Psychological and psychophysiological effects of music intensity and lyrics on simulated urban driving

Costas I. Karageorghis [a],[*], Garry Kuan [a],[b], William Payre [c], Elias Mouchlianitis [a],[d], Luke W. Howard [a], Nick Reed [e], Andrew M. Parkes [f]

[a] *Department of Life Sciences, Brunel University London, United Kingdom*
[b] *School of Health Sciences, Universiti Sains Malaysia, Malaysia*
[c] *Institute for Future Transport and Cities, Coventry University, United Kingdom*
[d] *Institute of Psychiatry, Psychology & Neuroscience, King's College London, United Kingdom*
[e] *Department of Engineering and Physical Sciences, University of Surrey, United Kingdom*
[f] *School of Art, Design and Architecture, Monash University, Australia*

A R T I C L E   I N F O

A B S T R A C T

The main aim of this study was to investigate the effect of musical characteristics (i.e., presence of lyrics and loudness) in the context of simulated urban driving. Previous work has seldom isolated musical characteristics and examined these both singularly and interactively. We investigated the potentially distracting effects of processing lyrics through exposing young drivers to the same piece of music with/without lyrics and at different sound intensities (60 dBA [soft] and 75 dBA [loud]) using a counterbalanced, within-subjects design ($N = 34$; $M_{age} = 22.2$ years, $SD = 2.0$ years). Six simulator conditions were included that comprised low-intensity music with/without lyrics, high-intensity music with/without lyrics, plus two controls – ambient in-car noise and spoken lyrics. Between-subjects variables of driving style (defensive vs. assertive) and sex (women vs. men) were explored. A key finding was that the no lyrics/soft condition yielded lower affective arousal scores when compared to the other music conditions. There was no main effect of condition for HRV data (SDNN and RMSSD). Exploratory analyses showed that, for assertive drivers, NASA-TLX Performance scores were lower in the no lyrics/soft condition compared to the lyrics/loud condition. Moreover, women exhibited higher mean heart rate than men in the presence of lyrics. Although some differences emerged in subjective outcomes, these were not replicated in HRV, which was used as an objective index of emotionality. Drivers should consider the use of soft, non-lyrical music to optimise their affective state during urban driving.

1. Introduction

The last two decades have seen increased interest in human factors, particularly driver psychology, in the causation of road accidents (Brodsky, 2015; Parnell, Rand, & Plant, 2020). Despite a general trend towards road safety improvement, younger road users are overrepresented in fatality statistics and remain at heightened risk (Brodsky & Slor, 2013; Department for Transport, 2020; World Health Organization, 2020). A burgeoning body of work links the psychological and psychophysiological effects of music to safety-relevant driving behaviours (Brodsky & Slor, 2013; Wiesenthal, Hennessy, & Totten, 2003), and directly implicates the use of

---

music in road accidents (FakhrHosseini & Jeon, 2019; Royal, 2003).

The presence of digital technologies, such as those used for music delivery, mediate perceptions, behaviours and practices in the vehicular environment or *autosphere*. In 2019, there were 153,315 reported road casualties in Great Britain, including 1,748 fatalities and the lives of 244 young road users (i.e., aged 17–24 years) were sadly lost (DfT, 2020). In particular, young, novice drivers are at a ten-time greater risk of non-fatal accidents than those over 20 years of age (see Brodsky & Slor, 2013). Currently, as we wait for the advance of autonomous vehicles, driver behaviour is the key determinant of road safety and thus a high priority in accident prevention (Department for Transport, 2010; IAM, 2010; Summala, 1996). Notably, human factors, as opposed to environmental or mechanical factors, are the most important contributors to road accidents (Khattak, Ahmad, Wali, & Dumbaugh, 2021).

In order to reduce casualty rates, a better understanding of the effects of various performance-influencing emotional factors, such as in-vehicle music listening, is of pivotal importance (Brodsky, 2015; Brodsky & Slor, 2013). Around 90% of drivers listen to music in-car (Brodsky & Kizner, 2012) and 91% of their musical exposure occurs during driving (Sloboda, O'Neill, & Ivaldi, 2001). For a significant proportion of young male drivers, music listening coupled with the fidelity and intensity of sound are integral to their driving experience (Pêcher, Lemercier, & Cellier, 2009).

The present study is predicated on Fuller's Driver Control Theory (Fuller, 2011), which posits that drivers adjust their behaviour based on their perceived capability to match the demands of the driving situation. Such demands may, in turn, be influenced by perceptual input, the behaviour of other drivers, perceived task difficulty, driving conditions, the experience level of the driver and affective/emotional information. Accordingly, music can be examined as a moderator of behaviour and performance in the context of Fuller's (2011) model.

Music has the potential to distract attention, reduce the capacity for information processing (e.g., traffic signals), force one perceptual modality (auditory system) to dominate attentional processing, and promote social interaction with passengers that presents additional safety threats (Brodsky, 2018; Shih, Huang, & Chiang, 2012). Wickens (2008) detailed how multiple tasks, even in different modalities, can produce performance deficits if the mental workload of a secondary task (e.g., listening to music) exceeds the "residual capacity" not used for the primary task (e.g., driving). Such deficits were shown by Reimer (2009), who demonstrated visual tunnelling (reduced gaze distributions) by drivers when engaged in a secondary auditory task. These consequences of music listening relate to Fuller's (2011) notion of perceived task difficulty, as they can result in the primary task of driving being compromised. Essentially, the distractive qualities of music can modulate actual demand, anticipated demand and perceived capability, leading to decrements in driving performance.

1.1. Music as a distractor and driver responses

Several informational properties of music may influence behaviour by consuming attentional resources: the complexity of rhythmical features, excessive volume, fast tempi and the presence of intelligible lyrics (Ünal, Steg, & Epstude, 2012). An upshot is that the presence of music can engender a sense of immersion and thus distract the driver from critical safety-related cues (Brodsky, 2015). The present study has a specific focus on lyrics and sound intensity. Extraneous in-vehicle sound, particularly at a high intensity (i.e., $\geq$ 75 dBA) can obscure the sound of other vehicles or road users, cause irritability, reduce attentional control, prompt risk-taking behaviours, and even harm a driver's hearing (Brodsky, 2015; Ünal et al., 2012). Lyrics can be distracting as the linguistic element causes active, syntactic processing that involves specific neural mechanisms (see Chien & Chan, 2015). From an information processing perspective (e.g., Consiglio, Driscoll, Witte, & Berg, 2006; Ho, Reed, & Spence, 2007), additional input such as music can negatively affect performance if the attentional demands exceed the driver's capacity to respond (e.g., in response to unexpected events; see Millet, Ahn, & Chattah, 2019).

Given that attention and arousal have a reciprocal relationship (see Vuilleumier, Armony, & Dolan, 2003), music that is calming in nature may enhance the attentiveness of overloaded drivers (e.g., in an urban environment). Brodsky and Kizner (2012) reviewed evidence documenting the use of in-car music and reported a marked, cross-cultural tendency for drivers to favour stimulative music, thus under-evaluating the risks of its use, directly in line with the predictions of Fuller's (2011) aforementioned model. The same authors designed a sedative music programme that was employed for an *in vivo* study. When compared to self-selected stimulative music, the sedative programme improved driving performance in terms of safety-relevant behaviours and positive affective states. Recent work by FakhrHosseini and Jeon (2019) has demonstrated the propensity of music to assuage anger, and thus promote safer driving behaviours. Furthermore, it is apparent that music variables such as volume and tempo are directly related to driving speed (Brodsky, 2002); high volume absorbs sensory channels and fast tempi upregulate psychomotor arousal.

The psychology literature has highlighted biological sex as a moderator that is worthy of study in the context of selective attention (e.g., Merritt et al., 2007). Investigation of how the sexes respond differently to musical stimuli in the driving context is important from a road safety perspective; there are many anecdotal reports suggesting that women are better able to multitask (Szameitat, Hamaida, Tulley, Saylik, & Otermans, 2015), but this is not borne out in experimental studies (e.g., Mäntylä, 2013). Accordingly, we adopted sex as an exploratory independent variable in all analyses in the present study.

