What makes background music distracting? Investigating the role of song lyrics using self-paced reading

Martin R. Vasilev[1*]

(ORCID: 0000-0003-1944-8828)

Licia Hitching[1]

Sophie Tyrrell[1]

[1]Bournemouth University, Department of Psychology, Poole, United Kingdom

Corresponding author* at:

Department of Psychology, Bournemouth University

Poole House, Talbot Campus, Fern Barrow

Poole, Dorset, BH12 5BB, United Kingdom

Phone: +441202 961558

Email: mvasilev@bournemouth.ac.uk

Word count: 10405 words

Abstract

It has been suggested that listening to music during reading may be distracting, but the empirical results have remained inconclusive. One limitation of previous studies is that they have often had limited control over the number of lyrics present in the songs. We report 4 experiments that investigated whether song lyrics make music distracting. Participants read short paragraphs in a self-paced reading paradigm in three sound conditions: 1) silence; 2) lyrical songs at ~150 words per minute; and 3) the instrumental version of the same songs. The results showed that listening to instrumental music either did not affect reading times or led to slightly faster reading times compared to silence. However, lyrical music led to an increase in reading times in three experiments. We conclude that instrumental music does not lead to distraction during reading. Song lyrics appear to be distracting, even if the observed distraction is quite mild.

Keywords: reading, music, distraction, lyrics, reading time

Word count: 149

People often listen to music in the background while doing everyday activities. For instance, 62% of university students report listening to music while studying (David, Kim, Brickman, Ran, & Curtis, 2015) and 80% of UK employees report listening to music at work (Haake, 2006). Because this is such a common occurrence, researchers and educators have long been interested in whether listening to music while studying causes distraction (e.g., Henderson et al., 1945; Miller, 1947). While there is some evidence to suggest that music may reduce reading comprehension accuracy (Kämpfe, Sedlmeier, & Renkewitz, 2011; Vasilev, Kirkby, & Angele, 2018), the results have remained mixed and inconclusive. As a result, it is still not well understood whether music is distracting, or which factors are responsible for the observed distraction. One limitation of previous studies is that they have often had limited control over the number of lyrics present in the songs. The present research attempted to find out whether song lyrics are a key contributor to distraction by music.

Distraction by Background Music During Reading

To study the effect of music on reading, researchers have typically presented background music to participants while they are engaged in a reading comprehension task. If participants show reduced comprehension when exposed to music compared to a silence baseline, this is then taken as evidence that music is distracting. While such studies have been conducted for more than 80 years (e.g., Fendrick, 1937; Henderson et al., 1945; L. R. Miller, 1947; A. H. Mitchell, 1949), it has remained frustratingly difficult to draw firm conclusions about what effect, if any, music has on reading comprehension. While some studies have shown certain types of music to be distracting (Anderson & Fuller, 2010; Avila, Furnham, & McClelland, 2012; Daoussis & Mc Kelvie, 1986; Doyle & Furnham, 2012; Etaugh & Michals, 1975; Etaugh & Ptasnik, 1982; Fendrick, 1937; Fogelson, 1973; Furnham & Bradley, 1997; Furnham & Strbac, 2002; Henderson et al., 1945; Johansson, Holmqvist, Mossberg, & Lindgren, 2012; Martin, Wogalter, & Forlano, 1988, Experiment 2; Perham & Currie, 2014;

Quan & Kuo, 2022), others have found that it either has no effect on reading (Cauchard, Cane, & Weger, 2012; Chitwood, 2018; Freeburne & Fleischer, 1952; Furnham & Allass, 1999; Furnham, Trew, & Sneade, 1999; Gillis, 2010; Kelly, 1994; Kou, McClelland, & Furnham, 2018; Madsen, 1987; Martin et al., 1988, Experiment 1; L. R. Miller, 1947; A. H. Mitchell, 1949; Tucker & Bushman, 1991), or that it actually improves reading performance (Falcon, 2017; Hall, 1952; Kiger, 1989; Mullikin & Henk, 1985; Que, Zheng, Hsiao, & Hu, 2020).

Reviews of the literature have often painted a similarly mixed picture. For instance, de la Mora Velasco and Hirumi (2020) conducted a systematic review on the effect of background music on learning and found the results to be inconclusive. They did, however, note the need to develop studies with more rigourous methods and to improve the overall reliability of measures. Hallam and MacDonald (2016) reviewed issues surrounding the literature and noted many structural, cultural, and associative influences that may play a role in explaining the effect of music on task performance. They proposed a theoretical framework that considers the characteristics of the music (e.g., genre, familiary, preference, complexity, level of stimuliation), individual characteristics (e.g., personality, musical expertese, frequency of music use), as well as different emotional, arousal, mood, task, and environmental charestistics. Clearly, taking all these issues into account is difficult, which may well explain why the research literature has been so inconsistent.

Nevertheless, there have been attempts to look at some of these factors in isolation. For example, previous studies haved considered the effect of music genres (Kallinen, 2002; C. Miller, 2014; L. K. Miller & Schyb, 1989; Mullikin & Henk, 1985; Tucker & Bushman, 1991), tempo (Kallinen, 2002; Thompson, Schellenberg, & Letnic, 2012), preference (Etaugh & Michals, 1975; Etaugh & Ptasnik, 1982; Johansson et al., 2012; Perham & Currie, 2014), and familiarity of the music (Chew, Yu, Chua, & Gan, 2016; Hilliard & Tolin, 1979). While such studies have provided interesting initial results, more evidence is required to reach firmer conclusions.

Reseach on individual differences, most notably differences between extraversion and introversion, has received far more attention (Avila et al., 2012; Daoussis & Mc Kelvie, 1986; Furnham & Allass, 1999; Furnham & Bradley, 1997; Furnham & Stephenson, 2007; Furnham & Strbac, 2002; Furnham et al., 1999; Gheewalla, McClelland, & Furnham, 2020; Kou et al., 2018; Lim, Furnham, & McClelland, 2022). Based on Eysenck's (1967) personality theory, it has been been predicted that extraverts will be distracted less by background music compared to introverts due to their higher cortical arousal. However, a recent review on the topic conlcuded that there is as much evidence in support of this hypothesis as there is against it (Küssner, 2017). In summary, the literature has suggested that music may cause distraction during reading, but the results have remained mixed and more evidence is required to understand when such distraction may occur.

The Effect of Lyrics on Distraction by Background Music

Meta-analyses have attempted to address some of these inconsistencies by pooling together all the available evidence and deriving a single estimate on the effect of music on reading. Kämpfe et al. (2011) reported an effect size of r= -11 (d= -0.22) based on 8 studies, indicating that music has a mild distracting effect. Vasilev et al. (2018) conducted a meta-analysis with 36 studies and found a similar result: the overall effect of music on reading comprehension was d= –0.19, again indicating mild distraction. Interestingly, however, a meta-regression analysis suggested that studies using lyrical music yielded much bigger distraction effects than studies using instrumental music (a mean difference of d= -19). While the effect size for lyrical music was d= -0.35, the effect size for instrumental music was effectively 0. Additionally, lyrical music was found to be just as distracting as intelligible background speech. In summary, Vasilev et al.'s (2018) findings suggest that lyrical music is distracting but that instrumental music does not cause any distraction.

Previous studies that have directly compared lyrical and instrumental music also lend some support to these results. For example, Martin et al. (1988) reported that the presence of sung or spoken lyrics (either accompanied by instrumentals or not) led to greater distraction compared to a no-lyrics condition. Additionally, Perham and Currie (2014) reported that both liked and dysliked lyrical music was more distracting compared to instrumental music. Similarly, C. Miller (2014) presented classical and rock music that was either instrumental or lyrical. There was a marginally significant main effect of lyrics (lower comprehension in lyrical compared to instrumental music) and a significant interaction with genre- with the means suggesting a bigger difference between the lyrical and instrumental conditions for classical music. However, Gillis (2010) reported no difference between instrumental classical music and lyrical pop music. Avila et al. (2012) also found no difference in comprehension between the lyrical and instrumental version of the same songs, though performance in both conditions was significantly worse than silence. Similarly, Furnham et al. (1999) also found no difference in comprehension between lyrical and instrumental music, but the two conditions also did not differ from silence. In contrast, Reed (2019) reported that lyrical music led to lower comprehension compared to an instrumental version of the same songs.

More recently, Kyoung (2020) found that lyrical music led to a reduction in the P2 and P600 amplitude compared to silence (corresponding to the orthographic and syntactic stages of text processing, respectively). However, no difference was found between lyrical music and the instrumental version of the same music, or between instrumental music and silence. This suggests that the difference in their study may not be entirely due to the lyrics, but to some combination of lyrics and instrumentals. Therefore, while the results are again far from conclusive, there is at least some indication in the literature that song lyrics may contribute to the observed distraction.

Theoretical Perspectives

The potential of lyrics to cause distraction is not surprising, given that irrelevant speech is well-known to disrupt reading (e.g., Baker & Madell, 1965; Hyönä & Ekholm, 2016; Martin et al., 1988; Sörqvist et al., 2010; Yan et al., 2018). In fact, recent evidence from eye-tracking has suggested that distraction by music may show very similar eye-movement signature to that of distraction by irrelevant speech (Zhang, Miller, Cleveland, & Cortina, 2018). If the meaning of lyrics is processed in a similar way to that of speech, distraction by lyrical music would also be expected.

There are different theories that can explain why lyrics may be distracting. For instance, the *interference-by-process* account (Jones & Tremblay, 2000; Marsh, Hughes, & Jones, 2008, 2009) predicts that distraction can occur if both the main task and the distractor are drawing on similar cognitive processes, thus leading to interference. Readers need to engage in semantic processing of the text in the main task to achieve sufficient comprehension. However, if the lyrics from the music distractor also undergo semantic processing, this would cause semantic interference and therefore distraction. *Phonological interference* theories (Salamé & Baddeley, 1982, 1987, 1989) predict similar interference, though the proposed mechanism is different: phonemes from the lyrics would gain obligatory access to the phonological loop and interfere with the phonological encoding of the text, thus causing distraction. Finally, the *changing-state* account (Hughes & Jones, 2001; Jones & Macken, 1993; Jones, Madden, & Miles, 1992) may also offer an insight. In this theory, changing-state sounds that exhibit greater acoustic variation are predicted to cause greater distraction compared to aperiodic, steady-state sounds. Because a song with lyrics would have more acoustic variation than the same song without the lyrics, this theory would also predict greater distraction by lyrical music. Therefore, while all three theories offer different explanations for why distraction occurs, they all agree that lyrical music should be more distracting than instrumental music.

Present Research

Vasilev et al.'s (2018) meta-regression results, as well as findings from previous studies (Kyoung, 2020; Martin et al., 1988; Perham & Currie, 2014; Reed, 2019), suggest that song lyrics may be an important contributor to distraction by music. Nevertheless, there are few well-controlled studies that have investigated the role of lyrics in distraction. More broadly, studies have often had limited control over the acoustical properties of the music conditions that are being compared and the number of lyrics present in the songs. Therefore, the aim of the present research was to test whether lyrics are a key contributor to distraction in a more controlled manner.

Participants read short passages in three sound conditions: silence, instrumental music, and lyrical music. We used the lyrical and instrumental version of the same songs (Avila et al., 2012; Furnham et al., 1999; Kyoung, 2020), thus ensuring that any difference between the two conditions can be attributed solely to the presence of lyrics. To maximise the amount of distraction, songs that have an average lyrics rate of about 140-150 words per minute were used, about the same rate as normal speech (Brysbaert, 2019). Additionally, participants were asked to provide ratings on the familiarity, preference, pleasantness, offensiveness, and perceived distractibility of the music, as well as their own daily music usage, so that the influence of these variables on the results can be examined.

The present research used a self-paced reading paradigm (Aaronson & Scarborough, 1976; Jegerski, 2014; Marsden, Thompson, & Plonsky, 2018), where participants pressed a button to reveal each new word in the text. This made it possible to calculate reaction times for each word in the text, as well as to measure overall comprehension accuracy at the end. This methodology was preferred because previous research has suggested that word-level reading times (measured with eye-tracking) may be a more sensitive predictor of distraction by irrelevant speech than comprehension accuracy (Cauchard et al., 2012; Hyönä & Ekholm, 2016; Meng, Lan, Yan, Marsh, & Liversedge, 2020; Vasilev, Liversedge, Rowan, Kirkby, & Angele, 2019; Yan et al., 2018). Thus, the self-paced reading paradigm made it possible to calculate word reading times in a more cost-effective way, particularly for research online during the Covid-19 pandemic. We expected word reading times to be a more sensitive measure of distraction than comprehension accuracy. Therefore, all key predictions are based on word reading times.

We report 4 experiments. Experiment 1a examined the effect of song lyrics when participants listened to familiar pop/rap songs in an online study. Experiment 1b repeated the same study in the lab. Experiments 2-3 examined the effect of unfamiliar pop/rap music in an online study. We expected that lyrical music will lead to an increase in reading times compared to instrumental music. If one takes Vasilev et al.'s (2018) results at face value, it can be predicted that there should be no difference between silence and the instrumental music condition. However, it is also possible that there could be a small difference between the two conditions if music instrumentals also contribute to the distraction.

Hypotheses

- **H1**: If the presence of lyrics makes background music distracting, lyrical music should result in longer self-paced reading times compared to instrumental music.

- **H2.1**: If lyrics are the only aspect of background music that causes distraction, then: 1) H1 should be supported; and 2) there should be NO difference in self-paced reading times between instrumental music and the silence baseline.
- **H2.2**: If instrumental music also causes at least some distraction, self-paced reading times should be longer in the instrumental music condition compared to the silence baseline.

