The recall and recognition of sonic logos.

Sebastian Silas[1,2], Megan Robinson[1,5], David John Baker[1,3], Peter Harrison[4,5], Nori Jacoby[4], & Daniel Müllensiefen[1,2,6]

[1] Goldsmiths, University of London, London, UK

[2] Hochschule für Musik, Theater und Medien, Hannover, Germany

[3] Institute for Logic, Language and Computation, University of Amsterdam, The Netherlands

[4] Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany

[5] University of Cambridge, Cambridge, UK

[6] Institute of Systematic Musicology, University of Hamburg

Author Accepted Manuscript

©American Psychological Association, 2026.

This paper is not the copy of record and may not exactly replicate the authoritative document published in the APA journal. The final article is available, upon publication, at:

https://doi.org/10.1037/aca0000870

Author note

The research was granted ethical consent by Goldsmiths, University of London and the Ethics Subcommittee of the Faculty of Music, University of Cambridge. It was supported by an InnovateUK grant (no. 96040) as part of a collaboration between Goldsmiths, University of London and the sonic branding agency SoundOut/Slicethepie Ltd. Sebastian Silas has been supported by a doctoral scholarship from the Studienstiftung des deutschen Volkes. Daniel Müllensiefen has been supported by an Anneliese Maier-Research Prize awarded by the Humboldt Foundation. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising.

The authors made the following contributions. Sebastian Silas: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing, Visualization, Formal Analysis, Data Curation, Investigation; Megan Robinson: Writing - Review & Editing, Data Curation; David John Baker: Writing - Original Draft Preparation, Writing - Review & Editing, Visualization, Formal Analysis, Data Curation; Peter Harrison: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing, Formal Analysis; Nori Jacoby: Writing - Review & Editing, Project Administration, Resources; Daniel Müllensiefen: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing, Formal Analysis, Funding Acquisition, Project Administration, Resources.

Correspondence concerning this article should be addressed to Daniel Müllensiefen, Institute of Systematic Musicology, University of Hamburg, Neue Rabenstraße 13, 20354, Hamburg, Germany. E-mail: D.Mullensiefen@gold.ac.uk

Abstract

Sonic logos are short musical signatures designed to support brand recognition and facilitate the transmission of brand identity. Yet little is known about how their melodic properties shape memorability, or how melodic recognition and sung recall relate to one another. Using two established paradigms—the continuous recognition task (Herff, Olsen, & Dean, 2018) and the melodic recall task (Silas & Müllensiefen, 2023; Sloboda & Parker, 1985)—we tested 90 sonic logos with a large number of participants (N = 1,860) and found that both recognisability and recallability vary reliably across representative commercial sonic logos. Melodic features indicating complexity predicted both forms of memory but in opposite directions: greater structural complexity enhanced recognisability while hindering sung recall, producing a negative correlation between recall and recognition. However, once both paradigm scores, melodic features, and contextual influences were taken into account in the same models, the apparent negative relationship between recognition and recall disappeared, indicating that the two forms of memory draw on different aspects of melodic structure rather than directly reflecting one another. These findings clarify the cognitive bases of melodic memorability and offer practical guidance for the design of effective sonic logos.

Keywords: sung recall, sonic logos, oral transmission, melodic memory, melodic recognition

The recall and recognition of sonic logos.

Sarah asks, “Do you like that new song on Beyoncé’s latest album?” “Which one?” Michael replies. Sarah responds by singing a short melodic fragment: “The one that goes…”. “Ah, *that* one!”, Michael says. This everyday exchange illustrates two fundamental processes of melodic memory: sung melodic recall (Anglada-Tort, Harrison, Lee, & Jacoby, 2023; Jacoby et al., 2019; Silas & Müllensiefen, 2023; Silas, Müllensiefen, & Kopiez, 2023) and melodic recognition (Bella, Peretz, & Aronoff, 2003; Cuddy, Cohen, & Miller, 1979; Dowling, 1986; Dowling, Kwak, & Andrews, 1995; Müllensiefen & Halpern, 2014). As this example suggests, both processes are essential for transmitting musical ideas in everyday life (Anglada-Tort et al., 2023).

To study the memorability of melodic material, prior research has employed a range of memory paradigms (Herff et al., 2018; e.g., Müllensiefen & Halpern, 2014; Sloboda & Parker, 1985). These paradigms provide complementary indices of memorability and may tap into different stages of encoding and retrieval. The present study focuses on two such paradigms—the *continuous recognition paradigm* (Herff et al., 2018) and the *melodic recall paradigm* (Silas & Müllensiefen, 2023; Sloboda & Parker, 1985)—using sonic logos as stimuli. Sonic logos are short auditory signatures used in real-world branding and thus offer high ecological validity (Krishnan Palghat, 2009; Krishnan & Kellaris, 2021; Krishnan, Kellaris, & Aurand, 2012).

While recall and recognition have been extensively compared in non-musical domains (e.g., Barber, Harris, & Rajaram, 2015; Koriat & Goldsmith, 1994), relatively few studies have examined *melodic* sung recall and recognition within the same experimental framework (for an exception, see Gardiner, Kaminska, Java, Clarke, & Mayer, 1990). Addressing this gap, we

examine the relationship between melodic recognition and sung recall and interpret the findings in both cognitive-psychological and marketing contexts. To further differentiate the two processes, we model recognition and recall performance using computational features of sonic logos, both separately and jointly. This approach allows for the decomposition of sonic logo memorability into constituent components and enables predictions about the memorability of novel logos based on their features alone.

Background

Memory and Recall in Cognitive Psychology

Recall and recognition have long been central topics in cognitive psychology (Allen, Hitch, & Baddeley, 2018; Gillund & Shiffrin, 1984; Li, Masitah, & Hills, 2020; Raaijmakers & Shiffrin, 1992; Rhodes, Greene, & Naveh-Benjamin, 2019; Tversky, 1973). Reproduction—a special case of recall that is especially relevant here—has been shown to enhance memory encoding. For instance, drawing pictures improves later memory, likely by integrating semantic, visual, and motor information, an effect known as the *drawing effect* (Fernandes, Wammes, & Meade, 2018; Meade, Ahmad, & Fernandes, 2020; Wammes, Meade, & Fernandes, 2016). Sung recall may engage similar mechanisms, as singing relies heavily on motor processes (Pfordresher et al., 2015; Silas et al., 2023). A comparable phenomenon has been observed in reading: producing words aloud improves both recall and recognition (the *production effect*), suggesting shared underlying working memory processes (Rumbaugh & Landau, 2018).

At the same time, recall and recognition are known to differ systematically. Both are context-dependent and subject to interference and false positives (e.g., Roediger & McDermott, 1995). In visual and verbal domains, similarity effects disproportionately impair recall relative to

recognition (Bestgen et al., 2017; Chubala, Guitard, Neath, Saint-Aubin, & Surprenant, 2020), and sentence-level advantages have been observed for recall but not recognition (Allen et al., 2018). Evidence surrounding the *Tulving–Wiseman law* (Gardiner et al., 1990; Tulving & Wiseman, 1975) further demonstrates that recall and recognition can dissociate: items may be recognised without being recallable, or vice versa. These findings suggest that, although related, recall and recognition constitute partially distinct cognitive mechanisms.

Several theoretical accounts attribute these differences to variations in cue availability and retrieval processes. For example, Tulving and Watkins (1973) proposed that recognition relies more on later, semanticised memory representations, whereas recall depends more strongly on earlier, episodic representations. Others have argued that recall relies primarily on recollection, while recognition draws on both recollection and familiarity (Uner & Roediger, 2022).

More recently, memorability has been reconceptualised as a stable stimulus-level property rather than a consequence of individual strategies or task demands. In vision, memory performance for images is highly reliable across observers and tasks (Bainbridge, Walther, Fukuda, & Goetschalckx, 2026). Comparable effects have been demonstrated for voices (Revsine, Goldberg, & Bainbridge, 2025), but this framework has not yet been applied to melodic materials. The present study addresses this gap by testing whether sonic logos exhibit reliable item-level differences in recall and recognition across listeners and paradigms.

Overall, research outside music suggests that recall and recognition are distinct yet related processes, and documents a *recall–recognition paradox*, whereby recognition is not always easier than recall (Karlsen & Snodgrass, 2004; Mandler, Goodman, & Wilkes-Gibbs, 1982). Such discrepancies may also arise in melodic memory.

Operationalising Melodic Recall vs. Recognition

In this paper, *melodic recall* is operationalised as the ability to *sing back* a melody (Silas & Müllensiefen, 2023; Silas et al., 2023; Sloboda & Parker, 1985), rather than merely judging familiarity. Sung recall thus reflects both memorability and singability, which may be linked for evolutionary reasons—for example, vocal constraints favour small intervals, which become more frequent and easier to remember (Anglada-Tort et al., 2023).

While cognitive research suggests that recall and recognition may share mechanisms and therefore correlate, substantial evidence indicates they are partially distinct. It remains unclear whether this pattern extends to melodic recall and recognition. These processes could be uncorrelated, positively correlated (reflecting shared memorability), or even negatively correlated, implying trade-offs between distinct cognitive demands.

As suggested by the opening example, recognition and recall can be conceptualised within a shared transmission–reception framework: recall involves transmitting a memory, whereas recognition involves receiving and interpreting one. Both rely on working memory—often described as a central "cognitive primitive" (Alloway & Alloway, 2013; Baddeley & Hitch, 1974)—to maintain melodic material online during perception, comparison, or production. However, the broader processes differ. Sung recall requires not only a mental representation but also accurate vocal reproduction, self-monitoring, and feedback (Pfordresher et al., 2015; Silas et al., 2023), resulting in higher working memory demands. Moreover, recognition typically yields binary judgements, whereas recall performance varies continuously (e.g., partial recall, tuning accuracy) (Silas et al., 2023).

One plausible hypothesis is that recall and recognition are positively correlated, with melodic complexity impairing both due to increased cognitive load (Müllensiefen & Halpern, 2014; Silas et al., 2023). However, consistent with the Tulving–Wiseman law (Gardiner et al., 1990; Tulving & Wiseman, 1975), the two may be weakly related or dissociated. For instance, simple stepwise melodies may be easy to recall but difficult to recognise due to their lack of distinctiveness, whereas highly distinctive melodies (e.g., the Nokia ringtone or *Flight of the Bumblebee*) may be easily recognised but difficult to reproduce. See Figure 1 for an example comparison of sonic logos which may be either easy to recognise or recall.

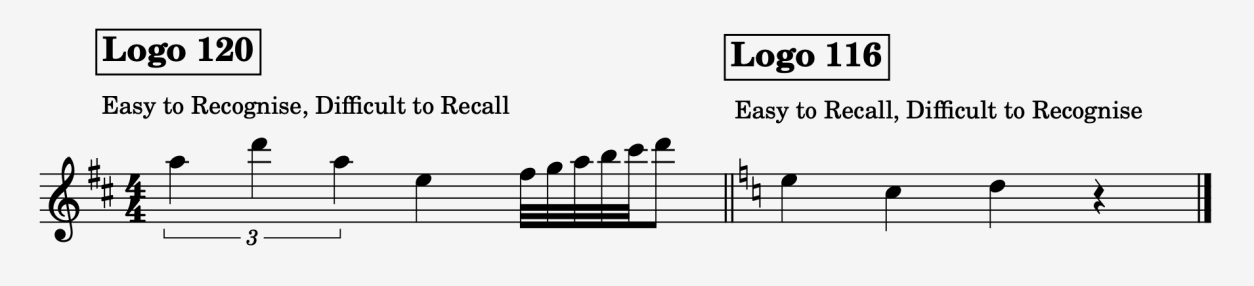


Figure 1. An example of a melody which is easy to recognise but difficult to recall (Logo 120) vs. a logo which is easy to recall but difficult to recognise (Logo 116). The former has larger intervallic leaps, more notes and overall information, and a more complicated rhythm. The latter has only three notes, is isorhythmic, and only moves in scale steps or skips.