1.2. Rationale for the present study

An implication of Fuller's (2011) model is that in-car music choices will moderate the effects of psychological state (emotion, attention and arousal) on speed and risk perception, dependent on situational and personal factors. Drivers may misperceive either the difficulty of the driving task and/or the distractive effect of the music (Brodsky, 2002); empirical evidence suggests that drivers generally overestimate their own driving skills and underestimate the risk of distractions (e.g., Amado, Arikan, Kaca, Koyunca, & Turkan, 2014).

Music may even alter perceptions of tasks by activating knowledge structures (see North & Hargreaves, 2008 for review) or priming affective states that predispose us to focus on information with either a positive or negative affective valence (Ignacio, Gerkens, Aguinaldo, Arch, & Barajas, 2019; Juslin, 2013). For example, aggressive forms of music, such as heavy metal (generally a negatively valenced artform), can predispose drivers towards driving with less regard for other road users (Arnett, 1991). Moreover, as a passive form of distraction, music can play an important part in unconscious decision-making processes (e.g., participants who were distracted by music outperformed individuals distracted by more difficult tasks, such as solving an anagram puzzle; McMahon, Sparrow, Chatman, & Riddle, 2011).

Ünal et al. (2012) reported that listening to lyrical music increased perceptions of mental effort during simulated driving irrespective of attentional load; nonetheless, this load did not lead to impaired driving performance. The researchers concluded that drivers attempt to regulate their mental effort as a cognitive compensatory strategy to deal with task demands. Accordingly, mental effort is a salient mediator of the relationship between music listening and driving performance.

1.3. Aims and hypotheses

The aim of the study was to investigate the effects of the lyrical (with/without) and loudness (i.e., loud/soft) dimensions of music on a range of psychological (e.g., affect), psychophysiological (e.g., heart rate variability [HRV]) and behavioural outcomes (e.g., pedal control). A secondary aim was to examine differences across the full range of dependent variables between two clusters of participants – defensive drivers and assertive drivers – identified through exploratory cluster analysis.

We hypothesised that a no lyrics/soft condition would engender the best affective state for simulated urban driving (i.e., positive affective valence coupled with moderate affective arousal) when compared to lyrics/loud, no lyrics/loud and lyrics/soft (H_1). We also expected that the no lyrics/soft condition would lead to the lowest scores for Rating Scale Mental Effort (RSME) scores and the NASA Task Load Index (TLX) items (H_2). In regard to the lyrics/loud condition, we hypothesised that higher mean speeds and shorter completion times in the lyrics/loud condition versus the other three experimental conditions (H_3). Similarly, we expected the most risk-taking behaviour to be evident through the video analysis of trigger incidents and simulator data (accelerator/brake pedals) for a subset of these trigger incidents (i.e., those incidents likely to entail the use of the pedals) in the lyrics/loud condition (H_4).

A number of exploratory analyses were conducted that entailed: (a) influence of the experimental conditions on heart rate (HR) and HRV; (b) cluster analyses in which we crossed the defensive/assertive drivers independent variable with four experimental conditions and two control conditions across a range of dependent variables; and (c) the use of sex as an independent variable in a range of analyses. Given the exploratory nature of these analyses we tested the null hypothesis in each case (H_5).

2. Method

2.1. Power analysis

To calculate the appropriate sample size, a power analysis was undertaken using G*Power3 (Faul, Erdfelder, Buchner, & Lang, 2009). Effect size was based on a conservative estimation in relation to the effect of auditory stimuli on core affect during simulated driving (van der Zwaag et al., 2012; $\eta^2 = 0.19$). With an effect size of $\eta^2 = 0.10$, at an alpha level of 0.05, and power at 0.80 to protect beta at four times the level of alpha (Cohen, 1988, pp. 4–6), the analysis indicated that 31 participants would be required to detect effects on affective responses on the Affect Grid. Three additional participants were recruited to protect the study against experimental attrition and deletions due to outliers.

2.2. Participants

With written informed consent, a convenience sample of 34 adult volunteers was recruited (17 women and 17 men; $M_{age} = 22.2$ years, $SD = 2.0$ years; $M_{BMI} = 24.9$, $SD = 5.2 \, \text{kg m}^{-2}$). Recruitment was conducted through word-of-mouth, promotional flyers, posters, social media posts and e-mail circulars at Coventry University, UK as well as in the surrounding area (i.e., the city of Coventry, UK). Ethical approval was granted by the Ethics Committee of Brunel University London, UK as well as the Ethics Committee of Coventry University, UK. Inclusion criteria were that participants: (a) were in apparently good health; (b) did not have a tendency to feel sick when playing immersive video games or suffer from motion sickness as a road passenger; (c) did not have a hearing deficiency and/or visual impairment; and (d) held a UK driver's licence. Details of a four-stage music selection procedure to establish appropriate playlists for the experimental participants can be found in Supplementary File 1.

2.3. Psychological measures

Several psychological measures were administered, full details of which are included in Supplementary File 2, with a brief outline provided herein. The Affect Grid (AG; Russell, Weiss, & Mendelsohn, 1989) is a single-item measure that facilitates a quick assessment of affect along the dimensions of pleasure–displeasure and arousal–sleepiness. The NASA Task Load Index (NASA-TLX; Hart & Staveland, 1988) is a tool that measures subjective mental workload. The Rating Scale Mental Effort (RSME; Zijlstra, 1993) is a unidimensional instrument used to assess subjective mental workload. The Multidimensional Driving Style Inventory (MDSI; Taubman-Ben-Ari, Mikulincer, & Gillath, 2004) is a self-report questionnaire that is widely used in the driving psychology literature to evaluate four driving styles.
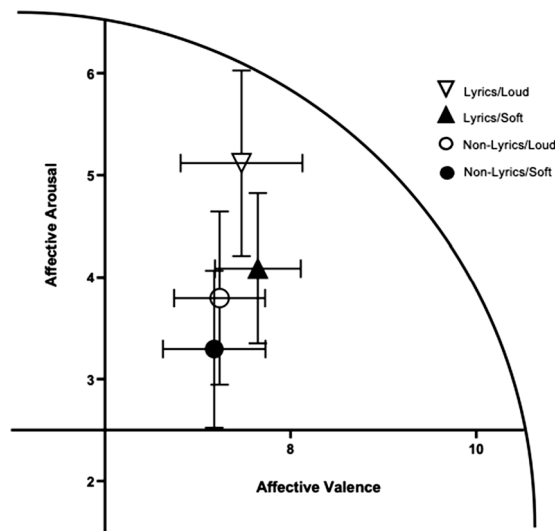
Fig. 1. The Lyrics × Loudness Interaction for Core Affect. *Note*. Error bars represent 95% confidence intervals.

2.4. Driving simulator

The fixed-based driving simulator consisted of a car buck (see Fig. 1 in Supplementary File 3) and a three-channel HD projection system. This system provided a full panoramic view rendered on a 220° curved projection screen, and a seamless output with 5760 × 1080 px display resolution at 60 Hz. The wing mirrors had integrated 10″ SVGA resolution LCD screens, whereas a 32″ LED HD screen was mounted at the rear of the buck to simulate the rear-view mirror. OpenDS 4.0 software was used to create the simulation environment, with the speedometer displayed on the projection screen. An outline of the driving simulator route can be seen in Fig. 2 (Supplementary File 3). The steering wheel was equipped with a force feedback steering control unit. The buck was fitted with two bass shakers that served to convert audio signals from an amplifier into physical vibrations that participants felt through the frame. Music was administered in stereo via two wireless 20-W portable speakers (JBL Charge 3) positioned to the left and right of the back of the buck.

2.5. Experimental procedures

In advance of visiting the simulator room, participants completed a set of online questionnaires. These included an informed consent form, demographic questions and idiographic items to gauge which radio stations they listened to while driving. Each participant visited the simulator room on one occasion for ~2 h. On arrival, the participant was offered ginger confectionary (nut-free), as this is known to mitigate the effects of motion sickness (Lien et al., 2003). Some initial/baseline measures were taken, and a full habituation to an urban driving simulation task ensued (i.e., 10 min in the simulator), coupled with an opportunity to ask questions. The participant was then led on a 10-min walk around the campus of Coventry University to refresh them prior to the experimental phase.