Experiment 1a

The study protocol was pre-registered prior to data collection (https://osf.io/4gw63). Experiment 1a was conducted online due to the Covid-19 pandemic. Previous research has suggested that lab-based and online-based studies of distraction should yield comparable results (Elliott, Bell, Gorin, Robinson, & Marsh, 2022).

Method

Participants. Participants were recruited from two sources: a local university pool and Prolific.co. All participants were UK adults who reported English as their first language, normal (or corrected-to-normal) vision, normal hearing, and no prior diagnosis of reading disorders. University pool participants received course credits and Prolific participants were compensated at £7/ hour. Overall, 204 participants[1] took part (65.68% female, 33.33% male, 0.98% other genders; *N*= 101 Prolific; *N*= 103 university pool). Participants' average age was 25.32 years (*SD*= 7.87 years; range= 18- 49 years). In the university pool, 94.2% participants had completed A-levels (≈ high school) and 5.8% indicated they had already studied for

[1] 11 more participants were tested but excluded based on the pre-registered criteria: 3 failed one or more of the listening comprehension "trap" trials, 3 admitted to not wearing headphones, and 5 were discarded due to missing or invalid data. Additionally, 5 more participants were excluded due to chance-level comprehension (<60%). While the comprehension accuracy criterion was not pre-registered, it was deemed necessary to ensure that participants were reading for comprehension.

(another) undergraduate degree. In the Prolific participants, 1% had completed primary school, 9.9% GCSEs, 26.73% A-levels, 46.53% an undergraduate degree, 12.87% a postgraduate degree, and 2.97% a PhD degree[2]. All experiments received ethical approval from the Bournemouth University Research Ethics Committee (ID: 36794). All participants provided informed electronic consent.

Prospective statistical power simulations using the *simr* R package v.1.0.5 (Green & Macleod, 2016) were done on a pilot dataset of 12 subjects (not included here). The simulation parameters were: 1) sample size that can detect a difference between the lyrical and instrumental music conditions with a 95% probability; 2) the expected effect size (a 12 ms difference) was reduced to 75% of that from the pilot data because small studies are known to overestimate the effect size (Albers & Lakens, 2018); 3) 10% random data loss was added to account for missing data and outliers. The results indicated that 156 participants are needed to reach 95% power. To be sure, the sample size was increased to 204 participants. The power simulations indicated no reliable difference between silence and instrumental music. A Bayesian model was found to be sufficiently precise with this sample size to find evidence in support of H0 for this effect.

Design and materials. The study had a within-subject design with *sound condition* (silence, instrumental music, lyrical music) as the only factor. The reading stimuli consisted of 15 short passages (see Figure 1a) from the Provo corpus (Luke & Christianson, 2018). The passages were on average 53.4 words long (*SD*= 4.42 words; range: 46-62 words). The words in each passage were presented one-by-one using a self-paced reading paradigm (Aaronson & Scarborough, 1976; Jegerski, 2014; D. C. Mitchell & Green, 1978). A non-cumulative

[2] Prolific participants were, on average, more educated than the university pool, though they also had more varied educational backgrounds. They were also older (M_{age}= 30.7 years) than the university pool participants (M_{age}= 20 years). The Prolific participants also had a more balanced gender representation (48.5% female) compared to the university pool participants (82.5% female).

presentation was used, meaning that only the current word was visible at any given time and all other words were masked. This made it possible to calculate a reaction time for each word, which roughly corresponds to the time participants spent processing it (including integrating it with previously read material). After each passage, participants answered 2 True/ False comprehension questions (see Figure 1c).

While reading, participants were exposed to the three background sound conditions. The music played in the experiment consisted of six pop/ rap songs. To avoid presenting any of the songs twice (once in the lyrical, and once in the instrumental condition), they were split into two sets. Half of the participants heard Set A in the lyrical music condition and Set B in the instrumental music condition; the other half heard Set A in the instrumental music condition and Set B in the lyrical music condition. Thus, the two sets were heard equally often across all participants and conditions, but participants heard each song only once. The songs in Set A were: 1) Eminem- The way I am; 2) Post Malone- WoW; 3) Nicki Minaj (feat. Rihanna)- Fly. The songs in Set B were: 1) Jessie J (feat. B.o.B)- Price tag; 2) Iggy Azalea (ft. Charli XCX)- Fancy; 3) Outkast- Ms. Jackson. The songs were always played in the same order. The songs were selected based on their high lyrics content and the availability of an officially released instrumental version that was identical to the original song. The songs in Set A had an average lyrics rate of 148.2 words per minute (*SD*=37.7) and the songs in Set B had an average lyrics rate of 145.8 words per minute (SD= 6.4). There were no significant differences in lyrics rate between the two sets, $t(2.11) = 0.11$, $p = 0.9219$.

The sound conditions were blocked, and the order of blocks was counterbalanced across participants. Within each block, the five passages were presented in random order. The assignment of sound conditions to the passages was counter-balanced with a full Latin square design. At the start and end of a block, participants were presented with a listening comprehension "trap" trial, which was designed to catch participants who were not listening to the audio (see Figure 1d). During those trials, participants heard a spoken statement (e.g., "A cat sits on a bed") and had to choose the picture corresponding to the statement (e.g., A cat sitting on a bed vs. a cat sitting on a table).

After the reading task, participants completed a short questionnaire about their listening habits and the music played in the experiment. First, they were asked about their preferred music genre(s) and their average daily time spent listening to music. Second, participants were presented with a 30s sample of all songs used in the experiment and were asked to rate them on their familiarity, preference, pleasantness, offensiveness, and perceived distractibility on a scale from 1 (not at all familiar/ likable/ pleasant, offensive/ distracting) to 10 (very familiar/ likable/ pleasant, offensive/ distracting) (see Perham & Currie, 2014; Perham & Sykora, 2012). The song samples always started at the first chorus. Participants were also asked to write down the name of the artist(s) and the song title (if they know them) to test their actual knowledge of the songs.

A) Example passage

```
I agree that California's "three strikes and you're out" law will be a financial
disaster for taxpayers who care about education and other vital services. But
it's far from clear that the law can even be credited with a reduction in crime
in California. While it's true that crime declined in California last year, crime
also dropped nationwide.
```

B) Self-paced reading illustration

```
I ----- ---- ------------ ------ ------- --- ------ ---- --- ---- -- - ---------
- agree ---- ------------ ------ ------- --- ------ ---- --- ---- -- - ---------
- ----- that ------------ ------ ------- --- ------ ---- --- ---- -- - ---------
- ----- ---- California's ------ ------- --- ------ ---- --- ---- -- - ---------
- ----- ---- ------------ "three ------- --- ------ ---- --- ---- -- - ---------
```

C) Comprehension assessment

```
1) The author agrees that the "three strikes and you're out" law will help
   taxpayers. TRUE/FALSE?
2) Last year, crime declined in California, but not nationwide. TRUE/FALSE?
```

D) Example trap trial

Click on the picture that describes what you hear:

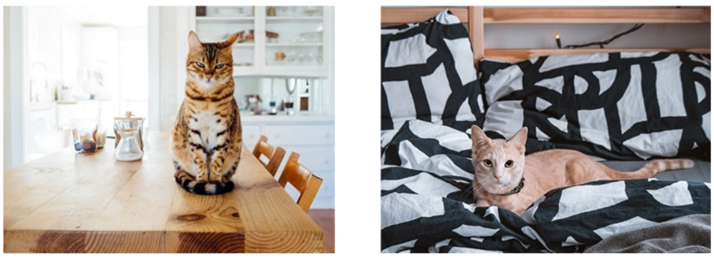

Figure 1. An illustration of the materials used in the experiments.

Apparatus. The experiment was programmed in Lab.js (Henninger, Shevchenko, Mertens, Kieslich, & Hilbig, 2022) and hosted online on Pavlovia.org. The passages were formatted in a Consolas mono-spaced font and appeared as black text over a white background. The width of each letter was set to be 2% of the width of the browser window size. The text was double-spaced and aligned to the left. Participants completed the experiment on their own laptop/ PC using headphones.

Procedure. Participants read the information sheet, provided electronic consent, and were forwarded to the online experiment, which started in full-screen mode on their browser. They were instructed to put on their headphones and perform a headphone screening and calibration procedure (Woods, Siegel, Traer, & McDermott, 2017). Participants were instructed to set the volume to a loud, but comfortable level. Following instructions, participants were presented with 2 practice trials, followed by the experimental trials (blocked by sound condition). In each block, the music started playing 15s before the first trial to allow participants to get used to it. Each trial started with only the first word visible and a prompt at the top of the screen reminding participants to press the SPACE bar to reveal the next word. Once the SPACE bar was pressed, the current word disappeared, and the next word was revealed. This procedure was repeated until the whole paragraph was presented (see Figure 1b). There were 2 comprehension questions after each paragraph, which were answered with a mouse click. Participants were given a maximum of 45s per question. There was a 7-second break between trials. Before and after each sound block, one listening comprehension "trap" trial was presented. After the reading task, participants completed the music questionnaire and were asked if they wore headphones for the whole duration of the experiment.

Data analysis. There were two dependent measures: word *reaction time* (RT; time taken to press the button to move to a new word) and *comprehension accuracy*. RT was the main measure of interest. Statistical analysis was done with (Generalised) Linear Mixed Models ((G)LLMs) using the "lme4" package v.1.1-29 (Bates, Machler, Bolker, & Walker, 2014) in R v.4.10 (R Core Team, 2022). Random intercepts were included for both participants and items (Baayen, Davidson, & Bates, 2008). Additionally, we attempted to include random slopes for sound condition for both participants and items (Barr, Levy, Scheepers, & Tily, 2013). If the models failed to converge, the slopes were removed one by one until convergence was achieved. Reaction times were log-transformed in the analysis. Successive differences contrast was used from the MASS package (Venables & Ripley, 2002), which compared instrumental music to silence and lyrical music to instrumental music. The results were considered statistically significant if the $|t|$ and $|z|$ values were ≥ 1.96. Effect sizes are reported in Cohen's *d* (Cohen, 1988).

Additionally, Bayesian (G)LMMs were fit using the same model structure to calculate Bayes factors (BF_{10}). This was done with the *brms* R package v.2.16.1 (Bürkner, 2017, 2018) using the Stan software (Carpenter et al., 2017). Four chains were run with 5000 iterations each and 500 samples burn-in. Bayes factors (BF_{10}) were calculated using the Savage-Dickey density ratio method (Dickey & Lientz, 1970; Morey, Rouder, Pratte, & Speckman, 2011). In the reaction time model, priors of Normal(0, 6) and Normal(0, 0.05) were used for the intercept and slopes, respectively. The slope prior roughly corresponds to a maximum expected difference of 20-30 ms on the log scale. In the accuracy model, priors of Normal(0, 2) and Normal(0, 0.75) were used for the intercept and slopes, respectively. The slope prior roughly corresponds to a maximum expected difference of 10-12% on the logit scale.

Results

During pre-processing, 0.32% of the data was excluded due to outliers (RTs < 100 or > 5000ms). Additionally, 4 trials (0.14%) were removed due to a lack of response on more than 5 words. This left 99.54% of the data for analysis. Descriptive statistics are shown in Table 1 and visualised in Figure 2. The results from the statistical analysis are shown in Tables 2-3.

Table 1

Descriptive Statistics for the Reaction Time and Comprehension Accuracy Measures

Sound	Mean reaction time in ms per word (SD)	Mean comprehension accuracy in % (SD)
Experiment 1a (familiar music, online)		
Silence	405 (241)	87.0 (33.6)
Instrumental music	403 (239)	85.1 (35.6)
Lyrical music	410 (257)	84.7 (36.0)
Experiment 1b (familiar music, lab replication)		
Silence	452 (255)	85.3 (35.4)
Instrumental music	451 (248)	84.0 (36.6)
Lyrical music	464 (262)	80.9 (39.3)
Experiment 2 (unfamiliar music, online)		
Silence	377 (248)	87.1 (33.5)
Instrumental music	365 (219)	87.1 (33.5)
Lyrical music	381 (240)	85.6 (35.1)
Experiment 3 (unfamiliar music + speech, online)		
Silence	370 (210)	87.7 (32.9)
Instrumental music	367 (201)	88.2 (32.2)
Lyrical music	376 (211)	85.9 (34.8)
Irrelevant speech	377 (219)	84.8 (35.9)

Note: *SD*: standard deviation.

There was no significant difference in RTs between instrumental music and silence, with the Bayesian model showing “substantial” evidence in support of the null hypothesis of no difference (Jeffreys, 1961; Wetzels et al., 2011). Similarly, there was also no difference in RTs between lyrical and instrumental music. The Bayesian model again favoured the null hypothesis of no difference, though the evidence for this was only anecdotal. In summary, neither music condition affected word RTs and there was no evidence of any distraction.

The comprehension accuracy measure revealed similar results, with no significant difference between instrumental music and silence or lyrical and instrumental music. The

Bayesian models again favoured the null hypothesis of no difference, though the evidence was substantial only in the comparison between lyrical and instrumental music.