Sonic Branding and Sonic Logos

Music is widely used in advertising as a means of transmitting brand information (Allan, 2021; Schönpflug, 2008). *Sonic branding* focuses on how sound conveys meaning in branding contexts (Krishnan & Kellaris, 2021), and sonic logos—short auditory signatures paired with brands—are a central tool (Krishnan Palghat, 2009). Sonic logos may be non-melodic (e.g.,

Netflix), melodic (e.g., *Intel*, *Nokia*), or lyrical (e.g., *McDonald's*: "I'm lovin' it"). They can function as attentional cues, mnemonic signatures, or vehicles for brand personality, depending on design goals.

As North and Hargreaves (2010) note, music in advertising ultimately serves profit. Sonic logos may contribute by supporting memory processes: brand names are better recalled when paired with musical jingles rather than spoken slogans (Yalch, 1991), suggesting a mnemonic function (Bellezza, 1984). Sonic logos may also evoke familiarity and positive affect (Fraedrich & King, 1998; Schubert, 2007; Wu, Fu, Huang, & Wang, 2010), increasing the likelihood of brand use (Huron, 1989). Accordingly, effective sonic logo design requires understanding which musical features enhance attention, memorability, and brand associations (Silas, Baker, & Müllensiefen, 2024). This paper focuses specifically on memorability, operationalised as recallability and recognisability.

Melodic Features and Sonic Logo Testing

Identifying which sonic logos are memorable is of clear practical value, but understanding *why* certain logos are memorable enables prediction and theory-building. From a cognitive-science perspective, this involves linking stimulus properties to memory outcomes. Prior research suggests that effective sonic logos are easy to hum, of moderate length, and feature simple contours and clear tonality (Krishnan & Kellaris, 2021; Krishnan et al., 2012; Müllensiefen, Egermann, & Burrows, 2015). However, these accounts omit several computational melodic features known to influence melodic memory, including tonal stability, rhythmic surprise, and statistical frequency (Dreyfus, Crawford, Müllensiefen, & Baker, 2016; Harrison, Musil, & Müllensiefen, 2016; Müllensiefen & Halpern, 2014; Silas & Müllensiefen, 2023; Silas et al., 2023).

We therefore extend prior sonic branding work by incorporating a broader feature set drawn from melodic memory research. To assess memorability and transmission, we apply two paradigms—the *continuous recognition paradigm* and the *melodic recall paradigm*—which have not yet been widely used in sonic logo testing.

The Present Study: Melodic Recognition and Sung Recall of Sonic Logos

This study advances research in three ways. First, it jointly examines melodic recognition and sung recall of sonic logos, providing a more precise account of memorability relevant to branding. Second, it broadens the set of predictive melodic features, drawing on melodic memory research and implementing them via the `musicassessr` *R* package[1], which builds on the *FANTASTIC* feature set (Müllensiefen, 2009), additional structural measures (Beaty et al., 2021), and information-theoretic metrics (Harrison, Bianco, Chait, & Pearce, 2020). Third, it demonstrates how latent variable modelling can integrate item features with multiple memory outcomes simultaneously.

Using sonic logos enhances ecological validity: these stimuli are intentionally designed to engage memory processes relevant to real-world outcomes, yet are not artificially constructed for experimental purposes.

Hypotheses

Two core premises motivate the hypotheses. First, recall and recognition are related but dissociable memory processes (Koriat & Goldsmith, 1994; Tulving & Wiseman, 1975; Uner &

[1] https://sebsilas.github.io/musicassessr/

Roediger, 2022). Recognition benefits more from familiarity and external cues, whereas recall requires reconstructive retrieval. Second, melodic complexity and distinctiveness influence memorability in different ways. Distinctive, information-rich melodies facilitate recognition (Dreyfus et al., 2016; Müllensiefen & Halpern, 2014), but may hinder recall due to increased cognitive and motor demands (Pfordresher et al., 2015; Silas & Müllensiefen, 2023; Silas et al., 2023).

H1: Melodic features predict memory performance, but complexity has opposing effects on recognition and recall

Melodic features reflecting structural complexity were predicted to influence both recognition and recall (Bailes, 2010; e.g., Müllensiefen & Halpern, 2014; Schulkind, Posner, & Rubin, 2003; Silas & Müllensiefen, 2023). However, complexity was expected to enhance recognition by providing distinctive cues, while impairing recall by increasing cognitive and motor demands.

H2: Recall and recognition rely on distinct but partially overlapping processes

Based on extensive memory research (Gardiner et al., 1990; Tulving & Wiseman, 1975; Uner & Roediger, 2022), it was predicted that item-level recall and recognition scores would be correlated, but only modestly, reflecting partial overlap between distinct underlying cognitive processes.

Methods

Materials

Melodic Sonic Logos.

The sonic logo dataset was a collection of 133 sonic logos provided by the *MassiveMusic*[2] creative music agency. These logos were chosen to be representative of commercial sonic branding assets but with the guarantee that they had never been released or used in commercial branding campaigns. This ensured our stimuli were ecologically valid in the context of sonic branding, yet unfamiliar to participants. The original audio versions of the logos varied in instrumentation, timbre, and dynamics; however, in order to maximise the comparability between the different logos, an expert musician transcribed and created versions of logos with a piano sound with no dynamic variation. Transcriptions were checked for accuracy by one of the authors of this paper. Each logo had a length between 2 and 9 seconds. Of the total 133 logos, 112 used salient melodic material as opposed to a single, timbral gesture or sound. We excluded the 21 non-melodic sonic logos to compare melodies in an unbiased way, since a single gestural logo would be much easier to discriminate in a recognition task among a set of predominantly melodic logos. Additionally, there was a separate block of melodies whose data we do not analyse here (see Appendix A).

Continuous Recognition Paradigm. We used the continuous recognition paradigm to test the recognisability of sonic logos. The paradigm works by presenting a continuous stream of melodies, which may or may not be repeated, and asking participants to rate how confident they are that each melody has appeared earlier in the experiment. Recognition is inferred from within-item increases in familiarity ratings across repetitions rather than from explicit old/new judgements, following Herff et al. (2018).

[2] https://massivemusic.com/

Nine online surveys implementing this paradigm were created and administered using the *Qualtrics* platform (*Qualtrics*, 2018). Participants were instructed: "You will now listen to a series of short musical extracts. Please listen to each extract carefully". Each survey comprised 40 trials: 10 target logos presented three times each (30 trials total) and 10 distractor logos presented once. Target and distractor items were randomly interleaved, and repetitions of the same target logo were separated by multiple intervening items, such that no logo was repeated consecutively and repetition spacing varied across the survey.

After each logo, participants answered two compulsory questions assessing recognisability ("How confident are you that you have heard this before?") and pleasantness using a 7-point Likert scale (1 = Not at all, 7 = Extremely)[3]. Each survey page automatically played the extract, and the Next button only appeared after playback, ensuring full stimulus exposure. The survey took approximately 15 minutes to complete.

Because distractor items constituted 25% of all trials and repetitions were temporally distributed, participants could not rely on a simple response bias toward recognising repeated items. Recognisability was operationalised as the logo-specific slope of familiarity ratings over repetitions, while logo-specific intercepts captured initial familiarity, allowing learning-related increases to be distinguished from ceiling effects for logos that were already recognised on first presentation. A third repetition was included (beyond the two used by Herff et al. (2018)) to reduce floor effects and to better estimate recognition trajectories.

Melodic Recall Paradigm. For the melodic recall test, sonic logos were presented through the *Singing Ability Assessment* (*SAA*; Silas et al., 2023) melodic recall test, which is built

[3] Note that the pleasantness data are not presented here due to showing very low reliability.

inside the *psychTestR* (Harrison, 2020) framework and was deployed on an *Amazon Web Services* EC2 server instance. We used the melodic recall paradigm as described in previous research (Silas & Müllensiefen, 2023; Sloboda & Parker, 1985). First, a quality check was carried out to test the noise level of the participant's environment and only participants with a sufficient signal-to-noise ratio were allowed to proceed. Then, participants' vocal range was computed in real time via singing low and high notes. All further stimuli were transposed into this range in order to minimise errors related to singing and focus on melodic memory. Participants sang five long tones. If they were detected not to be singing via a categoriser we built to detect whether someone passes a stable singing criterion, they could again be disqualified before proceeding further (see Appendix A for details on these points). Participants then sang 10 sonic logos, randomly sampled from the set of 112, as well as 10 other melodies, whose data are not used here, but published elsewhere (Silas et al., 2023).

Participants were required to sing back a melody or sonic logo without rehearsal. As in Sloboda and Parker (1985), the target melody could potentially be heard again for subsequent attempts, but only once per attempt with a maximum of three attempts. However, in this experiment, this depended on which experimental conditional participants were in, since testing was deployed across four different conditions with a 2 x 2 design (arrhythmic vs. rhythmic; single-attempt vs. multiple attempts), producing: 1) arrhythmic single attempt; 2) rhythmic single attempt; 3) arrhythmic multiple attempts; 4) rhythmic multiple attempts. The four conditions were released online in a staggered fashion, starting days and weeks apart, but then ran simultaneously.

Participants entered one of the four conditions to complete the recall test (i.e., each test instance only exposed the participant to one of these conditions). They were however able to

complete the test again under a different condition. Arrhythmic melodies presented melody notes isochronously (0.5 *s* per note); rhythmic melodies retained their rhythmicity (i.e., varying durations). For multiple attempt conditions, participants could optionally have up to three attempts per sonic logo, if they wanted. We only present sonic logo data here and because most participants did not opt for more than one attempt (only ~ 3% trials were a second or third attempt), we filtered the entire dataset across the four testing conditions to only model first attempts. For more information about the *SAA* procedures, see Silas et al. (2023) or the accompanying documentation[4].

Participants

For all test instances, participants were recruited via the *SliceThePie*[5] platform run by the UK sonic branding company, *SoundOut*. A small payment from *SliceThePie* incentivised participation. Across the test instances for the continuous recognition paradigm, 1,227 participants (Female = 891; Male = 519; Non-binary / third gender = 32; Prefer not to say = 11; Prefer to self-describe = 11) aged 18-74 (M = 32.91; SD = 12.02) were recruited. Participants were randomly redirected to one of the nine surveys, each containing 20 of the sonic logos. For the melodic recall paradigm, 633 participants aged 16-72 (M = 31.07, SD = 11.54; Female = 443;

[4] https://saa.musicassessr.com/

[5] https://www.slicethepie.com/. *SliceThePie* is a long-running paid review platform where users listen to and rate musical content in exchange for small payments.

Male = 226) were recruited across the four testing conditions ($N = 219$; $N = 249$; $N = 207$; $N = 227$). Eight participants' demographic data was missing (reason unknown)[6].

Note that, the recognition and the recall experiments were not run at the same time via *SliceThePie*, but weeks apart, and therefore, it is unlikely that there were any carry-over or learning effects for the few participants who took part in both studies.

Data Analysis

Melodic Feature Computation

For all sonic logos, we computed melodic features from the *FANTASTIC* (Müllensiefen, 2009) toolbox as well as other melodic features used in Beaty et al. (2021) and measures of total information content based on Harrison et al. (2020), resulting in a set of 16 melodic features (see Appendix B for a table describing these features). In general, all features are taken to indicate melodic complexity in some manner (e.g., longer length/durations, higher intervallic range, more information content, lower tonalness = higher complexity).

