Transportation Research Part F 60 (2019) 743–760

Contents lists available at ScienceDirect

Transportation Research Part F

journal homepage: www.elsevier.com/locate/trf

The impact of music on vehicular performance: A meta-analysis

Barbara Millet [a,*], Soyeon Ahn [b], Juan Chattah [c]

[a] School of Communication, University of Miami, Coral Gables, FL, USA
[b] School of Education and Human Development, University of Miami, Coral Gables, FL, USA
[c] Frost School of Music, University of Miami, Coral Gables, FL, USA

ARTICLE INFO

Article history:
Received 25 July 2018
Received in revised form 2 October 2018
Accepted 11 October 2018
Available online 19 December 2018

Keywords:
Music
Driving
Meta-analysis
Vehicular performance

ABSTRACT

Various studies offer insightful perspectives on the potential impact of music-listening on driving performance. These studies, however, present conflicting views on the effect of music as either hindering or enhancing driving performance and advance inconclusive claims regarding how and to what extent specific music parameters affect vehicular performance. In this study, therefore, we conducted a systematic review and meta-analysis of relevant studies. First, we identified experimental studies that measured the effects of music-listening on driving performance through database searches using multiple variants including "car", "driv*", "perf*", and "music*"; of the 118 publications reviewed, 12 met the inclusion criteria for the current meta-analysis. Second, we coded independent variables—i.e., tempo, volume, instrumentation, familiarity, musical style, the music's source, and whether music was selected by the researchers or the drivers—and dependent variables—i.e., vehicular longitudinal and lateral control, driver reaction time, traffic signal violations, collisions, and driving scores. Third, we ran mixed-effects and random-effects models to identify both general tendencies and more particular trends related to the effect of music-listening on driving performance—driving performance is here understood as the combination of vehicle manipulation and road navigation. Consistent with anecdotal evidence, the results of this meta-analysis show that music-listening has a statistically significant detrimental effect on driving performance, specifically for collisions and longitudinal control. In contrast with anecdotal evidence, however, the results of this meta-analysis show a detrimental effect associated with music-listening at soft volumes and no significant difference in driving performance associated with tempo. The study's findings contributed to the development of a process model, and the concluding discussion offers suggestions for future empirical investigations related to music and driving.

1. Introduction

Because music-listening while driving is a widespread behavior (Rentfrow & Gosling, 2003; Sloboda, 1999), traffic researchers, vehicle manufacturers, and insurance companies are beginning to draw their attention to the benefits and risks associated with it. This has propelled an upsurge of studies that provide insightful perspectives on the potential impact of music-listening on driving performance—driving performance is here understood as the combination of vehicle manipulation and road navigation. Existing research, however, offers conflicting views on whether music-listening enhances or hin-

* Corresponding author.
E-mail addresses: bmillet@miami.edu (B. Millet), s.ahn@miami.edu (S. Ahn), j.chattah@miami.edu (J. Chattah).

https://doi.org/10.1016/j.trf.2018.10.007
1369-8478/

ders driving performance, and conclusive empirical studies that attend to particular musical parameters and their effect on driving performance are scarce.

Studies suggesting a detrimental effect of music-listening point to the music masking critical sounds (e.g., sirens, horns, or engine sounds) that help drivers navigate through traffic (Bellinger, Budde, Machida, Richardson, & Berg, 2009; Dibben & Williamson, 2007; Ho & Spence, 2017), and to activities related to music-listening (e.g., changing tracks, searching stations or through MP3 files, changing volume) that may lead to distraction (Harvey & Carden, 2009; Lee, Roberts, Hoffman, & Angell, 2012; Lerner, Baldwin, Higgins, Lee, & Schooler, 2015; Young, Mitsopoulos-Rubens, Rudin-Brown, & Lenné, 2012). Attending more closely to musical parameters, Brodsky (2001) found a positive correlation between music tempo and both simulated driving speed and perceived speed estimates. Additionally, behavioral research not measuring driving performance showed that not only fast but also loud music (>85 dB) significantly correlate with physiological arousal (Dalton, Behm, & Kibele, 2007; Husain, Thompson, & Schellenberg, 2002), which is consistent with the notion that fast and loud music may cause drivers to increase acceleration, cruising speed, and in turn traffic violations (Konz & Mcdougal, 1968). Alternatively, several studies note that musical features (particularly volume and tempo) help improve performance, particularly during car-following tasks (Cummings, Koepsell, Moffat, & Rivara, 2001; Oron-Gilad, Ronen, & Shinar, 2008; Ünal, de Waard, Epstude, & Steg, et al., 2013; Ünal, Platteel, Steg, & Epstude, et al., 2013). Other studies acknowledge that fast and loud music may induce risky driving practices, but reveal that drivers adopt cognitive or behavioral compensatory strategies, such as slowing down or sitting upright (Hockey, 1997; Hughes, Rudin-Brown, & Young, 2013; Ünal, de Waard, et al, 2013; Ünal, Platteel, et al., 2013).

Given the mixed and inconclusive claims in the literature, this meta-analysis aims to understand the overall effect of music-listening on driving performance, and then further explore the extent to which music's various parameters moderate (i.e., enhance or interfere with) the vital processes of vehicle control and road navigation, such as speed maintenance, lane-keeping, reaction time, and gap negotiation. In an effort to synthesize research results, we focused our analysis on two primary questions: (1) Does listening to music influence driving performance? and, (2) Does performance differ as a consequence of specific music parameters?

2. Method

2.1. Search processes and study selection

As recommended by Cooper (2009), multiple sources were used to identify an exclusive set of studies for the current meta-analysis. First, all three authors independently conducted systematic searches of literature via bibliographic databases including PsycINFO, ERIC, and PubMed; systematic searches took place from February to March of 2018. The key terms used in these searches included different words for motor vehicle (i.e., *car*, *vehicle*, *automobile*), different words for operating motor vehicles (i.e., *driving*, *manipulating*, *operating*), and different words related to music (i.e., *music*, *rock*, *song*). Our search strategies also included truncation (e.g., *driv**, *auto**, *manipulat**, *music**, *perf**, etc.) and term combination using Boolean connectors. Second, we searched the online library catalogue at the authors' home institution (University of Miami) and Google Scholar using the aforementioned search terms—this is known as forward searches—and manually searched the reference lists of published studies to identify additional studies for inclusion—this is known as backward searches. The independent searches resulted in a cumulative total of 118 publications.

All three authors independently screened the titles and abstracts of all 118 publications, evaluated references for inclusion, and removed duplicates. This initial screening also accounted for the scholarly validity and availability of studies, and ensured studies shared a similar conceptual hypothesis. The inclusion/exclusion criteria used during this initial screening stage were: (1) the study must be written in English; (2) the study must be empirical; and (3) the study must examine the effect of music-listening while driving. Discrepancies in the intra-author results after this initial stage were discussed and re-evaluated collectively. This initial screening identified a total of 43 publications.

As a next step, each author independently reviewed the identified 43 publications. The specific inclusion/exclusion criteria at this stage were: (1) the study must include at least one pair of groups that differ in music stimuli while driving; (2) the study must evaluate the effect of music stimuli on at least one dependent variable related to direct measures of vehicular performance; (3) there must have been sufficient statistical information for the calculation of the primary effect size and its associated variance; and (4) the data must have been used only once in a study to avoid dependence among effect sizes. If the same subjects were used in multiple publications, independent effect size from the study with more participants was extracted. As a result, a total of 10 studies met inclusion criteria and were selected for inclusion in the current meta-analysis. A supplemental literature search conducted in June 2018 resulted in the inclusion of 2 additional studies, leading to a total of 12 independent studies. Fig. 1 summarizes the literature search process.

