

Sound Matters: An Experimental Study of Vocal Performance and Economic Behaviour

Research article

Björn A. Kuchinke and Jürgen Rösch*

Bauhaus-Universität Weimar

Received: May 25, 2025; Accepted: August 20, 2025

Abstract: Can subtle differences in vocal performance influence economic decision-making? This question has renewed relevance in a digital landscape shaped by audio-driven persuasion, from short-form videos to synthetic voices in marketing. One of the earliest studies using music as a stimulus in economic games is Oxoby (2009), who employed an ultimatum game to compare AC/DC songs featuring different vocalists. We replicate and extend this design by holding lyrics and instrumentation constant, using the same song performed by both Bon Scott and Brian Johnson. While we find no consistent effect across all treatments, results for one comparison suggest that vocals may influence decision efficiency, as reflected in participants' minimum acceptable offers and final payouts. In addition, exploratory acoustic analysis of the vocal tracks highlights measurable differences in performance features, such as loudness, key and timbre. These findings point to the underexplored role of vocal cues in shaping judgement, and they raise important questions for future research in an era of customisable voices generated by artificial intelligence.

1. Introduction

Can music shape economic decisions? As music becomes increasingly ubiquitous through streaming platforms and social media, understanding its influence on human behaviour is more important than ever. Platforms like TikTok have transformed the role of music; no longer merely a background feature, it now serves as a central cue in communication, persuasion and branding. Many TikTok videos, political messages and advertisements rely heavily on carefully selected musical segments, suggesting that music may subtly yet powerfully shape individual and collective judgements.

Research in marketing and psychology long suggests that music can influence consumer attention, emotion and memory. Acoustic stimuli are processed rapidly and instinctively (Beckerman 2020). Music can evoke nostalgia (Gustafsson, 2015), enhance message recall (Steiner 2018), and even affect brand trust and loyalty (Gustafsson 2015; Jackson 2003). Crucially, the effectiveness of music in influencing behaviour appears to depend not only on its presence but on its fit with the message and context (North & Hargreaves 2008).

However, the role of music in economic decision-making remains underexplored. While some studies find null effects in experimental games, others suggest that specific musical characteristics such as tempo or harmony may influence trust and generosity (e.g. García-Gallego et al. 2019; Greitemeyer 2009; Riedl et al. 2017).

A particularly underexamined dimension in the economics of music is the influence of vocal performance. Unlike instrumental music, vocal music engages listeners through emotional, social and evolutionary pathways. Prior research suggests that humans are especially attuned to vocal cues, with music involving the human voice, even wordless singing, eliciting stronger emotional and social responses than instrumental music alone (Trehub et al. 2015). Vocal music has been linked to social bonding, particularly in cooperation and caregiving contexts. Similarly, prosocial songs have been shown to increase generosity and empathy in experimental settings

*E-mail: juergen.roesch@uni-weimar.de

(García-Gallego et al. 2019; Greitemeyer 2009). The presence of lyrics further enhances music's narrative and affective dimensions, shaping how messages are framed, remembered and emotionally evaluated.

In digital media environments such as TikTok, where short vocal segments are repeatedly looped or remixed, voice's familiarity and human resonance may subtly influence how content is perceived and trusted. Rösch and Rauch (2025) provide direct evidence of this phenomenon in a decision-making context. Using Spotify data and emotional arc analysis of lyrics, they demonstrate that both music and lyrical structure have a measurable impact on voting outcomes in a national song contest.

Beyond lyrics and narrative, the expressive characteristics of vocal delivery, such as loudness, pitch variability, harmonic richness and timbral brightness, also play a critical role in how music is perceived and responded to. Banse and Scherer (1996) showed that specific acoustic cues in the human voice reliably convey emotional states like anger or sadness across cultures. Juslin and Laukka (2003) argue that both speech and singing draw on a shared 'emotional code', making vocal tone as psychologically impactful as verbal content. This implies that vocal performance alone may influence listener perception and behaviour.

Understanding how voice influences human decision-making is a growing concern in economics, marketing and digital communication. As society becomes increasingly shaped by video- and audio-based platforms, ranging from voice assistants to podcast ads to AI-generated audio, it is an important research question of how vocal delivery affects behaviour. Social media platforms like TikTok rely heavily on audio, not only for dance challenges but also for background music, potentially influencing how the content is perceived. Recent research indicates that even subtle acoustic features, such as loudness modulation and vocal presence, can significantly influence consumer responses to advertising (Barnes & Wang 2025). But can these same vocal elements influence economic decisions? This paper explores that question using a stylised economic experiment involving different performances of the same rock song.

In recent literature, there has been a noticeable increase in attention to the behavioural aspects of music. While it has not been explicitly stated, the surge of audio-centric platforms like TikTok might fuel this interest in the potential influence of music and sound on human behaviour. A systematic literature review by Anglada-Tort et al. (2023) has led to the proposition of a research programme termed the Behavioural Economics of Music. One of the earliest empirical attempts to test the influence of music on behaviour is Oxoby (2009), who used an ultimatum game to test preferences for AC/DC's vocalists. His design implicitly suggests that different vocal performances could alter economic decisions. However, Oxoby (2009) did not isolate whether it was the musical content, vocal delivery or participant biases that drove his result.

Our study addresses this gap. We reconstruct Oxoby's experiment and extend it by keeping the musical composition constant while varying only the singer. This allows us to assess the influence of vocals on decision-making directly. Our failure to replicate the original effect raises questions about the robustness of prior claims and offers new insight into how (and when) vocal music may affect economic behaviour. These findings are especially relevant given recent advances in AI-generated music, where both vocals and instrumentals can be crafted and manipulated to shape audience responses.

The paper proceeds as follows: First, we review relevant literature on music, emotion and economic behaviour, with particular attention to how vocal and acoustic features can shape perception. This is followed by a description of our experimental design, which isolates the effect of vocal performance by holding lyrics and instrumentation constant. We then present the main empirical results from the ultimatum game. An exploratory acoustic analysis of the vocal tracks provides further insight into measurable differences. The discussion addresses the broader implications of our findings, especially in light of emerging applications of AI-generated audio. The paper concludes with reflections on the role of voice in decision-making and potential directions for future research.