Dimension Reduction of Melodic Features

As a first analysis step, we reduced the pool of 16 melodic feature variables via dimension reduction. A parallel analysis (Horn, 1965) suggested four components. Hence, we conducted a four-dimensional principal components analysis (*PCA*) on the data using the *equamax* rotation method, which is an orthogonal rotation method and hence the components we obtain should be

[6] Note that, in this case, demographic data were provided by *SoundOut*, who only collect binary gender data.

fairly independent of each other. One feature had a relatively high uniqueness index of 0.95 and was removed. All remaining components had several loadings above .30 and all communality (h^2) values were greater than 0.50. The amount of variance explained by the four components was: 0.25, 0.25, 0.13, 0.12. We labelled the components *Intervallic Complexity*, *Breadth of Material*, *Tonality*, and *Durational Complexity* and computed aggregate scores for each of the four components, for each sonic logo. See Appendix C for the full *PCA* results.

Derivation of Paradigm Scores and By-Logo Aggregates

Continuous Recognition Paradigm.

To derive a *recognisability score*, we fitted a mixed-effects model with *repetition number* as fixed effect and participant intercept, sonic logo intercept, the interaction between participant and sonic logo intercept, and sonic logo slopes over repetition as random effects. The slope was estimated across all three repetitions of each logo, not just the first two. Estimating the slope across three time points produces a more stable slope parameter. We extracted the sonic logo slopes over repetition to represent a *recognisability* score for each sonic logo as well as the sonic logo random intercepts to serve as a control for *initial familiarity* in later modelling.

Melodic Recall Paradigm. For each trial, the *SAA* computes the melodic similarity of symbolic representations between a) the target sonic logo with b) the corresponding sung recall of that target sonic logo using the *opti3* measure of similarity via the `opti3` function from the *R* package `musicassessr` (see Appendix D for more information). The use of *opti3* in the context of melodic recall is discussed extensively across Silas and Müllensiefen (2023) and Silas et al. (2023) to which we refer the reader, as well as the articles in the online supplement to its

test environment, *musicassessr*[7]. To derive a score for each sonic logo, we fitted a random effects-only model with *opti3* as dependent variable and participant, sonic logo, and the interaction between participant and logo as random effects intercepts. We extracted the sonic logo intercepts to represent a *sung recallability* score for each sonic logo.

By-Logo Score Assessment. To assess variability in memory performance across individual sonic logos, the logo-level scores extracted in the previous steps were combined into a single dataset. This allowed for visual and statistical examination of the range and distribution of logo scores across paradigms to confirm that the logos differed sufficiently to support subsequent modelling. It also provides the basis for later correlational and regression analyses treating each logo as a unit of observation.

Split-Half Reliability Computation

In the memorability literature, split-half reliability is interpreted as evidence that memory performance reflects intrinsic stimulus properties rather than participant-specific effects (Bainbridge et al., 2026). Accordingly, reliable item-level scores indicate convergence across listeners on which stimuli are more or less memorable. We therefore treat split-half reliability not only as a psychometric check, but as a theoretically meaningful index of stimulus-level melodic memorability.

We computed a split-half reliability coefficient for each paradigm score using a Monte Carlo approach. For each paradigm score derived via the above methods, we split the dataset into two randomly selected participant groups, and for each split half dataset, computed sonic logo scores via the methods described above. We then computed a Pearson's *r* correlation for an

[7] https://sebsilas.github.io/musicassessr/articles/

estimate of the correspondence between the two halves and thus the reliability of the scoring procedure. For each paradigm, we computed this process on 1,000 randomly split datasets and took the average of the result as the paradigm's reliability value.

Hypothesis-Driven Modelling

The remaining analyses followed the previously stated hypotheses and examined: 1) the extent to which melodic features predict memory performance and differences in the directional effects of melodic features on recognition vs. recall; 2) the relationship between recall and recognition scores.

H1: Melodic features predict memory performance, but complexity has opposing effects on recognition and recall

To explain recognisability and recallability scores on a univariate basis, we regressed the by-logo paradigm scores onto the by-logo scores for the four melodic feature principal components, using two independent multiple linear regression models. Next, the direction of regression coefficients from the previous two univariate models were compared to assess potential opposing effects for certain melodic feature components.

H2: Recall and recognition rely on distinct but partially overlapping processes

2.1 Correlation Analysis

To estimate the strength and direction of the relationship between recognition and recall scores, we assessed their bivariate relationship with one another via a simple bivariate correlation and then a *correlation attenuation analysis* (Spearman, 1904). Correlation attenuation analysis

allows us to more accurately estimate the magnitude of the relationship by taking into account the respective reliability estimates for each paradigm.

2.2 Partial Correlation and Path Analyses

A limitation of the correlational analyses above is that they assess relationships only at the bivariate level. While informative descriptively (e.g., for sonic logo design), such analyses cannot distinguish direct relationships from those driven by shared dependencies. Partial correlation analysis addresses this limitation by modelling conditional dependencies among variables, thereby revealing networks of relationships beyond pairwise associations. Accordingly, to clarify whether recallability and recognisability are directly related, we conducted an exploratory partial correlation analysis (Pearl, 2000, 2009; Silas, Müllensiefen, Gelding, Frieler, & Harrison, 2022).

A core principle of this approach is to represent the full system of variables so that genuine associations—those not confounded by third variables—can be identified, while spurious relationships are ruled out (Pearl & Mackenzie, 2018). Specifically, we tested whether recallability and recognisability are directly linked or whether their association is mediated by melodic or contextual variables, as has been observed in non-musical memory research (e.g., Roediger & McDermott, 1995). To capture relevant contextual effects, we included two additional variables. First, initial familiarity was operationalised as the random-effects intercept for each sonic logo in the continuous recognition paradigm, indexing baseline recognisability prior to learning effects across repetitions. Modelling both intercepts and slopes allowed us to separate pre-existing familiarity from learning-related gains and ceiling effects. Second, mean inverted document frequency (*mean IDF*; see Appendix B) was included as a proxy for stimulus-level distinctiveness, capturing how much a logo stands out relative to the broader corpus.

Together, these variables represent contextual influences at the behavioural and stimulus levels and allow us to test whether apparent recall–recognition associations persist once such factors are controlled.

Initial familiarity, logo distinctiveness (*mean IDF*), and the four melodic feature components were submitted to the PC algorithm, which estimates networks of conditional dependencies among variables (Richardson, 2013), implemented in the *R* package *pcalg* (*v* 2.7.12). Conditional independence was assessed by testing whether partial correlations between variable pairs were statistically indistinguishable from zero. Using a threshold of $\alpha = .05$, edges were retained in the estimated skeleton when $p < .05$. All variables were included as nodes, and the algorithm searched the full space of conditional associations without additional constraints.

To quantify the strength of relationships identified in the partial correlation network, we subsequently fitted path models using *lavaan* (*v* 0.6.20). A model specified directly from the partial correlation structure showed inadequate fit. We therefore estimated a maximal model in which all melodic features predicted recallability, recognisability, and initial familiarity, pruned non-significant paths, and compared the resulting model to the structure suggested by the partial correlation analysis.

Results

Descriptive Statistics

For the *continuous recognition paradigm*, the average response value on repetition 1 was *M* = 3.55 (*SD* = 1.93), on repetition 2, *M* = 4.10 (*SD* = 1.95), and on repetition 3, *M* = 4.46 (*SD* = 1.95). Across participants, mean learning slopes were positive (mean slope = 0.43), indicating

systematic improvement in response values with repeated exposure. For the *melodic recall paradigm*, the average participant *opti3* score was 0.28 (*SD* = 0.12).

By-Logo Scores

As shown in Figure 2, scores aggregated at the level of sonic logo suggest that there is sufficient variability between sonic logos to be able to statistically distinguish many from one another, in terms of recall and recognisability paradigm scores. As shown in Figure 3, recognition ratings increase over repetitions in a statistically significant manner.

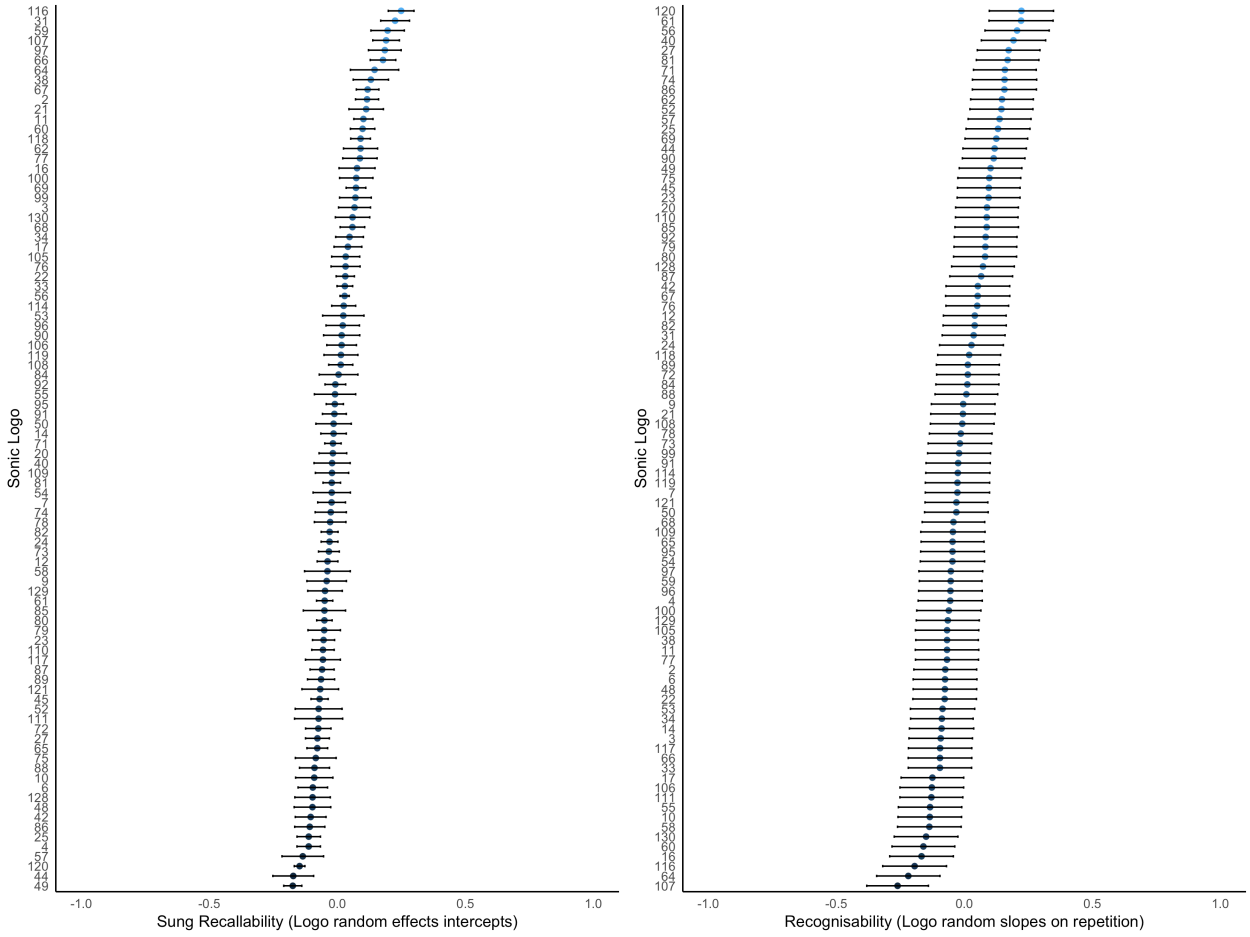


Figure 2. Recall and recognition scores for each sonic logo with 95% confidence intervals.