Table 2

LMM Results for Reaction Times in the Experiments

	Experiment 1a				Experiment 1b			
Fixed effects	b	SE	t	BF_{10}	b	SE	t	BF_{10}
Intercept	5.900	.0299	**196.8**		6.028	.0292	**206.5**	
Instrumental vs. Silence	-.006	.0076	-.774	**.2036**	-.0035	.0069	-.513	**.1503**
Lyrical vs. Instrumental	.0107	.0075	1.422	.3920	.0245	.0070	**3.505**	**52.17**
Random Effects	Var.	SD	Corr.		Var.	SD	Corr.	
Intercept (subjects)	.0663	.2575			.0504	.2245		
Instrumental vs. Silence (subjects)	.0110	.1048	.08		.0089	.0947	.11	
Lyrical vs. Instrumental (subjects)	.0107	.1034	- .02	- .45	.0093	.0963	.07	-.43
Intercept (items)	.0086	.0927			.0091	.0952		
Residual	.1040	.3225			.0959	.3097		
	Experiment 2				**Experiment 3**			
Fixed effects	b	SE	t	BF_{10}	b	SE	t	BF_{10}
Intercept	5.818	.0280	**207.9**		5.825	.0242	**240.32**	
Instrumental vs. Silence	-.0175	.0074	**-2.374**	2.145	-.0075	.0019	**-3.842**	**590.32**
Lyrical vs. Instrumental	.0323	.0071	**4.533**	**3.9 x 10^8**	.0187	.0019	**9.630**	**2.5 x10^{16}**
Speech vs Instrumental	N/A	N/A	N/A	N/A	.0011	.0019	0.561	**.0030**
Random Effects	Var.	SD.	Corr.		Var.	SD	Corr.	
Intercept (subjects)	.0687	.2621			.0629	.2508		
Instrumental vs. Silence (subjects)	.0103	.1014	-.22					
Lyrical vs. Instrumental (subjects)	.0096	.0979	.18	-.54				
Intercept (items)	.0067	.0818			.0057	.0754		
Residual	.0998	.3158			.1008	.3174		

Note: Statistically significant t-values are formatted in **bold**. N/A: speech condition was not present in Experiment 2. BF_{10}: Bayes factor comparing the alternative to the null hypothesis;

values <1 indicate evidence in support of the null hypothesis and values > 1 indicate evidence in support of the alternative hypothesis. Bayes factors of <1/3 or >3 are highlighted in **bold**.

Table 3

GLMM Results for Reading Comprehension in the Experiments

	Experiment 1a				Experiment 1b			
Fixed effect	b	SE	z	BF_{10}	b	SE	z	BF_{10}
Intercept	2.161	.2356	**9.173**		1.916	.2259	**8.483**	
Instrumental vs. Silence	-.1861	.0950	-1.959	.5766	-.0793	.1071	-.740	**.2141**
Lyrical vs. Instrumental	-.0327	.0917	-0.355	**.1291**	-.2591	.0968	**-2.677**	**3.967**
Random Effects	Var.	SD	Corr.		Var.	SD	Corr.	
Intercept (subjects)	.4238	.6510			.2220	.4712		
Instrumental vs. Silence (subjects)					.2370	.4868	.36	
Lyrical vs. Instrumental (subjects)					.0404	.2010	-.23	-.34
Intercept (items)	.7644	.8743			.7167	.8466		
	Experiment 2				Experiment 3			
Fixed effect	b	SE	z	BF_{10}	b	SE	z	BF_{10}
Intercept	2.305	.2730	**8.442**		2.308	.2145	**10.76**	
Instrumental vs. Silence	-.0085	.0996	-.085	**.1489**	.1027	.1017	1.010	**.1877**
Lyrical vs. Instrumental	-.1416	.0973	-1.455	.3781	-.2487	.0992	**-2.506**	1.128
Speech vs Instrumental	N/A	N/A	N/A	N/A	-.1296	.0944	-1.373	.3589
Random Effects	Var.	SD	Corr.		Var.	SD	Corr.	
Intercept (subjects)	.2723	.5218			.5005	.7074		
Intercept (items)	1.055	1.027			.8316	.9119		

Note: Statistically significant t-values are formatted in **bold**. N/A: speech condition was not present in Experiment 2. BF_{10}: Bayes factor comparing the alternative to the null hypothesis; values <1 indicate evidence in support of the null hypothesis and values > 1 indicate evidence in support of the alternative hypothesis. Bayes factors of <1/3 or >3 are highlighted in **bold**.

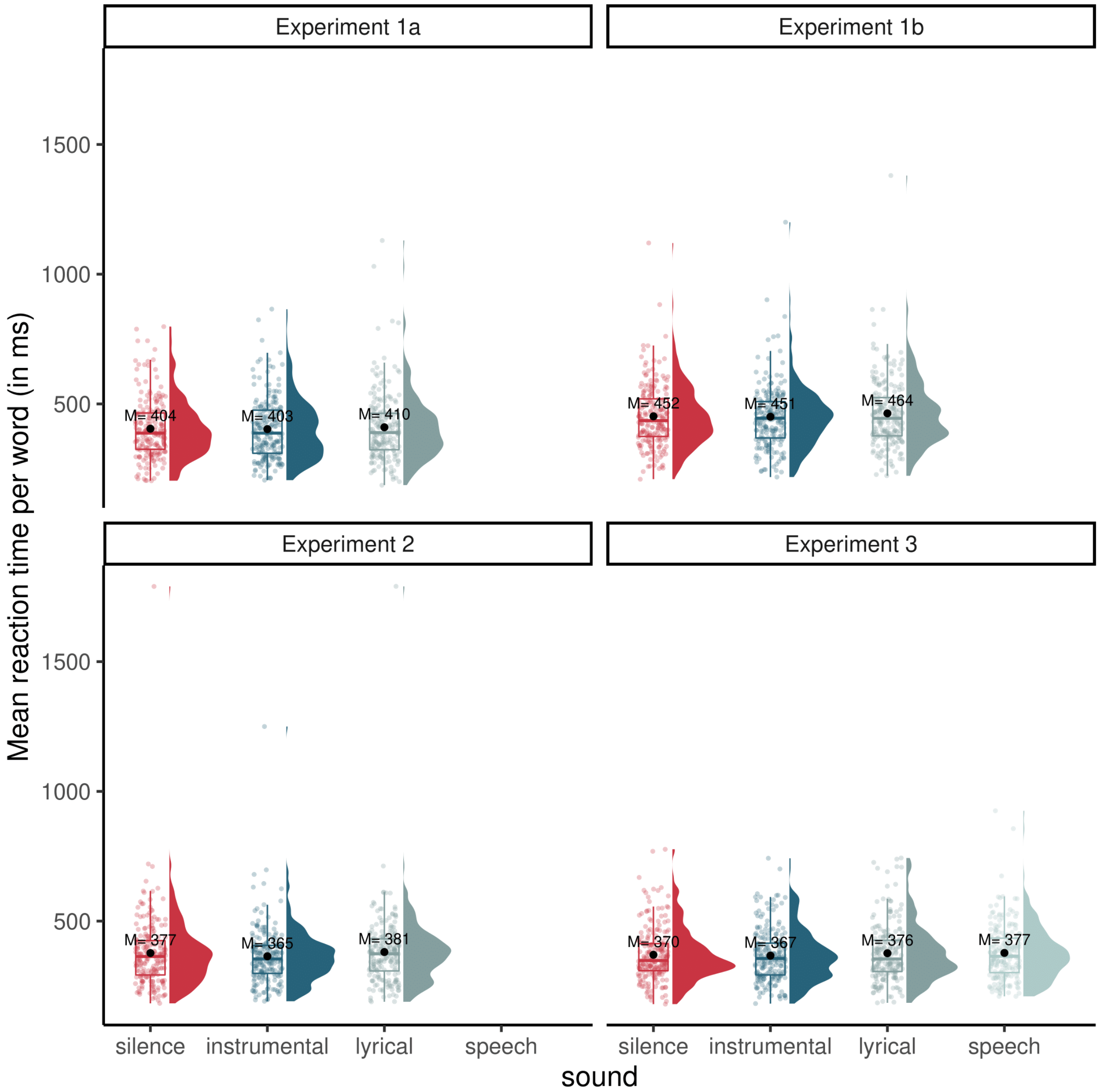

Figure 2. Distribution of word reaction times in the different sound conditions in all four experiments. Dots represent individual participant means for each condition.

Discussion

Experiment 1a did not show any evidence of distraction by either instrumental or lyrical music, thus failing to support any of the study hypotheses. The results generally favoured the null hypothesis of no difference, though the evidence for this was anecdotal in half of the comparisons. Therefore, while the results were not perfectly conclusive, they generally suggest that lyrical and instrumental music do not cause distraction during self-paced reading.

One possible explanation for the lack of difference in Experiment 1a is that the online data collection may have affected the validity of the results. While there is evidence that online testing yields similar (albeit smaller) auditory distraction effects compared to in-person testing (Elliott et al., 2022), this has not been studied in the present paradigm. To evaluate the same hypotheses in standardised conditions, the experiment was repeated in the lab.

Experiment 1b

Method

The study protocol was pre-registered prior to data collection (https://osf.io/6d3fj). The method was the same as Experiment 1a, except for the following differences.

Participants. A total of 204 Bournemouth University students[3] took part in return for course credits (80.88% female; 17.65% male; 0.98% other genders; 0.49% no answer). None of them took part in Experiment 1a. Participants' average age was 20.64 years (*SD*= 5.29 years; range= 18- 50 years). Most participants (92.15%) had completed A-levels, 0.98% had

[3] Another 6 participants were tested but excluded (2 because they failed one or more of the trap trials and 4 due to chance-level accuracy (<60%; accuracy criterion not pre-registered).

completed GCSEs, and 6.86% indicated they had already studied for (another) undergraduate degree.

Design, materials, apparatus, procedure, and data analysis. Participants completed the study in an individual lab cubicle at Bournemouth University. The music was played at 67 ±1.5 dB(A) via Bose QuietComfort 25 noise-cancelling headphones. The study was run on a Chrome web browser on a Hewlett-Packard EliteDesk 800 G1 SFF computer with 8GB RAM (running on Windows 7). The monitor was a 24" BENQ XL2411 with a 1920 x 1080-pixel resolution and a 60 Hz refresh rate. Participants sat about 60-70 cm from the monitor. All other aspects were identical to Experiment 1a.

Results

During pre-processing, 0.15% of the data was excluded due to outliers (RTs < 100 or > 5000ms). Additionally, 1 trial (0.03%) was removed due to a lack of response on more than 5 words. This left 99.85% of the data for analysis. Descriptive statistics are shown in Table 1 and the statistical analyses are shown in Tables 2-3.

Similar to Experiment 1a, there was no significant difference in RTs between instrumental music and silence. The Bayesian model showed substantial evidence in support of the null hypothesis of no difference. However, the lyrical music condition led to significantly longer RTs compared to instrumental music (d= 0.05), thus supporting **H1**. The Bayesian model showed very strong support for the alternative hypothesis that there is a difference between the two conditions. Therefore, lyrical music was more distracting than instrumental music, but instrumental music did not differ from the silence baseline, which supports **H2.1**.

The reading comprehension data showed the same pattern of results. While there was no significant difference between instrumental music and silence, lyrical music led to significantly lower comprehension accuracy compared to instrumental music (*d*= -0.082).

Discussion

Experiment 1b showed that lyrical music led to longer word RTs compared to instrumental music. This suggests that song lyrics caused distraction and increased overall word reading times. Additionally, comprehension accuracy was lower in lyrical compared to instrumental music, replicating previous findings (Martin et al., 1988, Experiment 2; C. Miller, 2014; Perham & Currie, 2014; Reed, 2019). However, no difference emerged in the comparison between instrumental music and silence, suggesting that the presence of instrumentals alone had no effect on self-paced reading. This is consistent with studies showing no difference between instrumental music and silence (e.g., Cauchard et al., 2012; Martin et al., 1988, Experiment 1).

Overall, Experiment 1b was successful in showing reliable distraction by lyrical music in the lab, even if the effect sizes were quite small. Because no significant difference was found in the online version of the study (Experiment1a), it may be tempting to attribute the conflicting results to the mode of testing. However, it is important to keep in mind that Experiment 1a also contained a mixed sample (one half was from a student pool and the other half was from Prolific). A post-hoc comparison of the RT measure between Experiment 1b and the student-pool sub-sample from Experiment 1a revealed an overall difference between lyrical and instrumental music in both experiments, but no interaction with Experiment (see the Supplemental Files). This suggests that the distraction effect was also present in the student sub-sample of Experiment 1a and it did not differ from the lab-based testing conditions in Experiment 1b. Therefore, these results corroborate Elliott et al.'s (2022) finding that lab-based and online-based distraction experiments yield similar results, with the caveat that the same sample population is used.

Experiment 2

Experiment 1 used music that was rated as familiar and roughly half of all participants could correctly identify the songs/ artists (see Figure 4 and Table 4 below). However, it is not known if similar results would be obtained with a set of unfamiliar songs. Therefore, Experiment 2 attempted to replicate the results from Experiment 1b, but with unfamiliar songs.

Song familiarity is an important, but little understood factor. To our knowledge, only two studies have directly examined it. Hilliard and Tolin (1979) presented participants with either "familiar" music (the same music piece presented 15 minutes before the test session) or unfamiliar music (a new, previously unheard piece). They reported that unfamiliar music led to lower comprehension test scores compared to familiar music. Chew et al. (2016) also manipulated the familiarity of the music (along with the language of the songs). However, they used a popular song in the "familiar" condition ("My Heart Will Go On" by Celine Dion) instead of repeating the same song twice. They found no difference between familiar and unfamiliar music in a reading comprehension task (though unfamiliar music reduced word memory test scores compared to familiar music). Therefore, the results are inconclusive as to whether song familiarity plays a role in distraction. Interestingly, some studies have actively avoided using familiar music (e.g., Furnham et al., 1999; Furnham & Allass, 1999; Kyoung, 2020), presumably because it was thought that unfamiliar music will yield stronger distraction. However, the actual impact of music familiarity remains poorly understood.