2. From Musical Features to Vocal Performance in Economic Contexts

Behavioural effects of music have been studied in experimental economic settings, particularly in games designed to elicit generosity, fairness and trust. Studies using the dictator game show that classical or relaxing music can increase generosity (Fukui & Toyoshima 2014; García-Gallego et al. 2019), while prosocial lyrics have been found to foster more altruistic decisions (Greitemeyer 2009). However, results from ultimatum and trust games are more mixed (Chung et al. 2016; Riedl et al. 2017), and much of the existing literature presents complete musical pieces as

undifferentiated stimuli. This raises questions about which musical elements— lyrics, melody, or vocal expression— drive observed behavioural effects.

Much of the literature to date has examined complete musical pieces, often combining multiple auditory elements, melody, instrumentation and vocals, into a single treatment. This makes it difficult to disentangle the specific influence of vocal performance. Prior research suggests that the human voice carries unique emotional and social salience, even when divorced from lyrical content (Trehub et al. 2015). Building on this, our study seeks to isolate vocal delivery as a behavioural stimulus. By adapting Oxoby's (2009) ultimatum game paradigm, we hold the musical composition and lyrics constant, varying only the singer. This design allows for a more targeted test of whether voice alone can influence fairness judgements and responder behaviour.

Research also discusses the role of contextual fit in music perception, showing that emotional alignment between music and surrounding content can influence engagement and evaluative judgements (Herget & Albrecht 2022; Moormann 2010). Yet little is known about whether congruence at the vocal level alone, such as variation in tone, delivery or timbre, can elicit similar effects in strategic decision-making contexts.

Our study builds on these insights by examining whether vocal performance alone, with all other musical elements held constant, can shape economic decisions in an ultimatum game. This design uniquely isolates the influence of the human voice as a decision-making cue.

2.1. Background: Oxoby (2009), Spotify Data and Emotional Arcs

Over four decades have passed since the death of the iconic AC/DC frontman, Bon Scott. Despite Scott's passing away, the band endured and thrived, with Brian Johnson taking over as lead singer. Since then, an ongoing discourse among fans and scholars has revolved around the question of which of the two is the better singer: Bon Scott or Brian Johnson. Oxoby (2009) made a novel contribution to this debate by using an ultimatum game in which two distinct groups of participants listened to either 'It's a Long Way to the Top' (Bon Scott) or 'Shoot to Thrill' (Brian Johnson) before making decisions.

While the study was presented in a tongue-in-cheek manner, it nonetheless suggested that different vocal performances could affect responder behaviour. Though Oxoby did not investigate the mechanism behind this effect, the design implicitly pointed to a more serious question: can the human voice influence economic decisions?

To assess the comparability of these two songs, we rely on two analytical tools: (1) Spotify audio features and (2) sentiment-based emotional arcs of lyrics. Spotify provides structured audio data, including acousticness, danceability, energy, instrumentalness, liveness, speechiness, valence, loudness and tempo. Prior work using these metrics has yielded mixed results regarding their role in music success (Al-Beitawi et al. 2020; Sciandra & Spera 2022), but together they provide a standardised, objective measure of musical structure.

Emotional arcs, first applied to books (Reagan et al. 2016), later to films (Del Vecchio et al. 2021) and music (Rösch & Rauch 2025), are introduced here to analyse lyrical content in music. However, new in this context, emotional arcs may be informative when decisions are affected by lyrical framing and emotional tone. For instance, sad or emotionally intense music is often paradoxically enjoyed more (Pannese et al. 2016).

As shown in Figure 1, both 'Shoot to Thrill' and 'It's a Long Way to the Top' are similar in their Spotify-derived audio features. However, their emotional arcs, as derived from sentiment analysis of lyrics, show considerable differences, suggesting that affective trajectory, not just vocal quality, may explain Oxoby's original finding.

To rigorously isolate the effect of vocal performance, we build on Oxoby's framework by using a single song, 'High Voltage', performed by both singers. These versions are musically and lyrically identical but differ in vocal delivery. This choice enables a cleaner test of the influence of voice, without confounds introduced by varying melody, lyrics or arrangement (see Table 1).

2.2. Vocal Performance, Superstardom and the Behavioural Economics of Music

Although Oxoby's (2009) design is often cited for its novelty, the idea that subtle vocal variations might affect behaviour finds broader support. Research on music competitions shows that subjective evaluations frequently outweigh objective performance metrics (Budzinski et al. 2021). Theoretical work on superstardom argues that small differences in talent or charisma can lead to outsized success in low-substitution markets (Borghans & Groot 1998; Rosen 1981).

However, evaluating vocalists from a band like AC/DC presents a challenge: most variables beyond the singer are held constant. Stream counts may reflect popularity but not isolate vocal effects. Our experiment offers a cleaner test by using the same song, with only the singer changed, thereby directly targeting vocal influence on responder decisions.

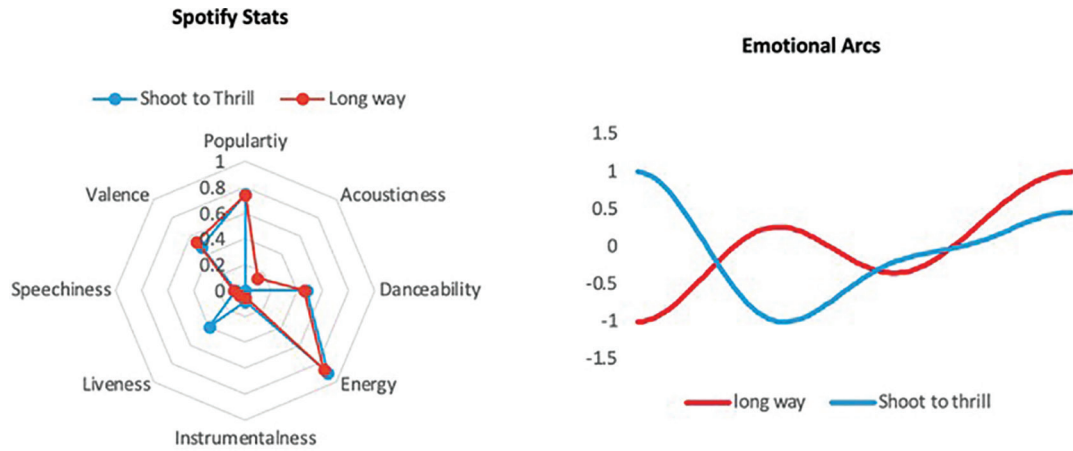


Figure 1. Comparison of shoot to thrill vs. long way to the top based on spotify stats and emotional arcs.