Reliability of Paradigm Scores

The estimated split-half reliability of the recognition paradigm was $\alpha = 0.55$ and $\alpha = 0.75$ for the recall paradigm.

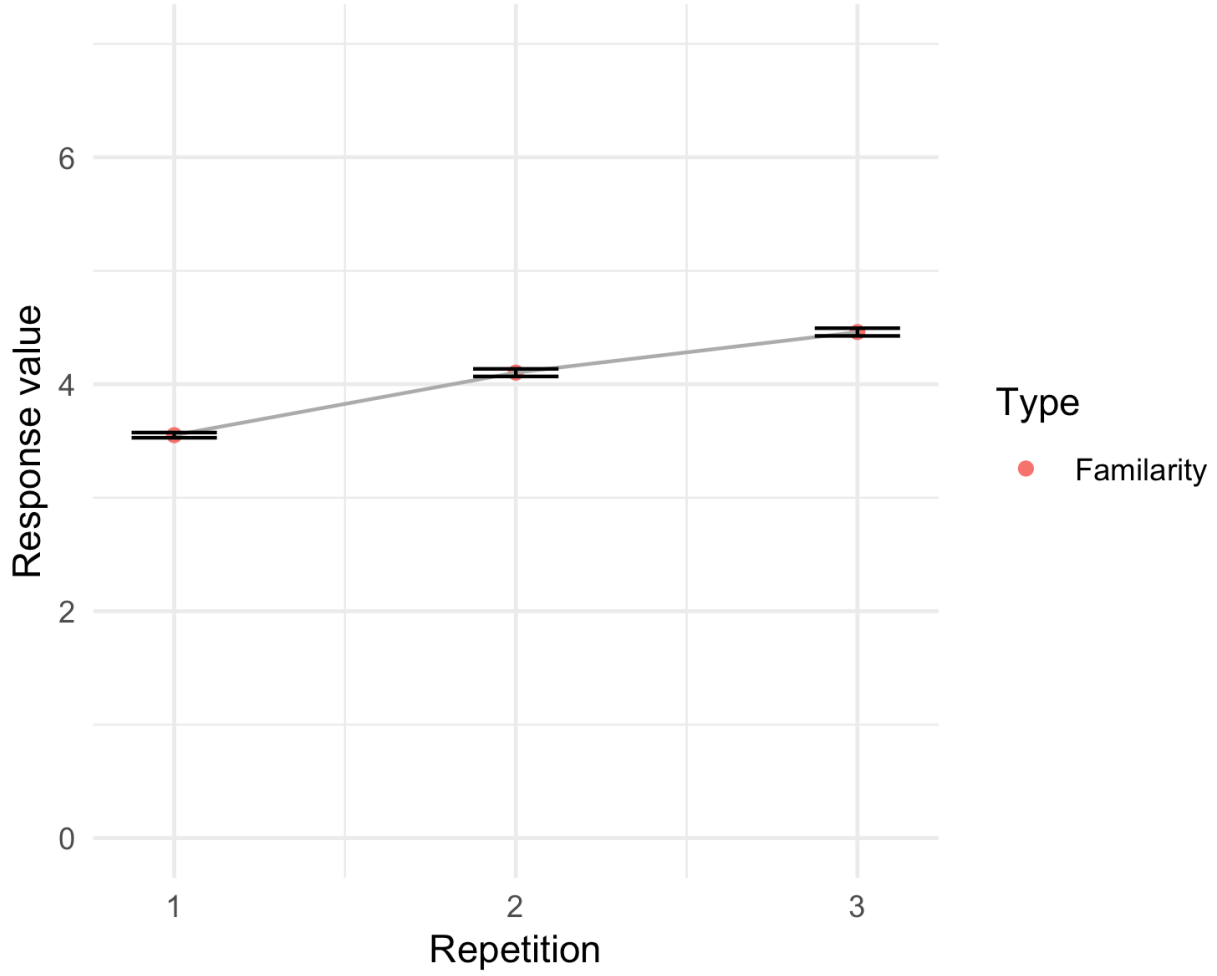


Figure 3. Average change in familiarity over repetitions. Error bars represent 95% confidence intervals.

Hypothesis-Driven Modelling

H1: Melodic Features Predict Memory Performance: Univariate Modelling.

To establish a baseline account of how melodic structure relates to the paradigm scores and thus memory, we first examined the independent predictive value of the four melodic feature components on each memory paradigm before introducing contextual or multivariate dependencies.

1.1 Univariate Multiple Regression Models. First, we investigated which melodic feature components were significant predictors of paradigm scores. For the recognition model, the R^2 of the model was 0.46 (adjusted = 0.44) suggesting that melodic features explain a moderate amount of variance in the data. In this model (*Model 1*), presented in Table 1, all predictor coefficients were positive. Tests of statistical significance suggest that greater breadth of melodic material ($B = 0.09$, $p < .001$), durational complexity ($B = 0.02$, $p = .03$), and possibly intervallic complexity ($B = 0.02$, $p = .056$) predict better recognition of a sonic logo. Tonality was not a significant predictor of recognition ($p = .62$).

Table 1

Model 1: Linear regression model regressing recognisability score onto melodic feature component scores

Predictor	b	95% CI	t	df	p
Intercept	0.01	[-0.01, 0.02]	0.83	85	.409
Intervallic Complexity	0.02	[0.00, 0.04]	1.95	85	.055
Breadth of Material	0.09	[0.07, 0.11]	7.49	85	< .001
Tonality	0.00	[-0.01, 0.02]	0.50	85	.618

Predictor	*b*	95% CI	*t*	*df*	*p*
Durational Complexity	0.02	[0.00, 0.04]	2.21	85	.030

For the recall model, the R^2 of the model was 0.49 (adjusted = 0.47) suggesting that melodic features explain a similar moderate amount of variance in recall scores as to recognition scores. In this model (*Model 2*), presented in Table 2, all predictor model coefficients which imply melodic complexity were negative. Tests of statistical significance suggest that *lower* intervallic complexity (B = -0.05, $p < .001$), breadth of melodic material (B = -0.04, $p < .001$) and durational complexity (B = -0.03, $p < .001$) reliably predict the sung recall of a sonic logo. As in Model 1, tonality was not a statistically significant predictor (p = .93).

Table 2

Model 2: Linear regression model regressing sung recallability score onto melodic features

Predictor	*b*	95% CI	*t*	*df*	*p*
Intercept	0.00	[-0.02, 0.01]	-0.20	85	.839
Intervallic Complexity	-0.05	[-0.07, -0.04]	-6.34	85	< .001
Breadth of Material	-0.04	[-0.06, -0.03]	-4.55	85	< .001
Tonality	0.00	[-0.01, 0.01]	0.09	85	.928
Durational Complexity	-0.03	[-0.04, -0.01]	-3.97	85	< .001

These initial models show that melodic features explain a large proportion of variance in both recognition and recall scores, providing a clear starting point for evaluating how melodic structure contributes to each memory process before considering how these relationships might differ or conflict across tasks.

H1.2: Opposing Effects of Melodic Complexity on Recognition vs. Recall

Having established that the same features predict recognition and recall, we next explicitly contrasted the coefficient patterns to clarify whether this asymmetry reflects a systematic trade-off rather than isolated effects. A direct comparison of the two regression models reveals a reversal in how melodic complexity predicts performance across the recognition and recall paradigms. In Model 1 (recognition), all three complexity-related components—breadth of melodic material, durational complexity, and intervallic complexity—show positive coefficients. In other words, sonic logos that make use of more varied melodic content, more intricate rhythmic structures, or more varied interval patterns tend to be recognised more successfully. This pattern suggests that added melodic richness and detail create distinctive perceptual cues that aid listeners when judging whether they have heard the logo before. In contrast, Model 2 (recall) displays a consistent set of negative coefficients for the same three components. Here, increased melodic complexity predicts worse performance: sonic logos with more breadth, intricate rhythms, or elaborate interval patterns are recalled less accurately when participants attempt to sing them back from memory. This opposing direction indicates that the kinds of features that support rapid, familiarity-based decisions may simultaneously hinder the ability to recall a logo. Figure 4 illustrates the directional differences.

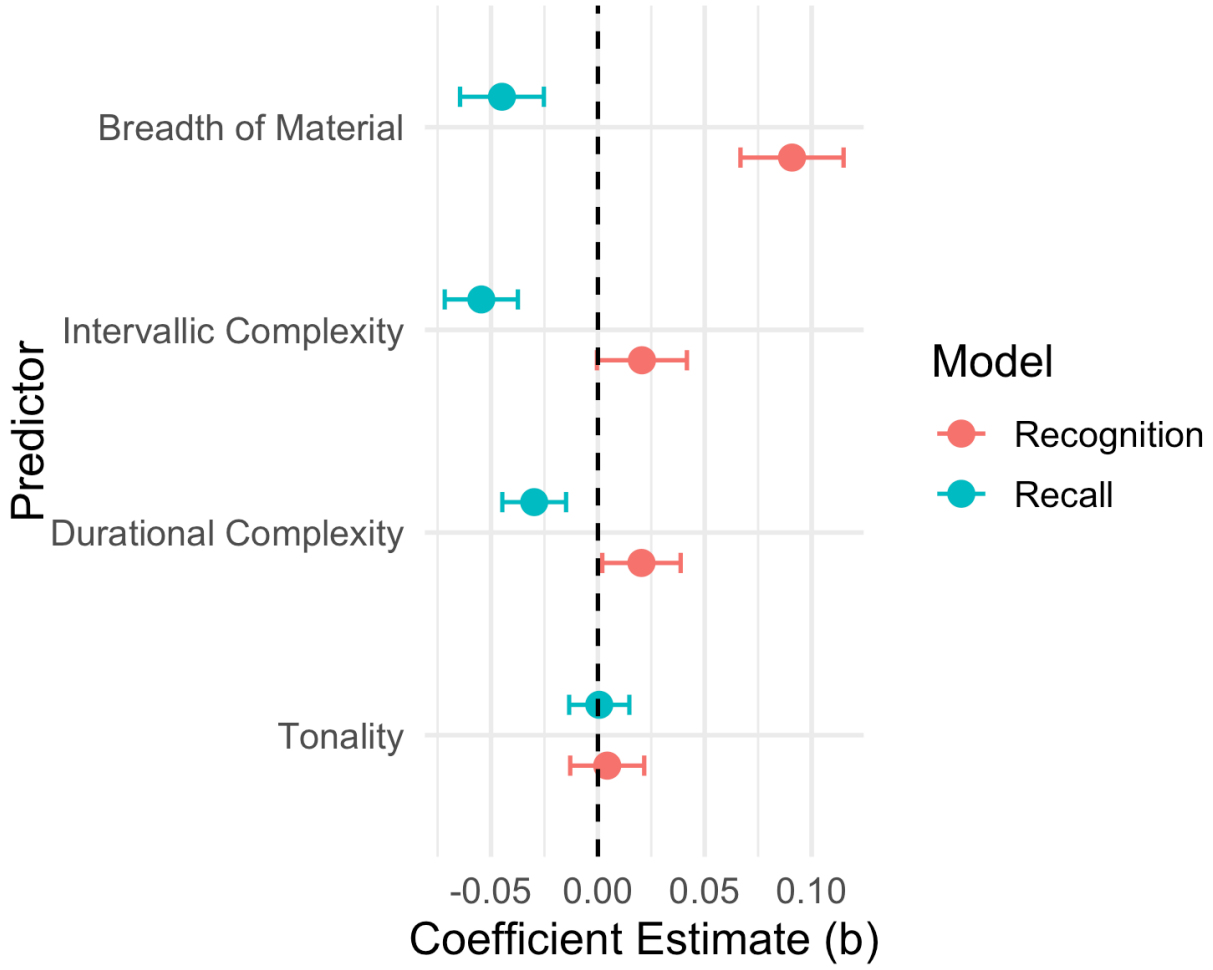


Figure 4. Coefficient estimates for recall vs. recognition models with 95% CIs ordered by magnitude of difference.