Experiment 2 used unfamiliar songs that had the same genre(s) and number of lyrics as those in Experiment 1. As a result, it was not a direct test of music familiarity, but an attempt to replicate and extend the results from Experiment 1b to a set of unfamiliar songs. Due to the constraints of the Covid-19 pandemic, Experiments 2-3 were run online.

Method

The study protocol was pre-registered prior to data collection (https://osf.io/b38hn).

Participants. A total of 204 UK adults[4] recruited from Prolific.co participated in return for compensation at £7/ hour (60.3% female; 39.2% male; 0.49% other genders). None of them took part in the previous experiments. Participants had an average age of 31.6 years (*SD*= 11 years; range: 18-50 years). In terms of education, 0.49% had completed primary school, 4.41% had completed GCSEs, 28.43% had completed A-levels, 44.61% had completed an undergraduate degree, 19.61% had completed a postgraduate degree, and 2.45% had completed a PhD degree.

Design, materials, apparatus, procedure, and data analysis. All aspects of the study were identical to Experiment 1a, except that a new set of (unfamiliar) songs was used. Set A contained the following songs: 1) Sa-Roc- Starseed; 2) Johnie Bee ft. Rasco- In My Prime; 3) Evidence- Throw It All Away. Set B contained: 1) The Four Owls- Old Earth; 2) Aesop Rock- Molecules; 3) Atmosphere- Just for show. Set A had an average lyrics rate of 148.9 words per minute (*SD*= 9.6) and set B had an average lyrics rate of 149.1 words per minute (*SD*= 24.1). There were no differences in lyrics rate between Set A and Set B ($t(2.624)$= -.017, p= .987) or between the songs used in Experiment 1 and Experiment 2 ($t(8.784)$= -0.171, p= .868). The songs were

[4] 5 more participants were tested but excluded (3 because they failed one or more trap trials and 2 due to missing or invalid data).

selected so that they are matched on lyrics rate and overall genre to those used in Experiment 1, but that they have a low likelihood of being known to participants (judged by their number of views on YouTube.com). The results confirmed that recognition of the songs was < 1% (see Table 4).

Results

During pre-processing, 0.32% of the data was excluded due to outliers (RTs < 100 or > 5000ms). Additionally, 3 trials (0.10%) were removed due to a lack of response on more than 5 words. This left 99.58% of the data for analysis. Descriptive statistics are shown in Table 1 and the results are presented in Tables 2-3.

Similar to Experiment 1b, lyrical music led to longer word RTs compared to instrumental music (d= 0.071). This supports **H1**. However, contrary to the other predictions, instrumental music led to significantly *lower* word RTs compared to silence (d= -0.052). Therefore, instrumental music led to an unexpected facilitation where reading was faster compared to the silence baseline. The difference between instrumental music and silence showed only “anecdotal” evidence (Jeffreys, 1961; Wetzels et al., 2011) in support of the alternative hypothesis in the Bayesian model, thus suggesting the result was reliable only in the frequentist model.

The comprehension accuracy analysis showed no significant differences between instrumental music and silence or between lyrical music and instrumental music. The Bayesian model supported the null hypothesis of no difference, though the evidence was “substantial” only in the comparison between instrumental music and silence.

Discussion

Experiment 2 replicated the key finding from Experiment 1b, where lyrical music led to longer word RTs compared to instrumental music. Therefore, there was more evidence to suggest that the presence of lyrics in songs leads to distraction. Because Experiment 2 used unfamiliar songs, the findings also show that this result extends to music that is unknown to participants. The effect size was similar to that of Experiment 1b, which suggests that the amount of distraction between the two studies was roughly comparable.

Experiment 2 also showed one unexpected finding: instrumental music led to *faster* word reading times compared to silence. While the source of this facilitation effect is unknown, there have been sporadic reports of classical (instrumental) music leading to improved reading performance compared to silence (e.g., Falcon, 2017; Mullikin & Henk, 1985). To ensure this facilitation effect is reliable, we attempted to replicate it in Experiment 3.

Experiment 3

The goal of Experiment 3 was to replicate and extend the results from Experiment 2. The study was identical, except that a new condition of irrelevant background speech was added. Experiments 1b and 2 demonstrated that lyrical music is more distracting that instrumental music, thus showing that the processing of lyrics in music interferes with reading efficiency. However, it is not known if lyrics lead to the same distraction as irrelevant speech. Vasilev et al.'s (2018) results suggest that lyrical music is just as distracting as intelligible background speech. However, their findings were only observational in nature, so this prediction has never been tested directly. Because the present research used songs with a rate of lyrics that approximates the rate of normal speech, it can be predicted that lyrical music and intelligible speech would cause the same amount of distraction. Therefore, the second goal of Experiment 2 was to test if lyrical music and irrelevant speech cause equivalent distraction when they are matched on language rate. However, it is also possible that the instrumentals present in songs may partially mask the distracting effect of the lyrics, thus leading to smaller distraction in lyrical music compared to irrelevant speech. As a result, two new hypotheses were formed:

- **H3.1**: If lyrical music yields the same distraction as spoken language (when language rate is controlled), there should be no difference in self-paced reading times between the irrelevant speech and lyrical music conditions.
- **H3.2**: If certain properties of the music (e.g., instrumentals) partially mask the distracting nature of the lyrics, the irrelevant speech condition should result in longer self-paced reading times compared to the lyrical music condition.

The study protocol was pre-registered prior to data collection (https://osf.io/ztpb6).

Method

Participants. A total of 208 UK adults[5] recruited from Prolific.co participated in return for compensation at £7/ hour (54.3% female; 44.2% male; 1.44% other genders). None of them took part in the previous experiments. Participants had an average age of 34.25 years (*SD*= 8.28 years; range: 18- 50 years). Participants' educational background was: 12% had completed GCSEs, 24.5% had completed A-levels, 43.8% had completed an undergraduate degree, 15.9% had completed a postgraduate degree, 3.8% had completed a PhD degree.

[5] 8 more participants were tested but excluded based on the pre-registered criteria (2 participants admitted to not wearing headphones, 3 participants failed one or more of the trap trials, 2 participants had missing or incomplete data). Additionally, 2 more participants were excluded due to chance-level comprehension (<60%; comprehension criterion was not pre-registered).

Design, materials, apparatus, procedure, and data analysis. The study was the same as Experiment 2, except that a new condition of irrelevant speech was added. This condition consisted of short spoken statements, concatenated together in Adobe Audition 2019 to create about 10 minutes of audio (e.g., "This theory has implications for spatial illusions such as the visual angle illusion", "They will take the Piccadilly Line to Covent Garden from Leicester Square", "Concentrated solar power uses molten salt energy storage in a tower or in trough configurations"). The speech files were taken from the LibriSpeech ASR corpus (Panayotov, Chen, Povey, & Khudanpur, 2015), available through the Open Speech and Language Resources project (https://www.openslr.org/12). The rate of speech in the irrelevant speech condition (M= 149.1; SD= 4.979) was matched to that of the unfamiliar songs (M= 149.01; SD= 16.391), $t(5.559)$= .0117, p= 0.991. To maintain the same statistical power as the previous experiments, 5 more passages were added from the Provo corpus (Luke & Christianson, 2018). Thus, 20 items were used in total (5 per condition).

Results

During pre-processing, 0.31% of the data was excluded due to outliers (RTs < 100 or > 5000ms). Additionally, 5 trials (0.12%) were removed due to a lack of response on more than 5 words. This left 99.57% of the data for analysis. Descriptive statistics are shown in Table 1 and the statistical results are presented in Tables 2-3.

Consistent with **H1**, lyrical music led to longer word RTs compared to instrumental music (d= 0.044). Additionally, consistent with Experiment 2, but contrary to predictions, instrumental music led to faster word RTs compared to the silence baseline (d= -0.016). This time, the Bayesian model showed "decisive" evidence (Jeffreys, 1961; Wetzels et al., 2011) in support of the alternative hypothesis. Finally, there was no significant difference between irrelevant speech and lyrical music; the Bayesian model showed “decisive” evidence for the null hypothesis of no difference. Therefore, this supports **H3.1** and suggests that distraction by lyrical music and irrelevant speech was equivalent.

In the comprehension accuracy measure, there was no difference in accuracy between silence and instrumental music; The Bayesian model showed “substantial” support for the null hypothesis. Lyrical music led to a significant decrease in comprehension accuracy compared to instrumental music (d= -0.016), though the Bayesian model showed inconclusive evidence for either the null or alternative. Finally, there was no difference in comprehension accuracy between speech and lyrical music; the Bayesian model favoured the null hypothesis, though the evidence was “anecdotal”. In summary, there was no reliable evidence for distraction in comprehension accuracy.

Discussion

Experiment 3 replicated the two key findings from Experiment 2: 1) lyrical music led to longer RTs compared to instrumental music; and 2) instrumental music led to *shorter* RTs compared to silence. Thus, the unexpected facilitation of instrumental music from Experiment 2 was confirmed in a new sample. We will return to this in the General Discussion.

Interestingly, irrelevant speech did not differ from lyrical music in RTs, which suggests that the amount of distraction was equivalent between the two conditions. This supports Vasilev et al.'s (2018) results that lyrical music is just as distracting as speech. We now turn to the covariate analyses looking at whether properties of the songs affected differences between the lyrical and instrumental music conditions.

Covariate analyses with Song Ratings, Song knowledge, and Daily Music Use

Participants’ music genre preferences are shown in Figure 3. Participants reported listening to music each day for an average of 2.7 hours in Experiment 1a (*SD*= 2.09; *range*= 0 - 12 hours), 2.92 hours in Experiment 1b (*SD*= 1.82; *range*= 0-14 hours), 2.39 hours in Experiment 2 (*SD*= 2.04; *range*= 0-15 hours), and 2.16 hours in Experiment 3 (*SD*= 1.89; *range*= 0-13 hours). Participants’ ratings of the songs are shown in Figure 4 and their correlations are visualised in Figure 5. Participants actual knowledge of the songs is shown in Table 4.

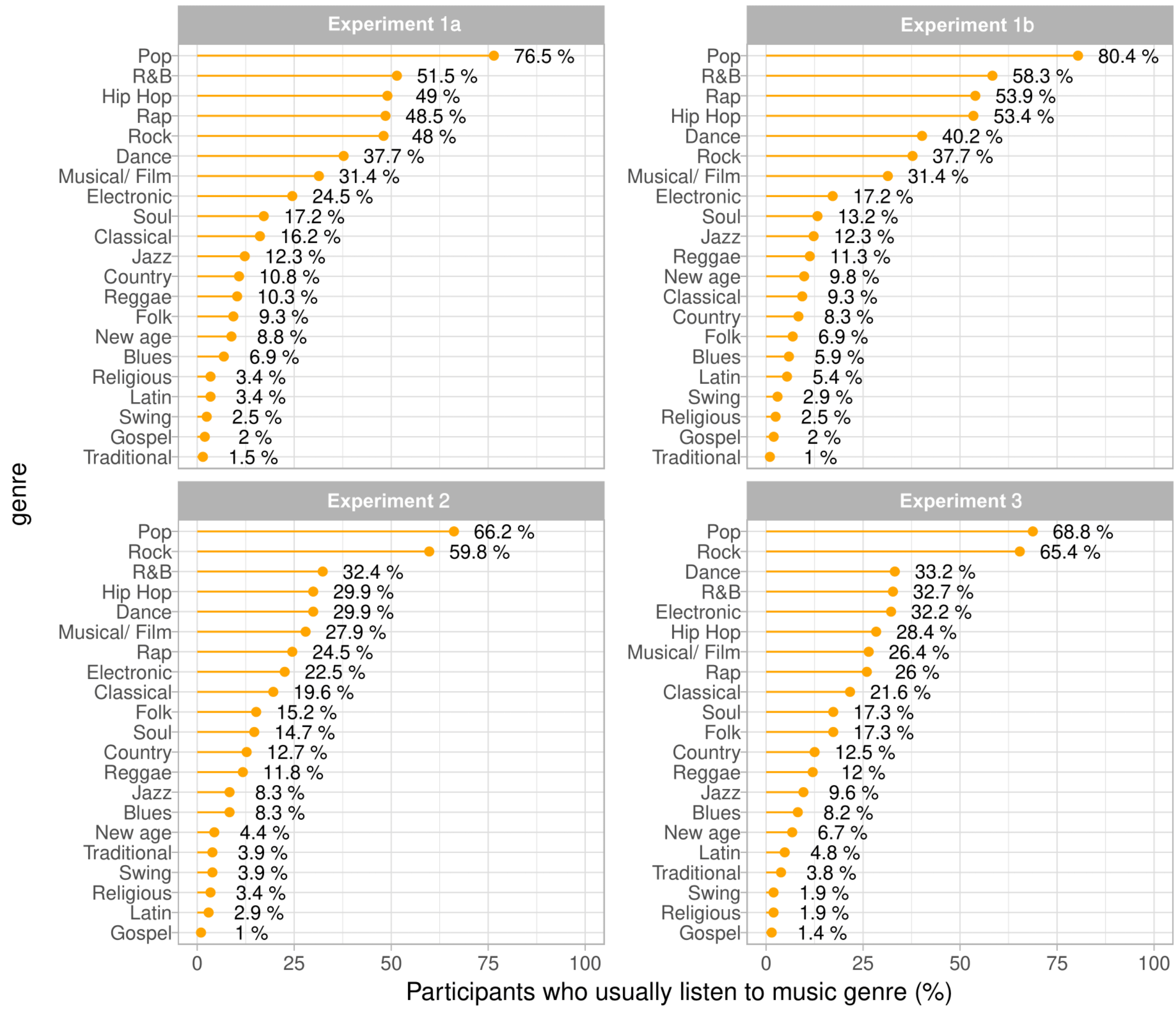

Figure 3. Music genre preference of participants in the four experiments. Participants were asked to indicate *all* genres that they usually listen to. The percentages show the proportion of participants who selected a given genre and thus the numbers do not add up to 100%.

