *Psychonomic Bulletin & Review* 25, 1 (2018), 178–206. doi:10.3758/s13423-016-1221-4
[45] John B Krygier. 1994. Sound and geographic visualization. In *Modern Cartography Series*. Vol. 2. Elsevier, Amsterdam, The Netherlands, 149–166. doi:10.1016/B978-0-08-042415-6.50015-6
[46] Xingyu Lan, Yang Shi, Yanqiu Wu, Xiaohan Jiao, and Nan Cao. 2021. Kineticharts: Augmenting affective expressiveness of charts in data stories with animation design. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2021), 933–943. doi:10.1109/TVCG.2021.3114775
[47] Joseph Lanza. 2004. *Elevator music: a surreal history of muzak, easy-listening, and other moodsong; revised and expanded edition*. University of Michigan Press, Ann Arbor, MI, USA.
[48] Nilli Lavie. 2005. Distracted and confused?: Selective attention under load. *Trends in Cognitive Sciences* 9, 2 (2005), 75–82. doi:10.1016/j.tics.2004.12.004
[49] Teresa Lesiuk. 2010. The effect of preferred music on mood and performance in a high-cognitive demand occupation. *Journal of Music Therapy* 47, 2 (2010), 137–154. doi:10.1093/jmt/47.2.137
[50] Guowei Li, Monica Taljaard, Edwin R Van den Heuvel, Mitchell AH Levine, Deborah J Cook, George A Wells, Philip J Devereaux, and Lehana Thabane. 2017. An introduction to multiplicity issues in clinical trials: the what, why, when and how. *International Journal of Epidemiology* 46, 2 (2017), 746–755. doi:10.1093/ije/dyw320
[51] Wenchao Li, Yun Wang, Haidong Zhang, and Huamin Qu. 2020. Improving engagement of animated visualization with visual foreshadowing. In *Proceedings of the IEEE Visualization Conference*. IEEE, Piscataway, NJ, USA, 141–145. doi:10.1109/VIS47514.2020.00035
[52] Jens Madsen, Elizabeth Hellmuth Margulis, Rhimmon Simchy-Gross, and Lucas C Parra. 2019. Music synchronizes brainwaves across listeners with strong effects of repetition, familiarity and training. *Scientific Reports* 9, 1 (2019), 3576. doi:10.1038/s41598-019-40254-w
[53] Anna Matamala. 2018. Voice-over: practice, research and future prospects. In *The Routledge Handbook of Audiovisual Translation*. Routledge, Abingdon, UK, 64–81.
[54] Richard E Mayer. 2005. Cognitive theory of multimedia learning. *The Cambridge Handbook of Multimedia Learning* 41, 1 (2005), 31–48. doi:10.1017/CBO9781139547369.005
[55] Richard E. Mayer. 2009. *Multimedia Learning* (2 ed.). Cambridge University Press, Cambridge, UK. doi:10.1017/CBO9780511811678
[56] Richard McElreath. 2018. *Statistical rethinking: A Bayesian course with examples in R and Stan*. Chapman and Hall/CRC, Boca Raton, FL, USA. doi:10.1201/9780429029608
[57] Benjamin Meltzer, Chagit S Reichenbach, Chananel Braiman, Nicholas D Schiff, AJ Hudspeth, and Tobias Reichenbach. 2015. The steady-state response of the cerebral cortex to the beat of music reflects both the comprehension of music and attention. *Frontiers in Human Neuroscience* 9 (2015), 436. doi:10.3389/fnhum.2015.00436
[58] Alan P Merriam and Valerie Merriam. 1964. *The anthropology of music*. Northwestern University Press, Evanston, IL, USA. doi:10.2307/942218

[59] MetaBallStudios. 2018. Number of deaths in WW2 per country. https://www.youtube.com/watch?v=7cgRwDkP6vk. Accessed: 2025-04-09.

[60] Ronald E Milliman. 1982. Using background music to affect the behavior of supermarket shoppers. *Journal of marketing* 46, 3 (1982), 86–91. doi:10.2307/1251706

[61] Jon D Morris and Mary Anne Boone. 1998. The effects of music on emotional response, brand attitude, and purchase intent in an emotional advertising condition. *Advances in Consumer Research* 25 (1998).

[62] Adrian North and David Hargreaves. 2008. *The social and applied psychology of music*. Oxford University Press, Oxford, UK. doi:10.1093/acprof:oso/9780198567424.001.0001

[63] Brian A Nosek, Anthony G Greenwald, and Mahzarin R Banaji. 2007. The Implicit Association Test at age 7: A methodological and conceptual review. In *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes*. Psychology Press, New York, NY, USA, 265–292.