Several studies were filtered by the inclusion/exclusion criteria—e.g., studies that used driving-related tasks (e.g., Beh & Hirst, 1999; Consiglio, Driscoll, Witte, & Berg, 2003), that lacked experimental control conditions (e.g., Dalton et al., 2007; Santoso, Maulina, Adystia, & Oei, 2013), that lacked direct measures of driving performance (e.g., Brodsky & Kizner, 2012), or that did not provide the statistical information necessary for calculating effect size (e.g., Wiesenthal, Hennessy, & Totten, 2003; van der Zwaag et al., 2012). These studies, while remaining outside of the meta-analysis, provided keen

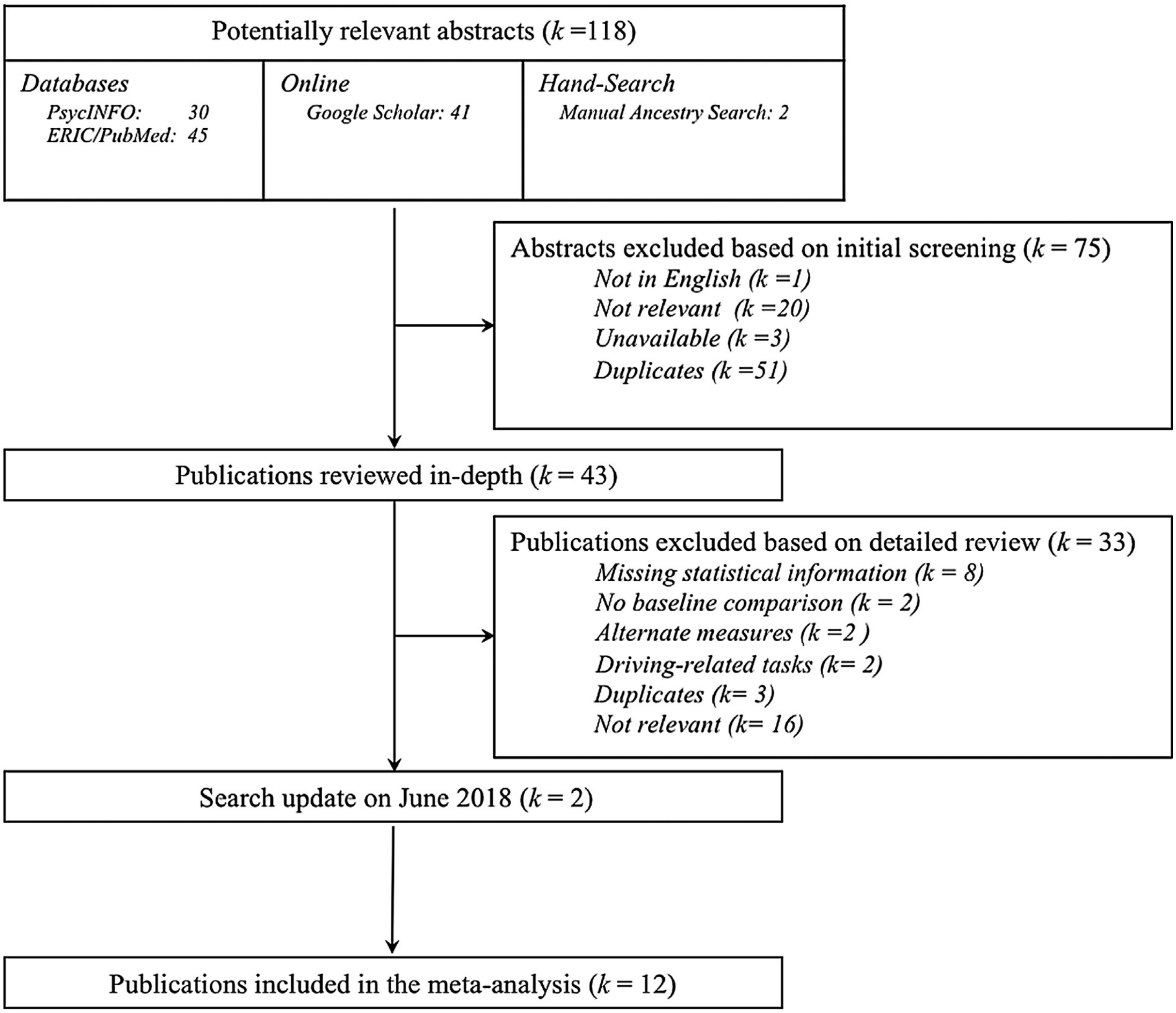

Fig. 1. Search flow and selection of publications included in the meta-analysis.

insights on the impact of music on driving performance and hence were included in our systematic review and informed the arguments put forth in the discussion section.

2.2. Coding of the included studies

A systematic coding scheme was developed to identify salient characteristics among all studies that would contribute to the general conceptual hypothesis. These characteristics included: (1) study design, (2) participant characteristics, (3) experimental conditions, (4) music characteristics, and (5) outcomes/measures. Each of these characteristics was independently coded by the authors, and any coding discrepancies were resolved upon discussions and/or consultations among the authors. Table 1 provides study authors, year of publication, experimental setting, participant characteristics, and the independent and dependent variables. The coded information from each study will provide the general context, whereas the current meta-analysis can be generalizable. Table 2 attends to the music component, and includes the music experimental conditions, participants familiarity with the music, the source of the music, as well as the coded independent variables for music directly related to the posited research questions.

Study design characteristics. A number of study design characteristics were coded, including (1) where the study was conducted (United States vs. outside of United States), (2) publication type (published vs. unpublished), (3) sampling method (probability based, non-probability based), (4) assignment method (probability based, non-probability based, and not applicable because the same participants were measured after being exposed to different experimental conditions), (5) study

Table 1
Study authors, year of publication, experimental setting, participant characteristics, and independent and dependent variables. Studies included in the meta-analysis are listed first, additional studies included in the systematic review are listed below the double line.

Study	Pub. year	Setting	N	Ages and gender (SD)	Music Conditions (# of conditions)	Measures (not included in analysis)
Brodsky Experiment 2	2001	Simulator	28	Ages: MN = 25 (4.58) M 10, F 18	(4) Base, Tempo (Slow, Med., & Fast)	Sig., Lat., Coll., Long. (Phys.)
Brodsky and Slor	2013	Roadways	85	Ages: MN = 17.6 (0.41) M 49, F 36	(3) Base, DSel, RSel	Sig., Lat., Long. (Score, Sub.)
Cassidy and MacDonald*	2009	Simulator	125	Ages: MN = 21.2 (2.1) M 60, F 65	(5) Base, CarS., CarS + DSel, CarS + RSel (Low & High Arousal)	Coll., Long. (Sub.)
Cassidy and MacDonald	2010	Simulator	70	Ages: MN = 20.5 (2.1) M 38, F 32	(7) Base, CarS., CarS + DSel, CarS + RSel (Low & High Arousal; Slow & Fast Tempo)	Coll., Long. (Sub.)
Febriandirza, Chaozhong, Zhong, Hu, and Zhang	2017	Simulator	98	Ages: MN = 24.47 (2.03) M 49, F 49	(4) Base, NatS, Hard Rock, Classical	Lat., Long. (Phys., Sub.)
Jimison**	2014	Simulator	165	Ages: MD 21 M 47, F 118	(4) Base, Workload (Low & High), Music: DSel, RSel-Ufam, or RSel-Fam	RT, Sig., Lat., Coll., Long.
Konz and McDougal	1968	Roadways	24	Ages: 18–23 M 24, F 0	(3) Base, Slow & Tijuana Brass	Lat., Long.
Mizoguchi and Tsugawa Experiment 1	2012	Simulator	5	Ages: early 20s M 5, F 0	Base, Volume (Low, Med., & High)	Coll., Long.
Navarro et al. Experiment 1	2018	Simulator	24	Ages: MN = 22.7 (3.5) M 13, F 11	(4) Base, DSel, DSel + Tempo, DSel -Tempo	RT, Long. (Phys., Sub.)
Oron-Gilad et al.	2008	Simulator	12	Ages: 32–53 M 12, F 0	(2) Base, DSel	Lat., Long. (Phys., Sub.)
Ünal et al.*	2012	Simulator	69	Ages: MN = 21.04 (1.96) M 23, F 46	(2) Base, DSel	RT, Long. (Sub.)
Ünal et al.**	2013	Simulator	47	Ages: MN = 20.7 (1.34) M 26, F 21	(2) Base, DSel (Volume: Med. or High)	Lat. (Phys., Sub.)
Beh and Hirst*	1999	Laboratory	45	Ages: 18–24 M 21, F 24	Base, Volume (Low, High)	RT
Brodsky and Kizner Study A	2012	Roadways	22	Ages: MN = 26.3 (1.83) M 14, F 8	DSel, RSel	Sub.
Study B			31	Ages: MN = 25.5 (2.07) M 11, F 20	RSel	
Consigilo et al.	2003	Lab	22	Ages: MN = 21 (2.1) M 11, F 11	Base, RSel, Concurrent Tasks	RT
Dalton et al.	2007	Simulator	12	M 6, F 6	Music Type × Volume: Classical, Hard Rock, IndN, High, Low	Coll., Long., Phys.
Fairclough, van der Zwaag, Spiridon, and Westerink*	2014	Simulator	100	Ages: MN = 21.2 (4.7) M 49, F 51	Base, Music: Arousal (Low or High] & Valence (Pos. or Neg.)	Phys., Sub.
Groene and Barrett	2012	Simulator	23	Ages: MN = 16.3 (0.47) M 5, F 5	Base, Music & Driving Suggestion (Y/N)	Score, Sub.
Hughes et al.	2013	Simulator	21	Ages: MN = 35.05 (13.75) M 1, F 20	Base, Music, & Singing	RT, Lat., Long., Sub.
Mizoguchi and Tsugawa Experiment 2 Experiment 3	2012	Simulator	5	Ages: early 20s M 5, F 0	Base, RSel (Tempo: Slow, Med., Fast) Base, Fav. & Slow Tempo, NFav. & Fast Tempo	Coll., Long.