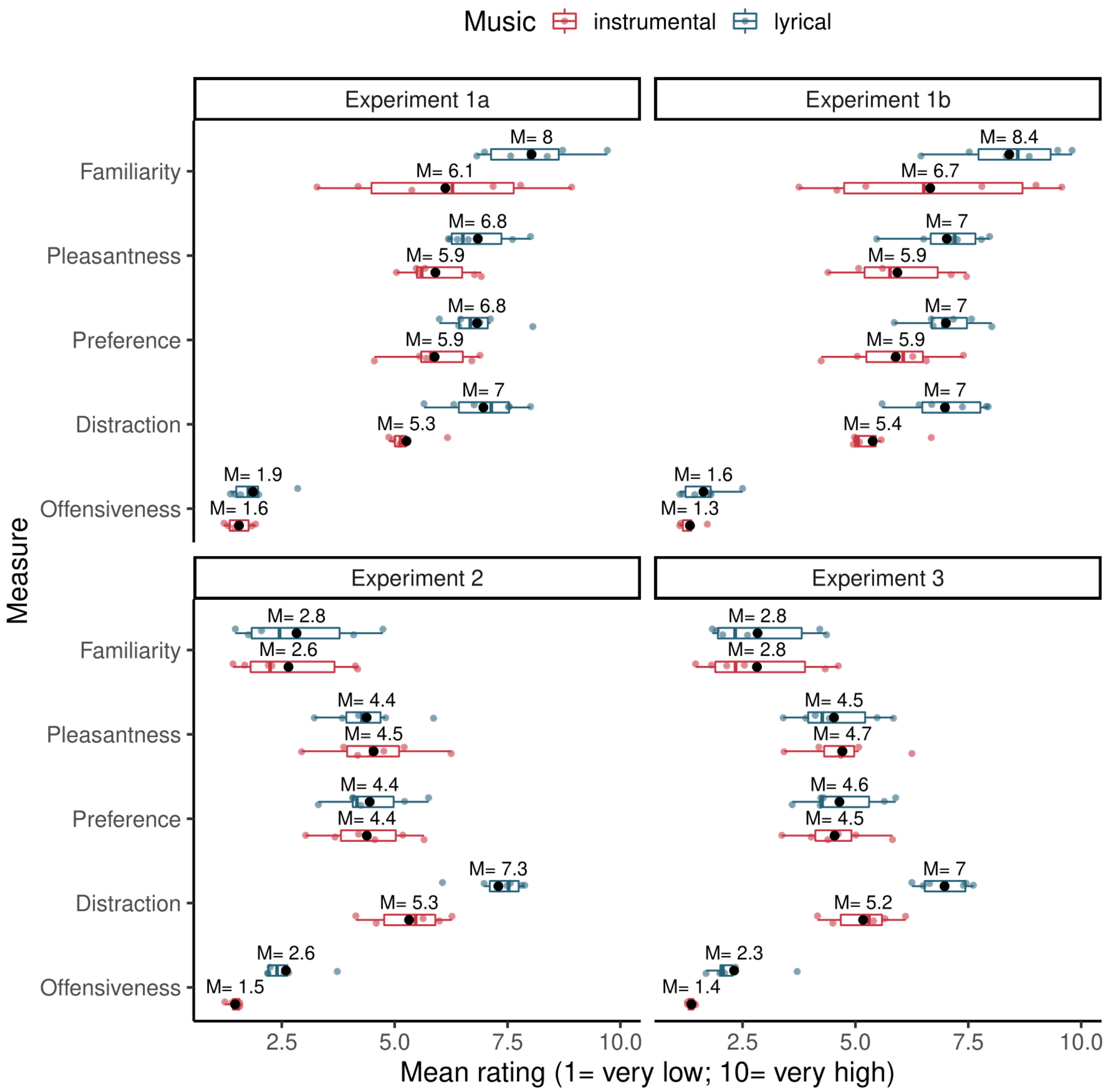

Figure 4. Participants' rating of the songs split by music type (instrumental vs lyrical version of the songs). The means are plotted and shown by a black dot.

Table 4

Percentage of the Experimental Songs for which Participants Could Correctly Identify the Artist(s) and Song Title

Experiment	Song version	Artist accuracy (%)	Song accuracy (%)
1a	Lyrical	67.8 (46.7)	61.1 (48.8)
1a	Instrumental	38.4 (48.7)	37.3 (48.4)
1b	Lyrical	65.5 (47.6)	54.9 (49.8)
1b	Instrumental	42.8 (49.5)	41.5 (49.3)
2	Lyrical	0.33 (5.71)	0.33 (5.71)
2	Instrumental	0.16 (4.04)	0.16 (4.04)
3	Lyrical	0.80 (8.92)	0.48 (6.92)
3	Instrumental	0 (0)	0 (0)

Note: Participants were given a 30s sample of each song after the experiment and were asked to write down the artist(s) and song title, if they know them.

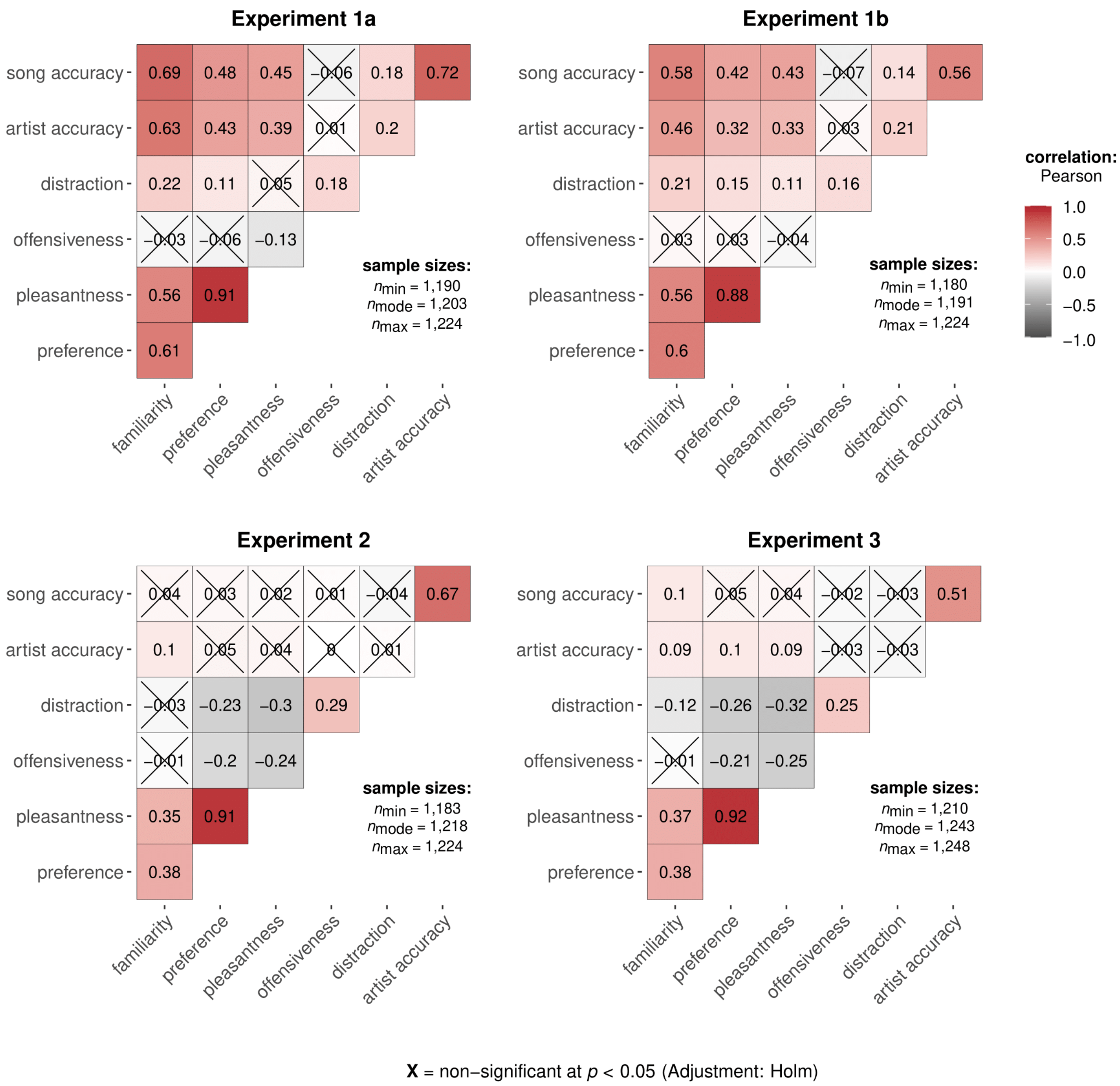

Figure 5. Correlation matrix plot of the music rating and song/artist accuracy variables in the experiments. Experiments 1a-1b contained familiar music and Experiments 2-3 contained unfamiliar music.

Covariate analysis. The goal of the pre-registered co-variate analysis was to test if the difference between lyrical and instrumental music is still significant after adjusting for the effect of the covariates. In this analysis, the silence condition was excluded from the data, thus leaving only the comparison between lyrical and instrumental music (the speech condition was also excluded from Experiment 3). This is because only the two music conditions received ratings of the songs that could be used in the analysis. The following covariates were then added to the model: music familiarity, preference, offensiveness, perceived distractibility, song knowledge (composite measure of artist accuracy and song title accuracy), and daily music use frequency. Music pleasantness was not included because the measure was almost perfectly correlated with the music preference ratings (see Figure 5). Additionally, song knowledge was removed as a covariate in Experiments 2-3 because almost no participants knew the songs, so the model parameters could not be reliably estimated. All covariates were converted into z-scores to deal with multi-collinearity issues and improve the scaling of the models. The results are visualised in Figure 6.

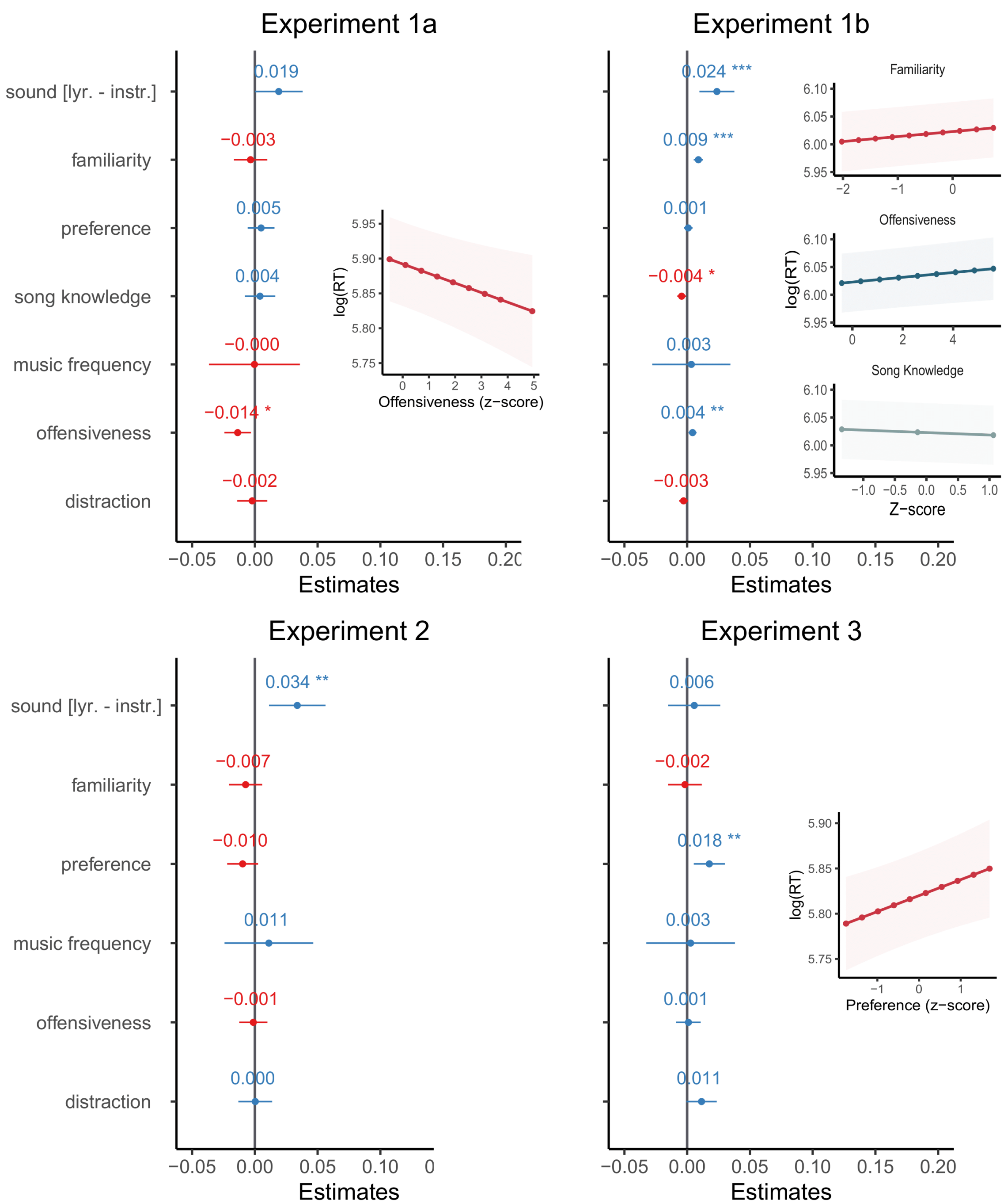

Figure 6. Results from the pre-planned covariate analyses using participants' ratings of the songs, their song knowledge (a composite measure of song title and artist accuracy), and daily music use frequency. Note that only the lyrical and instrumental music conditions are included, as no ratings were possible in the silence and speech conditions. Plotted are the LMM estimates for each predictor in the model. The subplots on the right of each panel show a visualisation of the significant effects for that model. * p<0.05; ** p<0.01; *** p<0.001

In Experiment 1a, the difference between lyrical and instrumental music in RTs was still not significant (even though it was just under the .05 threshold). Therefore, the conclusions from the main analysis remained unchanged. Interestingly, music offensiveness reached statistical significance. This result was counter-intuitive, as greater offensiveness ratings of the music were associated with faster reading times. In Experiment 1b, the difference between lyrical and instrumental music was still significant and thus the model results also remained unchanged. Interestingly, however, familiarity, offensiveness and song knowledge reached significance. Reading times were longer when the music was rated as more familiar and more offensive. Thus, the offensiveness effect was in the opposite direction to that of Experiment 1a. Additionally, greater knowledge of the song that was playing was associated with slightly lower RTs (i.e., faster reading).