Each participant completed four experimental conditions and two control conditions in a counterbalanced order using the high-grade simulator. Participants completed an 8-min urban driving simulation on the left-hand side of the road (i.e., as required on UK roads). Each condition was separated by a 5-min break that incorporated a 2-min filler task (a wordsearch). The simulation included five triggers or events: (a) A pedestrian who walked at 5 km/h (from the left-hand side) across a zebra crossing; (b) a garbage truck that moved slowly in the left-hand lane and prompted an overtaking manoeuvre; (c) traffic lights that changed to red; (d) a slow vehicle on a stretch of road where overtaking was prohibited; and (e) a vehicle that cut across unexpectedly (from the right-hand side) at a four-way intersection.

The five triggers were used in the collation of observational data from a video recording of each simulation and a meaningful subset (i.e., wherein participant responses were most consistent; pedestrian on zebra crossing and vehicle cutting) was used for pedal data (accelerator and brake). One member of the research team viewed and risk-rated the subset of triggers following discussion regarding standardisation of the process with the first author. Once the risk ratings had been completed, a random selection of 25% of ratings were checked with the first author for accuracy. No adjustments were deemed necessary.

A within-subjects experimental design was adopted. Participants were instructed to have adequate sleep the night before testing, avoid alcohol consumption and refrain from ingesting caffeine on the day of the visit. Four experimental (music) conditions (lyrics/loud, no lyrics/loud, lyrics/soft and no lyrics/soft) and two control conditions (urban traffic noise and spoken lyrics [using the same lyrics as in the experimental conditions]) were administered to identify the psychological, psychophysiological and behavioural effects of music during a simulated urban driving task (60 dBA [soft] and 75 dBA [loud]).

Conditions were presented in a counterbalanced order with each experimental trial lasting ~8 min. Experimental manipulations
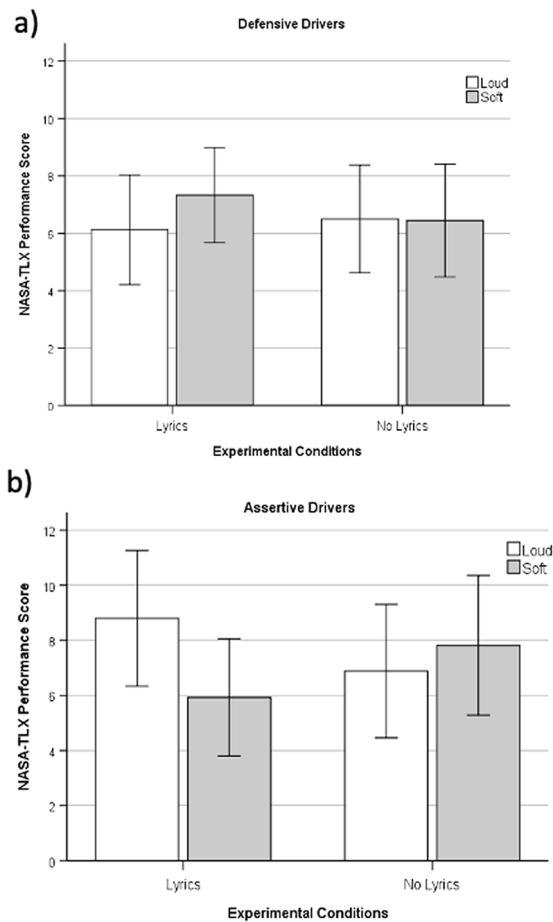
Fig. 2. Higher-order Cluster × Lyrics × Loudness Interaction for Defensive Drivers (a) and Assertive Drivers (b) for the NASA-TLX Performance Variable ($p = 0.049$). *Note.* Error bars represent 95% confidence intervals. NASA-TLX = NASA Task Load Index.

coincided with the 8-min urban driving phase. Participants were asked to maintain a speed of 30 mph and observe all traffic signals/ road signs in accord with the UK Highway Code. The driving simulator was equipped with an automatic transmission system and so participants were instructed to leave the gearstick in "drive", and use only the steering wheel and accelerator/brake pedals. They were also asked to use the rear-view and side-view mirrors as normal.

2.6. Heart rate variability

HRV data were recorded throughout each trial. Data were imported into Kubios software and the signal broken down into four samples, each lasting 5 min. This epoch was used in order to avoid anomalies from the beginning/end of the trial and to account for the fact that drivers performed the simulation task at slightly different speeds. Two time-domain indices were extracted from the cardiac electrical signal. Standard deviation of normal-to-normal RR intervals (SDNN) was used as an index of global activity of the sympathetic–parasympathetic system and root mean square of successive RR interval differences (RMSSD) was used as an index of parasympathetic activity (Acharya, Joseph, Kannathal, Lim, & Suri, 2006).

2.7. Data analysis

Data were checked for inputting errors then screened for univariate outliers using standardised scores ($z > ± 3.29$) and for multivariate outliers using the Mahalanobis distance test ($p < 0.001$; Tabachnick & Fidell, 2019). Data were also examined for the parametric assumptions that underlie (multivariate) analysis of covariance (M)ANCOVA (Tabachnick & Fidell, 2019). The behavioural data derived from video analysis were not checked for the parametric assumption as these data were predicated on frequency counts (i. e., nominal level data). Where violations of the sphericity assumption were observed, Greenhouse–Geisser-adjusted F tests were used. Checks were also made for the eight assumptions that underlie the inclusion of covariates (e.g., homogeneity of variance, linearity of regression; see Tabachnick & Fidell, 2019).

Initial analyses employed repeated-measures (RM) 2 (Lyrics) × 2 (Loudness) (M)ANCOVAs, with scores from the urban traffic noise

and spoken lyrics control conditions entered into the model as covariates. This enabled us to establish the interactive and singular effects of the two musical qualities under investigation. Thereafter, exploratory analyses were conducted using a mixed-model approach that embraced the between-subject factors of driving style (defensive vs. assertive) and sex (women vs. men). A cluster analysis was conducted using the *K*-means algorithm applied to participants' MDSI scores. Two expected clusters were defined *a priori*, relating to defensive and assertive driving styles, that would be informed by the eight dimensions of the MDSI (i.e., dissociative, anxious, risky, angry, high-velocity, distress reduction, patient and careful). Based on each participant's scores, the clustering algorithm extracted the two groups of drivers as expected (i.e., defensive and assertive).

Behavioural data were collated under two categories: (a) video data pertaining to five triggers (pedestrian, garbage truck, traffic lights, slow vehicle, vehicle cut-in); (b) and data from the simulator, derived from the accelerator and brake pedals position (i.e., 0 = no pressure applied, 1 = maximal acceleration/braking). These variables were analysed using the same RM and mixed-model (M) ANCOVAs as the affective, perceptual and cardiac variables. Significant *F* tests were followed up with checks of 95% confidence intervals to identify where differences lay.

Where parametric assumptions were not met and transformations would not serve to normalise the distribution, rank-based nonparametric three-way and two-way ANOVAs (same factors as in the parametric analyses) were conducted using the nparLD package (Noguchi, Gel, Brunner, & Konietschke, 2012) of data analysis software R. Such analyses do not enable the inclusion of a covariate, and so for the within-subjects factors of lyrics and intensity, we first ran analyses containing only data from the experimental conditions to single out the two-way interaction of main interest, and second with the inclusion of the two control conditions to ascertain whether differences would emerge among the four experimental and two control conditions. These analyses were followed-up with the inclusion of the between-subject factors of cluster (defensive vs. assertive) and sex (women vs. men). The Wald-Type Statistic (WTS) and ANOVA-Type Statistic (ATS) were computed in these nonparametric analyses.

3. Results

Two univariate outliers were modified to be one unit larger or smaller than the next most extreme score in the distribution, until the corresponding *z*-scores fell within the range −3.29 to 3.29 (Tabachnick & Fidell, 2019). No multivariate outliers were identified. Q–Q plots used to examine the distributional properties of interval and ratio data in each cell of the analysis revealed minor violations in six out of 240 cells. Covariates were only included when relevant assumptions were met (see Tabachnick & Fidell, 2019). Descriptive statistics are presented in Table 1.