Table 1 (*continued*)

Study	Pub. year	Setting	N	Ages and gender (SD)	Music Conditions (# of conditions)	Measures (not included in analysis)
Navarro et al. Experiment 2	2018	Simulator	32	Ages: 20–43 M 12, F 20	Base, RSel (Tempo: Slow, Med., Fast)	Long. N = 31
North and Hargreaves	1995	Lab	96	Ages:18.99 (1.57) M 48, F 48	Concurrent Task (Y/N) × RSel (Low & High Arousal)	Long., Sub.
Santoso et al.	2013	Roadways	60	Ages: MN = 30.15 (3.38) M 60, F 0	Dangdut, Rock	Long.
Turner et al.**	1996	Lab	90	Ages: 18–49 M 45, F 45	Base, Volume (Low Med., & High)	RT, MT, Tot.
Ünal et al.** Study 2	2013a	Simulator	46	Ages: MN = 21.83 (2.44) M 21, F 25	Traffic (Low & High) × Music (Base, Radio)	Other
van der Zwaag et al.	2012	Simulator	19	Ages: MN = 27.5 (5.2) M 13, F 6	Music Type × Lane Width Base, Pos., Neg. Wide, Narrow	Lat., Long., Phys., Sub.
Wang et al.**	2015	Simulator	165	Ages: MD 21 M 47, F 118	Base, Workload (Low & High), Music: DSel, RSel-Ufam, or RSel-Fam	RT, Sig., Lat., Coll., Long.
Wiesenthal et al.*	2000	Roadways	40	Ages: MN = 26.2 M 20, F 20	Base, DSel	Sub.
Wiesenthal et al.*	2003	Roadways	40	Ages: MN = 25.60 M 20, F 20	Base, DSel	Sub.

Notes. Pub. is publication, N is number of participants, MN is mean, MD is median, M is male, F is female, Y is yes, N is no, DSel is driver-selected music, RSel is researcher-selected music, CarS is car sounds, NatS is natural sounds, IndN is industrial noise, ± is an increase/decrease to specific music component. Base is a baseline or control condition (e.g. no music or silence), Pos. is positive, Neg. is negative, Fav. is favorite, NFav. is non-favorite, Med. is medium, RT is reaction time (e.g., lead vehicle braking), Lat. is lateral control (e.g., standard deviation of lane position [SDLP], lane excursions, etc.), Long. is longitudinal control (e.g., speed, speed variance, time or distance headway), Sig. is traffic signal violations (e.g., disobeying traffic lights, signs, or signals), Coll. is number of collisions, Score includes measures of driving deficiency scores based on mechanical, behavioral, or combination of these data types, Phys. is physiological response, Sub. is subjective measures, MT is movement time, Tot. is total time, and Other is other measure (e.g. % of recalled content). Unless otherwise. All other studies used within-subjects experimental designs.

* Between-groups design for music conditions.

** Mixed-groups design across conditions.

design (posttest only control group design [between-subject design], a single group pretest-posttest design [within-subject design], and pretest-posttest control group design [mixed-design]), and (6) publication year.

Participant characteristics. The coded participants characteristics included: (1) number of male/female participants, (2) participants' mean age (with standard deviation or range), (3) years of driving experience, and (4) whether participants had prior car accidents/collisions.

Experimental conditions. The experimental setting was categorized into roadways or simulation. The roadways category included studies that were conducted on public roads, which frequently used in-vehicle data recorders. The simulation category included studies using driving simulators ranging from video game consoles (without standard driving controls) to fixed based driving simulators with up to 240° view of the traffic environment and force feedback steering and pedals; the most commonly used test tasks in the simulated environment was monotonous driving and car-following.

Music characteristics. Coding the music characteristics involved discerning *extrinsic* musical variables (e.g., familiarity) versus *intrinsic* musical variables (e.g., tempo). The coded extrinsic music variables included: (1) music experimental conditions (researcher-selected music and driver-selected music), (2) driver's familiarity with the music selections (familiar, probably familiar, and unfamiliar), (3) source of music selections (pre-existing music and original music). The coded intrinsic musical variables included: (1) tempo (slow < 80 bpm, medium 80–120 bpm, fast > 120 bpm), (2) volume (low ≤ 60 dB*, medium 61–79 dB*, high ≥ 80 dB*), (3) volume-weighting (dB, dBA, dBC), (4) instrumentation (vocals, acoustic, electronic), and (5) musical style (classic rock/pop, classical, country, easy listening, electronica, film/TV soundtrack, hard rock/heavy metal, hip-hop/rap, jazz/blues, R&B/soul/funk, reggae, world/folk, and other).

Coding intrinsic musical variables involved four steps: (1) identifying music parameters considered across all studies—tempo, volume (and volume-weighting), instrumentation, and style/genre emerged as the common parameters considered across most (if not all) studies; (2) identifying subcategories represented within studies toward constructing generalized category scales; (3) coding all data provided in the studies; (4) combining subcategories to obtain a more representative cross-section. Steps #2 and #4 in this process entail making decisions that may affect the outcomes of the meta-analysis; further details about the rationale behind those decisions are provided below.

Tempo. Although there is converging evidence regarding the limits of tempo perception (London, 2012), only few studies venture to suggest exact bpm rates to define a category scale. Brodsky (2001) proposed the most specific category scale for

Table 2
Study authors, year of publication, music experimental conditions, familiarity, source, and coded independent music variables. Studies included in meta-analysis are listed first, additional studies included in the systematic review are listed below the double line.

Study	Pub. year	Music experimental conditions	Fam.	Source	Coded independent (music) variables			
					Tempo	Volume (weighting)	Instr.	Styles
Brodsky	2001	RSel	P	PE	S, M, F	H (dBA)	A, E	Rk, Co, Jz, RB
Brodsky and Slor	2013	RSel	N	O	S, M, F	H (dBC)	A	Ea, Jz, O
		DSel	Y	PE	NG	H (dBC)	V, A, E	Rk, Ea, El, TV, HR, Rp, RB, Re, Wo, O
Cassidy and MacDonald	2009	DSel	Y	PE	S, M, F	L (dBA)	V, A, E	Rk
Cassidy and MacDonald	2010	RSel	N	O	S, F	L (dBA)	V	NG
		DSel	Y	PE	S, M, F	L (dBA)	V, A, E	Rk, Cl, HR, Jz
Febriandirza et al.	2017	RSel	NG	PE	S, M, F	H (dBA)	V, A, E	Cl, HR
Jimison	2014	RSel	Y, N	PE	S, M, F	M (dB)	V, A, E	Rk
		DSel	Y	PE	NG	M (dB)	NG	NG
Konz and McDougal	1968	RSel	P	PE	S, M, F	NG	A	Wo, O
Mizoguchi and Tsugawa Experiment 1	2012	RSel	Y	PE	M	L, M (dB)	V, A, E	Rk
Navarro et al. Experiment 1	2018	DSel	Y	PE	S, M, F	M (dB)	NG	NG
Oron-Gilad et al.	2008	DSel	Y	PE	NG	NG	NG	NG
Ünal et al.	2012	DSel	Y	PE	NG	H (dBA)	NG	NG
Ünal et al.	2013b	RSel	P	PE	NG	M (dB)	V, A, E	Rk, El, Jz, RB, O
Beh and Hirst	1999	RSel	NG	PE	NG	M (dBA)	NG	HR
Brodsky and Kizner	2012	RSel	N	O	S, M, F	H (dB)	A	Ea, Jz, O
		DSel	Y	PE	NG	H (dB)	NG	Rk, Cl, TV, Rp, Re, Wo
Consigilo et al.	2003	NG	NG	PE	NG	NG	V	Rk
Dalton et al.	2007	RSel	NG	PE	S, M, F	L, H (dBA)	V, A, E	HR, Wo
Fairclough et al.	2014	RSel	P	PE	S, M, F	NG	V, A, E	Rk, Cl, El, TV, HR, O
Groene and Barrett	2012	RSel	N	O	S	NG	V, A, E	O
Hughes et al.	2013	RSel	P	PE	S, M	NG	V, A, E	Rk
Mizoguchi and Tsugawa Experiment 2	2012	RSel	P	PE	S, M, F	M (dB)	V, A, E	Rk
Experiment 3		DSel	Y	PE	S, M, F	M (dB)	V, A, E	Rk
Navarro et al. Experiment 2	2018	RSel	P	PE	S, M, F	M (dB)	A, E	Rk, Co, Jz, RB
North and Hargreaves	1995	RSel	N	O	M, F	L, M, H (dB)	E	El
Santoso et al.	2013	RSel	N	PE	F	NG	V, A, E	Rk, Wo
Turner et al.	1996	RSel	Y	PE	S, F	L, M, H (dBA)	V, A, E	Rk
Ünal et al.	2013a	DSel	Y	PE	NG	M, H, (dB)	NG	NG
van der Zwaag et al.	2012	DSel	Y	PE	NG	NG	V	NG
Wang et al.	2015	RSel	P	PE	NG	M (dB)	NG	NG
		DSel	Y	PE	NG	M (dB)	NG	NG
Wiesenthal et al.	2000	DSel	Y	PE	NG	NG	NG	NG
Wiesenthal et al.	2003	DSel	Y	PE	NG	NG	NG	O