In Experiment 2, the significant difference between lyrical and instrumental music remained unchanged after adjusting for the covariates. However, in Experiment 3, the difference between lyrical and instrumental music was no longer statistically significant. As shown in Figure 6, the difference could be explained by a significant music preference effect, where music that was rated as more preferred by participants resulted in longer reading times. This suggests that the difference between the two conditions in Experiment 3 may not be explained by the presence of lyrics per se, but by a general slow-down in reading when listening to preferred music. In summary, the results from the first three Experiments remained unchanged after adjusting for the covariates; the result from Experiment 4 appeared to be explainable by how much participants preferred the music.

Analysis of music ratings as function of music type (instrumental vs lyrical). Finally, we analysed the music ratings as dependent variables to understand how participants rated the songs based on whether they heard the lyrical or the instrumental version of them. The ratings were collapsed across experiments based on whether participants were rating the "familiar" music (Experiment 1a-1b) or the "unfamiliar" music (Experiments 2-3). In the "familiar" music dataset, participants rated the lyrical version of songs as significantly more familiar (b=1.835, SE= 0.524, t= 3.50), more preferred (b= 1.024, SE= 0.196, t= 5.224), more offensive (b= 0.309, SE= 0.1520, t= 2.031), and more distracting (b= 1.661, SE= 0.352, t= 4.725) than the instrumental version of songs (see Figure 4). Additionally, they were significantly more likely to correctly recall the artist(s) (b= 1.637, SE= 0.4433, t= 3.692) and song name (b= 1.384, SE= 0.132, z=10.480) when they heard the lyrical compared to the instrumental version of songs. Clearly, these results suggest that participants partly derive the identity of songs (as well their familiarity, preference, perceived offensiveness, and distraction) from the lyrics.

In the "unfamiliar" music dataset, the lyrical version of songs was rated as significantly more offensive (b=1.028, SE= 0.258, t= 3.976) and more distracting (b= 1.888, SE= 0.275, t= 6.876) than the instrumental version. However, there were no significant differences in the other variables (all $|t|$s and $|z|$s $\leq$ 1.21). Therefore, this suggests that participants' preference for the unfamiliar songs was not confounded by the presence of lyrics, but participants still perceived lyrical music to be more distracting and offensive.

General Discussion

The present study used self-paced reading to test whether song lyrics play a key role in distraction by background music. The results from three out of four experiments showed that lyrical music led to slower word reading times, thus indicating that the presence of lyrics in songs caused distraction and reduced overall reading efficiency. Despite this increase in reading times, there was no associated decrease in comprehension in most of the experiments (only Experiment 1b indicated a decrease in comprehension in lyrical compared to instrumental music). The reading time data generally support previous findings showing that lyrical music is more distracting than instrumental music (Martin et al., 1988; C. Miller, 2014; Perham & Currie, 2014; Reed, 2019; Vasilev et al., 2018) but contradict others that have shown no such difference (Avila et al., 2012; Furnham et al., 1999; Kyoung, 2020). Still, it is important to keep in mind that Experiment 1a showed no overall evidence of distraction by lyrical music. Therefore, while the present results were also “mixed”, on balance, the evidence seems to suggest that lyrics can give rise to distraction.

Because the present study used short texts that are arguably not a very strong test of comprehension, it remains to be seen whether the distraction observed in reading times can be replicated in standardised reading comprehension tests. Nevertheless, the current research clearly demonstrates that lyrics can interfere with word-level reading processes. This is consistent with evidence from eye-tracking, where distraction at the word level was also observed in eye fixations when readers were exposed to irrelevant speech (Cauchard et al., 2012; Hyönä & Ekholm, 2016; Meng et al., 2020; Yan et al., 2018) and music (Zhang et al., 2018; but see Johansson et al., 2012).

Why Are Lyrics Distracting?

The increase in reading times in lyrical compared to instrumental music can be readily explained by both semantic (Jones & Tremblay, 2000; Marsh et al., 2008, 2009; Martin et al., 1988) and phonological interference theories (Salamé & Baddeley, 1982, 1987, 1989), which assume that either the semantic or phonological content of the lyrics is processed inadvertently and causes interference with the main task due to the use of shared processes. Critically, both theories assume that this interference is language-related. On the other hand, the changing state-theory (Jones & Macken, 1993; Jones et al., 1992) does not share this assumption as lyrics can cause distraction only because they increase the acoustic variation on top of the instrumental track. Therefore, distraction should mostly occur due to the vocal and acoustic changes associated with sung lyrics.

Language-related distraction fits well with established findings, such as the fact that irrelevant speech interferes with reading processes and that this interference appears to be mostly semantic in nature (Hyönä & Ekholm, 2016; Martin et al., 1988; Meng et al., 2020; Vasilev et al., 2019). This suggests that language (either spoken or sung) may undergo obligatory processing (Crinion, Lambon-Ralph, Warburton, Howard, & Wise, 2003; Marsh & Jones, 2010) and interfere with the task at hand. This interpretation is consistent with the results of Experiment 3, where intelligible speech was just as distracting as lyrical music when the two were matched on language rate. At present, it is not clear if phonological or semantic information from the lyrics was responsible for the observed distraction. However, future studies comparing the same music sung in different languages (e.g., Chew et al., 2016) may possibly adjudicate between the two views.

The question of whether lyrics may also cause acoustic-related distraction is an interesting one. Language-related theories reduce distraction to the processing of the language within the lyrics, but ignore other factors such as the musical prosody of the lyrics and the way they are sung. To our knowledge, Martin et al. (1988, Experiment 2) is the only study to consider this question. They found no difference between sung and spoken lyrics, which led them to believe that the musicality of the sung lyrics played no role in distraction. However, lyrics clearly contain other information as well, such as the voice, vocal characteristic, and identity of the singer. This information could in turn influence participants' memory and perception of the music.

The song ratings demonstrated this very clearly. The "familiar" set of songs used in Experiments 1a-1b were more recognisable and were rated as more preferred, pleasant, and familiar when heard in the lyrical compared to the instrumental condition. No such difference was observed for the "unfamiliar" set of songs used in Experiments 2-3, which virtually no participants could recognise. These results suggest that, for "familiar" songs, participants' perception and recollection of the music is intrinsically linked to the lyrics, thereby introducing potential confounds when trying to isolate the unique role of language. These results agree with previous research showing that the recognition of melodies is better when they are presented vocally rather than instrumentally (Weiss, Schellenberg, Trehub, & Dawber, 2015; Weiss, Trehub, & Schellenberg, 2012; Weiss, Vanzella, Schellenberg, & Trehub, 2015), even if the melody is sung in a different voice (Weiss, Schellenberg, & Trehub, 2017). Therefore, it is not surprising that participants partly derive the identity of the songs from their lyrics.

Clearly, this poses a problem as any performance differences between lyrical and instrumental music could occur simply because the two conditions are perceived differently by participants. One way to avoid such confounds is to use only unfamiliar music, as some studies have done in the past (e.g., Furnham et al., 1999; Furnham & Allass, 1999; Kyoung, 2020). Another way is to statistically control for such variables, which was the approach taken here. The covariate analysis suggested that the results remained unchanged for the first three experiments after accounting for the effect of song knowledge and music ratings. However, surprisingly, the lyrical distraction effect in Experiment 3 did not "survive" the covariate correction, in part because it appeared to be explained by participants' preference for the music. More research is needed to explain why this may be the case. However, the present study highlights the importance of tracking such variables.

Is Instrumental Music Distracting?

One interesting finding in the present research was that instrumental music did not cause any distraction. This agrees with Vasilev et al.'s (2018) meta-analysis finding that instrumental music does affect reading comprehension. Nevertheless, the primary literature has been less conclusive. For instance, while some studies have reported no difference between instrumental music and silence (Cauchard et al., 2012; Martin et al., 1988; Perham & Currie, 2014), others have reported that instrumental music causes distraction (Avila et al., 2012), and yet others have reported that instrumental music improves performance compared to silence (Falcon, 2017; Mullikin & Henk, 1985). The present research showed a combination of no effects and positive effects, but crucially no hint of any distraction. This suggests that music instrumentals are not sufficient on their own to negatively affect reading performance. Therefore, given that students (Calderwood, Ackerman, & Conklin, 2014; David et al., 2015) and office employees (Haake, 2006) often report listening to music while studying or doing work, it seems prudent to recommend listening to instrumental rather than lyrical music when reading.

One unexpected finding was that instrumental music improved performance compared to the silence baseline, though this was statistically reliable only in Experiment 3 (a similar trend in the data was also present in Experiment 2). One possible explanation for this finding is that instrumental music may lead to an increase in arousal (Dillman Carpentier & Potter, 2007; Furnham et al., 1999), which could temporarily boost performance. We speculate that such improvements may be more difficult to sustain with tasks that involve reading longer texts. However, more research is needed to better understand this issue.

Limitations and Future Directions

The present study also had a few limitations. First, the reading stimuli consisted of passages that were short and easy to read. This was done to ensure that the stimuli can be read quickly in an online study format, as longer experiment times can negatively affect data quality (Sauter, Draschkow, & Mack, 2020). However, one consequence of this is that the stimuli may not have been very challenging for our participants, potentially leading to smaller distraction effects.

Additionally, because previous words in the text were masked, participants could not go back to re-read them (i.e., make regressions). We chose to mask previous words because it prevents participants from pressing the button in quick bursts to reveal the whole text, before they actually start reading it (Just, Carpenter, & Woolley, 1982). This was especially important as most participants completed the task in an unsupervised environment at home. However, this had the consequence that the task deviated from "natural" reading. We argue that the present paradigm is still useful in understanding reading processes in an online environment where more complex methodology (e.g., eye-tracking) cannot be used. The fact that we were able to observe distraction after all shows that such effects occur even in the absence of regressions. Future studies could address this limitation by using bi-directional self-paced reading (Paape & Vasishth, 2022), where participants can move both forward and backwards in the text.

Finally, the present study also differed in that it used more heterogeneous samples compared to the typical university student population. This arguably made it more difficult to interpret the results and compare them to those of previous studies. We do not necessarily view this as a limitation because any distraction effects that are meaningful in practice should be replicable in different populations, paradigms, and testing conditions. Clearly, online data testing has the potential to reach more diverse populations of readers and we believe that this will prove important for understanding how distraction occurs in the real world.

Conclusion

The present study tested whether song lyrics are a key component of what makes background music distracting. In three out of four experiments, we observed that lyrical music led to longer self-paced reading times compared to instrumental music. This suggests that lyrics interfere with reading by making it less efficient. Despite this, the observed effects were quite mild, suggesting that readers were mostly able to overcome the music distraction. On the other hand, instrumental music did not lead to any distraction, which seems to suggest that it has no negative influence on how participants process the text. Finally, the study also uncovered that "familiar" lyrical songs are both more recognisable and rated differently compared to the instrumental version of the same songs. This suggests that future studies need to take these differences into account. In summary, the present research provides some evidence that lyrics can cause distraction, but more research is needed to better understand why this is the case.

Funding

This study was funded in part by QR funds from the Department of Psychology, Bournemouth University awarded to M.R.V.

Data availability

The data and materials from this study are available at: https://osf.io/8zw4x/

Disclosure statement

The authors report there are no competing interests to declare.