[64] Heather L O'Brien and Elaine G Toms. 2010. The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology* 61, 1 (2010), 50–69.

[65] Daniel M Oppenheimer, Tom Meyvis, and Nicolas Davidenko. 2009. Instructional manipulation checks: Detecting satisficing to increase statistical power. *Journal of Experimental Social Psychology* 45, 4 (2009), 867–872. doi:10.1016/j.jesp.2009.03.009

[66] Martin T Orne. 1962. On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. *American Psychologist* 17, 11 (1962), 776. doi:10.1037/h0043424

[67] Joan Orpella, Daniel Liu Bowling, Concetta Tomaino, and Pablo Ripollés. 2025. Effects of music advertised to support focus on mood and processing speed. *PloS One* 20, 2 (2025), e0316047. doi:10.1371/journal.pone.0316047

[68] Fred Paas, Juhani E Tuovinen, Huib Tabbers, and Pascal WM Van Gerven. 2003. Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist* 38, 1 (2003), 63–71. doi:10.1207/S15326985EP3801_8

[69] Anshul Vikram Pandey, Anjali Manivannan, Oded Nov, Margaret Satterthwaite, and Enrico Bertini. 2014. The persuasive power of data visualization. *IEEE Transactions on Visualization and Computer Graphics* 20, 12 (2014), 2211–2220. doi:10.1109/TVCG.2014.2346419

[70] Agustín Perez Santangelo, Casimir JH Ludwig, Joaquín Navajas, Mariano Sigman, and María Juliana Leone. 2022. Background music changes the policy of human decision-making: Evidence from experimental and drift-diffusion model-based approaches on different decision tasks. *Journal of Experimental Psychology: General* 151, 9 (2022), 2222–2236. doi:10.1037/xge0001189

[71] Van Thanh Phan. 2023. Effects of Background Music in Instructional Videos on Learners' Retention. *Journal of Educators Online* 20, 3 (2023), n3. doi:10.9743/JEO.2023.20.3.4

[72] Jeran Poehls, Monique Meuschke, Nuno Carvalhais, and Kai Lawonn. 2025. Either Or: Interactive Articles or Videos for Climate Science Communication. *Computer Graphics Forum* 44, 3 (2025), e70129. doi:10.1111/cgf.70129

[73] Marta Garcia Quinones, Anahid Kassabian, and Elena Boschi. 2016. *Ubiquitous musics: The everyday sounds that we don't always notice*. Routledge, New York, NY, USA. doi:10.4324/9781315549453

[74] Alexander Rind, Kajetan Enge, Michael Iber, Niklas Rönnberg, Sara Lenzi, Elias Elmquist, Valentina Caiola, Fangfei Lan, Robert Höldrich, and Wolfgang Aigner. 2024. Integrating Sonification and Visualization - But Why?. In *Proceedings of the Panel of Eurographics / IEEE VGTC Conference on Visualization*, Hans-Jörg Schulz and Tobias Isenberg (Eds.). Eurographics Association, Goslar, DEU. doi:10.2312/evt.20241100

[75] Holly Rogers. 2014. *Music and sound in documentary film*. Routledge, New York, NY, USA. doi:10.4324/9781315851556

[76] Jen Rogers, Leixian Shen, Ab Mosca, Evan Peck, Mingwei Li, Anzu Hakone, Kristi Potter, and Remco Chang. 2025. House Advantage or House of Cards? Stacking the Deck for Data Videos Leads to Null Results. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–14. doi:10.1145/3706599.3720256

[77] Yumiko Sakamoto, Samar Sallam, Aaron Salo, Jason Leboe-McGowan, and Pourang Irani. 2022. Persuasive Data Storytelling with a Data Video during Covid-19 Infodemic: Affective Pathway to Influence the Users' Perception about Contact Tracing Apps in less than 6 Minutes. In *Proceedings of the Pacific Visualization Symposium*. IEEE, Piscataway, NJ, USA, 176–180. doi:10.1109/PacificVis53943.2022.00028

[78] Samar Sallam, Yumiko Sakamoto, Jason Leboe-McGowan, Celine Latulipe, and Pourang Irani. 2022. Towards design guidelines for effective health-related data videos: an empirical investigation of affect, personality, and video content. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–22. doi:10.1145/3491102.3517727

[79] Ulrich Schimmack. 2008. The structure of subjective well-being. *The Science of Subjective Well-being* 54, 1 (2008), 97–123.