Notes: Pub. is publication, RSel is researcher-selected, DSel is driver-selected, Fam. is familiarity, Y is yes, P is probable, N is no, NG is not-given, PE is pre-existing, O is original, S is slow, M is medium, F is fast, L is low, H is high, Instr. is instrumentation, V is vocal, A is acoustic, E is electronic, Rk is classic rock/pop, Cl is classical, Co is country, Ea is easy listening, El is electronica, TV is TV/film, HR is hard rock/heavy metal, Rp is rap/hip-hop, Jz is jazz/blues, RB is R&B/soul/funk, Re is reggae, Wo is world/folk, O is other.

tempo: slow is 40–70 bpm, medium is 85–110 bpm, and fast is 120–140 bpm. Although arbitrary, Brodsky's category scale broadly follows conventionalized musical descriptions of speed and their associated metronome markings, and the boundaries of Brodsky's scale reside closely within the thresholds offered by Westergaard (1975), which include 80 bpm and 120 bpm, to define the perception of speed in music (see also London, 2012). We therefore adopted Brodsky's category scale and normalized it according to Westgaard thresholds, resulting in the following: slow is <80 bpm, medium is 80–120, and fast is >120 bpm. The perception of the music's speed (in terms of fast, medium, or slow), however, is highly subjective and contingent upon the interaction of multiple musical parameters including meter, harmonic rhythm, inter-onset-intervals, drum-beat patterns, and texture. Furthermore, perceived pulses may fall within several (hyper-)metric levels, and the resulting rate of pulses would therefore fall within different tempo categories; this is a common phenomenon when approaching the outer boundaries of tempo perception, as listeners are likely to switch between metrical levels at very fast or very slow tempi (Madison & Paulin, 2010). Fortunately, within the selected styles of music, tempo was by-and-large stable, without sudden or gradual fluctuations. When the tempo of a music selection was categorized by the researcher or the driver in conflict with our coding strategy, we overrode the researcher's categorization and adjusted accordingly.[1]

Volume. For a category scale for volume we drew on Turner, Fernandez, and Nelson (1996), a study widely cited among the studies included in the systematic review (Bellinger et al., 2009; Brodsky, 2001; Brodsky & Kizner, 2012; Cassidy & MacDonald, 2009; Dalton et al., 2007; Groene & Barrett, 2012; Jimison, 2014; Ünal, Steg, & Epstude, 2012; Ünal, de Waard, et al, 2013; Ünal, Platteel, et al., 2013; Wang, Jimison, Richard, & Chuan, 2015), with the assumption that such category scale was adopted and embedded within their studies' design. Turner et al. (1996) suggests a continuous category scale divided by two thresholds, labeled as "softer (60 dBA)" and "louder (80 dBA)" in relation to a 70 dBA as a medium volume level; this is consistent with Staum and Brotons (2000) contextualization of volume levels within music-listening environments, which includes comfortable listening (70 dBA), symphony concert (76–100 dBA), Walkman headphone (90 dBA), bar and dance-club (100 dBA), and amplified rock concert (>112 dBA). Hence, we formulated the following category scale: low is $\leq$60 dB*, medium is 61–79 dB*, high is $\geq$80 dB*. Analogously to the perception of tempo, the perception of volume (in terms of soft, medium, or loud) is relative. This relative perception of volume is captured, in part, by the different volume-weighting systems.

Volume-weighting. The studies included in the meta-analysis varied in the volume-weighting (dB, dBA, and dBC), and a few studies omitted volume and volume-weighting altogether and offered a loudness rating—loudness is understood as an auditory sensation to sounds expressed on a scale of categorical units extending from soft to loud. Different volume-weightings indicate different loudness functions which are not analogous: dB is the standard unit of sound pressure level; an 'A' filter adjusts the number of decibels to account for the human ear's response to various frequencies; a 'C' filter is similar to the 'A' filter but measures uniformly within the 30 Hz to 10 kHz. Unfortunately, there is no function to map equivalencies between these volume-weightings; coding of volume, therefore, did not account for differences related to volume-weighting.

Genre/style. Genre and style classifications in music are highly problematic, as these are fluid cultural constructs that vary over time and over geographic location. Furthermore, music selections may generally fall within two or more style/genre categories—e.g., film/TV soundtracks may cross multiple categories—and the listeners' perception of the genre and style is contingent upon their conversance with all genres and styles included. We therefore adopted Brodsky and Slor (2013) categorization, which is the most detailed among all reviewed studies, and subsequently purged/merged categories when duplications were observed (e.g., the 'movies/TV shows' and the 'movies/TV shows in Hebrew' categories).

Instrumentation. In contrast to coding other intrinsic music variables, coding instrumentation was straightforward. While many studies provided a detailed list of musical instruments included in the music selections, other studies listed the particular music selections by title of the piece/song and name of the performer(s). In the latter case, we retrieved the music selections and noted the included instrumentation. Given the large number of possible instruments, we identified broader, more salient categories: vocals, acoustic instruments, and electronic instruments.

Outcome characteristics. The outcome variables consisted of direct measures of driving performance, with a coding categorization based on "operational definitions" provided by Caird and Horrey (2011), Caird, Johnston, Willness, Asbridge, and Steel (2014). The categories were (1) vehicular longitudinal control (e.g. speed, speed variance, coherence), (2) vehicular lateral control (e.g. lateral position, standard deviation of lateral position), (3) driver reaction time (e.g. brake response time, time to contact), (4) traffic signal violations (e.g. disobeying traffic lights, signs, or signals), and (5) collisions (when a vehicle collides with another vehicle, pedestrian, animal, or stationary or displaceable objects).

2.3. Effect sizes

The primary effect size for the current meta-analysis is the standardized mean difference (d), which quantifies mean difference on outcome variable between two comparison groups ($g1$: control group and $g2$: experimental group). The standardized mean difference (d) was computed as

$$d = \frac{M_{g1} - M_{g2}}{\sqrt{S_{pooled}/(n_{g1} + n_{g2} - 2)}},$$

[1] Two cases emerged as non-conforming with our coding strategy: Konz and Mcdougal (1968) and Cassidy and MacDonald (2009). Konz and Mcdougal (1968) indicates musical tempo using the 'slow' and 'fast' qualifiers, without providing a quantitative measure corresponding to these categories; we therefore retrieved all music selections included in the study and calculated the tempi, resulting in the following: Konz's 'slow' tempi comprise selections at 67–120 bpm, and Konz's 'fast' tempi comprise selections at 82-170bpm. Given the significant overlap of values between the two categories, our coding of Konz and Mcdougal (1968) reflects their inclusion of music at all three tempi ('slow', 'medium', and 'fast'). Cassidy and MacDonald (2009) only provides two categories ('slow' and 'fast') within the 'participant selected' condition, yet identifies a numerical threshold (98 bpm) located at the center of our 'medium' category. Therefore, our coding of Cassidy and MacDonald (2009), within the 'participant selected' condition, reflects their inclusion of music at all three tempi ('slow', 'medium', and 'fast').

where $S_{pooled} = SD_{g1}^2(n_{g1}-1)+SD_{g2}^2(n_{g2}-1)$, where *SD* is the standard deviation of the outcome variable for each comparison group, *M* is mean of outcomes for comparison groups, and *n* is sample size for each comparison group.