3.1. Core affect

A 2 (Lyrics) × 2 (Loudness) MANCOVA was computed for the core affect variables of valence and arousal, with inclusion of the urban noise and spoken lyrics control-condition valence and arousal scores as covariates. Omnibus statistics showed a marginally non-significant two-way interaction of Lyrics × Loudness, Pillai's = 0.184, $F(2, 28) = 3.171$, $p = 0.057$, $\eta_p^2 = 0.19$. Given the large effect size, this interaction was further investigated using step-down *F* tests, which showed a significant two-way interaction for arousal, $F(1, 29) = 5.31$, $p = 0.029$, $\eta_p^2 = 0.16$. Examination of confidence intervals indicated that arousal scores in the no lyrics/soft condition were lower than in the lyrics/loud condition (see Fig. 1).

Table 1
Descriptive statistics across conditions for all variables.

Condition	Lyrics/Loud		Lyrics/Soft		No-lyrics/Loud		No-lyrics/Soft		Traffic noise		Spoken lyrics	
Variable	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Affect Grid												
Affective valence	7.47	1.83	7.65	1.32	7.24	1.43	7.18	1.60	3.50	1.93	4.18	2.44
Affective arousal	5.12	2.53	4.09	2.18	3.79	2.50	3.29	2.14	5.88	2.55	4.18	2.51
Mental effort	46.94	26.59	44.74	24.36	41.47	23.94	40.80	26.56	51.59	26.56	49.82	26.81
NASA-TLX												
Mental Demand	7.79	4.36	6.74	4.52	6.91	4.60	6.68	4.48	7.85	4.59	7.82	4.90
Physical Demand	5.47	4.27	4.59	3.23	4.85	3.43	4.35	3.15	5.09	3.62	5.38	4.18
Temporal Demand	5.38	4.08	4.79	3.40	5.32	3.83	4.76	3.64	6.53	4.47	5.62	4.12
Performance	7.15	4.74	6.79	4.80	6.65	4.38	6.97	5.35	7.21	4.26	6.65	4.02
Effort	8.53	5.12	6.82	4.52	7.41	4.59	6.91	4.96	8.47	4.89	7.62	4.89
Frustration	5.09	4.03	5.65	4.39	4.82	4.25	5.44	4.49	6.76	4.75	7.26	4.89
Cardiac variables												
Mean heart rate	80.96	10.54	80.04	10.33	79.02	9.18	79.73	12.67	78.27	10.64	78.21	3.99
HRV SDNN	38.04	12.11	38.26	14.65	37.71	14.38	38.01	14.71	40.37	16.22	40.56	15.85
HRV RMSSD	35.25	14.69	36.39	16.49	36.90	17.19	37.59	17.04	38.83	18.17	38.19	19.81
Behavioural												
Pedestrian crossing	1.03	1.25	1.06	1.28	0.71	1.02	1.09	1.29	0.94	1.19	0.79	1.05
Garbage truck	0.62	0.91	0.47	0.81	0.50	0.85	0.35	0.76	0.71	0.96	0.71	0.92
Traffic light	3.85	0.55	3.56	1.14	3.71	0.96	3.74	0.74	3.76	0.73	3.76	3.79
Slow car	1.38	0.64	1.41	0.55	1.50	0.50	1.38	0.59	1.47	0.55	1.53	0.55

Note. The means and standard deviations displayed above are not covariate adjusted.

3.2. Rating scale mental effort

A 2 (Lyrics) × 2 (Loudness) ANCOVA with urban traffic noise control and spoken lyrics as covariates did not show any significant interactions or main effects (all $Fs < 1.5$, $ps > 0.05$; see Table 2).

Table 2
Inferential statistics of within-subjects (M)ANCOVA analyses.

	Pillai's Trace	F	df	p	η_p^2
Core Affect[a]					
Lyrics × Loudness	0.185	3.18	2, 28	0.057	0.19
Lyrics	0.039	0.57	2, 28	0.570	0.04
Loudness	0.001	0.01	2, 28	0.988	0.02
Affective Valence[b]					
Lyrics × Loudness	—	0.01	1, 29	0.952	0.00
Lyrics	—	0.02	1, 29	0.886	0.00
Loudness	—	0.01	1, 29	0.931	0.00
Affective Arousal[b]					
Lyrics × Loudness	—	5.31	1, 29	0.029	0.16
Lyrics	—	1.82	1, 29	0.876	0.00
Loudness	—	0.02	1, 29	0.893	0.00
Rating Scale Mental Effort					
Lyrics × Loudness	—	0.69	1, 31	0.414	0.02
Lyrics	—	1.48	1, 31	0.233	0.05
Loudness	—	0.27	1, 31	0.614	0.01
NASA-TLX Mental Demand					
Lyrics × Loudness	—	0.11	1, 31	0.748	0.00
Lyrics	—	0.58	1, 31	0.453	0.02
Loudness	—	0.35	1, 31	0.560	0.01
NASA-TLX Physical Demand					
Lyrics × Loudness	—	0.80	1, 31	0.380	0.03
Lyrics	—	0.03	1, 31	0.856	0.00
Loudness	—	0.63	1, 31	0.803	0.00
NASA-TLX Temporal Demand					
Lyrics × Loudness	—	0.73	1, 31	0.401	0.02
Lyrics	—	0.32	1, 31	0.578	0.01
Loudness	—	0.64	1, 31	0.802	0.00
NASA-TLX Performance					
Lyrics × Loudness	—	0.00	1, 31	0.996	0.00
Lyrics	—	0.02	1, 31	0.996	0.00
Loudness	—	0.35	1, 31	0.559	0.01
NASA-TLX Effort					
Lyrics × Loudness	—	1.05	1, 31	0.313	0.03
Lyrics	—	0.14	1, 31	0.707	0.00
Loudness	—	0.91	1, 31	0.346	0.03
NASA-TLX Frustration					
Lyrics × Loudness	—	0.59	1, 31	0.448	0.02
Lyrics	—	2.09	1, 31	0.159	0.06
Loudness	—	0.05	1, 31	0.825	0.00
Mean heart rate[c]					
Lyrics × Loudness	—	1.34	1, 31	0.255	0.04
Lyrics	—	0.91	1, 31	0.345	0.03
Loudness	—	2.43	1, 31	0.129	0.07
RMSSD[c]					
Lyrics × Loudness	—	2.08	1, 31	0.160	0.06
Lyrics	—	0.39	1, 31	0.539	0.01
Loudness	—	0.66	1, 31	0.423	0.02
SDNN[c]					
Lyrics × Loudness	—	0.20	1, 31	0.656	0.00
Lyrics	—	1.28	1, 31	0.267	0.04
Loudness	—	0.55	1, 31	0.464	0.02

Note. All analyses include urban noise control and spoken lyrics control as covariates. NASA-TLX = NASA Task Load Index; RMSSD = Root mean square of successive RR interval differences; SDNN = Standard deviation of NN intervals.

[a] MANCOVA for core affect, with affective valence and affective arousal as dependent variables.
[b] Step-down ANCOVA analyses for the affective valence and affective arousal dimensions of the Affect Grid.
[c] Nonparametric rank-based factorial analyses, F values represent ANOVA-Type Statistic (ATS).

3.3. NASA task load index (TLX)

A 2 (Lyrics) × 2 (Loudness) ANCOVA with urban traffic noise control and spoken lyrics as covariates conducted for each of the six NASA-TLX dimensions (Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, Frustration) did not show any significant interactions or main effects (all $ps > 0.05$; see Table 2).

3.4. Total time to complete simulated driving task

A 2 (Lyrics) × 2 (Loudness) ANCOVA on total time to complete task with urban traffic noise and spoken lyrics scores as covariates did not show any significant interaction or main effects (all $ps > 0.05$; see Table 2).

3.5. Cardiac variables

The 2 (Lyrics) × 2 (Loudness) nonparametric factorial analyses for each of the cardiac variables (i.e., mean HR, SDNN and RMSSD) revealed no interactions or main effects (all $ps > 0.05$; see Table 2).