Treatment 1	Treatment 2	Treatment 3	Treatment 4
Song: High voltage	Song: High voltage	Song: It's a long way [...]	Song: Shoot to thrill
Singer: Bon Scott	Singer: Brian Johnson	Singer: Bon Scott	Singer: Brian Johnson
Year: 1976	Year: 1991	Year: 1976	Year: 1980
Album: If you want blood you've got it	Album: Live	Album: High voltage	Album: Back in black
Popularity: 0.35	Popularity: 0.38	Popularity: 0.74	Popularity: 0.75
Acoustics: 0.00	Acoustics: 0.00	Acoustics: 0.133	Acoustics: 0
Danceability: 0.55	Danceability: 0.58	Danceability: 0.456	Danceability: 0.475
Energy: 0.919	Energy: 0.944	Energy: 0.863	Energy: 0.904
Instrumentalness: 0.602	Instrumentalness: 0.777	Instrumentalness: 0.054	Instrumentalness: 0.088
Liveness: 0.935	Liveness: 0.952	Liveness: 0.055	Liveness: 0.396
Speechiness: 0.06	Speechiness: 0.046	Speechiness: 0.089	Speechiness: 0.075
Valence: 0.564	Valence: 0.546	Valence: 0.532	Valence: 0.48

Table 1. Overview of treatments and song characteristics.

This perspective aligns with the emerging field of the Behavioural Economics of Music (Anglada-Tort et al. 2023), which explores how heuristics and biases shape music perception. For example, the peak-end rule implies that musical evaluations hinge on the emotional climax and ending rather than the entire experience (Schäfer et al. 2014). Other heuristics, such as affect and availability, may also play roles in how listeners evaluate singers, songs or decisions.

Our contribution is to extend this behavioural lens to a controlled economic setting. By experimentally varying only the vocal element, we offer a novel test of whether vocal expression, a powerful but often underappreciated stimulus, can shift economic behaviour.

3. Experimental Design

Aligned with Oxoby's (2009) methodology, we employ the ultimatum game as a tool to examine decision-making in the context of different vocalists. Widely adopted in numerous studies across diverse settings, the ultimatum game's simplicity facilitates participant comprehension and adherence to instructions. Two participants are randomly paired in this game and designated proposers or responders. In the proposer role, an individual is endowed with \$10, which they may distribute between themselves and the responder. Subsequently, the responder becomes aware of the proposed allocation and decides to either accept, leading to both participants receiving rewards based on the

proposed distribution, or reject, resulting in no rewards for either party. This game structure provides a straightforward yet effective means of exploring decision dynamics.

Oxoby (2009) justified the use of the ultimatum game based on the game-theoretic expectation that proposers should offer the minimum amount, while responders should accept any positive offer. In this theoretical framework, responders benefit from any offer above zero, making the outcome efficient without any money lost due to rejected offers. However, empirical observations deviate from this theoretical prediction. In practice, offers below 30% are often declined, and proposers typically select amounts ranging from 20–50% of the initial endowment (Camerer 2011; Oxoby 2009). Consequently, the ultimatum game is frequently used to explore themes, such as fairness, inequity aversion, cooperation and reciprocity, as evidenced in studies by (Bolton & Ockenfels 2000; Charness & Rabin 2002; Fehr & Schmidt 1999; Thielmann et al. 2021a).

Our approach closely adheres to the methodology outlined by Oxoby (2009). Like the referenced study, the participants were randomly paired, and the study's structure and procedures were communicated to them. The proposer, endowed with \$10, could allocate any value from zero to ten to share with the responder. Crucially, the participants had to pre-determine two decisions before learning their assigned role: first, the offer they would extend as proposers, and second, the minimum amount they would accept as responders. The participants were informed of their assigned roles and the game's outcome after this decision-making phase. If the proposed amount exceeded or equalled the minimum acceptance threshold, both players received payouts based on their choices.

The experimental setup involved different AC/DC songs as treatments. All participants listened to the songs before the game began and again during the decision-making process. The game instructions were based on standard formulations (Thielmann et al. 2021b) and slightly adapted to fit the experimental context.

However, our study departs significantly in several critical aspects. Most notably, we introduced two additional treatments, using 'High Voltage' by both Bon Scott and Brian Johnson, in addition to 'Shoot to Thrill' and 'It's a Long Way'. This expansion enables both a replication of Oxoby's (2009) experiment and a cleaner analysis of the distinction between the two vocalists in a controlled musical setting. 'High Voltage' was selected because it is not among the most popular AC/DC tracks, unlike 'Highway to Hell' or 'You Shook Me All Night Long', which helps mitigate potential biases due to familiarity or emotional salience (see Table 1).¹

The experiment was conducted online via oTree (Chen et al. 2016) and Prolific. To ensure engagement with the musical stimuli, participants were introduced to the procedure in advance and instructed to listen to each song for at least 1 min from a predefined start point, emphasising vocal sections. This approach ensured equal exposure length and avoided biases stemming from differing song structures or track durations. We did not normalise the audio or alter the recordings, preserving the original live characteristics, such as vocal tone, crowd interaction and recording dynamics. While this introduces some variability in acoustic quality, we consider it essential to capture the authentic vocal differences between the performers, which are central to the treatment. After this listening phase, participants received the ultimatum game instructions and listened to the song again while making their two decisions. A brief questionnaire followed, probing knowledge of the genre, band and singer. Participants who identified familiarity were asked to provide their Prolific ID. Role assignments and final payouts were disclosed only at the end. Each participant received a fixed €1 compensation, aligned with standard practice for Prolific-based studies.