The parallel structure of the models makes this opposition especially clear. Each component that enhances recognition does so with a positive coefficient, yet the very same components predict lower recallability with a negative coefficient. Tonality stands out as the only feature whose coefficients do not differ meaningfully between models, as it fails to reach significance in either case. Collectively, these results highlight a dissociation: melodic complexity appears to facilitate perceptual recognition while impeding accurate recall, pointing

to different cognitive (and motoric, for recall) demands underlying these two forms of memory for sonic logos. This explicit comparison suggests a true reversal in how complexity influences each task, suggesting that while complexity provides the discriminative cues needed for recognition, it simultaneously overloads the production-based demands of sung recall—an early sign of a structural "recall–recognition paradox" in melodic memory analogous to those documented in verbal memory research

H2: Relationship Between Recall and Recognition

Before examining conditional independence structure, we first quantified the direct association between recall and recognition to assess whether they behave as two expressions of a general memorability factor or alternatively as largely independent or opposing properties.

2.1 Correlation Analysis. As a first step to associate melodic recognition and sung recall scores, we correlated scores aggregated at the level of sonic logo with one another, yielding a Pearson's correlation of $r = -0.46$ ($p < .001$). When applying correlation disattenuation, which takes into account the respective reliabilities of the test scores ($\alpha_{Recognition} = 0.55$; $\alpha_{SungRecallability} = 0.75$) the r correlation value was estimated to be $r = -0.72$ ($p < .001$). The moderate negative correlation—strengthened to a strong negative association when corrected for reliability—is not surprising given the directional differences of feature predictors found in the univariate modelling stage. However, it clearly indicates that the two paradigms are not parallel indicators of a single (memory) ability; instead, they systematically trade off, underscoring the need for models capable of distinguishing shared from task-specific determinants.

2.2 Partial Correlation and Path Analyses. To disentangle genuinely shared memory influences from those inflated by contextual or stimulus-level confounds, we applied

partial correlation (conditional independence) modelling, which tests whether recallability and recognisability remain related once melodic features, distinctiveness (mean inverted document frequency), and initial familiarity are all taken into account. The results of this analysis, shown in Figure 5, reveal that, conditional on the four melodic feature components, distinctiveness, and initial familiarity, recallability and recognisability become statistically unrelated, with two feature sets tapping in differentially to each cognitive process. Recallability is related to intervallic and durational complexity, but recognisability is related to breadth of material and initial familiarity. In turn, melodic breadth is related to distinctiveness, which is in turn related to tonality. In general, then, recognisability is more sensitive to context effects (initial familiarity, distinctiveness), as found in non-musical studies of recognition (Uner & Roediger, 2022).

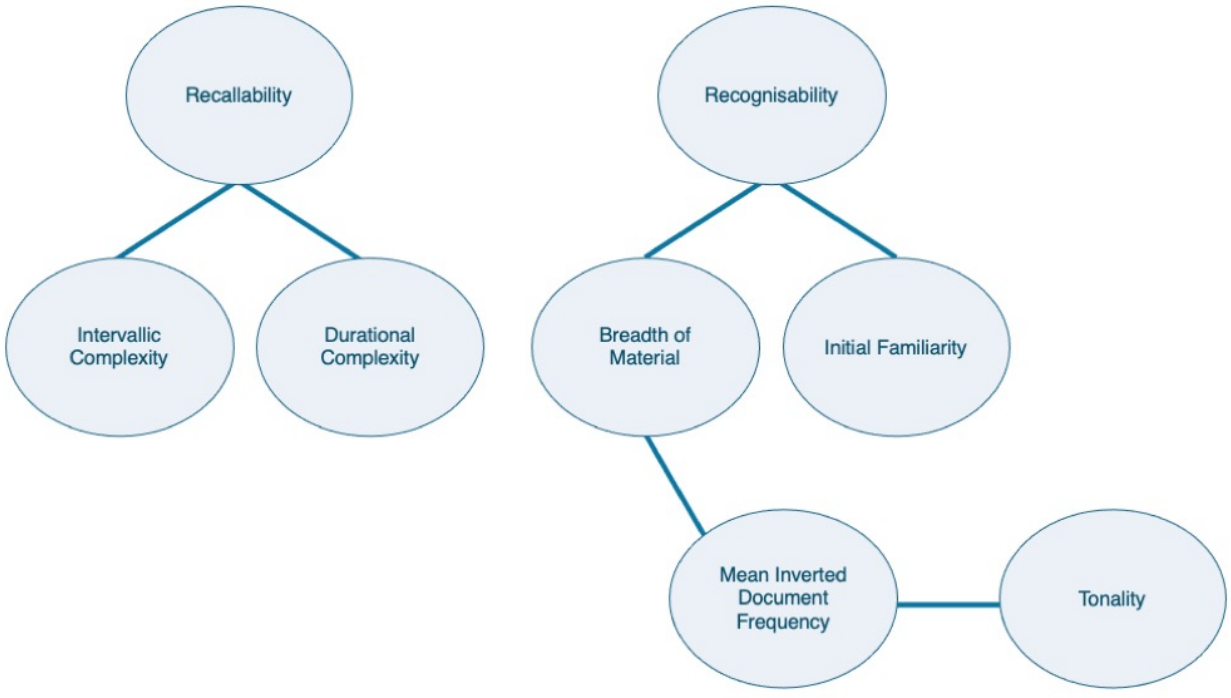


Figure 5. Model 5: Conditional independence analysis of melodic feature components as well as Initial Familiarity and Mean Inverted Document Frequency (to represent distinctiveness)

Formalising and clarifying these effects, the final path analysis model, represented in Figure 6, suggested a similar pattern of results. The model suggests that breadth of material, durational complexity, and intervallic complexity are all negatively associated with recallability, with standardised β path coefficients of -0.36, -0.31, and -0.49, respectively. These findings suggest that as the amount of melodic content and rhythmic or intervallic difficulty of material increase, individuals' ability to freely recall that information diminishes, likely due to heightened cognitive and motoric load. In contrast, recognisability is shaped by a different set of influences. Initial familiarity with the material shows a strong negative association with recognisability (β = -0.57), implying that content perceived as already familiar has less room to increase, consistent with a ceiling effect or because due to decreased attentional engagement or reduced encoding of distinctive features when material is deemed familiar. Conversely, intervallic complexity is positively associated with recognisability (β = 0.56), suggesting that more complex materials may offer richer or more distinctive cues that support recognition, even if they simultaneously hinder recall. A weak and statistically non-significant path was observed from recallability to recognisability (β = -0.08), indicating that the ability to recall information does not reliably predict one's ability to recognise it. The model demonstrated excellent fit indices with a comparative fit index (*CFI*) of 0.99, a standardised root mean square residual (*SRMR*) of 0.04 and a root mean square error of approximation (*RMSEA*) of 0.09 (robust = 0.06).

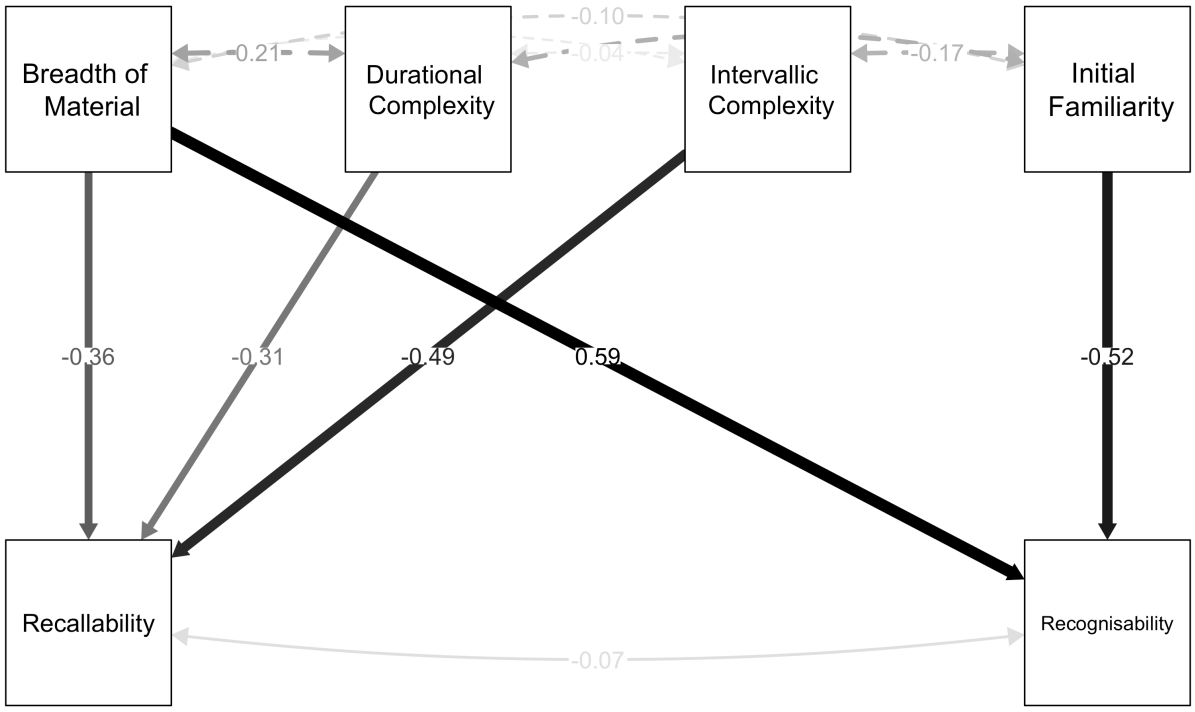


Figure 6. Model 6: Final path analysis model of melodic feature components, initial familiarity, recognisability, and recallability scores

Summary of Modelling Results. The modelling results underscore the dissociable cognitive mechanisms underpinning recall and recognition, and highlight the nuanced effects of material characteristics on different forms of memory performance. Both the partial correlation and path analysis models suggest that recall and recognition become statistically independent once feature-level and contextual factors are controlled, indicating that their surface-level correlation reflects overlapping influences rather than a direct link; the final path model further demonstrated that each task generally draws on a different subset of melodic features, but which partly correlate with one another. Only tonalness contributes to a shared "general memorability" component and/or harms neither property.

Discussion

Firstly, this research separately validated two paradigms—the continuous recognition paradigm (Herff et al., 2018) and the melodic recall paradigm (Silas & Müllensiefen, 2023; Silas et al., 2023)—in the context of sonic logos, a unique kind of musical material designed to facilitate the recollection and transmission of brand information. Both paradigms showed that they were able to distinguish between different sonic logos and had acceptable item-level reliability. This establishes the continuous recognition and the melodic recall paradigms as practical new devices to quantify the recognition and recall of short melodies and sonic logos based on musical features, which extends their use beyond the original scope.