The associated variance was computed as

$$v_d = \frac{n_{g1}+n_{g2}}{n_{g1}n_{g2}} + \frac{d^2}{2(n_{g1}+n_{g2})}$$

For studies using pretest-posttest design, where difference in pretest and posttest was compared, standardized mean difference between pretest and posttest was computed as

$$d_{gain} = \frac{M_{post}-M_{pre}}{\sqrt{(SD_{post}^2+SD_{pre}^2)/2}/\sqrt{2(1-r_{pre,post})}},$$

where *SD* is the standard deviation of the outcome variable measured at each time point, *M* is the mean of outcomes measured at each time point, and *r* is a correlation between pretest and posttest measures. When *r* is not reported, a correlation value of 0.5 was used to compute the effect size.

The associated variance was computed as

$$v_{d_{gain}} = \frac{2(1-r)}{n} + \frac{d_{gain}^2}{2n}$$

As suggested by Hedges and Olkin (1985), the computed standardized mean difference and its associated variance were corrected for the small sample bias by multiplying them by a correcting factor (J), $J = 1–3/(4 * df - 1)$, where *df* is the degrees of freedom. The unbiased effect size (g) and its associated variance (v_g) were computed using $g = d * J$ and $v(g) = v(d) * J^2$, which are used in the current meta-analysis. Herein, the unbiased effect size (g) is denoted as *d*.

2.4. Dependencies of effect sizes

As many studies used multiple dependent variables to evaluate the effect of music-listening while driving, multiple effect sizes were extracted from a single study, leading to the violation of independence assumption for the meta-analysis (Borenstein, Hedges, Higgins, & Rothstein, 2009; Gleser & Olkin, 2009). For example, Groene and Barrett (2012) evaluated the effect of music on one outcome variable (i.e., the number of driving errors made), while Brodsky and Slor (2013) used multiple measures of driving outcomes in their study. Of several available methods that deal with dependency issues (Cooper, 2009), effect sizes were first separated by type of outcome variable. Then, within each category of the outcome variables, dependent effect sizes were averaged (Borenstein et al., 2009) and thus effect sizes were no longer dependent for performing the overall- and moderator-analysis.

2.5. Publication bias

The potential publication bias—the tendency that studies with significant effect in a favorable direction are more likely to be published—is assessed using multiple indicators including (1) Begg and Mazumdar's rank correlation test for funnel plot asymmetry, (2) Egger's regression test of intercept, and (3) funnel plot. When Begg and Mazumdar's rank correlation test for funnel plot asymmetry and Egger's regression test of intercept were found to be statistically significant, sufficient evidence that publication bias might exists.

2.6. Statistical analyses

The metafor package (Viechtbauer, 2010) in R statistical software (R Core Team, 2013) was used to perform meta-analysis. In particular, the current study was based on the meta-analytic methods suggested by Hedges and Olkin (1985), which is to compute an overall effect of individual effect sizes weighted by the inverse of their associated variance.

Overall effect. The overall effect of music listening on driving was estimated in the current meta-analysis. In particular, the overall homogeneity in individual effects was first assessed by testing the significance of *Q*-statistics under the assumption that all effects were from the same population. If *Q*-statistics with a number of effects - 1 degrees of freedom was found to be statistically significant, which indicates that the observed individual effects come from different populations, the overall effect was estimated under the random-effects model, where between-study variance was estimated using the Restricted Maximum Likelihood (REML) estimation method (Veroniki et al., 2016) was incorporated into the estimation of the overall effect. However, when the *Q*-statistics was found not to be statistically significant, indicating that the observed individual effects come from the same population, the fixed-effects model was used to estimate the overall effect.

Moderator effect. When significant between-study variation in effect sizes was found to be salient, and thus effect sizes were from the different population, a series of moderator analyses were performed to find sources of variation in effect sizes using various coded variables. In particular, the mixed-effects model with a moderator was adopted when a sizeable between-study variation after controlling for the moderator existed. When the mixed-effects model was used, the additional

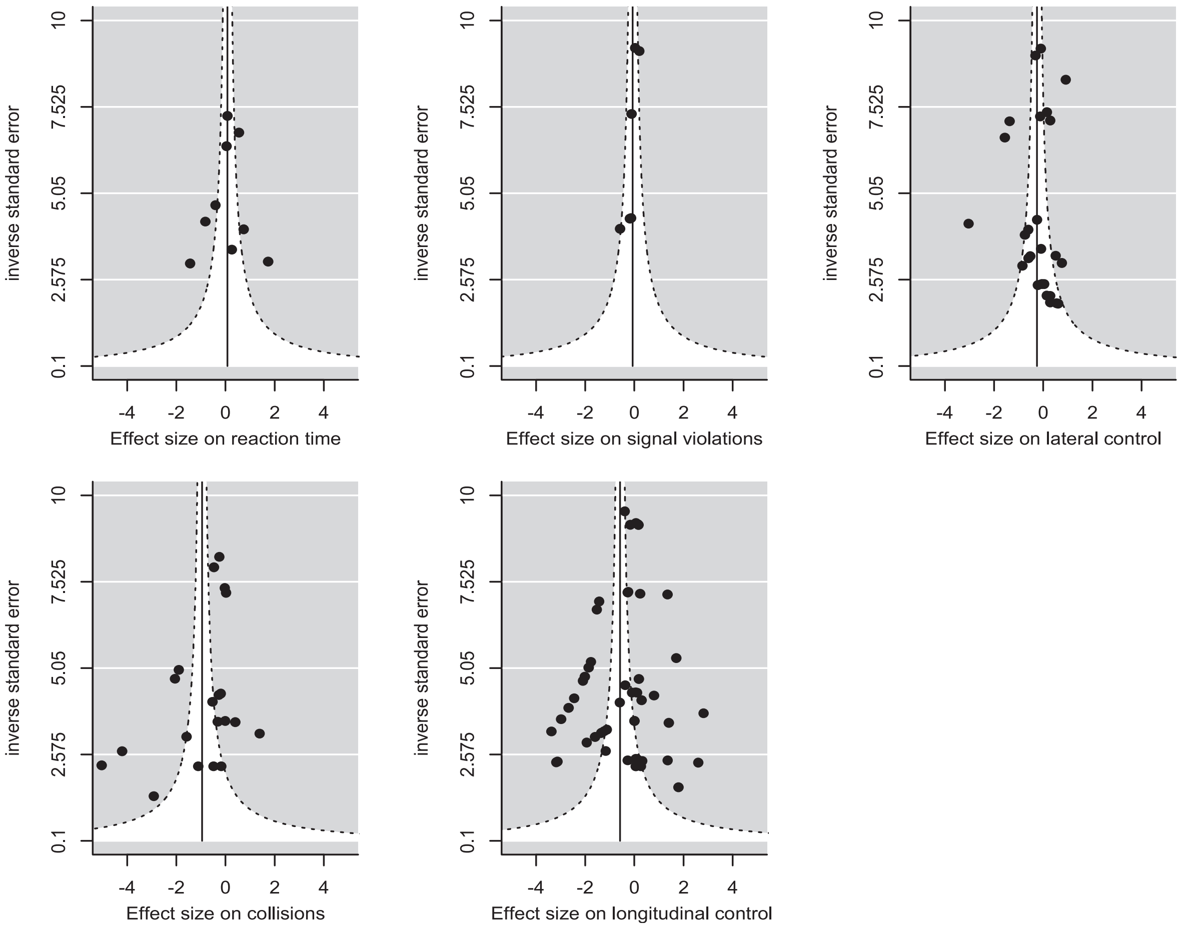

Fig 2. Funnel plots, separately by outcome.

between-study variance after controlling for moderator (estimated using REML method) was incorporated (Veroniki et al., 2016). Otherwise, the fixed-effects analysis with a moderator was used to compare individual effects by the different groups.

3. Results

3.1. Description of studies and sample included in the studies

Tables 1 and 2 display the coded information from the included studies. Of the 12 studies, one study was published before 2000, while all other studies were published between 2002 and 2018. Most of the studies ($k = 10$) used simulators for driving for their experiment, while a relatively low number of studies ($k = 2$) asked participants to drive on roadways. The total sample size used in all 12 studies was 752, with individual ranges from 5 to 165. Across all 12 studies, participants were (near-) equally distributed by gender (47% for males).

3.2. Publication bias

Fig. 2 displays funnel plots that are scatterplots of effect sizes from individual studies against the inverse values of their associated standard errors, separately by five outcomes measuring driving performance. These funnel plots suggest that publication bias might not be a concern on any of five outcome measures. Similar conclusions were drawn from both Egger's regression test for intercept ($z = -0.03$, $p = .98$ for reaction time; $z = -2.58$, $p = .009$ for signal violation; $z = 1.34$, $p = .18$ for lateral control; $z = -2.06$, $p = .04$ for collisions; $z = 1.58$, $p = .11$ for longitudinal control) and Begg and Mazumdar's rank test ($p = .76$ for reaction time; $p = .14$ for signal violation; $p = .51$ for lateral control; $p = .11$ for collisions; $p = .30$ for longitudinal control), where the null hypothesis stating no relationship between effect sizes and their associated variance was not rejected.