4. Results

The study involved 184 participants recruited via Prolific. No restrictions were placed on participant characteristics except for the requirement that they had access to audio playback. The average age was 31.23 years (SD = 7.26), with 47% identifying as female and 53% as male. In terms of ethnicity, 2% identified as Asian, 56% as Black, 4% as Mixed, 2% as Other and 35% as White. Language diversity was intentionally maintained, and the sample included participants beyond the English-speaking demographic; however, 64% of participants reported English as their first language. This aligns with our objective of analysing the broad impact of different singers. Each treatment involved playing the song for 1 min before the participants entered the study and again for a minimum of 1 min during the decision-making process. To maintain anonymity, Prolific's platform was employed, preventing the participants from

¹ We use live versions because Brian Johnson performs Bon Scott songs live, but has not re-recorded them in the studio; hence, comparable studio versions are not available.

identifying or communicating with each other. This approach adheres to ethical standards, safeguarding the privacy and confidentiality of participants throughout the experiment.

Consistent with Oxoby (2009), we present our results following a similar structure, beginning with the original treatments ('Shoot to Thrill' and 'It's a Long Way') and subsequently addressing the treatment involving the same song, 'High Voltage'. Similar to Oxoby's (2009) methodology, we compare the minimum acceptable offer (MAO) between treatments, representing the lowest offer a participant would accept. An efficient outcome is characterised by a low MAO in its assessment, implying a greater likelihood of offer acceptance and, therefore, more 'deals' being established. Conversely, an inefficient outcome is denoted by offer rejections. This analysis allows us to evaluate the impact of the different treatments on participants' willingness to accept offers and, therefore, on the comparative effectiveness of the singer's influence on decision-making.

In comparing 'High Voltage' performed by Bon Scott and Brian Johnson, Treatment 1 (T1) and Treatment 2 (T2), respectively (see Table 2), we observe minimal differences in both the mean amount offered and the closely aligned standard deviations. Consequently, the non-parametric Wilcoxon rank-sum test does not indicate a statistically significant difference. In contrast to Oxoby (2009), the analysis of Treatment 3 (T3) and Treatment 4 (T4), 'It's a Long Way [...]' performed by Bon Scott and 'Shoot to Thrill' performed by Brian Johnson, respectively, shows almost identical results. Despite variations in the mean amount offered and standard deviation, the Wilcoxon rank-sum test indicates no statistically significant distinction between these treatments regarding the offers. The graphical representations in Figure 2 further illustrate these findings, emphasising that although observable distinctions exist, they are not significantly different from each other.

Regarding MAOs, the treatment involving Bon Scott's version of 'High Voltage' exhibits the lowest mean, as indicated in Table 3. This suggests that fewer offers would be rejected when the participants listen to Bon Scott. This finding contradicts Oxoby's (2009) original results, in which more offers were expected to be rejected in the Bon Scott treatments. Notably, the difference in mean MAOs for the Bon Scott 'High Voltage' treatment is statistically significant, with a p -value of 0.0091, determined by the non-parametric Wilcoxon rank-sum test.

For the original songs 'It's a Long Way [...]' (T3) and 'Shoot to Thrill' (T4), we observe a lower mean acceptance amount for Brian Johnson, aligning with Oxoby's (2009) initial findings. However, this difference is not statistically significantly different from the song performed by Bon Scott, according to the Wilcoxon rank-sum test. Figure 3 visually represents this outcome, illustrating the nuanced distinctions in the MAOs between the two singers.

Examining the number of pairs resulting in a payout for participants, when the offered amount equalled or exceeded the minimum acceptable amount, reveals a distinct pattern (refer to Table 4). Notably, the payout for T2 (Brian Johnson) is lower than the payout for T1 (Bon Scott). This difference is statistically significant, as measured by the non-parametric rank-sum test. Therefore, we reject the null hypothesis that the distributions of T1 and T2 are equal.

Conversely, T3 and T4 do not exhibit statistically significant differences based on the Wilcoxon rank-sum test. However, in this case, the Bon Scott treatment (T3) achieves a lower mean compared to the Brian Johnson treatment (T4). In essence, in T1 and T2, Bon Scott leads to a more efficient outcome, while in T3 and T4, Brian Johnson, at least in absolute terms, yields a higher outcome. This nuanced distinction sheds light on the differential impact of the two singers on decision-making outcomes in varying treatments.

Combining T1 and T3, as well as T2 and T4, enables a comprehensive comparison of the effects of Bon Scott and Brian Johnson across all four songs and treatments. Table 5 presents the summary statistics, along with the results of the Wilcoxon rank-sum test. The means are generally rather close in most cases. However, the non-parametric Wilcoxon rank-sum test indicates a potential statistical difference for the MAO. In this context, listeners of Bon Scott, on average, demanded a lower offer from the proposer, implying a more efficient choice conducive to closing more 'deals'.

Offers							
Treatment		#	Mean	Std. dev.	Min	Max	Rank-sum
High voltage	T1 Bon Scott 'High Voltage'	49	5.694	2.2565	1	10	$p = 0.9776$
	T2 Brian Johnson 'High Voltage'	50	5.700	2.1689	2	10	
Replica of Oxoby (2009)	T3 Bon Scott 'It's a Long Way [...]'	45	6.244	2.3660	3	10	$p = 0.2760$
	T4 Brian Johnson 'Shoot to Thrill'	40	5.425	1.9857	0	10	

Table 2. Summary statistics and Wilcoxon rank-sum tests for offers.