Additionally, this indicates that sonic logos differ systematically in memorability indices across listeners. From a memorability perspective, this suggests that memory performance is not necessarily dominated by individual differences or task idiosyncrasies, but reflects stable stimulus-level properties (Bainbridge et al., 2026). However, the present results also show that memorability is not unitary: recallability and recognisability are both reliable yet dissociable, aligning with recent findings in verbal memory that recall and recognition capture distinct aspects of stimulus structure (Aka, Bhatia, & McCoy, 2023). Thus, melodic memorability appears to comprise multiple stable dimensions rather than a single property.

Melodic Feature Predictors of Recognisability and Recallability

We then proceeded to model such scores in terms of melodic features. To model recognisability, we constructed models which regressed recognition scores onto a battery of melodic features after they were reduced into four parsimonious dimensions. The model suggested that higher breadth of material, durational complexity, and intervallic complexity significantly predict the recognition score. Specifically, *higher* breadth of material, durational

complexity, and intervallic complexity contribute to a *higher* recognition score. This suggests that complex features contribute to a sonic logo's potential to be recognised on multiple hearings and stand out from a set, similar to effects seen in previous research (Dreyfus et al., 2016; Harrison et al., 2016; Müllensiefen & Halpern, 2014).

To model sung recallability, we constructed models which regressed *recallability* scores onto the four melodic feature components. The final model suggested that breadth of material, intervallic, and durational complexity are significant predictors of the recallability score we derived. Specifically, they contribute to a *lower* sung recallability score. In previous research, this pattern has been interpreted as higher melodic complexity causing greater working memory and vocal-motor demands and hence difficulty in sung recall (Silas et al., 2023). Tonality was not a significant predictor in modelling both recognisability and sung recallability. This may be because there is not a great amount of variance in sonic logos ' tonalities, which are often set to be clearly in a major key (Müllensiefen et al., 2015). Taken together, these univariate results show that melodic features explain a substantial amount of variance in both recognition and recall. They provide the foundation on which to examine whether the same features support both processes, or whether they diverge once we take the structure of the feature space—and contextual influences—into account.

Bringing It Together

We explored how recognition and recall scores for melodic sonic logos are related, first as simple bivariate correlations, and then second, in relation to melodic feature predictors. When assessing the relationship between recall and recognition scores aggregated at the level of sonic logo, a negative correlation with a medium effect size (r = -0.46; $r_{attenuated}$ = -0.72) was found. This suggests that, as a sonic logo's average score for recognition goes up, its ability to be

recalled through singing goes down. At first glance, this negative association could imply that the two memory processes are intrinsically opposed. However, as classic work in cognitive psychology suggests (e.g., Tulving & Wiseman, 1975), bivariate correlations between recall and recognition can mask more complex relationships.

In order to clarify what might be driving this correlation, we related melodic features to paradigm scores. The conditional independence analysis and the follow-up path analysis suggest that recallability and recognisability are more or less distinct processes which tap into different melodic feature components—or tap into the same ones in opposing ways. Higher complexity (in terms of intervallic, durational, and general breadth of information) predicts lower recall scores. Importantly, some of these effects may reflect production constraints inherent to sung recall rather than memory demands alone: features such as large intervallic leaps or irregular rhythmic patterns are known to increase vocal motor difficulty, meaning that the recall paradigm partly indexes the singability of a logo as well as the memorability of its underlying pitch structure. Accordingly, the sung recall paradigm should not be interpreted as a pure measure of memory retrieval, but as a composite index reflecting both mnemonic representation and the production constraints involved in vocal reproduction. This is not a limitation of the paradigm, but a defining feature: sung recall indexes how melodies are retrieved when they must be externally reproduced, a condition that differs fundamentally from recognition and is known to dissociate from perceptual judgements in music cognition. Crucially, recallability scores are derived as item-level random effects aggregated over many participants, rather than individual performance measures. This aggregation reduces the influence of individual differences in pitch accuracy or vocal control, such that the reported effects predominantly capture melody-level production constraints and memory demands shared across listeners, rather than singer-specific limitations.

Higher breadth of material allows better recognition of a sonic logo. However, if a melody already feels familiar on the first trial, the ability to recognise it later is impaired and/or higher initial familiarity simply means there is less of a potential for response values to increase over repetitions. This aligns with previous findings in non-musical studies of recognition and recall—that recall relies more heavily on recollection, whereas recognition relies on both recollection and familiarity (Uner & Roediger, 2022).

Critically, once contextual variables and the full set of melodic features are accounted for, recall and recognition become statistically unrelated in the conditional independence analysis. This suggests that the negative correlation observed at the surface level does not reflect a direct relationship between the two memory processes; instead, it emerges because the two tasks rely systematically on different types of melodic information and are differentially constrained by perceptual versus production mechanisms.

Importantly, in this study, recall and recognition were investigated using the same set of stimuli (sonic logos), but largely different participant samples. The present analyses therefore address relationships at the level of melodies (i.e., which structural properties predict higher recallability versus recognisability), rather than at the level of individual differences. To determine whether recall and recognition might be positively related within persons—that is, whether a broader "melodic memory ability" underlies performance across both paradigms—a follow-up study testing the same participants on both tasks will be required. It remains entirely plausible that melodic memory ability constitutes an individual-differences construct with separable but related facets of recall and recognition. However, taking the stimulus-level approach (Aka et al., 2023; Bainbridge et al., 2026), our findings indicate that memorability is

not a unitary feature construct of melodies themselves; rather, it dissociates into recognisability and recallability once melodic structure and task demands are taken into account.

This distinction is theoretically important rather than problematic. Melodic recognition is a purely perceptual judgement task, whereas sung recall necessarily involves an additional production stage that requires sensorimotor planning, vocal control, and online self-monitoring. A growing body of research demonstrates that perceptual and production processes in music can dissociate, such that accurate perception does not guarantee accurate reproduction, and vice versa (Anglada-Tort et al., 2023; Jacoby et al., 2019). From this perspective, the two paradigms used here should not be interpreted as parallel measures of a single underlying "melodic memorability" construct at the stimulus level. Instead, they index partially overlapping but functionally distinct stages of musical cognition: recognition primarily reflects perceptual encoding and familiarity-based access to stored representations, whereas sung recall additionally reflects the constraints imposed by vocal-motor production.

The present findings therefore align with prior work showing systematic divergences between perception and production in musical memory, and indicate that apparent trade-offs between recall and recognition at the stimulus level are an expected consequence of engaging different cognitive and sensorimotor subsystems rather than an artefact of task design. To intuitively understand the trade-off between melodic recognition and recall, we refer the reader to the following companion application: https://musicassessr.com/sonic-logo-mem-explorer/.

Practical Marketing Implications

The main finding in this paper is that melodic recognition and melodic recall are negatively correlated when aggregated at the level of sonic logo. Generally, maximising for one

through appropriate selection of melodic content will decrease the other. This is not because the cognitive processes are directly related to one another (i.e., that better recallers are worse recognisers). Instead, the negative bivariate correlation arises because different melodic features drive recognition and recall differentially and in a generally opposing manner.

However, the modelling presented here suggested that stable tonality may help both properties. Therefore, atonal, or unstable, tonal melodies might be avoided in sonic logo design. These patterns may be explained by broader cognitive principles. Strongly tonal melodies provide a predictable pitch framework and a clear hierarchical structure, reducing cognitive load and supporting rapid perceptual anchoring. By contrast, atonal or tonally ambiguous melodies lack these organisational cues, increasing uncertainty and making both recognition and recall more difficult.

In terms of wider practical consequences, this yields the question: is it more important for a sonic logo to be recognisable or recallable? Perhaps the answer to this question depends on what the primary objective is. A practical consequence of this would be that designers of sonic logos may need to prioritise what function matters most for a given a sonic logo or campaign, since both good recognisability and recallability may not be easy to obtain simultaneously. Perhaps recallability is more important if the goal is to facilitate the memory and recollection of a brand name or create a viral quality that allows a slogan to be transmitted easily (e.g., the *McDonald*'s sonic logo, "Ba da da da da - I'm lovin' it"). Conversely, recognisability may be more important for simply creating basic associations with a brand and a sense of brand familiarity. Alternatively, perhaps an optimum balance can be struck. Do sensational sonic logos strike an optimum balance between recallability and recognisability, explaining their success?

Conclusion

In summary, the melodic recognition and melodic sung recall of sonic logos appear to be negatively correlated. However, this negative correlation only exists at the bivariate level. Further modelling shows that melodic recall and recognition are distinct processes tapping into melodic features differentially. These relationships have important implications for cognitive science as well as sonic branding experts.

References

Aka, A., Bhatia, S., & McCoy, J. (2023). Semantic determinants of memorability. *Cognition*, *239*, 105497. https://doi.org/10.1016/j.cognition.2023.105497

Allan, D. (2021). Commercial Sound. In *The Oxford Handbook of Music and Advertising*. https://doi.org/10.1093/oxfordhb/9780190691240.013.29

Allen, R. J., Hitch, G. J., & Baddeley, A. D. (2018). Exploring the sentence advantage in working memory: Insights from serial recall and recognition. *Quarterly Journal of Experimental Psychology*, *71*(12), 2571–2585. https://doi.org/10.1177/1747021817746929

Alloway, T. P., & Alloway, R. G. (2013). *Working memory: The connected intelligence*. Psychology Press.

Anglada-Tort, M., Harrison, P. M. C., Lee, H., & Jacoby, N. (2023). Large-scale iterated singing experiments reveal oral transmission mechanisms underlying music evolution. *Current Biology: CB*, *33*(8), 1472–1486.e12. https://doi.org/10.1016/j.cub.2023.02.070

Baddeley, A. D., & Hitch, G. (1974). Working Memory. In G. H. Bower (Ed.), *Psychology of Learning and Motivation* (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Bailes, F. (2010). Dynamic melody recognition: Distinctiveness and the role of musical expertise. *Memory & Cognition*, *38*(5), 641–650. https://doi.org/10.3758/MC.38.5.641

Bainbridge, W. A., Walther, D. B., Fukuda, K., & Goetschalckx, L. (2026). Memorability of visual stimuli and the role of processing efficiency. *Nature Reviews Psychology*, *5*(1), 47–58. https://doi.org/10.1038/s44159-025-00512-3

Baker, D. (2021). MeloSol Corpus. *Empirical Musicology Review*, *16*, 106–113. https://doi.org/10.18061/emr.v16i1.7645

Barber, S. J., Harris, C. B., & Rajaram, S. (2015). Why two heads apart are better than two heads together: Multiple mechanisms underlie the collaborative inhibition effect in memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(2), 559–566. https://doi.org/10.1037/xlm0000037

Beaty, R. E., Frieler, K., Norgaard, M., Merseal, H. M., MacDonald, M. C., & Weiss, D. J. (2021). Expert musical improvisations contain sequencing biases seen in language production. *Journal of Experimental Psychology*. https://doi.org/10.1037/xge0001107

Bella, S. D., Peretz, I., & Aronoff, N. (2003). Time course of melody recognition: A gating paradigm study. *Perception & Psychophysics*, *65*(7), 1019–1028. https://doi.org/10.3758/bf03194831

Bellezza, F. S. (1984). The self as a mnemonic device: The role of internal cues. *Journal of Personality and Social Psychology*, *47*, 506–516. https://doi.org/10.1037/0022-3514.47.3.506

Berkowitz, S., Fontrier, G., Goldstein, P., & Smaldone, E. (2017). *A new approach to sight singing* (Sixth edition). New York: W. W. Norton & Company.