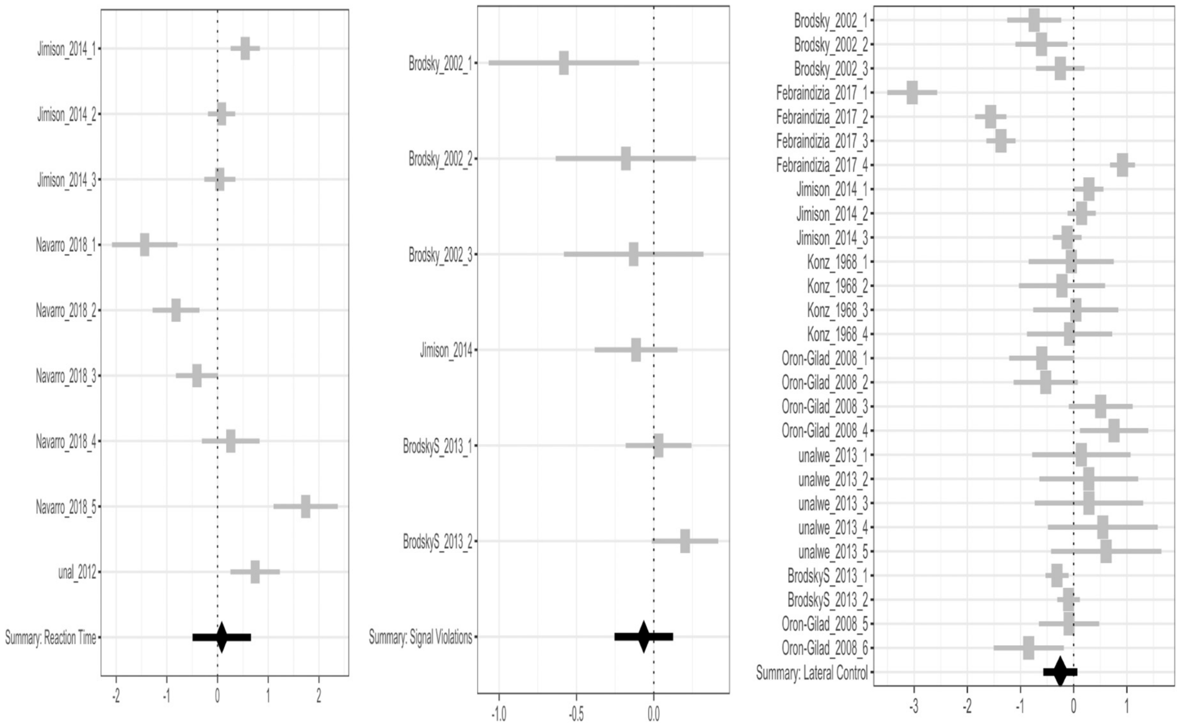

Fig. 3. Forest plots for effects related to reaction time, signal violations, and lateral control.

3.3. Effect of music on driving performance

The effect of music on driving performance was evaluated separately by the following five outcome variables: reaction time ($k = 9$), signal violations ($k = 6$), lateral control ($k = 23$), collisions ($k = 20$), and longitudinal control ($k = 55$). On all outcome measures, except signal violations, between-study variations were all statistically significant and substantially larger in their magnitudes ($Q(8) = 84.56$, $p < .01$, $I^2 = 94.53$ for reaction time, $Q(22) = 430.81$, $p < .01$, $I^2 = 93.63$ for lateral control, $Q(19) = 412.80$, $p < .01$, $I^2 = 97.84$ for collisions, and $Q(54) = 1792.07$, $p < .01$, $I^2 = 98.63$ for longitudinal control). Figs. 3 and 4 display forest plots of individual study effects with the estimated overall effect separately by the five outcome measures.

Under the random-effects model, where between-study variation estimated using Restricted Maximum Likelihood Estimation (REML) was incorporated, the statistically significant effect of listening to music while driving was found on the dependent variables of collisions ($z = -2.78$, $p = .01$) and longitudinal control ($z = -2.36$, $p = .02$), suggesting that driving with music leads to a negative relation with driving performance.

As shown in Fig. 5, the estimated standardized mean differences between control and experimental groups were −0.96 ($SE = 0.34$, 95% CI: −1.63, −0.28) for collisions and −0.58 ($SE = 0.25$, 95% CI: −1.06, −0.10) for longitudinal control. These results suggest that subjects in the experimental group who listened to music while driving showed significantly more collisions and poorer longitudinal control, when compared to subjects in the control group. The magnitudes of mean difference between the two comparison groups were large on both collisions and longitudinal control. No statistically significant difference between experimental and control groups, however, was found on other outcomes, including reaction time, signal violations, and lateral control.

3.4. Moderating effect of music stimuli

For each of the five outcome variables: reaction time ($k = 9$), signal violations ($k = 6$), lateral control ($k = 23$), collisions ($k = 20$), and longitudinal control ($k = 55$), a series of the mixed-effects model were conducted to examine whether the effect of music-listening on driving performance vary by the different features of music stimuli used in the experiment. Features of music stimuli examined in moderator analyses included (1) whether music was selected by the researcher or driver (researcher vs. driver), (2) whether music is familiar to participant (familiar, probably familiar, unfamiliar), (3) the source of music (pre-existing vs. original), (4) tempo of music (slow, medium, fast, mixed), and (5) volume of music (low, medium, high). Table 3 summarizes results from the moderator analyses, which were evaluated separately for each outcome measure.

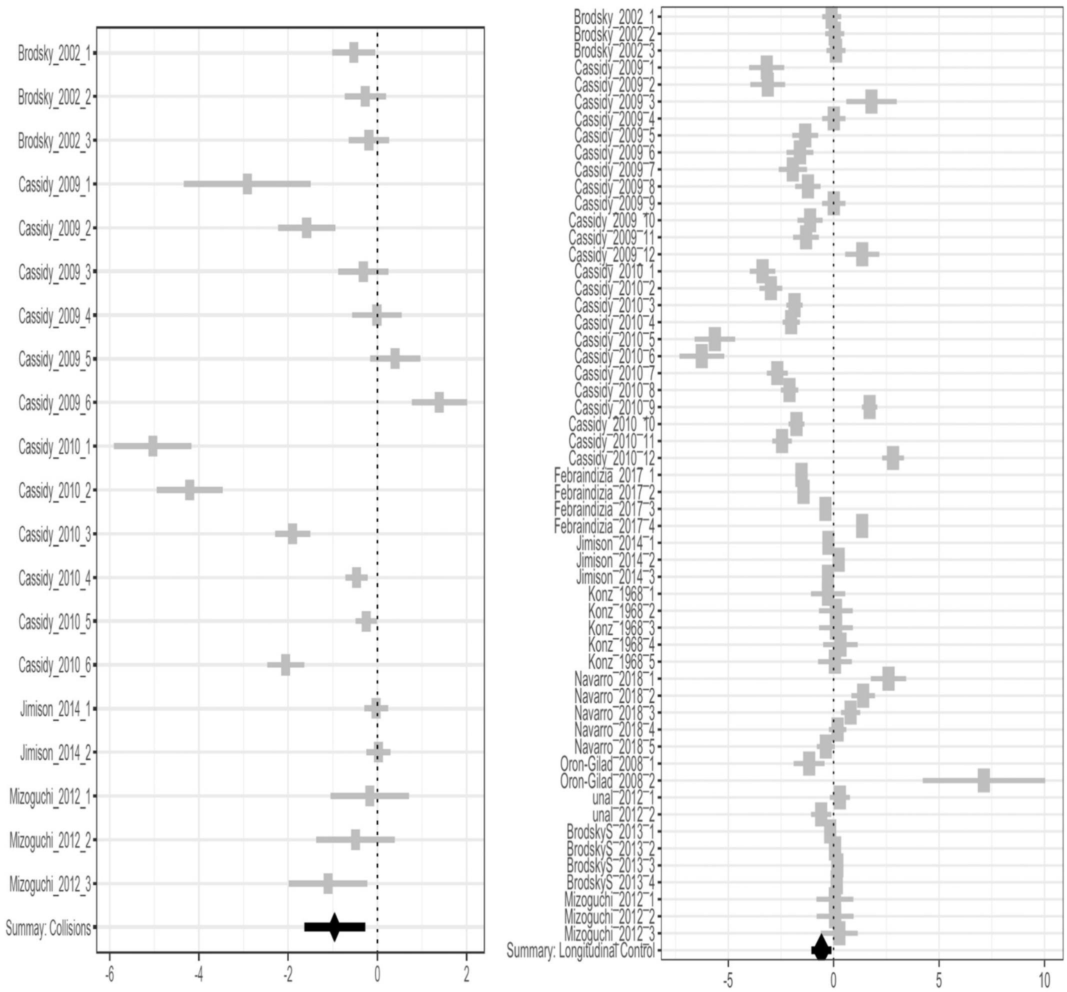

Fig. 4. Forest plots for effects related to collisions, and longitudinal control.