3.6. Exploratory analyses

All of the aforementioned variables were also investigated with inclusion of the between-subject factors of driving style (assertive vs. defensive) and sex (women vs. men). In terms of driving style, we used a cluster analysis predicated on responses to the Multi-dimensional Driving Style Inventory (MDSI) and this extracted two clusters of cases that were labelled "defensive drivers" (Cluster 1: $n = 21$; $n_{women} = 12$) and "assertive drivers" (Cluster 2: $n = 13$; $n_{women} = 5$). In terms of mean scores, the two clusters broke down as follows: Dissociative Driving Style – $M = 1.00$ vs. $M = 2.38$, $p < 0.001$; Anxious Driving Style – $M = 2.00$ vs. $M = 2.71$, $p = 0.046$; Risky Driving Style $M = 1.00$ vs. $M = 2.00$, $p = 0.001$; Angry Driving Style – $M = 1.20$ vs. $M = 4.60$, $p < 0.001$; High Velocity Driving Style – $M = 1.33$ vs. $M = 4.83$, $p < 0.001$; High Distress Reduction Driving Style – $M = 1.50$ vs. $M = 1.25$, $p = 0.067$; Patient Driving Style – $M = 5.75$ vs. $M = 3.50$, $p < 0.001$; Careful Driving Style – $M = 6.00$ vs. $M = 4.00$, $p < 0.001$.

A 2 (Cluster) × 2 (Lyrics) × 2 (Loudness) ANCOVA for the NASA Performance variable showed a significant three-way interaction with a medium effect size, $F(1, 30) = 4.22$, $p = 0.049$, $\eta_p^2 = 0.12$. Examination of 95% CIs indicated that Performance scores in the no lyrics/soft condition were lower than in the lyrics/loud condition for assertive drivers (see Fig. 2). A 2 (Sex) × 2 (Lyrics) × 2 (Loudness) ANCOVA for mean HR showed a significant two-way interaction of Sex × Lyrics with a medium-to-large effect size, $F(1, 30) = 4.47$, $p = 0.043$, $\eta_p^2 = 0.13$. Examination of 95% CIs indicated that, in the presence of lyrics, women exhibited higher mean HR than men.

3.7. Simulator data

Video data pertaining to five triggers (pedestrian crossing, garbage truck, traffic lights, slow vehicle and vehicle cut in) and data derived from the accelerator and brake pedals position analysed using nonparametric tests. The 2 (Lyrics) × 2 (Loudness) nonparametric factorial analyses did not show any interaction or main effects for any of the trigger or pedal data measures (all $ps > 0.05$). Moreover, no interaction or main effects emerged after including the between-subject factors of sex and cluster (all $ps > 0.05$).

4. Discussion

The main purpose of the present study was to investigate the effects of the musical dimensions of lyrics and loudness on a broad range of psychological, psychophysiological and behavioural outcomes. The work was grounded in Fuller's (2011) Driver Control Theory, with explicit emphasis on the in-car, auditory environment during simulated urban driving. The first hypothesis (H_1) that the no lyrics/soft condition would engender the best affective state for simulated driving was accepted (see Fig. 1). The second hypothesis (H_2), that the no lyrics/soft condition would lead to the lowest scores for the Rating Scale Mental Effort and NASA-TLX items was not supported.

The third hypothesis (H_3), that higher mean speeds and shorter completion times would be observed in the lyrics/loud condition, was not supported. Similarly, the fourth hypothesis (H_4), that the highest incidence of risk-taking behaviour would be observed in the lyrics/loud condition was not supported. The series of exploratory analyses (H_5) that entailed comparisons of defensive and assertive drivers as well as the two sexes, yielded very few differences, and so the null hypothesis was accepted.

There was a higher-order, Cluster × Lyrics × Loudness interaction for the NASA-TLX Performance item, which indicated that assertive drivers recorded higher scores than defensive drivers in the lyrics/loud condition when compared to no lyrics/soft (see Fig. 2). The defensive drivers exhibited the converse: higher scores in the lyrics/soft condition. There was also a sex difference in mean HR, wherein women exhibited higher HR than men when exposed to lyrical music (see Fig. 3).

Overall, it is clear that the auditory manipulations under investigation had little bearing on psychophysiological and behavioural outcomes. This is not in accord with the findings of a recent meta-analysis on the effects of music on driving behaviours (Millet et al., 2019). Nonetheless, the findings of other studies concur with the present findings in terms of a lack of effect (Hughes, Rudin-Brown, & Young, 2013; Ünal, de Waard, Epstude, & Steg, 2013; Ünal, Platteel, Steg, & Epstude, 2013). It is clear that drivers adopt cognitive and behavioural strategies to overcome excesses in certain musical qualities, such as loudness (Ünal, de Waard, Epstude, & Steg, 2013). Such strategies pertain to Fuller's (2011) notion of *task difficulty homeostasis* wherein "…the driver needs to continuously create and
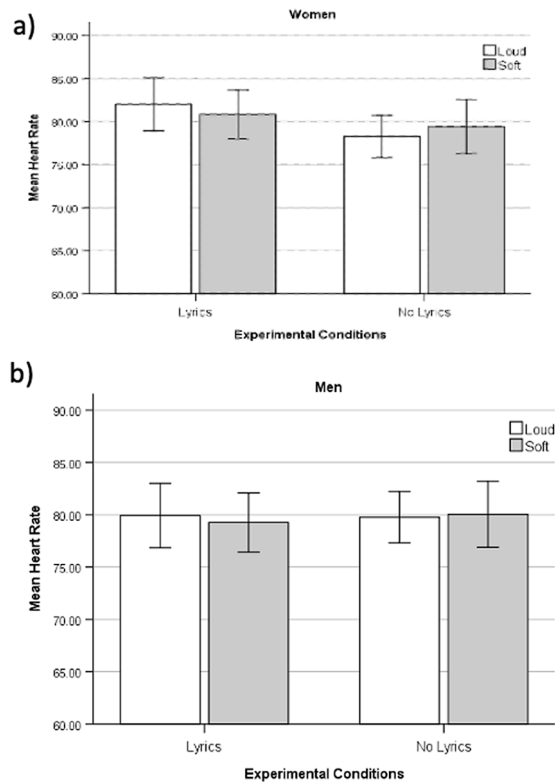
Fig. 3. Significant Sex × Lyrics × Loudness Interaction for mean heart rate ($p = 0.043$). *Note.* Error bars represent 95% confidence intervals.

maintain conditions for driving within these limitations." (p. 13).

The most marked effect of musical manipulation was observed in the arousal dimension of affect (see Fig. 1). From an attentional processing perspective (Fuller, 2011; Ünal et al., 2012), this finding is salient for the context of urban driving, wherein there is a range of potential stressors that might include dense traffic, frequent junctions, roadworks, interweaving motorcyclists/cyclists, pedestrians engaging in risk-taking behaviour, aggressive eye contact with other drivers, frequent speed checks and buses blocking the road (Brodsky, 2002, 2015; Wen, Sze, Zeng, & Hu, 2019). Excessive levels of physiological arousal during driving, just as in any perceptual-motor skill, have the potential for errors of under-inclusion, with related consequences for road safety (Brodsky, 2015; Fuller, 2011).

4.1. Affective responses

The present findings support the notion that the lyrical and intensity dimensions of music can engender a meaningful effect on drivers' affective state, in support of numerous previous studies (e.g., Brodsky, 2002; Chien & Chan, 2015). Interestingly, the music manipulations that we employed had little bearing on the valence dimension of affect – only the arousal dimension (see Fig. 1). Note that we explored the arousal dimension at univariate level, in light of a large multivariate effect ($\eta_p^2 = 0.19$) that was marginally non-significant ($p = 0.057$). As expected, all four experimental conditions elicited affect ratings that fell in the top-right quadrant of Russell's (1980) Circumplex Model of Affect. It appears that this model can be a useful tool both for drivers and those who compile playlists for driving to optimise selections.

Although we went to some lengths to ensure that music selections matched participants' musical predilections (i.e., through drawing on tracks from their preferred radio stations), we did not afford participants individual choice (i.e., they could not choose their favourite tracks). This would suggest that autonomy in music selection might lead to more positive affective valence, for which there are parallels in the music-in-exercise literature (e.g., Hutchinson et al., 2018). The decision not to allow a free choice in music was taken in order that: (a) there could be some standardisation of musical qualities across participants; and (b) we did not confound our findings through the bias associated with self-selection (i.e., an experimenter-type effect).