Bestgen, A.-K., Edler, D., Müller, C., Schulze, P., Dickmann, F., & Kuchinke, L. (2017). Where Is It (in the Map)? Recall and Recognition of Spatial Information. *Cartographica: The International Journal for Geographic Information and Geovisualization*, *52*(1), 80–97. https://doi.org/10.3138/cart.52.1.3636

Chubala, C. M., Guitard, D., Neath, I., Saint-Aubin, J., & Surprenant, A. M. (2020). Visual similarity effects in immediate serial recall and (sometimes) in immediate serial recognition. *Memory & Cognition*, *48*(3), 411–425. https://doi.org/10.3758/s13421-019-00979-5

Cuddy, L. L., Cohen, A. J., & Miller, J. (1979). Melody recognition: The experimental application of musical rules. *Canadian Journal of Psychology / Revue Canadienne de Psychologie*, *33*(3), 148–157. https://doi.org/10.1037/h0081713

Dowling, W. J. (1986). Context Effects on Melody Recognition: Scale-Step versus Interval Representations. *Music Perception*, *3*(3), 281–296. https://doi.org/10.2307/40285338

Dowling, W. J., Kwak, S., & Andrews, M. W. (1995). The time course of recognition of novel melodies. *Perception & Psychophysics*, *57*(2), 136–149. https://doi.org/10.3758/bf03206500

Dreyfus, L., Crawford, T., Müllensiefen, D., & Baker, D. (2016). Recognition of Leitmotives in Richard Wagner's Music: An Item Response Theory Approach. In A. F. X. Wilhelm & H. A. Kestler (Eds.), *Analysis of Large and Complex Data* (pp. 473–483). Springer International Publishing. Retrieved from https://www.springer.com/gb/book/9783319252247

Fernandes, M. A., Wammes, J. D., & Meade, M. E. (2018). The Surprisingly Powerful Influence of Drawing on Memory. *Current Directions in Psychological Science*, *27*(5), 302–308. https://doi.org/10.1177/0963721418755385

Fraedrich, J. P., & King, M. F. (1998). Marketing Implications of Nonmusical Sounds. *Journal of Business and Psychology*, *13*(1), 127–139. https://doi.org/10.1023/A:1022931302341

Gardiner, J. M., Kaminska, Z., Java, R. I., Clarke, E. F., & Mayer, P. (1990). The Tulving-Wiseman law and the recognition of recallable music. *Memory & Cognition*, *18*(6), 632–637. https://doi.org/10.3758/BF03197105

Gillund, G., & Shiffrin, R. M. (1984). A retrieval model for both recognition and recall. *Psychological Review*, *91*(1), 1–67. https://doi.org/10.1037/0033-295X.91.1.1

Harrison, P. M. C. (2020). psychTestR: An R package for designing and conducting behavioural psychological experiments. *Journal of Open Source Software*, *5*(49), 2088. https://doi.org/10.21105/joss.02088

Harrison, P. M. C., Bianco, R., Chait, M., & Pearce, M. T. (2020). PPM-Decay: A computational model of auditory prediction with memory decay. *PLOS Computational Biology*, *16*(11), e1008304. https://doi.org/10.1371/journal.pcbi.1008304

Harrison, P. M. C., Musil, J. J., & Müllensiefen, D. (2016). Modelling melodic discrimination tests: Descriptive and explanatory approaches. *Journal of New Music Research*, *45*(3), 265–280. https://doi.org/10.1080/09298215.2016.1197953

Herff, S. A., Olsen, K. N., & Dean, R. T. (2018). Resilient memory for melodies: The number of intervening melodies does not influence novel melody recognition. *The Quarterly Journal*

of Experimental Psychology, *71*, 1150–1171. https://doi.org/10.1080/17470218.2017.1318932

Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*(2), 179–185. https://doi.org/10.1007/BF02289447

Huron, D. (1989). Music in Advertising: An Analytic Paradigm. *Musical Quarterly*, *73*, 557–574. https://doi.org/10.1093/mq/73.4.557

Jacoby, N., Undurraga, E. A., McPherson, M. J., Valdés, J., Ossandón, T., & McDermott, J. H. (2019). Universal and Non-universal Features of Musical Pitch Perception Revealed by Singing. *Current Biology*, *29*(19), 3229–3243.e12. https://doi.org/10.1016/j.cub.2019.08.020

Karlsen, P. J., & Snodgrass, J. G. (2004). The word-frequency paradox for recall/recognition occurs for pictures. *Psychological Research*, *68*(4), 271–276. https://doi.org/10.1007/s00426-003-0138-5

Kim, J. W., Salamon, J., Li, P., & Bello, J. P. (2018). Crepe: A Convolutional Representation for Pitch Estimation. *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 161–165. https://doi.org/10.1109/ICASSP.2018.8461329

Koriat, A., & Goldsmith, M. (1994). Memory in naturalistic and laboratory contexts: Distinguishing the accuracy-oriented and quantity-oriented approaches to memory assessment. *Journal of Experimental Psychology: General*, *123*(3), 297–315. https://doi.org/10.1037/0096-3445.123.3.297

Krishnan Palghat, V. (2009). *Hearing, Remembering, and Branding: Guidelines for Creating Sonic Logos* (PhD thesis, University of Cincinnati). University of Cincinnati. Retrieved from https://etd.ohiolink.edu/apexprod/rws_olink/r/1501/10?clear=10&p10_accession_num=ucin1242848683

Krishnan, V., & Kellaris, J. J. (2021). Hearing, Remembering, and Branding. In *The Oxford Handbook of Music and Advertising*. https://doi.org/10.1093/oxfordhb/9780190691240.013.26

Krishnan, V., Kellaris, J., & Aurand, T. (2012). Sonic logos: Can sound influence willingness to pay? *Journal of Product & Brand Management*, *21*, 275–284. https://doi.org/10.1108/10610421211246685

Krumhansl, C. (1990). *Cognitive Foundations of Musical Pitch*. Oxford University Press, USA.

Li, Y., Masitah, A., & Hills, T. T. (2020). The Emotional Recall Task: Juxtaposing recall and recognition-based affect scales. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(9), 1782–1794. https://doi.org/10.1037/xlm0000841

Mandler, G., Goodman, G. O., & Wilkes-Gibbs, D. L. (1982). The word-frequency paradox in recognition. *Memory & Cognition*, *10*(1), 33–42. https://doi.org/10.3758/BF03197623

Meade, M. E., Ahmad, M., & Fernandes, M. A. (2020). Drawing pictures at encoding enhances memory in healthy older adults and in individuals with probable dementia. *Aging, Neuropsychology, and Cognition*, *27*(6), 880–901. https://doi.org/10.1080/13825585.2019.1700899

Mongeau, M., & Sankoff, D. (1990). Comparison of musical sequences. *Computers and the Humanities*, *24*(3), 161–175. https://doi.org/10.1007/BF00117340

Müllensiefen, D. (2009). *FANTASTIC: Feature ANalysis Technology Accessing STatistics (In a Corpus; Technical report)*. 37. Retrieved from http://www.doc.gold.ac.uk/isms/m4s/FANTASTIC_docs.pdf

Müllensiefen, D., Egermann, H., & Burrows, S. (2015). *Radio Station Jingles: How statistical learning applies to a special genre of audio logos*. https://doi.org/10.5771/9783845260815-53

Müllensiefen, D., & Frieler, K. (2004). Melodic similarity: Approaches and applications. In S. D. Lipscombe, R. Ashley, R. O. Gjerdingen, & P. Webster (Eds.), *Proceedings of the 8th ICMPC* (pp. 283–289). IL.

Müllensiefen, D., & Frieler, K. (2006). Evaluating Different Approaches to Measuring the Similarity of Melodies. In V. Batagelj, H.-H. Bock, A. Ferligoj, & A. Žiberna (Eds.), *Data Science and Classification* (pp. 299–306). Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-34416-0_32

Müllensiefen, D., & Frieler, K. (2007). Modelling experts' notions of melodic similarity. *Musicae Scientiae*, *11*(1_suppl), 183–210. https://doi.org/10.1177/102986490701100108

Müllensiefen, D., & Halpern, A. (2014). The role of features and context in recognition of novel melodies. *Music Perception*, *31*(5), 418–435. https://doi.org/10.1525/mp.2014.31.5.418

North, A. C., & Hargreaves, D. J. (2010). Music and Marketing. In P. N. Juslin (Ed.), *Handbook of Music and Emotion: Theory, Research, Applications* (p. 0). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199230143.003.0032

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. Cambridge University Press.

Pearl, J. (2009). *Causality*. Cambridge University Press.

Pearl, J., & Mackenzie, D. (2018). *The book of why: The new science of cause and effect*. Penguin UK.

Pfordresher, P. Q., Demorest, S. M., Bella, S. D., Hutchins, S., Loui, P., Rutkowski, J., & Welch, G. F. (2015). Theoretical perspectives on singing accuracy: An introduction to the special issue on singing accuracy (part 1). *Music Perception*, *32*(3), 227–231. https://doi.org/10.1525/mp.2015.32.3.227

Qualtrics. (2018). Provo, Utah, USA: Qualtrics.

Raaijmakers, J. G. W., & Shiffrin, R. M. (1992). Models for Recall and Recognition. *Annual Review of Psychology*, *43*(1), 205–234. https://doi.org/10.1146/annurev.ps.43.020192.001225

Revsine, C., Goldberg, E., & Bainbridge, W. A. (2025). The memorability of voices is predictable and consistent across listeners. *Nature Human Behaviour*, *9*(4), 758–768. https://doi.org/10.1038/s41562-025-02112-w

Rhodes, S., Greene, N. R., & Naveh-Benjamin, M. (2019). Age-related differences in recall and recognition: A meta-analysis. *Psychonomic Bulletin & Review*, *26*(5), 1529–1547. https://doi.org/10.3758/s13423-019-01649-y

Richardson, T. S. (2013). A discovery algorithm for directed cyclic graphs. *arXiv:1302.3599 [Cs]*. Retrieved from http://arxiv.org/abs/1302.3599

Roediger, H. L., & McDermott, K. B. (1995). Creating false memories: Remembering words not presented in lists. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(4), 803–814.

Rumbaugh, C. M., & Landau, J. D. (2018). Recognition and recall performance both benefit from the production effect with content-specific information. *Reading Psychology*, *39*(1), 29–40. https://doi.org/10.1080/02702711.2017.1361494

Schönpflug, U. (Ed.). (2008). *Cultural Transmission: Psychological, Developmental, Social, and Methodological Aspects*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511804670

Schubert, E. (2007). The influence of emotion, locus of emotion and familiarity upon preference in music. *Psychology of Music*, *35*(3), 499–515. https://doi.org/10.1177/0305735607072657

Schulkind, M. D., Posner, R. J., & Rubin, D. C. (2003). Musical Features That Facilitate Melody Identification: How Do You Know It's "Your" Song When They Finally Play It? *Music Perception*, *21*(2), 217–249. https://doi.org/10.1525/mp.2003.21.2.217

Silas, S., Baker, D., & Müllensiefen, D. (2024). Musical Manipulation of Visual Scenes In Video, Film and TV Advertisements: A Large-Scale Investigation into the Implicit Effects of Sonic Branding. *Journal of Advertising Research.*

Silas, S., & Müllensiefen, D. (2023). Learning and recalling melodies: A computational investigation using the melodic recall paradigm. *Music Perception*. https://doi.org/10.1525/mp.2023.41.2.77

Silas, S., Müllensiefen, D., Gelding, R., Frieler, K., & Harrison, P. M. C. (2022). The associations between music training, musical working memory, and visuospatial working memory: An opportunity for causal modeling. *Music Perception*, *39*(4), 401–420. https://doi.org/10.1525/mp.2022.39.4.401

Silas, S., Müllensiefen, D., & Kopiez, R. (2023). Singing Ability Assessment: Development and validation of a singing test based on item response theory and a general open-source software environment for singing data. *Behavior Research Methods*. https://doi.org/10.3758/s13428-023-02188-0

Sloboda, J. A., & Parker, D. H. H. (1985). Immediate recall of melodies. In R. West, P. Howell, & I. Cross (Eds.), *Musical structure and cognition* (pp. 143–167). London: Academic Press.