References

Aaronson, D., & Scarborough, H. S. (1976). Performance theories for sentence coding: Some qualitative observations. *Journal of Experimental Psychology: Human Perception and Performance*, *2*(1), 42–55. https://doi.org/10.1037/0096-1523.2.1.42

Albers, C., & Lakens, D. (2018). When power analyses based on pilot data are biased: Inaccurate effect size estimators and follow-up bias. *Journal of Experimental Social Psychology*, *74*(September 2017), 187–195. https://doi.org/10.1016/j.jesp.2017.09.004

Anderson, S. a., & Fuller, G. B. (2010). Effect of music on reading comprehension of junior high school students. *School Psychology Quarterly*, *25*(3), 178–187. https://doi.org/10.1037/a0021213

Avila, C., Furnham, A., & McClelland, A. (2012). The influence of distracting familiar vocal music on cognitive performance of introverts and extraverts. *Psychology of Music*, *40*(1), 84–93. https://doi.org/10.1177/0305735611422672

Baayen, H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, *59*(4), 390–412. https://doi.org/10.1016/j.jml.2007.12.005

Baker, R. W., & Madell, T. O. (1965). A continued investigation of susceptibility to distraction in academically underachieving and achieving male college students. *Journal of Educational Psychology*, *56*(5), 254–258. https://doi.org/10.1037/h0022467

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278. https://doi.org/10.1016/j.jml.2012.11.001

Bates, D. M., Machler, M., Bolker, B. M., & Walker, S. C. (2014). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Brysbaert, M. (2019). How many words do we read per minute? A review and meta-analysis of reading rate. *Journal of Memory and Language*, *109*(April). https://doi.org/10.1016/j.jml.2019.104047

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28. https://doi.org/10.18637/jss.v080.i01

Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*(1), 395. https://doi.org/10.32614/RJ-2018-017

Calderwood, C., Ackerman, P. L., & Conklin, E. M. (2014). What else do college students "do" while studying? An investigation of multitasking. *Computers and Education*, *75*(2014), 19–29. https://doi.org/10.1016/j.compedu.2014.02.004

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., … Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of Statistical Software*, *76*(1), 1–32. https://doi.org/10.18637/jss.v076.i01

Cauchard, F., Cane, J. E., & Weger, U. W. (2012). Influence of background speech and music in interrupted reading: An eye-tracking study. *Applied Cognitive Psychology*, *26*(3), 381–390. https://doi.org/10.1002/acp.1837

Chew, A. S.-Q., Yu, Y.-T., Chua, S.-W., & Gan, S. K.-E. (2016). The effects of familiarity and language of background music on working memory and language tasks in Singapore. *Psychology of Music*, *44*(6), 1431–1438. https://doi.org/10.1177/0305735616636209

Chitwood, M. R. (2018). Cognitive Performance and Sounds: The Effects of Lyrical Music and Pink Noise on Performance. *The NKU Journal of Student Research*, *1*, 1–7.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Crinion, J. T., Lambon-Ralph, M. A., Warburton, E. A., Howard, D., & Wise, R. J. S. (2003). Temporal lobe regions engaged during normal speech comprehension. *Brain*, *126*(5), 1193–1201. https://doi.org/10.1093/brain/awg104

Daoussis, L., & Mc Kelvie, S. J. (1986). Musical preferences and effects of music on a reading comprehension test for extraverts and introverts. *Perceptual and Motor Skills*, *62*(1), 283–289. https://doi.org/10.2466/pms.1986.62.1.283

David, P., Kim, J.-H., Brickman, J. S., Ran, W., & Curtis, C. M. (2015). Mobile phone distraction while studying. *New Media & Society*, *17*(10), 1661–1679. https://doi.org/10.1177/1461444814531692

de la Mora Velasco, E., & Hirumi, A. (2020). The effects of background music on learning: a systematic review of literature to guide future research and practice. *Educational Technology Research and Development*, *68*(6), 2817–2837. https://doi.org/10.1007/s11423-020-09783-4

Dickey, J. M., & Lientz, B. P. (1970). The weighted likelihood ratio, sharp hypotheses about chances, the order of a Markov Chain. *The Annals of Mathematical Statistics*, *41*(1), 214–226. Retrieved from https://www.jstor.org/stable/2239734

Dillman Carpentier, F. R., & Potter, R. F. (2007). Effects of music on physiological arousal: Explorations into tempo and genre. *Media Psychology*, *10*(3), 339–363. https://doi.org/10.1080/15213260701533045

Doyle, M., & Furnham, A. (2012). The distracting effects of music on the cognitive test performance of creative and non-creative individuals. *Thinking Skills and Creativity*, *7*(1), 1–7. https://doi.org/10.1016/j.tsc.2011.09.002

Elliott, E. M., Bell, R., Gorin, S., Robinson, N., & Marsh, J. E. (2022). Auditory distraction can be studied online! A direct comparison between in-Person and online experimentation. *Journal of Cognitive Psychology*, *34*(3), 307–324. https://doi.org/10.1080/20445911.2021.2021924

Etaugh, C., & Michals, D. (1975). Effects on reading comprehension of preferred music and frequency of studying to music. *Perceptual and Motor Skills*, *41*(2), 553–554. https://doi.org/10.2466/pms.1975.41.2.553

Etaugh, C., & Ptasnik, P. (1982). Effects of studying to music and post-study relaxation on reading comprehension. *Perceptual and Motor Skills*, *55*(1), 141–142. https://doi.org/10.2466/pms.1982.55.1.141

Eysenck, H. (1967). *The biological basis of personality*. Springfield, IL: Thomas.

Falcon, E. (2017). *The relationship between background classical music and reading comprehension on 7th and 8th grade students (Unpublished doctoral dissertation)*. St. Thomas University, Florida, USA.

Fendrick, P. (1937). The influence of music distraction upon reading efficiency. *Journal of Educational Research*, *31*(4), 264–271. https://doi.org/10.1007/s13398-014-0173-7.2

Fogelson, S. (1973). Music as a distractor on reading-test performance of eighth grade students. *Perceptual and Motor Skills*, *36*, 1265–1266. https://doi.org/10.2466/pms.1973.36.3c.1265

Freeburne, C. M., & Fleischer, M. S. (1952). The effect of music distraction upon reading rate and comprehension. *Journal of Educational Psychology*, *43*, 101–109.

https://doi.org/10.1037/h0054219

Furnham, A., & Allass, K. (1999). The influence of musical distraction of varying complexity on the cognitive performance of extroverts and introverts. *European Journal of Personality*, *13*(1), 27–38. https://doi.org/10.1002/(SICI)1099-0984(199901/02)13:1<27::AID-PER318>3.0.CO;2-R

Furnham, A., & Bradley, A. (1997). Music while you work: The differential distraction of background music on the cognitive test performance of introverts and extraverts. *Applied Cognitive Psychology*, *11*(5), 445–455. https://doi.org/10.1002/(SICI)1099-0720(199710)11:5<445::AID-ACP472>3.0.CO;2-R

Furnham, A., & Stephenson, R. (2007). Musical distracters, personality type and cognitive performance in school children. *Psychology of Music*, *35*(3), 403–420. https://doi.org/10.1177/0305735607072653

Furnham, A., & Strbac, L. (2002). Music is as distracting as noise: The differential distraction of background music and noise on the cognitive test performance of introverts and extraverts. *Ergonomics*, *45*(3), 203–217. https://doi.org/10.1080/00140130210121932

Furnham, A., Trew, S., & Sneade, I. (1999). The distracting effects of vocal and instrumental music on the cognitive test performance of introverts and extraverts. *Personality and Individual Differences*, *27*(2), 381–392. https://doi.org/10.1016/S0191-8869(98)00249-9

Gheewalla, F., McClelland, A., & Furnham, A. (2020). Effects of background noise and extraversion on reading comprehension performance. *Ergonomics*, *64*(5), 593–599. https://doi.org/10.1080/00140139.2020.1854352

Gillis, A. (2010). The effect of background music on reading comprehension and self-report of college students. *Florida State Libraries. Electronic Theses, Treatises and*

Dissertations. The Graduate School.

Green, P., & Macleod, C. J. (2016). SIMR: An R package for power analysis of generalized linear mixed models by simulation. *Methods in Ecology and Evolution*, *7*(4), 493–498. https://doi.org/10.1111/2041-210X.12504

Haake, A. B. (2006). Music listening practices in workplace settings in the UK: An exploratory survey of office-based settings. In *Proceedings of the Ninth International Conference on Music Perception and Cognition*.

Hall, J. C. (1952). The effect of background music on the reading comprehension of 278 eighth and ninth grade students. *The Journal of Educational Research*, *45*(6), 451–458. https://doi.org/10.1080/00220671.1952.10881962

Hallam, S., & MacDonald, R. (2016). The Effects of Music in Community and Educational Settings. In S. Hallam, I. Cross, & M. Thaut (Eds.), *The Oxford Handbook of Music Psychology* (pp. 1–18). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780198722946.013.46

Henderson, M. T., Crew, A., & Barlow, J. (1945). A study of the effect of music distraction on reading efficiency. *Journal of Applied Psychology*, *29*(4), 313–317. https://doi.org/10.1037/h0056128

Henninger, F., Shevchenko, Y., Mertens, U. K., Kieslich, P. J., & Hilbig, B. E. (2022). lab.js: A free, open, online study builder. *Behavior Research Methods*, *54*(2), 556–573. https://doi.org/10.3758/s13428-019-01283-5

Hilliard, O. M., & Tolin, P. (1979). Effect of familiarity with background music on performance of simple and difficult reading comprehension tasks. *Perceptual and Motor Skills*, *49*(3), 713–714. https://doi.org/10.2466/pms.1979.49.3.713

Hughes, R. W., & Jones, D. M. (2001). The intrusiveness of sound: Laboratory findings and their implications for noise abatement. *Noise & Health*, *4*(13), 51–70.

Hyönä, J., & Ekholm, M. (2016). Background speech effects on sentence processing during reading: An eye movement study. *PloS One*, *11*(3), e0152133. https://doi.org/10.1371/journal.pone.0152133

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford, UK: Oxford University Press.

Jegerski, J. (2014). Self-paced reading. In J. Jegerski & B. VanPatten (Eds.), *Research methods in second language psycholinguistics* (pp. 20–49). New York, USA: Routledge.

Johansson, R., Holmqvist, K., Mossberg, F., & Lindgren, M. (2012). Eye movements and reading comprehension while listening to preferred and non-preferred study music. *Psychology of Music*, *40*, 339–356. https://doi.org/10.1177/0305735610387777

Jones, D. M., & Macken, W. J. (1993). Irrelevant tones produce an irrelevant speech effect: Implications for phonological coding in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *19*(2), 369–381. https://doi.org/10.1037/0278-7393.19.2.369

Jones, D. M., Madden, C., & Miles, C. (1992). Privileged access by irrelevant speech to short-term memory: The role of changing state. *The Quarterly Journal of Experimental Psychology*, *44*(4), 645–669. https://doi.org/10.1080/14640749208401304

Jones, D. M., & Tremblay, S. (2000). Interference in memory by process or content? A reply to Neath (2000). *Psychonomic Bulletin & Review*, *7*(3), 550–558. https://doi.org/10.3758/BF03214370

Just, M. A., Carpenter, P. A., & Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology. General*, *111*(2), 228–238.

https://doi.org/10.1037/0096-3445.111.2.228

Kallinen, K. (2002). Reading news from a pocket computer in a distracting environment: Effects of the tempo of background music. *Computers in Human Behavior*, *18*(5), 537–551. https://doi.org/10.1016/S0747-5632(02)00005-5

Kämpfe, J., Sedlmeier, P., & Renkewitz, F. (2011). The impact of background music on adult listeners: A meta-analysis. *Psychology of Music*, *39*(4), 424–448. https://doi.org/10.1177/0305735610376261

Kelly, S. N. (1994). A comparison of the effects of background music on the reading comprehension of university undergraduate music majors and nonmusic majors. *Southeastern Journal of Music Education*, *5*, 86–97.

Kiger, D. M. (1989). Effects of music information load on a reading comprehension task. *Perceptual and Motor Skills*, *69*(2), 531–534. https://doi.org/10.2466/pms.1989.69.2.531

Kou, S., McClelland, A., & Furnham, A. (2018). The effect of background music and noise on the cognitive test performance of Chinese introverts and extraverts. *Psychology of Music*, *46*(1), 125–135. https://doi.org/10.1177/0305735617704300

Küssner, M. B. (2017). Eysenck's theory of personality and the role of background music in cognitive task performance: A mini-review of conflicting findings and a new perspective. *Frontiers in Psychology*, *8*(NOV), 1–6. https://doi.org/10.3389/fpsyg.2017.01991

Kyoung, E. (2020). The effect of lyrical and non-lyrical background music on different types of language processing - An ERP study. *Korean Journal of Cognitive Science*, *31*(4), 155–178. https://doi.org/10.19066/cogsci.2020.31.4.003

Lim, W., Furnham, A., & McClelland, A. (2022). Investigating the effects of background noise and music on cognitive test performance in introverts and extraverts: A cross-cultural study. *Psychology of Music*, *50*(3), 709–726. https://doi.org/10.1177/03057356211013502

Luke, S. G., & Christianson, K. (2018). The Provo Corpus: A large eye-tracking corpus with predictability norms. *Behavior Research Methods*, *50*(2), 826–833. https://doi.org/10.3758/s13428-017-0908-4

Madsen, C. K. (1987). Background music: Competition for focus of attention. In C. K. Madsen & C. A. Prickett (Eds.), *Applications of research in music behavior* (pp. 315–325). Tuscaloosa, USA: The University of Alabama Press.