4.2. Mental load

In terms of the Rating Scale Mental Effort and NASA-TLX, no interaction or main effects emerged, with the exception of a higher-order Cluster × Lyrics × Loudness interaction for the NASA-TLX Performance dimension in the exploratory analyses (see Fig. 2). The task load-related findings suggest that in-vehicle use of music generally does not exert a high attentional demand and that listening to music while driving is a form of parallel processing that does not threaten the primary task (i.e., driving). This concurs with the

findings of several previous studies in the human factors literature (e.g., Hughes et al., 2013; Ünal, de Waard, Epstude, & Steg, 2013; Ünal, Platteel, Steg, & Epstude, 2013) but stands in contrast to the findings of a recent meta-analysis (Millet et al., 2019). Clearly, we did not employ excessively loud music in the present study; nonetheless, we did administer loud music (75 dBA) in two of the experimental conditions. There is evidence that music > 80 dBA (Millet et al., 2019) can be disruptive but we did not want to place participants at any risk of temporary hearing loss.

Examining the interaction effect for Performance in more detail, this NASA-TLX item taps a subjective assessment of performance efficacy. The finding that defensive drivers reported lower scores in the lyrics/loud vs. lyrics/soft condition, while assertive drivers reported the converse. This provides a vista into the attentional style of defensive drivers insofar as their focus on the road, and by extension their perceived efficacy, is degraded by the presence of loud, lyrical music. Contrastingly, the assertive drivers are apparently given a boon by such music. There are clearly consequences for risk-taking and ignorance to such behaviours for assertive drivers (Brodsky, 2015; Brodsky & Kizner, 2012). Ostensibly, under the influence of loud, lyrical music, assertive drivers can become a greater threat to themselves and to other road users. Fuller (2011) refers to the notion of "poorly calibrated drivers" (p. 17); these are drivers who either overestimate their capability or underestimate task demand. They tend to operate with less spare capacity and thus visit the boundary where task demand meets capability more frequently.

4.3. Mean speed and completion time

The experimental manipulations had no bearing on mean speed or course completion time (see Table 2). Given that excessive speed is known to be one of the biggest contributors to road accidents (e.g., head-on collisions; Taylor, Baruya, & Kennedy, 2002), it appears that the music conditions employed in the present study did not have either a deleterious or beneficial effect. It is notable that although the lyrics/loud condition elicited the highest scores for subjective arousal (see Fig. 1), this did not manifest in any increase in driving speed (see Table 2). There is the possibility that, as the participants had been instructed to comply with rules of the road, they knew that their driving performance was being observed and monitored (i.e., a Hawthorne-type effect), therefore they would endeavour to comply with the speed restrictions (30 mph). In a real-life driving environment, they might well behave differently and therein lies one of the major limitations of simulator-based research (Brodsky, 2015).

4.4. Psychophysiological data

The differences observed in subjective arousal in response to the experimental conditions (see Fig. 1) were not reflected in HRV data that tracked the function of the sympathetic and parasympathetic nervous systems. Accordingly, self-report measures of affect and psychophysiological measures do not tally in the present study. There are noted difficulties in the application of HRV measures in tasks such as simulated driving (Berntson et al., 1997) but in the present study, we did record baseline measures and only extracted 5 min of HRV data from the middle epoch of the ~8-min simulation task. For greater consistency, we could have kept participants in the simulator in between trials (i.e., for a ~90-min period in total) but this would have been rather uncomfortable for them and so we chose to incorporate a respite period/filler task in between trials. This is also consistent with the protocol adopted in previous simulation studies (see e.g., Ünal et al., 2012; Ünal, de Waard, Epstude, & Steg, 2013).

A two-way Sex × Lyrics interaction emerged for mean HR wherein women exhibited higher mean HR in the presence of lyrics compared to no lyrics. Contrastingly, for men, no such difference emerged. This might relate to women's greater propensity to use music for emotional regulation, while men use music more for social identity (North & Hargreaves, 2008). Despite the medium-to-large effect evident in this interaction ($\eta_p^2 = 0.13$), we should not discount that the significant difference could also be a chance/erroneous finding given the large number of analyses conducted in the present study (i.e., a "false positive"). It is notable that biological sex did not moderate any other effects of music evident in the present study.

4.5. Simulator data

The behaviours that were monitored in the simulator (i.e., through video analysis and pedal pressure readings) showed that the experimental manipulations had no bearing on how participants dealt with the five simulator triggers. This reinforces the trend evident in the subjective data, as the presence or absence of music did not influence how participants drove. It could be that five triggers are insufficient to provide sensitivity in a continuous perceptual-motor task to identify changes in behaviour (Baker et al., 2020). Moreover, albeit we did include a familiarisation trial, the simulation environment is novel/disconcerting and thus task-related variability can mask any effect of experimental manipulation (Davis-Stober, Dana, & Rouder, 2018).

4.6. Strengths and limitations

The present simulation study included four experimental conditions that facilitated an analysis of the interactive effects of the musical dimensions of lyrics and loudness, as well as two control conditions that were used in parametric analyses as covariates. The design was robust and extends this lineage of work in the human factors literature but, at the same time, the use of covariates does have the propensity to yield Type II errors (see Tabachnick & Fidell, 2019). In terms of hardware, a high-grade driving simulator was used to maintain a level of ecological validity.

Extensive efforts were made to standardise the music conditions and their psycho-acoustic properties (see Supplementary File 1). This was achieved through a four-stage process that entailed surveying participants' radio-listening habits, the gathering by the

research team of a large pool of suitable tracks from radio playlists, and a music-rating exercise using a representative panel of participants that assessed the suitability of each track for urban driving. Moreover, a broad range of measures was taken to gain comprehensive understanding of the subject matter that encompassed psychological, psychophysiological and behavioural outcomes. We cannot be sure that participants had a genuine aesthetic appreciation for the tracks used in the experimental conditions. Nonetheless, we made every effort to ensure that the tracks would be fully representative of what each participant would typically listen to while driving (see Supplementary File 1).

Notably, albeit we found that the no-lyrics/soft condition elicited the lowest scores for affective arousal (see Table 1), we did not pre-establish participants' optimal affective arousal for the task of urban driving. Accordingly, it might be that some drivers can drive safely at moderate-to-high levels of arousal in an urban setting. Nonetheless, given the attentional demands of an urban setting, a high-arousal state is likely to be undesirable for most drivers (Wen et al., 2019). Also, although we asked participants to refrain from consuming alcohol and caffeine on the day of testing, we did not assess their alcohol/caffeine consumption habits and how abstention might have impacted their driving performance. A further limitation is that a singular rater was used and not a team of raters to conduct the risk ratings in relation to a subset of simulator triggers, with ensuing calculations of interrater reliability.

There are some general limitations pertaining to simulator-based research that also warrant brief mention. For example, the driving controls (e.g., steering wheel, pedals) do not feel identical to those of an on-road vehicle. Lengthy or repeated use of a simulator can lead to the phenomenon of simulator sickness (Sawada et al., 2020). We attempted to mitigate this through the administration of ginger confectionary (Lien et al., 2003) and a familiarisation process. In examining safety-relevant behaviours, drivers know that they are being observed, and so simulator-based performance may not be fully representative of real-world driving performance (Zöller, Abendroth, & Bruder, 2019). Moreover, to facilitate some standardisation from an analytical perspective, triggers appeared in a rather predictable manner throughout the simulation, and thus their presentation was not randomised.

4.7. Conclusions and recommendations

The present study attempted to shine a light on the interactive effects of the musical dimensions of lyrics and loudness on a range of psychological, psychophysiological and behavioural outcomes. The overarching finding is that, in accord with Fuller's (2011) notion of task difficulty homeostasis, music with or without lyrics and at soft and loud intensities did not have a bearing or safety-relevant behaviours (i.e., participants made real-time decisions to maintain the perceived difficulty of the driving task within certain boundaries). By way of illustration, the experimental manipulations had no bearing on how participants responded to triggers during simulated urban driving (see Table 2). The most marked effect was in the subjective measure of affective arousal, wherein the non-lyrical music at a low intensity yielded the lowest scores for this variable (i.e., the music had a mildly calming effect; see Fig. 1).