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, *15*(1), 72–101. https://doi.org/10.2307/1412159

Tulving, E., & Watkins, M. J. (1973). Continuity between Recall and Recognition. *The American Journal of Psychology*, *86*(4), 739–748. https://doi.org/10.2307/1422081

Tulving, E., & Wiseman, S. (1975). Relation between recognition and recognition failure of recallable words. *Bulletin of the Psychonomic Society*, *6*(1), 79–82. https://doi.org/10.3758/BF03333153

Tversky, B. (1973). Encoding processes in recognition and recall. *Cognitive Psychology*, *5*(3), 275–287. https://doi.org/10.1016/0010-0285(73)90037-6

Uitdenbogerd, A. (2002). *Music information retrieval technology*.

Uner, O., & Roediger, H. L. (2022). Do Recall and Recognition Lead to Different Retrieval Experiences? *The American Journal of Psychology*, *135*(1), 33–43. https://doi.org/10.5406/19398298.135.1.03

Wammes, J. D., Meade, M. E., & Fernandes, M. A. (2016). The drawing effect: Evidence for reliable and robust memory benefits in free recall. *Quarterly Journal of Experimental Psychology*, *69*(9), 1752–1776. https://doi.org/10.1080/17470218.2015.1094494

Wu, W.-Y., Fu, C.-S., Huang, H.-S., & Wang, H.-C. (2010). Effects of sound stimuli applied in branding: An empirical study of its antecedents and consequences. *Asian Journal of Business and Accounting*, *3*(1), 29–59.

Yalch, R. F. (1991). Memory in a jingle jungle: Music as a mnemonic device in communicating advertising slogans. *Journal of Applied Psychology*, *76*, 268–275. https://doi.org/10.1037/0021-9010.76.2.268

Appendices

Appendix A

Additional experiment details for melodic recall test

Algorithm to correct vocal range as R code.
SAA signal-to-noise ratio and long note singing exclusion criteria.

At the beginning of the *SAA*, a signal-to-noise ratio is computed for the participant's test environment. Based on previous research indicating performance accuracy dropoff in the audio transcription process (Kim, Salamon, Li, & Bello, 2018), participants with a *SNR* ratio < 14 were not allowed into the main test.

The long note singing categoriser we built is based on the following features computed on the fundamental frequency curve: the mean autocorrelation value of this curve, the dynamic time warp distance between the target tone and the sung pitch curve, the max sung frequency, the minimum sung frequency, the precision of singing (i.e., variance) across all trials and within each trial, a run's test of randomness, and the accuracy of the sung pitch curve with respect to the target note. Signal detection theory analysis suggested that the following way to disqualify participants was optimum: if two or more are were failed on a majority of trials (here, 3), disqualify the participant. This criteria was determined based on a ground truth dataset whereby 26 participants were asked to either deliberately make noise/not sing vs. truly sing long notes. 225 test instances were disqualified on this basis.

Second melodic stimuli set: The Berkowitz Corpus. The second set of melodies (not analysed here) were drawn from (Berkowitz, Fontrier, Goldstein, & Smaldone, 2017) (see also Baker (2021)), a corpus of melodies designed to improve sight-singing (singing from musical notation). The book consists of a collection of itemised melodic sequences. We took the first 629 sequences in the book and split them into N-grams of length 3 to the the length of the melody - 1, resulting in a stimulus item bank of 627,021 unique N-grams.

Appendix B

Table B1. Melodic features used in the experiment.

Feature	Description	Equation	Reference
target_melody_length	Number of notes in the target melody.	N	Müllensiefen (2009)
i.entropy	Interval entropy: uncertainty of successive pitch intervals.	$-\frac{\sum_i f_i \log_2 f_i}{\log_2 23}$	Müllensiefen (2009)
span	Pitch span: difference between highest and lowest MIDI pitch.	$\max(p) - \min(p)$	Müllensiefen (2009)
tonalness	Strength of implied key: maximum correlation with major/minor key profiles.	$\max\big(\text{corr}(\text{key profiles}, melody)\big)$	Krumhansl (1990); Müllensiefen (2009)
tonal.clarity	Tonal clarity: ratio between strongest and second-strongest key correlations.	A_0/A_1	Temperley (2007); Müllensiefen (2009)
tonal.spike	Tonal spike: dominance of strongest key correlation over all positive	$A_0/\sum_{i:A_i>0} A_i$	Müllensiefen (2009)

Feature	Description	Equation	Reference
	correlations.		
step.cont.glob.var	Global variation of step contour (standard deviation).	$\sqrt{\\sum_{i=1}^{N-1}(x_i - \bar{x})^2/(N-1)}$	Müllensiefen (2009)
step.cont.glob.dir	Global direction of step contour (correlation with time).	$\mathrm{cor}(x, 1{:}N)$	Müllensiefen (2009)
step.cont.loc.var	Local variation of step contour (mean absolute adjacent difference).	$\frac{1}{N-1}\\sum_{i=1}^{N-1}\\lvert x_{i+1}-x_i \\rvert$	Müllensiefen (2009)
mean_int_size	Mean absolute pitch interval size.	$\frac{1}{N}\\sum_{i=1}^{N}\\lvert \Delta p_i \\rvert$	Müllensiefen (2009)
int_range	Maximum absolute pitch interval.	$\max \\lvert \Delta p \\rvert$	Müllensiefen (2009)
d.entropy	Duration entropy: uncertainty of rhythmic values.	$-\frac{\sum_i f_i \log_2 f_i}{\log_2 24}$	Müllensiefen (2009)
d.eq.trans	Proportion of equal-duration transitions.	$\frac{1}{\lvert R \rvert}\sum_i \mathbb{1}[r_i = 1]$	Steinbeck (1982); Müllensiefen (2009)

Appendix C

Principal components analysis of melodic features solution

Table C1. Four-component principal components analysis solution for the melodic feature variables. Component loadings less than 0.30 are presented in grey to facilitate component interpretation.

Melodic Feature	Intervallic Complexity	Breadth of Material	Tonality	Durational Complexity	h2	u2
N	-0.03	0.90	-0.07	-0.05	0.82	0.18
tonalness	0.32	0.20	0.79	0.06	0.76	0.24
tonal.clarity	-0.07	-0.10	0.75	-0.08	0.58	0.42
tonal.spike	-0.08	-0.14	0.83	0.05	0.72	0.28
step.cont.glob.var	0.92	0.21	0.09	0.04	0.90	0.10
step.cont.loc.var	0.67	0.67	-0.04	-0.02	0.89	0.11
i.entropy	0.30	0.77	0.22	0.15	0.76	0.24
d.entropy	0.11	0.21	0.08	0.93	0.93	0.07
d.eq.trans	0.03	0.18	0.04	-0.94	0.91	0.09
mean_int_size	0.90	-0.17	0.04	0.13	0.87	0.13
int_range	0.91	0.12	0.06	0.06	0.85	0.15
span	0.85	0.38	0.08	0.00	0.87	0.13
total_duration	0.06	0.75	-0.08	0.38	0.71	0.29
total_information_content	0.23	0.89	-0.03	-0.06	0.85	0.15
mean_idf	0.20	0.58	0.33	0.20	0.52	0.48

Appendix D

Similarity Measures (opti3)

Having obtained numerical representations of both the sung recalls and the target melodies, the similarities between a target melody and sung recall of that melody, for each trial of each participant, can be calculated using the algorithmic similarity measures described in Müllensiefen and Frieler (2004). The similarity measures we employed complied with two main points already raised by Sloboda and Parker (1985) in the discussion of their methodology of melodic comparison: in most cases - that is especially true for the earlier trials - participants only recall a smaller part of the original melody, which may not even start with the beginning of the original. Thus, a similarity measure (or algorithm) must be chosen that automatically looks for the best alignment of the (short) melodic sequence of the sung recall with the original melodic sequence. An algorithm for the optimal alignment of two symbol sequences that has been widely used in domains such as text retrieval or bio-computing as well as music information retrieval is the so-called edit distance or Levenshtein distance (e.g., Mongeau and Sankoff (1990)). Just like the hand-made scoring techniques employed by Sloboda and Parker (1985), the edit distance calculates the similarity between two symbolic sequences by taking the number of edits (i.e. additions, deletions or replacements) that are necessary to transform one of the sequences into the other, and dividing this figure by the number of symbols in the longer sequence. It thus could be argued that Sloboda and Parker intuitively used a version of the edit distance, evaluating the similarity between the recalls of their participants on the original melody.

We here use the edit distance to calculate similarities with respect to various musical dimensions (i.e., similarities between the pitches sequences, pitch contours, rhythm sequences

and sequences of implied harmonies; see Müllensiefen and Frieler (2004) for technical details). Specifically, we employ the *opti3* measure of melodic similarity (Müllensiefen & Frieler, 2004) as our main dependent variable. *opti3* is a hybrid measure derived from the weighted sum of three individual measures which represent different aspects of melodic similarity. The similarity in interval content is captured by the *ngrukkon* measure that measures the difference of the occurrence frequencies of short pitch sequences (e.g., length 3-8) contained within two melodies (Uitdenbogerd, 2002). Harmonic similarity is measured by the *harmcore* measure. This measure is based on the chords implied by a melodic sequence, taking pitches and durations into account. Implied harmonies are computed using the Krumhansl-Schmuckler algorithm (Krumhansl, 1990) and the harmonic sequences of the two melodies are compared by computing the number of operations necessary to transform one sequence into the other sequence (i.e., the so-called edit distance; Mongeau and Sankoff (1990)). Finally, rhythmic similarity is computed by first categorizing the durations of the notes of both melodies (known as "fuzzification") and then applying the edit distance to measure the distance between the two sequences of durations. The resulting measure of rhythmic similarity is called *rhythfuzz* (Müllensiefen & Frieler, 2004). Based on the perception data collected by Müllensiefen and Frieler (2004), the three individual measures are weighted and combined to form a single aggregate measure of melodic similarity, *opti3*. Hence, *opti3* is sensitive to similarities and differences in three important aspects of melodic perception (pitch intervals, harmonic content, rhythm). We note that all three individual measures (*ngrukkon*, *harmcore*, *rhythfuzz*) can take values between 0 (=no similarity) and 1 (=identity) and are length-normalized by considering the number of elements of the longer melody. See Appendix B to compare the similarity measures and Appendix C for notated examples of development in sung recall performance and a qualitative description of their change in similarity.

It is important to understand that *opti3* is dependent on the length of the two comparison melodies, meaning, it is then also dependent on number of notes to some extent. However, it is distinct from number of note events in that it explicitly evaluates similarity as a weighted average of symbolic a) intervallic, b) rhythmic and c) harmonic representations of a melody. For example, the smaller the count of similar interval patterns, the lower the similarity score. Thus, *opti3* clearly captures perceptually relevant information specific to intervallic, harmonic as well as rhythmic structure and beyond length. The ecological validity of the aggregate similarity measure has been established in several perceptual experiments (Müllensiefen & Frieler, 2004, 2006, 2007).