3.4.1. Selection of music

The random-effects model was first conducted to gather effect sizes that compare driving performance of participants who listened to researcher-selected or driver-selected music while driving. Results from the random-effects model suggests that the driving performance of participants who drove while listening driver-selected music did not differ significantly, compared to those who drove with researcher-selected music ($\bar{d} = -0.01$, $SE = 0.06$, $z = -0.19$, $p = .85$, $k = 15$).

Results from the mixed-effects model, however, suggest that the effect of listening to music while driving significantly varied depending on whether music was selected by researchers vs. drivers. Such moderating effect was found on lateral control ($Q(1) = 4.50$, $p = .03$) and collisions ($Q(1) = 3.95$, $p = .04$). Specifically, the effects of music on lateral control and collisions were statistically significant when participants were driving with researcher-selected music ($\bar{d} = -0.61$, $SE = 0.23$ for lateral control; $\bar{d} = -1.53$, $SE = 0.43$ for collision), suggesting a detrimental effect of researcher-selected music on driving performance.

3.4.2. Familiarity of music

Results from the mixed-effects model indicate that the effect of listening to music while driving significantly differs depending on the extent to which music is familiar to participants. Such effect was found on three outcomes: signal violations ($Q(3) = 7.82$, $p = .02$), lateral control ($Q(3) = 9.41$, $p = .02$), and collisions ($Q(2) = 7.78$, $p = .02$). Specifically, participants in the experimental group showed significantly higher mean on collisions when they listened to unfamiliar music while driving ($\bar{d} = -1.99$, $SE = 0.48$). In addition, a detrimental effect of music was found on signal violation when participants were less familiar with the music ($\bar{d} = -0.28$, $SE = 0.14$).

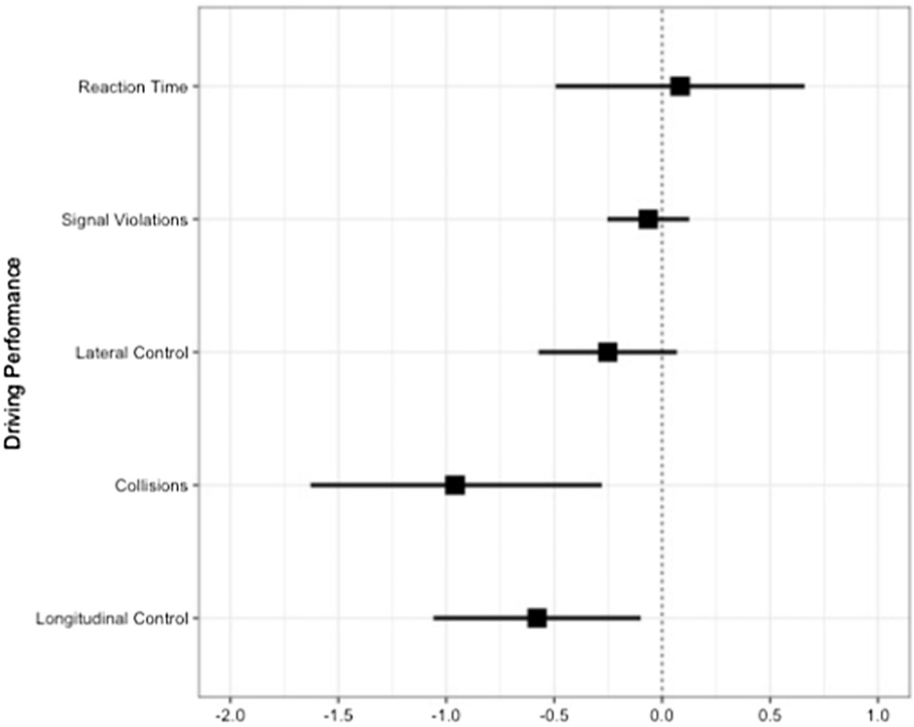

Fig 5. Estimated mean difference between control and experimental groups by outcomes.

Table 3
Moderator analysis.

Music feature	RT	Sig	Lat.	Coll.	Long.
Selection	NA	$Q_{(1)} = .10$	$Q_{(1)} = 4.50^{**}$	$Q_{(1)} = 3.96^{*}$	$Q_{(1)} = .05$
Researcher		−.11 (.15)	$-.61^{**}$ (.23)	-1.53^{**} (.43)	−.67 (.35)
Driver		−.04 (.18)	.05(.21)	−.25 (.48)	−.55 (.41)
Familiarity	NA	$Q_{(3)} = 7.82^{*}$	$Q_{(3)} = 9.41^{*}$	$Q_{(2)} = 7.78^{*}$	$Q_{(3)} = 2.03$
Yes		−.03 (.08)	.05 (.20)	−.25 (.45)	−.48 (.37)
Probably		$-.28^{*}$ (.14)	−.29 (.30)	−.33 (.77)	.05 (.65)
Not probable		.20 (.11)	−.10 (.71)	-1.99^{**} (.48)	-1.02^{*} (.44)
Not given		–	-1.23^{**} (.36)	–	−.50 (.91)
Source	NA	NA	$Q_{(1)} = .04$	$Q_{(1)} = 11.98^{**}$	$Q_{(1)} = 3.32$
Pre-existing			−.26 (.17)	−.47 (.30)	−.37 (.27)
Original			−.10 (.82)	-2.81^{**} (.61)	-1.51^{**} (.56)
Tempo	$Q_{(1)} = .11$	$Q_{(4)} = 3.73$	$Q_{(4)} = 2.33$	$Q_{(4)} = 8.36$	$Q_{(4)} = 2.08$
Slow		−.13 (.30)	−.09 (.52)	−.89 (.62)	−1.19 (.54)
Medium		−.18 (.30)	−.10 (.53)	−.27 (1.38)	.05 (1.84)
Fast		−.58 (.31)	.02 (.38)	-2.50^{**} (.63)	−.37 (.56)
Mixed	.008 (.39)	.03 (.22)	$-.59^{*}$ (.28)	−.33 (.53)	−.52 (.37)
Not given	.22 (.53)	.05 (.16)	−.11 (.35)	−.002 (.97)	.006 (.82)
Volume	$Q_{(1)} = .59$	$Q_{(1)} = .03$	$Q_{(2)} = 2.74$	$Q_{(2)} = 1.71$	$Q_{(3)} = 15.89^{**}$
<60	–	–	–	-1.28^{**} (.43)	-1.52^{**} (.33)
61–79	.001 (.32)	−.07 (.12)	.14 (.37)	−.16 (.89)	.47 (.54)
>80	.74 (.91)	−.05 (.28)	$-.53^{*}$(.23)	−.52 (.77)	−.13 (.43)
Not given	–	–	−.11 (.27)	–	.49 (.64)

Notes. NA: Not applicable due to no variation; RT is reaction time (e.g., lead vehicle braking), Lat. is lateral control (e.g., standard deviation of lane position [SDLP], lane excursions, etc.), Long. is longitudinal control (e.g., speed, speed variance, time or distance headway), Sig. is traffic signal violations (e.g., disobeying traffic lights, signs, or signals) and Coll. is number of collisions.

* $p < .05$.

** $p < .01$.

3.4.3. Music source

Results from the mixed-effects model indicate that the effect of music-listening differs depending on the music's source, specifically on collisions ($Q(1) = 11.98$, $p < .01$). These results suggest that participants perform significantly worse when music was original ($\bar{d} = -2.81$, $SE = 0.61$), suggesting that the effect of music is negative when participants listened to original music while driving.

3.4.4. Music tempo

None of the mixed-effects models show that the effect of music on driving performance significantly differs by the tempo of music used in the experimental design.

3.4.5. Music volume

Results from the mixed-effects model show that music volume significantly moderates the effect of music on longitudinal control ($Q(3) = 15.89$, $p < .01$). Specifically, a negative, but significant effect of music-listening while driving, was found when volume was less than 60 dB* on longitudinal control ($\bar{d} = -1.52$, $SE = 0.33$).