We can deduce from the present findings that listening to familiar music (e.g., classic hits on a radio station) does not result in any significant decrements in simulated driving performance. Music that is less arousing and requires less syntactic processing (i.e., non-lyrical/soft) might engender an optimal affective state – from a safety perspective – in the context of urban driving. Interestingly, assertive drivers, as identified on the MDSI, perceived that their performance was superior with loud, lyrical music but there was no objective evidence (e.g., simulator data) in support of this notion. Accordingly, drivers should be wary of using loud, lyrical music in urban environments, given the propensity for such music to elevate levels of activation and even aggression (Brodsky & Kizner, 2012). The upshot is an overestimation of capability or underestimation of task demand – particularly so in the case of young, inexperienced drivers (De Craen, Twisk, Hagenzieker, Elffers, & Brookhuis, 2011; Fuller, 2011).

Future research might seek to use longer simulations than the present ~8-min course and perhaps separate the administration of trials over several days. The limitation, however, in such an approach is that re-familiarisation is needed each time a participant visits a simulator facility, as they are likely to have been driving their own vehicle beforehand. It would also be valuable to test higher intensities of music than 75 dBA given the anecdotal reporting of very loud music at the site of road traffic accidents; particularly those involving young or inexperienced drivers (Brodsky & Slor, 2013). From an ecological validity perspective, a worthwhile contribution would entail examination of the interactive effects of music and car passengers on driver behaviour (i.e., how music influences in-car social dynamics).

CRediT authorship contribution statement

Costas I. Karageorghis: Conceptualisation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Supervision, Writing – original draft, Writing – review & editing. **Garry Kuan:** Formal analysis, Investigation, Methodology, Project administration, Resources, Writing – review & editing. **William Payre:** Investigation, Methodology, Resources, Software, Visualisation, Writing – original draft, Writing – review & editing. **Elias Mouchlianitis:** Formal analysis, Software, Visualisation, Writing – review & editing. **Luke W. Howard:** Data curation, Formal analysis, Resources, Writing – review & editing. **Nick Reed:** Conceptualisation, Funding acquisition, Methodology, Writing – review & editing. **Andrew M. Parkes:** Conceptualisation, Funding acquisition, Methodology, Project administration, Resources, Supervision, Writing – review & editing.

Acknowledgement

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trf.2021.05.022.

References

Acharya, U. R., Joseph, K. P., Kannathal, N., Lim, C. M., & Suri, J. S. (2006). Heart rate variability: A review. *Medical & Biological Engineering & Computing, 44*(12), 1031–1051. https://doi.org/10.1007/s11517-006-0119-0.

Amado, S., Arikan, E., Kaca, G., Koyunca, M., & Turkan, B. N. (2014). How accurately do drivers evaluate their own driving behaviour? An on-road observational study. *Accident Analysis & Prevention, 63*, 65–73. https://doi.org/10.1016/j.aap.2013.10.022.

Arnett, A. (1991). Heavy metal music and reckless behavior among adolescents. *Journal of Youth and Adolescence, 20*(6), 573–592. https://doi.org/10.1007/BF01537363.

Baker, D. H., Vilidaite, G., Lygo, F. A., Smith, A. K., Flack, T. R., Gouws, A. D., & Andrews, T. J. (2020). Power contours: Optimising sample size and precision in experimental psychology and human neuroscience. *Psychological Methods*. Advance online publication. http://doi.org/10.1037/met0000337.

Berntson, G. G., Thomas Bigger, J., Eckberg, D. L., Grossman, P., Kaufmann, P. G., Malik, M., et al. (1997). Heart rate variability: Origins methods, and interpretive caveats. *Psychophysiology, 34*, 623–648. https://doi.org/10.1111/j.1469-8986.1997.tb02140.x.

Brodsky, W. (2002). The effect of music tempo on simulated driving performance and vehicular control. *Transportation Research, Part F: Traffic Psychology and Behaviour, 4*(4), 219–241. https://doi.org/10.1016/S1369-8478(01)00025-0.

Brodsky, W. (2015). *Human factors in road and rail transport. Driving with music: Cognitive-behavioural implications.* Ashgate.

Brodsky, W. (2018). A performance analysis of in-car music engagement as an indication of driver distraction and risk. *Transportation Research Part F: Traffic Psychology and Behaviour, 55*, 210–218. https://doi.org/10.1016/j.trf.2018.02.034.

Brodsky, W., & Kizner, M. (2012). Exploring an alternative in-car music background designed for driver safety. *Transportation Research, Part F: Traffic Psychology and Behaviour, 15*(2), 162–173. https://doi.org/10.1016/j.trf.2011.12.001.

Brodsky, W., & Slor, Z. (2013). Background music as a risk factor for distraction among young-novice drivers. *Accident Analysis & Prevention, 59*, 382–393. https://doi.org/10.1016/j.aap.2013.06.022.

Chien, P. J., & Chan, S. H. (2015). Old songs can be as fresh as new: An ERP study on lyrics processing. *Journal of Neurolinguistics, 35*, 55–67. https://doi.org/10.1016/j.jneuroling.2015.02.002.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* Academic press.

Consiglio, W., Driscoll, P., Witte, M., & Berg, W. P. (2006). Effect of cellular telephone con-versations and other potential interface on reaction time in a braking response. *Accident Analysis & Prevention, 15*, 495–500. https://doi.org/10.1016/S0001-4575(02)00027-1.

Davis-Stober, C. P., Dana, J., & Rouder, J. N. (2018). Estimation accuracy in the psychological sciences. *PLoS ONE, 13*(11). https://doi.org/10.1371/journal.pone.0207239. Article 0207239.

De Craen, S., Twisk, D. A., Hagenzieker, M. P., Elffers, H., & Brookhuis, K. A. (2011). Do young novice drivers overestimate their driving skills more than experienced drivers? Different methods lead to different conclusions. *Accident Analysis & Prevention, 43*(5), 1660–1665. https://doi.org/10.1016/j.aap.2011.03.024.

Department for Transport (2020). Reported road casualties in Great Britain: provisional results 2019. Author. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/904698/rrcgb-provisional-results-2019.pdf.

FakhrHosseini, S. M., & Jeon, M. (2019). How do angry drivers respond to emotional music? A comprehensive perspective on assessing emotion. *Journal on Multimodal User Interfaces, 13*, 137–150. https://digitalcommons.mtu.edu/etdr/725.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160. https://doi.org/10.3758/BRM.41.4.1149.

Fuller, R. (2011). Driver control theory: From task difficulty homeostasis to risk allostasis. In B. Porter (Ed.), *Handbook of the 4th International Conference on Traffic and Transport Psychology* (pp. 13–26). Academic Press.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock, & N. Meshkati (Eds.), *Human mental workload.* North Holland Press.

Ho, C., Reed, N., & Spence, C. (2007). Multisensory in-car warning signals for collision avoidance. *Human Factors: The Journal of the Human Factors and Ergonomic Society, 49*(6), 1107–1114. https://doi.org/10.1518/001872007X249965.

Hughes, G. M., Rudin-Brown, C. M., & Young, K. L. (2013). A simulator study of the effects of singing on driving performance. *Accident Analysis & Prevention, 50*, 787–792. https://doi.org/10.1016/j.aap.2012.07.001.

Hutchinson, J. C., Jones, L., Vitti, S. N., Moore, A., Dalton, P. C., & O'Neil, B. J. (2018). The influence of self-selected music on affect-regulated exercise intensity and remembered pleasure during treadmill running. *Sport, Exercise, and Performance Psychology, 7*(1), 80–92. https://doi.org/10.1037/spy0000115.

Ignacio, D. A., Gerkens, D. R., Aguinaldo, E. R., Arch, D., & Barajas, R. (2019). Music-like stimuli affect the emotional processing of words and facial expressions. *Psychology of Music, 49*(3), 479–497. https://doi.org/10.1177/0305735619875800.