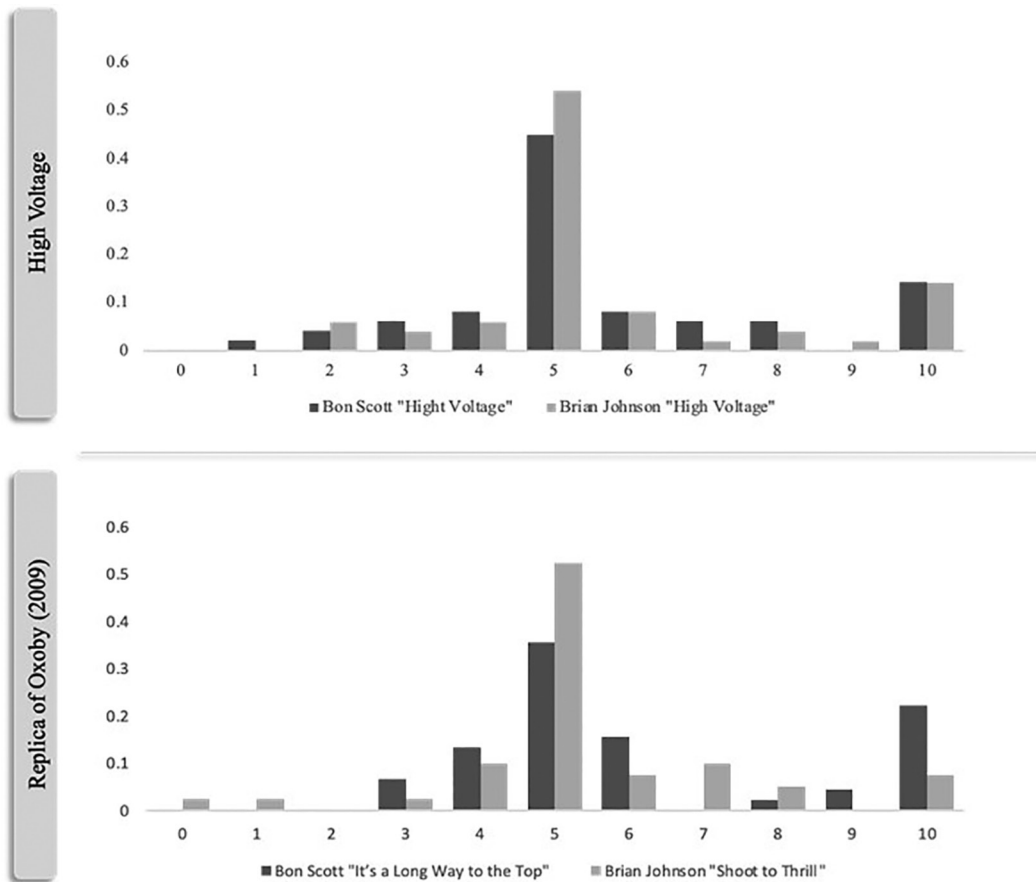


Figure 2. The distributions of offers in the 'High Voltage' treatments (upper part) and the replica of Oxoby's (2009) (lower part).

MAO						
Treatment	#	Mean	Std. dev.	Min	Max	Rank-sum
T1 Bon Scott 'High Voltage'	49	4.0408	1.5937	1	7	$p = 0.0091$
T2 Brian Johnson 'High Voltage'	50	5.3200	2.4532	1	10	
T3 Bon Scott 'It's a Long Way'	45	5.7333	2.5172	0	10	$p = 0.9655$
T4 Brian Johnson 'Shoot to Thrill'	40	5.5500	1.7090	2	10	

MAO, minimum acceptable offer.

Table 3. Summary statistics and Wilcoxon rank-sum tests for MAO

Interestingly, this difference in MAO does not translate into a statistically significant variance in the final payout. The distinction in this aspect remains statistically non-significant at an acceptable level. This aggregated analysis underscores that while listeners of Bon Scott may exhibit a preference for lower offers (a potentially more efficient choice), this preference does not significantly impact the ultimate financial outcome.

4.1. Direct comparison with Oxoby (2009)

In the original study, the sample consisted of 36 participants, divided evenly across two groups (specifically, treatments T3 and T4 in our analysis). To evaluate the replicability of Oxoby's (2009) findings within our dataset, we simulated the original experiment with 18 observations per treatment. For this analysis, we randomly selected 18 participants for each treatment and replicated the selection process 100 times. Subsequently, we assessed the differences in the songs using the same methodology as before. However, the scope of this analysis was limited to

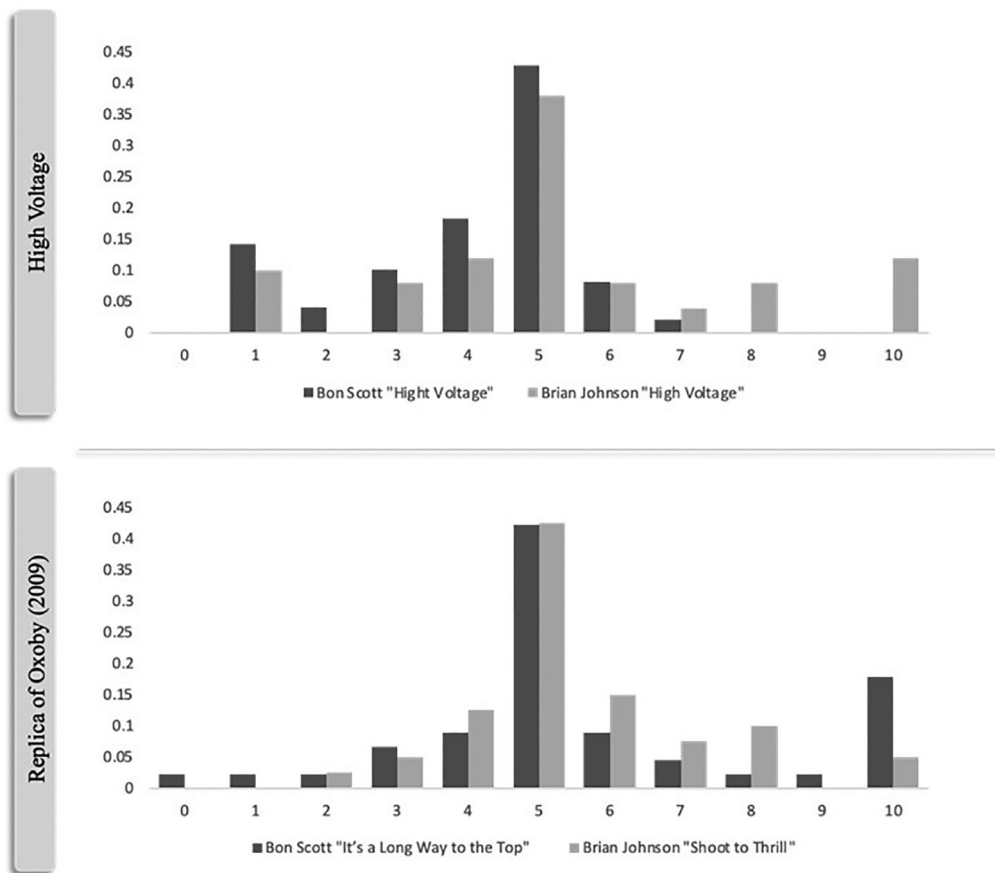


Figure 3. The distributions of MAOs in the Bon Scott and Brian Johnson treatments for 'High Voltage' (upper part) and the replica of Oxoby (2009) lower part. MAOs, minimum acceptable offer.