Marsden, E., Thompson, S., & Plonsky, L. (2018). A methodological synthesis of self-paced reading in second language research. *Applied Psycholinguistics*, *39*(5), 861–904. https://doi.org/10.1017/S0142716418000036

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2008). Auditory distraction in semantic memory: A process-based approach. *Journal of Memory and Language*, *58*(3), 682–700. https://doi.org/10.1016/j.jml.2007.05.002

Marsh, J. E., Hughes, R. W., & Jones, D. M. (2009). Interference by process, not content, determines semantic auditory distraction. *Cognition*, *110*(1), 23–38. https://doi.org/10.1016/j.cognition.2008.08.003

Marsh, J. E., & Jones, D. M. (2010). Cross-modal distraction by background speech: What role for meaning? *Noise & Health*, *12*(49), 210–216. https://doi.org/10.4103/1463-1741.70499

Martin, R. C., Wogalter, M. S., & Forlano, J. G. (1988). Reading comprehension in the presence of unattended speech and music. *Journal of Memory and Language*, *27*(4), 382–398. https://doi.org/10.1016/0749-596X(88)90063-0

Meng, Z., Lan, Z., Yan, G., Marsh, J. E., & Liversedge, S. P. (2020). Task demands modulate the effects of speech on text processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(10), 1892–1905. https://doi.org/10.1037/xlm0000861

Miller, C. (2014). *The differentiated effects of lyrical and non-lyrical music on reading comprehension (Unpublished Master's thesis)*. Rowan University, New Jersey, USA.

Miller, L. K., & Schyb, M. (1989). Facilitation and interference by background music. *Journal of Music Therapy*, *26*(1), 42–54. https://doi.org/10.1093/jmt/26.1.42

Miller, L. R. (1947). Some effects of radio-listening on the efficiency of reading-type study activities. *Journal of Educational Psychology*, *38*(2), 105–118. https://doi.org/10.1037/h0062228

Mitchell, A. H. (1949). The effect of radio programs on silent reading achievement of ninety-one sixth grade students. *The Journal of Educational Research*, *42*(6), 460–470. https://doi.org/10.1080/00220671.1949.10881709

Mitchell, D. C., & Green, D. W. (1978). The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology*, *30*(4), 609–636. https://doi.org/10.1080/14640747808400689

Morey, R. D., Rouder, J. N., Pratte, M. S., & Speckman, P. L. (2011). Using MCMC chain outputs to efficiently estimate Bayes factors. *Journal of Mathematical Psychology*, *55*(5), 368–378. https://doi.org/10.1016/j.jmp.2011.06.004

Mullikin, C., & Henk, W. A. (1985). Using music as a background for reading: An exploratory study. *Journal of Reading*, *28*(4), 353–358.

Paape, D., & Vasishth, S. (2022). Is reanalysis selective when regressions are consciously controlled? *Glossa Psycholinguistics*, *1*(1). https://doi.org/10.5070/G601139

Panayotov, V., Chen, G., Povey, D., & Khudanpur, S. (2015). Librispeech: An ASR corpus based on public domain audio books. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 5206–5210). IEEE. https://doi.org/10.1109/ICASSP.2015.7178964

Perham, N., & Currie, H. (2014). Does listening to preferred music improve reading comprehension performance? *Applied Cognitive Psychology*, *28*(2), 279–284. https://doi.org/10.1002/acp.2994

Perham, N., & Sykora, M. (2012). Disliked Music can be Better for Performance than Liked Music. *Applied Cognitive Psychology*, *26*(4), 550–555. https://doi.org/10.1002/acp.2826

Quan, Y., & Kuo, Y. (2022). The effects of Chinese and English background music on Chinese reading comprehension. https://doi.org/10.1177/03057356221101647

Que, Y., Zheng, Y., Hsiao, J. H., & Hu, X. (2020). Exploring the effect of personalized background music on reading comprehension. In *Proceedings of the ACM/IEEE Joint Conference on Digital Libraries* (pp. 57–66). https://doi.org/10.1145/3383583.3398543

R Core Team. (2022). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.r-project.org/

Reed, A. (2019). Background Music : The Effects of Lyrics and Tempo on Reading Comprehension and Speed. Retrieved from https://digitalcommons.brockport.edu/psh_theses

Salamé, P., & Baddeley, A. D. (1982). Disruption of short-term memory by unattended speech: Implications for the structure of working memory. *Journal of Verbal Learning* *and Verbal Behavior*, *21*(2), 150–164. https://doi.org/10.1016/S0022-5371(82)90521-7

Salamé, P., & Baddeley, A. D. (1987). Noise, unattended speech and short-term memory. *Ergonomics*, *30*(8), 1185–1194. https://doi.org/10.1080/00140138708966007

Salamé, P., & Baddeley, A. D. (1989). Effects of background music on phonological short-term memory. *The Quarterly Journal of Experimental Psychology*, *41*(1), 107–122. https://doi.org/10.1080/14640748908402355

Sauter, M., Draschkow, D., & Mack, W. (2020). Building, hosting and recruiting: A brief introduction to running behavioral experiments online. *Brain Sciences*, *10*(4), 1–11. https://doi.org/10.3390/BRAINSCI10040251

Sörqvist, P., Halin, N., & Hygge, S. (2010). Individual differences in susceptibility to the effect of speech on reading comprehension. *Applied Cognitive Psychology*, *24*(1), 67–76. https://doi.org/10.1002/acp.1543

Thompson, W. F., Schellenberg, E. G., & Letnic, A. K. (2012). Fast and loud background music disrupts reading comprehension. *Psychology of Music*, *40*(6), 700–708. https://doi.org/10.1177/0305735611400173

Tucker, A., & Bushman, B. J. (1991). Effects of rock and roll music on mathematical, verbal, and reading comprehension performance. *Perceptual and Motor Skills*, *72*(3), 942–942. https://doi.org/10.2466/pms.1991.72.3.942

Vasilev, M. R., Kirkby, J. A., & Angele, B. (2018). Auditory distraction during reading: A Bayesian meta-analysis of a continuing controversy. *Perspectives on Psychological Science*, *13*(5), 567–597. https://doi.org/10.1177/1745691617747398

Vasilev, M. R., Liversedge, S. P., Rowan, D., Kirkby, J. A., & Angele, B. (2019). Reading is disrupted by intelligible background speech: Evidence from eye-tracking. *Journal of*

Experimental Psychology: Human Perception and Performance, *45*(11), 1484–1512. https://doi.org/10.1037/xhp0000680

Venables, W. N., & Ripley, B. D. (2002). *Modern applied statistics with S* (4th ed.). New York, USA: Springer.

Weiss, M. W., Schellenberg, E. G., & Trehub, S. E. (2017). Generality of the memory advantage for vocal melodies. *Music Perception*, *34*(3), 313–318. https://doi.org/10.1525/MP.2017.34.3.313

Weiss, M. W., Schellenberg, G. E., Trehub, S. E., & Dawber, E. J. (2015). Enhanced processing of vocal melodies in childhood. *Developmental Psychology*, *51*(3), 370–377. https://doi.org/10.1037/a0038784

Weiss, M. W., Trehub, S. E., & Schellenberg, E. G. (2012). Something in the way she sings: Enhanced memory for vocal melodies. *Psychological Science*, *23*(10), 1074–1078. https://doi.org/10.1177/0956797612442552

Weiss, M. W., Vanzella, P., Schellenberg, E. G., & Trehub, S. E. (2015). Pianists exhibit enhanced memory for vocal melodies but not piano melodies. *Quarterly Journal of Experimental Psychology*, *68*(5), 866–877. https://doi.org/10.1080/17470218.2015.1020818

Wetzels, R., Matzke, D., Lee, M. D., Rouder, J. N., Iverson, G. J., & Wagenmakers, E.-J. (2011). Statistical evidence in experimental psychology: An empirical comparison using 855 t tests. *Perspectives on Psychological Science*, *6*(3), 291–298. https://doi.org/10.1177/1745691611406923

Woods, K. J. P., Siegel, M. H., Traer, J., & McDermott, J. H. (2017). Headphone screening to facilitate web-based auditory experiments. *Attention, Perception, and Psychophysics*, *79*(7), 2064–2072. https://doi.org/10.3758/s13414-017-1361-2

Yan, G., Meng, Z., Liu, N., He, L., & Paterson, K. B. (2018). Effects of irrelevant background speech on eye movements during reading. *Quarterly Journal of Experimental Psychology*, *71*(6), 1270–1275. https://doi.org/10.1080/17470218.2017.1339718

Zhang, H., Miller, K., Cleveland, R., & Cortina, K. (2018). How listening to music affects reading: Evidence from eye tracking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(11), 1778–1791. https://doi.org/10.1037/xlm0000544

Supplemental files

Comparison Between Experiment 1b and the Student Sub-sample in Experiment 1a

Experiment 1a was conducted online and found no difference in word RTs between the lyrical and instrumental music conditions. Experiment 1b was done in standard lab settings and found a significant difference in word RTs between lyrical and instrumental music. While this difference could be due to the mode of testing (online vs in-lab), a second explanation is that the two samples differed between the two studies. While Experiment 1a used a mixed sample (approx. half taken from a student pool and the other half taken from Prolific [wider UK public]), Experiment 1b used only a student-pool sample. Thus, the difference in the two samples is a possible confounding variable. To eliminate this confounding factor, we compared the student sub-sample from Experiment 1a (N=103) to the full sample of Experiment 1b (both of which came from the same University subject pool). Thus, the only thing that differed in this comparison was the mode of testing. The results are shown in Table S1 below.

Consistent with the results from Experiment 1b, there was a significant difference between lyrical and instrumental music. This was due to word RTs being longer in lyrical compared to instrumental music. This main effect shows that the difference was present in both experiments. Additionally, there was a main effect of Experiment, which was due to RTs being longer in Experiment 1b (lab-based) compared to Experiment 1a (online-based). Critically, however, there was no interaction between Experiment and the lyrical vs. instrumental music comparison. This suggests that the difference was the same in both experiments (see Figure S1 for an illustration). In summary, the difference between lyrical and instrumental music was found both in Experiment 1b and the student sub-sample of

Experiment 1a. Therefore, there was no evidence to suggest that the mode of testing (online vs lab-based) influenced the results.

Table S1

LMM Results for Reaction Times, Comparing the Sample from Experiment 1b (students) and the student Sub-sample in Experiment 1a

Fixed effects	b	SE	t	
Intercept	5.986	0.029	**209.5**	
Instrumental vs. Silence	-0.009	0.006	-1.549	
Lyrical vs. Instrumental	0.026	0.006	**4.305**	
Experiment (1b vs 1a [students])	0.042	0.014	**3.064**	
Instrumental vs. Silence x Experiment	0.006	0.006	0.977	
Lyrical vs. Instrumental x Experiment	-0.002	0.006	-0.305	
Random Effects	Var.	SD	Corr.	
Intercept (subjects)	0.0529	0.2301		
Instrumental vs. Silence (subjects)	0.0097	0.0985	0.07	
Lyrical vs. Instrumental (subjects)	0.0096	0.0978	0.05	-0.45
Intercept (items)	0.0093	0.0966		
Residual	0.0979	0.3129		

Note: Statistically significant effects are formatted in **bold**.

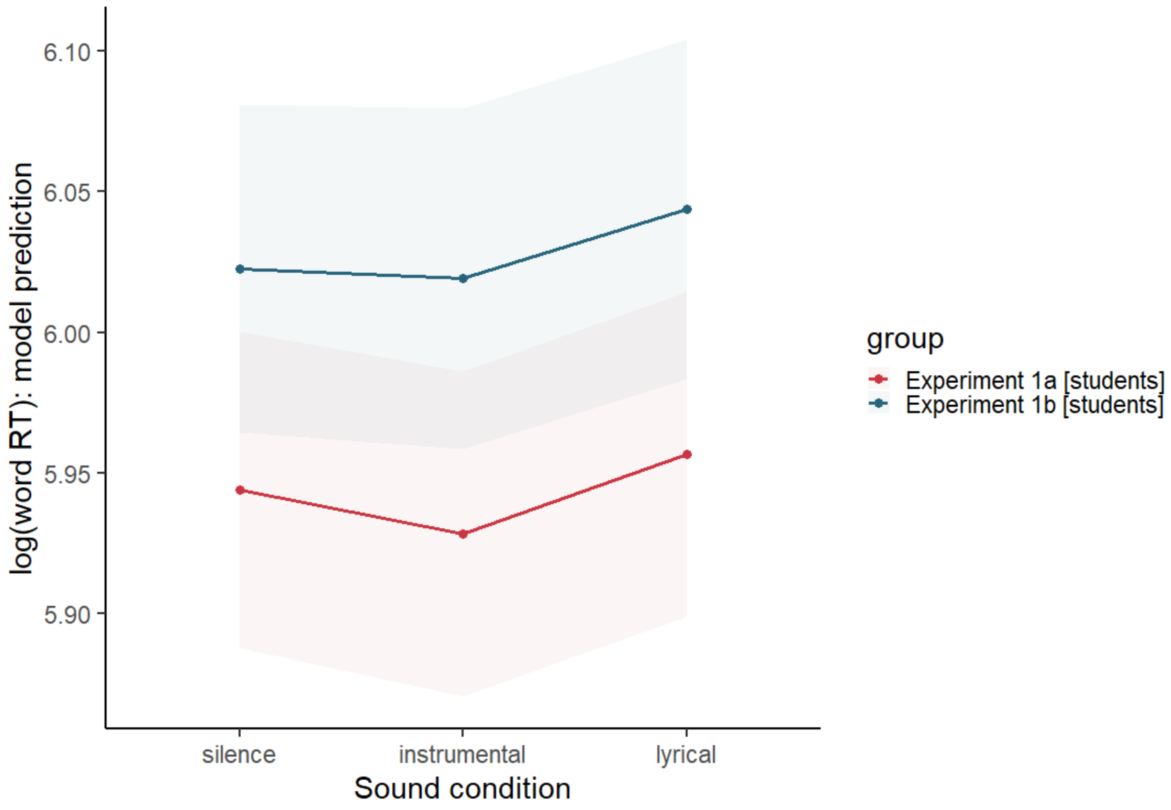

Figure S1. Model predictions from the analysis in Table S1. Light shading indicates 95% confidence intervals.