Juslin, P. N. (2013). From everyday emotions to aesthetic emotions: Towards a unifying theory of musical emotions. *Physics of Life Reviews, 10*(3), 235–266. https://doi.org/10.1016/j.plrev.2013.05.008.

Khattak, A. J., Ahmad, N., Wali, B., & Dumbaugh, E. (2021). A taxonomy of driving errors and violations: Evidence from the naturalistic driving study. *Accident Analysis & Prevention, 151*, Article 105873. https://doi.org/10.1016/j.aap.2020.105873.

Lien, H. C., Sun, W. M., Chen, Y. H., Kim, H., Hasler, W., & Owyang, C. (2003). Effects of ginger on motion sickness and gastric slow-wave dysrhythmias induced by circular vection. *American Journal of Physiology: Gastrointestinal and Liver Physiology, 284*(3), 481–489. https://doi.org/10.1152/ajpgi.00164.2002.

Mäntylä, T. (2013). Gender differences in multitasking reflect spatial ability. *Psychological Science, 24*(4), 514–520. https://doi.org/10.1177/0956797612459660.

McMahon, K., Sparrow, B., Chatman, L., & Riddle, T. (2011). Driven to distraction: The impact of distracter type on unconscious decision making. *Social Cognition, 29*(6), 683–698. https://doi.org/10.1521/soco.2011.29.6.683.

Merritt, P., Hirshman, E., Wharton, W., Stangl, B., Devlin, J., & Lenz, A. (2007). Evidence for gender differences in visual selective attention. *Personality and Individual Differences, 43*(3), 597–609. https://doi.org/10.1016/j.paid.2007.01.016.

Millet, B., Ahn, S., & Chattah, J. (2019). The impact of music on vehicular performance: A meta-analysis. *Transportation Research Part F: Traffic Psychology and Behaviour, 60*, 743–760. https://doi.org/10.1016/j.trf.2018.10.007.

Noguchi, K., Gel, Y. R., Brunner, E., & Konietschke, F. (2012). nparLD: An R software package for the nonparametric analysis of longitudinal data in factorial experiments. *Journal of Statistical Software, 50*(12). https://doi.org/10.18637/jss.v050.i12.

North, A. C., & Hargreaves, D. J. (2008). *The social and applied psychology of music.* Oxford University Press.

Parnell, K. J., Rand, J., & Plant, K. L. (2020). A diary study of distracted driving behaviours. *Transportation Research Part F: Traffic Psychology and Behaviour, 74*, 1–14. https://doi.org/10.1016/j.trf.2020.08.003.

Pêcher, C., Lemercier, C., & Cellier, J. (2009). Emotions drive attention: Effects on driver's [sic.] behaviour. *Safety Science, 47*(9), 1254–1259. https://doi.org/10.1016/j.ssci.2009.03.011.

Reimer, B. (2009). Impact of cognitive task complexity on drivers' visual tunneling. *Transportation Research Record, 2138*(1), 13–19. https://doi.org/10.3141/2138-03.

Royal, D. (2003). National survey of distracted and drowsy driving attitudes and behaviors: 2002. *National Highway Traffic Safety Administration*. http://www.nhtsa. gov/people/injury/drowsy_driving1/survey-distractive03/Technical_Page1.htm.

Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178. https://doi.org/10.1037/h0077714.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect Grid: A single-item scale of pleasure and arousal. *Journal of Personality and Social Psychology, 57*(3), 493–502. https://doi.org/10.1037/0022-3514.57.3.493.

Sawada, Y., Itaguchi, Y., Hayashi, M., Aigo, K., Miyagi, T., Miki, M., et al. (2020). Effects of synchronised engine sound and vibration presentation on visually induced motion sickness. *Scientific Reports, 10*(1), 1–10. https://doi.org/10.1038/s41598-020-64302-y.

Shih, Y. N., Huang, R. H., & Chiang, H. Y. (2012). Background music: Effects on attention performance. *Work, 42*(4), 573–578. https://doi.org/10.3233/WOR-2012-1410.

Sloboda, J. A., O'Neill, S. A., & Ivaldi, A. (2001). Functions of music in everyday life: An exploratory study using the experience sampling method. *Musicae Scientiae, 5*(1), 9–32. https://doi.org/10.1177/102986490100500102.

Summala, H. (1996). Accident risk and driver behaviour. *Safety Science, 22*(1–3), 103–117. https://doi.org/10.1016/0925-7535(96)00009-4.

Szameitat, A. J., Hamaida, Y., Tulley, R. S., Saylik, R., & Otermans, P. C. J. (2015). "Women are better than men"–Public beliefs on gender differences and other aspects in multitasking. *PLoS ONE, 10*(10), Article e0140371. https://doi.org/10.1371/journal.pone.0140371.

Tabachnick, B. G., & Fidell, L. S. (2019). *Using multivariate statistics* (7th ed.). Pearson Education.

Taubman-Ben-Ari, O., Mikulincer, M., & Gillath, O. (2004). The multidimensional driving style inventory—scale construct and validation. *Accident Analysis & Prevention, 36*(3), 323–332. https://doi.org/10.1016/S0001-4575(03)00010-1.

Taylor, M. C., Baruya, A., & Kennedy, J. V. (2002). The relationship between speed and accidents on rural single-carriageway roads. Prepared for Road Safety Division, Department for Transport, Local Government and the Regions. http://www.safespeed.org.uk/TRL511.pdf.

Ünal, A. B., de Waard, D., Epstude, K., & Steg, L. (2013). Driving with music: Effects on arousal and performance. *Transportation Research Part F: Traffic Psychology and Behaviour, 21*, 52–65. https://doi.org/10.1016/j.trf.2013.09.004.

Ünal, A. B., Platteel, S., Steg, L., & Epstude, K. (2013). Blocking-out auditory distractors while driving: A cognitive strategy to reduce task-demands on the road. *Accident Analysis & Prevention, 50*, 934–942. https://doi.org/10.1016/j.aap.2012.07.020.

Ünal, A. B., Steg, L., & Epstude, K. (2012). The influence of music on mental effort and driving performance. *Accident Analysis & Prevention, 48*, 271–278. https://doi.org/10.1016/j.aap.2012.01.022.

van der Zwaag, M. D., Dijkesterhuis, C., de Waard, D., Mulder, B. L. J. M., Westerink, J. H. D. M., & Brookhuis, K. A. (2012). The influence of music on mood and performance while driving. *Ergonomics, 55*(1), 12–22. https://doi.org/10.1080/00140139.2011.638403.

Vuilleumier, P., Armony, J., & Dolan, R. (2003). Reciprocal links between emotion and attention. In R. S. J. Frackowiak (Ed.), *Human brain function* (2nd ed., pp. 419–444). Academic Press.

Wen, H., Sze, N. N., Zeng, Q., & Hu, S. (2019). Effect of music listening on physiological condition, mental workload, and driving performance with consideration of driver temperament. *International Journal of Environmental Research and Public Health, 16*, Article 2766. https://doi:10.3390/ijerph16152766.

Wickens, C. D. (2008). Multiple resources and mental workload. *Human Factors: The Journal of the Human Factors and Ergonomic Society, 50*(3), 449–455. https://doi.org/10.1518/001872008X288394.

Wiesenthal, D. L., Hennessy, D. A., & Totten, B. (2003). The influence of music on mild driver aggression. *Transportation Research Part F: Traffic Psychology and Behaviour, 6*(2), 125–134. https://doi.org/10.1016/S1369-8478(03)00020-2.

World Health Organization: WHO. (2020). Road traffic injuries. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries.

Zijlstra, F. R. H. (1993). Efficiency in work behaviour: a design approach for modern tools (Publication No. 90-6275-918-1) [Doctoral dissertation, Delft University]. https://repository.tudelft.nl/islandora/object/uuid%3Ad97a028b-c3dc-4930-b2ab-a7877993a17f.

Zöller, I., Abendroth, B., & Bruder, R. (2019). Driver behaviour validity in driving simulators – Analysis of the moment of initiation of braking at urban intersections. *Transportation Research Part F: Traffic Psychology and Behaviour, 61*, 120–130. https://doi.org/10.1016/j.trf.2017.09.008.