Final payout						
Treatment	#	Mean	Std. dev.	Min	Max	Rank-sum
T1 Bon Scott 'High Voltage'	49	4.1837	2.690055	0	10	$p = 0.0014$
T2 Brian Johnson 'High Voltage'	50	2.4000	3.238795	0	10	
T3 Bon Scott 'It's a Long Way'	45	2.4444	3.166268	0	10	$p = 0.1860$
T4 Brian Johnson 'Shoot to Thrill'	40	3.3250	3.276861	0	10	

Table 4. Summary statistics and Wilcoxon rank-sum tests for final payouts.

the sent amounts and MAOs, as each participant independently made these choices prior to being matched with a partner and before learning of the partner's decisions.

Table 6 presents the outcomes of this analysis. Significantly, for the MAO associated with the 'High Voltage' treatment, we observed a greater number (26) of significant outcomes, consistently showing a lower MAO for the Bon Scott version of 'High Voltage'. Following Oxoby's (2009) rationale, this suggests that Bon Scott's version leads to more efficient outcomes, making him the superior vocalist in this context. In the case of offers, we found only one significant difference, similar to the MAO for treatments T3 and T4. In the comparative analysis of offers for T3 vs. T4, the Wilcoxon Ranksum test indicated higher offerings for the Brian Johnson song in 8 out of 100 instances.

Observations		Bon Scott	Brian Johnson	Rank-sum
		94	90	
Offers	Mean	5.9574	5.5778	$p = 0.4209$
	Std. Dev	2.3137	2.0824	
	Min.	1	0	
	Max.	10	10	
MAO	Mean	4.8511	5.4222	$p = 0.0503$
	Std. Dev	2.2431	2.1462	
	Min.	0	1	
	Max.	10	10	
Payout	Mean	3.3511	2.8111	$p = 0.1660$
	Std. Dev	3.0399	3.2702	
	Min.	0	0	
	Max.	10	10	

MAO, minimum acceptable offer

Table 5. Summary statistics and Wilcoxon rank-sum tests for all Bon Scott vs. Brian Johnson treatments.

Comparison	Measure	# of Significant Results ($p < 0.05$)	Direction of Difference	Singer associated with greater Economic efficiency (following Oxoby 2009)
T1 vs. T2 'High Voltage'	Offers	1	Bon Scott < Brian Johnson	Brian Johnson
	MAO	26	Bon Scott < Brian Johnson	Bon Scott
T3 vs. T4 Original songs	Offers	8	Bon Scott > Brian Johnson	Bon Scott
	MAO	1	Bon Scott > Brian Johnson	Brian Johnson

MAO, minimum acceptable offer.

Table 6. Results of randomly chosen 18 participants in each treatment with 100 repetitions—differences were tested with the Wilcoxon rank-sum test².

High voltage performed by	Loudness	Energy	Tempo	Key	Harmony	Flatness	Spectral bandwidth	Speech rate proxy
Brian Johnson	0.026608	0.204217	143.55	3	0.27718	0.07267	2,359.38	4,502.99
Bon Scott	0.04949	0.204189	136	5	0.274313	0.079435	2,276.34	4,502.38

Table 7. Acoustic feature comparison of vocal performances by Brian Johnson and Bon Scott on High Voltage, extracted using Demucs for source separation and Librosa for signal analysis.

Several methodological factors may explain these differences. Unlike Oxoby's original study, which involved only Canadian students in a controlled laboratory setting, our experiment was conducted online with a globally recruited sample via Prolific. This introduces variability in cultural background, musical familiarity and listening context, as well as in how participants perceive and interpret vocal performance. While we ensured that each participant listened to a 1-min excerpt beginning at the main vocal line, we could not standardise playback devices or environments. These real-world differences, combined with a broader sample composition, likely increased heterogeneity in responses.

² Notes on Table 6: MAO = Minimum Acceptable Offer. Each result is based on 100 random draws of 18 participants per treatment group. Lower MAO values indicate greater economic efficiency (Oxoby, 2009).

Furthermore, standard economic games like the ultimatum game can yield widely varying results depending on participants' socio-cultural backgrounds (Henrich 2000; Henrich et al. 2001),³ which might be especially true when music is involved. This broader heterogeneity, with a limitation in terms of direct comparability, also increases the ecological validity of our findings by reflecting the diversity of real-world, digital music experiences.

4.2. The potential influence of vocal features

In addition to the original experiment, we analysed the acoustic properties of isolated singing voices by separating the vocal track (vocals.wav) using the *htdemucs* model in the Demucs source separation framework. Audio features were extracted using Python libraries such as *librosa* and *soundfile*, focussing exclusively on the vocal stem. The extracted features capture well-established perceptual dimensions: loudness (root-mean-square amplitude), which reflects vocal intensity and emotional arousal (Banse & Scherer 1996); energy (zero-crossing rate), a marker of vocal sharpness and articulatory dynamics (Peeters et al. 2011); tempo, reflecting vocal rhythmic pacing; and key (dominant chroma bin), which approximates tonal centre (Juslin & Laukka 2003). We also computed harmonic content (chroma richness), spectral flatness (distinguishing tonal vs. noisy timbres), spectral bandwidth (related to vocal brightness) and a speech rate proxy based on zero-crossing rate (ZCR) and sampling rate (Peeters et al. 2011). Together, these features provide a multi-dimensional acoustic profile of vocal performance, grounded in prior work on affective voice science and music cognition.