[80] Oliver Shao. 2023. *Composing Aid: Music, Refugees, and Humanitarian Politics*. Indiana University Press, Bloomington, IN, USA. doi:10.2307/jj.5590551

[81] Leixian Shen, Haotian Li, Yun Wang, and Huamin Qu. 2025. Reflecting on Design Paradigms of Animated Data Video Tools. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–21. doi:10.1145/3706598.3713449

[82] Leixian Shen, Leni Yang, Haotian Li, Yun Wang, Yuyu Luo, and Huamin Qu. 2025. How Does Empirical Research Facilitate Creation Tool Design? A Data Video Perspective. arXiv:2507.15244 [cs.HC] doi:10.48550/arXiv.2507.15244 Preprint submitted to arXiv.

[83] Leixian Shen, Yizhi Zhang, Haidong Zhang, and Yun Wang. 2023. Data player: Automatic generation of data videos with narration-animation interplay. *IEEE Transactions on Visualization and Computer Graphics* 30, 1 (2023), 109 – 119. doi:10.1109/TVCG.2023.3327197

[84] Yung-Cheng Shen and Ting-Chen Chen. 2006. When East meets West: The effect of cultural tone congruity in ad music and message on consumer ad memory and attitude. *International Journal of Advertising: The Quarterly Review of Marketing Communications* 25, 1 (2006), 51–70. doi:10.1080/02650487.2006.11072951

[85] Mark Shevy and Kineta Hung. 2013. Music in television advertising and other persuasive media. *The Psychology of Music in Multimedia* 315 (2013), 338. doi:10.1093/acprof:oso/9780199608157.003.0014

[86] Yang Shi, Xingyu Lan, Jingwen Li, Zhaorui Li, and Nan Cao. 2021. Communicating with motion: A design space for animated visual narratives in data videos. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–13. doi:10.1145/3411764.3445337

[87] Ki-Ho Shin and In-Kwon Lee. 2017. Music synchronization with video using emotion similarity. In *Proceedings of the IEEE International Conference on Big Data and Smart Computing*. IEEE, Piscataway, NJ, USA, 47–50. doi:10.1109/BIGCOMP.2017.7881714

[88] Hafiva Hanum Siregar and Miza Nina Adlini. 2022. Discovery Learning Assisted by Audiovisual Learning Media: Its Effect on Student Biology Learning Outcomes. *Jurnal Pembelajaran Dan Biologi Nukleus* 8, 2 (2022), 284–292. doi:10.36987/jpbn.v8i2.2885

[89] Stuart Smith. 1991. Data sonification: issues and challenges. In *Final Program and Paper Summaries 1991 IEEE ASSP Workshop on Applications of Signal Processing to Audio and Acoustics*. IEEE, Piscataway, NJ, USA, 0_63–0_65. doi:10.1109/ASPAA.1991.634106

[90] John Sweller. 2011. Cognitive load theory. In *Psychology of Learning and Motivation*. Vol. 55. Elsevier, Amsterdam, The Netherlands, 37–76. doi:10.1016/B978-0-12-387691-1.00002-8

[91] Joey A Tan. 2019. Social Activism in the US Music Industry. *Backstage Pass* 2, 1 (2019), 24.

[92] Pei Pei Tan, Jernih Abdul Rahman, Mat Noh Sarah, Siti Az-Syazni Amirah Mohamad Azlin, Norsazila Zambree, and Sabariah Mohd Noor. 2022. Song for Blood Donor's Recruitment and Promotion during COVID-19 Pandemic. *The Malaysian Journal of Medical Sciences* 29, 6 (2022), 170. doi:10.21315/mjms2022.29.6.17

[93] Siu-Lan Tan, Annabel J. Cohen, Scott D. Lipscomb, and Roger A. Kendall. 2013. *The Psychology of Music in Multimedia*. Oxford University Press, Oxford, UK. doi:10.1093/acprof:oso/9780199608157.001.0001

[94] Junxiu Tang, Lingyun Yu, Tan Tang, Xinhuan Shu, Lu Ying, Yuhua Zhou, Peiran Ren, and Yingcai Wu. 2020. Narrative transitions in data videos. In *Proceedings of the IEEE Visualization Conference*. IEEE, IEEE, Piscataway, NJ, USA, 151–155. doi:10.1109/VIS47514.2020.00037

[95] Tan Tang, Junxiu Tang, Jiewen Lai, Lu Ying, Yingcai Wu, Lingyun Yu, and Peiran Ren. 2022. SmartShots: An optimization approach for generating videos with data visualizations embedded. *ACM Transactions on Interactive Intelligent Systems* 12, 1 (2022), 1–21. doi:10.1145/3484506

[96] John Thompson, Zhicheng Liu, Wilmot Li, and John Stasko. 2020. Understanding the design space and authoring paradigms for animated data graphics. *Computer Graphics Forum* 39, 3 (2020), 207–218.