The version performed by Bon Scott exhibited greater loudness (0.049 vs. 0.026), slightly lower spectral flatness (0.079 vs. 0.073) and a higher key centre (F vs. Eb), suggesting a more assertive and brighter vocal performance. Interestingly, Brian Johnson's version showed a slightly higher spectral bandwidth (2359 Hz vs. 2276 Hz), indicating a more open or resonant timbral quality. Both versions had near-identical speech rate proxies and harmonic structure, suggesting comparable vocal articulation and overall musical arrangement. While these acoustic differences cannot be directly linked to our behavioural results, they demonstrate that the singers' voices differ in measurable ways. This analysis points towards future research opportunities in which specific verbal or non-verbal vocal features could be systematically manipulated as stimuli.

5. Discussion

Although Oxoby's (2009) original study was presented with a degree of irony and sparked debate about the breadth of economic experiments, it inadvertently opened the door to a more serious line of inquiry. Specifically, it raised the question of whether music, and more precisely, vocal performance, can be used as a strategic tool to influence behaviour.

This question is especially relevant in today's digital media environments, where speech and background music are routinely combined in humour, advertising and political messaging content. Platforms like TikTok, Instagram and YouTube often rely on audio-visual cues to shape users' emotional responses, perceptions of credibility and willingness to engage. In such settings, the integration of voice and sound is not merely aesthetic; it forms the foundation of persuasive communication strategies employed by influencers, political actors and brands alike.

Previous research has shown that both musical and lyrical elements can shape decision-making. For example, lyrics and music each influenced voting behaviour in national song contests (Rösch & Rauch 2025), while songs with positive emotional valence were found to increase prosocial behaviour in economic games (Greitemeyer 2009). However, the specific role of vocals has not yet been systematically investigated. Our results, though based on a limited sample, suggest that vocal delivery may play a meaningful role in shaping behaviour. Importantly, this study replicates one of the earliest investigations in this space (Oxoby 2009), even though vocals were not explicitly considered in the original analysis. Future research should extend this line of inquiry by isolating and testing individual vocal characteristics, just as recent work has done with musical structure.

The relevance of vocal influence extends beyond academic interest. Advances in generative AI now allow for the design not only of musical features but also of voice characteristics. AI-generated voices, as demonstrated by tools such as Notebook LM, can convincingly simulate natural conversational flow, for instance, in podcast formats. This

³ However, because our participants were recruited via Prolific, a widely used online research platform, we expect that they are somewhat more culturally aligned than the highly diverse, cross-cultural samples in studies like Henrich (2000) and Henrich et al. (2001).

opens up the possibility of designing synthetic voices that are not only pleasant but are also intentionally shaped to sound more trustworthy, persuasive or emotionally resonant. Such capabilities present intriguing opportunities for marketers and content creators. Still, they also raise normative questions for policymakers, particularly regarding persuasion and manipulation in digital environments, even if the effects are subtle.

This study, however, comes with important limitations. Our focus on replicating Oxoby's (2009) original experiment necessarily centres on a narrow genre (hard rock) and two specific AC/DC songs. These may elicit different responses across demographic and cultural groups, potentially limiting the generalisability of the findings. Moreover, the perception of AC/DC may vary substantially across countries, influenced by cultural context and personal familiarity. While we held the song itself constant across vocalists, musical perception may still be shaped by production quality, performance nuances and recording conditions. In this sense, music and voice are not fully separable: they co-exist as part of a unified auditory experience.

Beyond genre-specific and cultural considerations, several further limitations should be acknowledged. First, the use of an online sample may limit generalisability. Participants recruited via Prolific, while being demographically diverse, may not fully represent the broader population, particularly in terms of age, cultural background or musical familiarity. Second, the experiment did not include a direct manipulation check to assess how participants perceived the vocal performances. Without such data, it remains unclear to what extent listeners consciously distinguished between vocal styles or attributes, such as warmth, clarity or emotionality. Finally, while using real musical recordings enhances ecological validity, it limits experimental control. Vocal renditions may differ in voice quality and subtle production or mixing choices, which could unintentionally influence outcomes. Future research might leverage synthesised or AI-generated voices to isolate specific vocal parameters with greater precision.

6. Conclusion

This study set out to revisit and refine one of the earliest experiments investigating the influence of music on economic decision-making. By replicating Oxoby's (2009) design with updated methodology and a cleaner identification strategy, we isolated vocal performance as a treatment variable, holding lyrics and instrumentation constant. Across two different experimental comparisons, our results offer mixed but suggestive evidence: while vocal variation did not consistently alter offers or outcomes, the MAOs and payout patterns in the 'High Voltage' treatments indicate that vocals may influence decision efficiency under specific conditions.

These findings raise broader questions about how subtle acoustic features, particularly voice, may affect economic behaviour. As audio content becomes more central to digital media environments, and as AI-generated voices become increasingly customisable, understanding the psychological and strategic role of vocal cues is both timely and important. Our study makes a small but novel contribution towards this goal, providing a framework for future research to explore how designed or manipulated voices may influence trust, fairness and cooperation in more complex and realistic settings.

References

- Al-Beitawi, Z., Salehan, M., & Zhang, S. (2020). What makes a song trend? Cluster analysis of musical attributes for Spotify top trending songs. *Journal of Marketing Development and Competitiveness*, 14(3), 79-91.
- Anglada-Tort, M., Masters, N., Steffens, J., North, A., & Müllensiefen, D. (2023). The Behavioural Economics of Music: Systematic review and future directions. *Quarterly Journal of Experimental Psychology*, 76(5), 1177-1194. DOI: 10.1177/17470218221113761
- Banse, R., & Scherer, K. R. (1996). Acoustic profiles in vocal emotion expression. *Journal of personality and social psychology*, 70(3), 614. DOI: 10.1037//0022-3514.70.3.614
- Barnes, S. J., & Wang, W. (2025). "I like the sound of that": understanding the effectiveness of audio in ads. *Internet Research*, 35(3), 1104-1125.
- Beckerman, J. (2020). Scoring the World: A Systems-Thinking Approach to Sonic Branding and Design. *Music and the Moving Image*, 13(1), 3-20.
- Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *American economic review*, 91(1), 166-193.