[97] Takahiko Tsuchiya, Jason Freeman, and Lee W. Lerner. 2015. Data-to-Music API: Real-Time Data-Agnostic Sonification with Musical Structure Models. In *Proceedings of the 21st International Conference on Auditory Display (ICAD 2015)*. Graz, Austria, 244–251.

[98] Visual Capitalist. 2019. The World's Largest 10 Economies in 2030. https://www.youtube.com/watch?v=L5v9jt4__ho. Accessed: 2025-04-09.

[99] Vox. 2018. Why America needs automatic voter registration. https://www.youtube.com/watch?v=Hd5Qs0fc_I0. Accessed: 2025-04-09.

[100] Vox. 2020. Why beef is the worst food for the climate. https://www.youtube.com/watch?v=3lrJYTsKdUM. Accessed: 2025-04-09.

[101] Vox. 2021. Why you can't compare Covid-19 vaccines. https://www.youtube.com/watch?v=K3odScka55A. Accessed: 2025-08-31.

[102] Vox. 2025. Vox Media — The Modern Media Company. https://corp.voxmedia.com. Accessed: 2025-08-31.

[103] Yun Wang, Zhutian Chen, Quan Li, Xiaojuan Ma, Qiong Luo, and Huamin Qu. 2016. Animated narrative visualization for video clickstream data. In *Proceedings*

of the SIGGRAPH Asia Symposium on Visualization. Association for Computing Machinery, New York, NY, USA, 1–8. doi:10.1145/3002151.3002155

[104] Zheng Wei, Yuelu Li, Wenchuan Lu, Qiming Gu, Huamin Qu, and Xian Xu. 2025. At the Peak: Empirical Patterns for Creating Climaxes in Data Videos. *IEEE Transactions on Visualization and Computer Graphics* 31, 10 (2025), 8696–8710. doi:10.1109/TVCG.2025.3576597

[105] Zheng Wei, Huamin Qu, and Xian Xu. 2025. Telling Data Stories with the Hero's Journey: Design Guidance for Creating Data Videos. *IEEE Transactions on Visualization and Computer Graphics* 31, 1 (2025), 962–972. doi:10.1109/TVCG.2024.3456330

[106] Xian Xu, Aoyu Wu, Leni Yang, Zheng Wei, Rong Huang, David Yip, and Huamin Qu. 2023. Is it the end? guidelines for cinematic endings in data videos. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–16. doi:10.1145/3544548.3580701

[107] Xian Xu, Leni Yang, David Yip, Mingming Fan, Zheng Wei, and Huamin Qu. 2022. From 'wow'to 'why': Guidelines for creating the opening of a data video with cinematic styles. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. Association for Computing Machinery, New York, NY, USA, 1–20. doi:10.1145/3491102.3501896

[108] Leni Yang, Aoyu Wu, Wai Tong, Xian Xu, Zheng Wei, and Huamin Qu. 2023. Understanding 3d data videos: From screens to virtual reality. In *Proceedings of the IEEE Pacific Visualization Symposium*. IEEE, Piscataway, NJ, USA, 197–206. doi:10.1109/PacificVis56936.2023.00029

[109] Leni Yang, Xian Xu, XingYu Lan, Ziyan Liu, Shunan Guo, Yang Shi, Huamin Qu, and Nan Cao. 2021. A design space for applying the freytag's pyramid structure to data stories. *IEEE Transactions on Visualization and Computer Graphics* 28, 1 (2021), 922–932. doi:10.1109/TVCG.2021.3114774

[110] Qinghua Yang and Cong Li. 2013. Mozart or metallica, who makes you more attractive? A mediated moderation test of music, gender, personality, and attractiveness in cyberspace. *Computers in Human Behavior* 29, 6 (2013), 2796–2804. doi:10.1016/j.chb.2013.07.026

[111] Ying Yang, Tim Dwyer, Michael Wybrow, Benjamin Lee, Maxime Cordeil, Mark Billinghurst, and Bruce H Thomas. 2022. Towards immersive collaborative sensemaking. *Proceedings of the ACM on Human-Computer Interaction* 6, ISS (2022), 722–746. doi:10.1145/3567741

[112] Zhonggen Yu and Mingle Gao. 2022. Effects of video length on a flipped English classroom. *Sage Open* 12, 1 (2022), 21582440211068474. doi:10.1177/21582440211068474

[113] Chengbo Zheng, Tian Gao, Shunan Guo, Yang Shi, Xiaojuan Ma, and Nan Cao. 2025. A design space of animating data-driven transitions in data videos. *Journal of Visualization* 28 (2025), 1–18. doi:10.1007/s12650-025-01066-5