- Borghans, L., & Groot, L. (1998). Superstardom and monopolistic power: Why media stars earn more than their marginal contribution to welfare. *Journal of Institutional and Theoretical Economics (JITE)/Zeitschrift für die gesamte Staatswissenschaft*, 546-571.
- Budzinski, O., Kohlschreiber, M., Kuchinke, B. A., & Pannicke, J. (2021). Does music quality matter for audience voters in a music contest? *Creative Industries Journal*, 14(2), 109-129.
- Camerer, C. F. (2011). Behavioral game theory: Experiments in strategic interaction. Princeton university press.
- Charness, G., & Rabin, M. (2002). Understanding social preferences with simple tests. *The Quarterly Journal of Economics*, 117(3), 817-869.
- Chen, D. L., Schonger, M., & Wickens, C. (2016). oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, 9, 88-97.
- Chung, H., Lee, E. J., Jung, Y. J., & Kim, S. H. (2016). Music-induced mood biases decision strategies during the ultimatum game. *Frontiers in Psychology*, 7, 453. DOI: 10.3389/fpsyg.2016.00453
- Del Vecchio, M., Kharlamov, A., Parry, G., & Pogrebna, G. (2021). Improving productivity in Hollywood with data science: Using emotional arcs of movies to drive product and service innovation in entertainment industries. *Journal of the Operational Research Society*, 72(5), 1110-1137.
- Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817-868.
- Fukui, H., & Toyoshima, K. (2014). Chill-inducing music enhances altruism in humans. *Frontiers in Psychology*, 5, 1215. DOI: 10.3389/fpsyg.2014.01215
- García-Gallego, A., Georgantzis, N., & Ruiz-Martos, M. J. (2019). The heaven dictator game: costless taking or giving. *Journal of Behavioral and Experimental Economics*, 82, 101449.
- Greitemeyer, T. (2009). Effects of songs with prosocial lyrics on prosocial behavior: Further evidence and a mediating mechanism. *Personality and social psychology bulletin*, 35(11), 1500-1511.
- Gustafsson, C. (2015). Sonic branding: A consumer-oriented literature review. *Journal of brand management*, 22, 20-37.
- Henrich, J. (2000). Does culture matter in economic behavior? Ultimatum game bargaining among the Machiguenga of the Peruvian Amazon. *American economic review*, 90(4), 973-979.
- Henrich, J., Boyd, R., Bowles, S., Camerer, C., Fehr, E., Gintis, H., & McElreath, R. (2001). In search of homo economicus: behavioral experiments in 15 small-scale societies. *American economic review*, 91(2), 73-78.
- Herget, A.-K., & Albrecht, J. (2022). Soundtrack for reality? How to use music effectively in non-fictional media formats. *Psychology of Music*, 50(2), 508-529.
- Jackson, D. M. (2003). Sonic branding. *Palgrave Macmillan Books*.
- Juslin, P. N., & Laukka, P. (2003). Communication of emotions in vocal expression and music performance: Different channels, same code? *Psychological bulletin*, 129(5), 770. DOI: 10.1037/0033-2909.129.5.770
- Moormann, P. (2010). Subtile Manipulation? Zur Musikgestaltung von Politikmagazinen. *Musik im Fernsehen: Sendeformen und Gestaltungsprinzipien*, 83-90.
- North, A., & Hargreaves, D. (2008). *The social and applied psychology of music*. OUP Oxford.
- Oxoby, R. J. (2009). On the Efficiency of AC/DC: Bon Scott versus Brian Johnson. *Economic Inquiry*, 47(3), 598-602.
- Pannese, A., Rappaz, M.-A., & Grandjean, D. (2016). Metaphor and music emotion: Ancient views and future directions. *Consciousness and Cognition*, 44, 61-71. DOI: 10.1016/j.concog.2016.06.015
- Peeters, G., Giordano, B. L., Susini, P., Misdariis, N., & McAdams, S. (2011). The timbre toolbox: Extracting audio descriptors from musical signals. *The Journal of the Acoustical Society of America*, 130(5), 2902-2916. DOI: 10.1121/1.3642604
- Reagan, A. J., Mitchell, L., Kiley, D., Danforth, C. M., & Dodds, P. S. (2016). The emotional arcs of stories are dominated by six basic shapes. *EPJ Data Science*, 5(1), 1-12.
- Riedl, R., Javor, A., Gefen, D., Felten, A., & Reuter, M. (2017). Oxytocin, trust, and trustworthiness: The moderating role of music. *Journal of Neuroscience, Psychology, and Economics*, 10(1), 1.
- Rösch, J., & Rauch, M.-J. (2025). 'Do songs tell stories?' An empirical analysis of the effect of emotional arcs on success in a national song contest. *Journal of Cultural Economics*, 1-35.
- Rosen, S. (1981). The economics of superstars. *The American Economic Review*, 71(5), 845-858.
- Schäfer, T., Zimmermann, D., & Sedlmeier, P. (2014). How we remember the emotional intensity of past musical experiences. *Frontiers in Psychology*, 5, 911.

- Sciandra, M., & Spera, I. C. (2022). A model-based approach to Spotify data analysis: a Beta GLMM. *Journal of Applied Statistics*, 49(1), 214-229. DOI: 10.1080/02664763.2020.1803810
- Steiner, P. (2018). *Sound branding*. Springer.
- Thielmann, I., Böhm, R., Ott, M., & Hilbig, B. E. (2021a). Economic games: An introduction and guide for research. *Collabra: Psychology*, 7(1), 19004.
- Thielmann, I., Böhm, R., Ott, M., & Hilbig, B. E. (2021b). Standardized Game Instructions - English.
- Trehub, S. E., Becker, J., & Morley, I. (2015). Cross-cultural perspectives on music and musicality. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1664), 20140096. DOI: 10.1098/rstb.2014.0096