4. Discussion

Research on the effects of music-listening on driving performance offers conflicting views, with results that show both detrimental and beneficial effects. This meta-analysis provides the first comprehensive understanding of the relationship between music-listening while driving and driving performance by integrating the quantitative findings of relevant independent studies, estimating the overall effect, and identifying potential moderators that would explain the observed variations in study findings. This discussion fleshes out the results by attending to each moderator in the relationship between music-listening and driving performance, by providing related background information that helps contextualize the results.

4.1. Music and driving behavior

Selection. Drivers may use music to regulate their mood and arousal level, which in turn may enhance driving performance (Brodsky & Slor, 2013; Dibben & Williamson, 2007; Wiesenthal, Hennessy, & Totten, 2000); Wiesenthal et al. (2000), for instance, found that driver-selected music limited the driver's stress when driving through irritating traffic congestion. Investigating the effect of driver-selected versus researcher-selected music would therefore offer a contextualized understanding of the effect of music on driving performance. Various studies are consistent with the premise that self-selected music helps drivers regulate their mood and arousal level, which translates into better performance. Cassidy and MacDonald (2009) compared the effect of driver-selected versus researcher-selected music on the performance of a driving game; their results show that researcher-selected music had a detrimental effect in task performance; in a later study, Cassidy and MacDonald (2010) also found that researcher-selected music resulted in deficient performance. In contrast, Jimison (2014) found no significant difference in the effect of driver- or researcher-selected music on driving performance. Two models run in this meta-analysis resulted in slightly different results: the results from the random-effects model are consistent with Jimison (2014), i.e., there is no difference between driver-selected or researcher-selected music; the results from the mixed-effects model, on the other hand, are partly consistent with Cassidy and MacDonald (2009, 2010) and Wiesenthal et al. (2000), showing a detrimental effect of researcher-selected music on lateral control and collisions.

Familiarity. The independent variables of familiarity and driver-selected/researcher-selected music are often confounded, wrongly suggesting that researcher-selected music is unfamiliar to drivers; within experiments where familiarity is not part of the conceptual hypothesis, for instance, researchers generally seek to increase ecological validity by selecting familiar music. To disentangle these confounded variables, Jimison (2014) designed three music conditions: driver-selected/familiar, researcher-selected/familiar, researcher-selected/unfamiliar; the results of that study, however, suggest no significant difference in driving performance under the familiar or unfamiliar music conditions. In contrast, Cassidy and MacDonald (2009) found that listening to unfamiliar music is detrimental to driving performance and speculate that their results stem from driver-selected music being "most appropriate for the task context, liked most, and perceived to be least distracting, being context-specific, goal appropriate and familiar in nature" (p. 375). From such conjecture transpires that familiarity and preference may also be confounded. To avoid the conflation of familiarity and preference, Mizoguchi and Tsugawa (2012) investigated the effects of 'favorite' rather than 'familiar' music, thus embedding a layer of enjoyment that may not exist in familiar music; the results of their study are nonetheless consistent with other studies focusing on familiarity and on preference (Navarro, Osiurak, & Reynaud, 2018; North & Hargreaves, 1995, 1999), suggesting that listening to 'non-favorite' music has a detrimental effect as measured by the number of accidents. The results of this meta-analysis are likewise consistent with prior findings, showing a detrimental effect (particularly on collisions and signal violation) when listening to unfamiliar music while driving.

Source. The possible motivations for including original music, according to our systematic review, include: (1) to control for intrinsic music variables (such as tempo, instrumentation, timbre, range, and volume) and thus "provide a level of per-

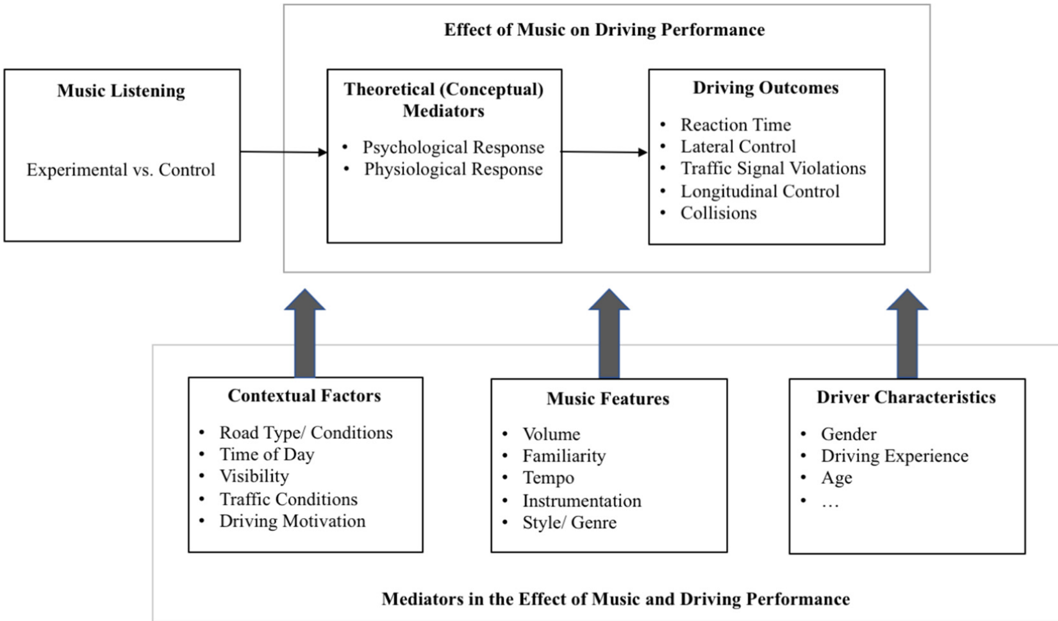

Fig. 6. The proposed process model for the effect of music on vehicular performance.

ceptual complexity that does not divert mental resources while driving" (Brodsky & Slor, 2013, p. 385)[2]; and (2) to control for familiarity effects (Brodsky & Kizner, 2012; Cassidy & MacDonald, 2010). Since originally composed music can only be researcher-selected, this variable lends itself to be confounded with familiarity. The results of our meta-analysis are consistent with Cassidy and MacDonald (2010), suggesting that participants perform significantly worse when listening to original music.

Tempo. Tempo is a music variable discussed in nearly all studies, arguably for the following reasons: (1) an understanding of tempo, as compared to instrumentation, texture, harmonic rhythm, meter, etc., is not contingent upon a sophisticated understanding of music; (2) a relatively early study (Brodsky, 2001) showed that tempo has an effect on simulated driving speed and on frequency of traffic violations, including speeding, collisions, lateral control, and red-light crossings[3]; (3) studies rely on survey data to advance broad claims and to conclude, for instance, that "the majority [of drivers] listen to highly energetic aggressive music consisting of a fast-tempo and accentuated beat played at strong intensity levels" (Brodsky & Kizner, 2012, p. 172). Additionally, numerous studies assume a detrimental effect of high tempi by drawing on indirect variables that result from it, including elevated heart rate (Brodsky, 2001; Dalton et al., 2007), high arousal potential (Cassidy & MacDonald, 2009; Dalton et al., 2007; Navarro et al., 2018; van der Zwaag et al., 2012), musical complexity (Brodsky, 2001; Brodsky & Kizner, 2012; Hughes et al., 2013), and cognitive load (Dalton & Behm, 2007; North & Hargreaves, 1995, 1999).[4] Aside from Brodsky (2001), the only other study that shows a positive correlation, albeit related to high arousal music rather than fast tempi, is Cassidy and MacDonald (2010), who found the "greatest overall inaccuracy in high-arousal music amplified by an increase in the tempo of the music" (462). In contrast, other studies found a negative correlation between tempo (categorized as high arousal music) and lap times (North & Hargreaves, 1995, 1999), and between tempo and task completion times (Cassidy & MacDonald, 2009). The results of our meta-analysis are consistent with studies that found no effect of tempo on accuracy and speed (Cassidy & MacDonald, 2009), on driving behavior (Navarro et al., 2018), and on car-following and braking for unexpected obstacles (Wang et al., 2015).

Volume. The literature offers conflicting views on the effect of volume on driving performance. Numerous studies report that loud volumes have a detrimental effect in the overall driving performance (Mizoguchi & Tsugawa, 2012; Spinney, 1997), in speeding (Brodsky, 2001), vigilance (Beh & Hirst, 1999; Dalton & Behm, 2007), and in response to stop-light signals (Beh & Hirst, 1999; Dalton & Behm, 2007). Similarly, North and Hargreaves (1995) suggested that combining high volume and high tempo results in cognitively demanding listening situations conducive to a higher frequency of driving errors and increased lap times. In contrast to studies that associate loud volumes with detrimental effects, Ünal, de Waard, et al. (2013), Ünal, Platteel, et al. (2013) found no difference in driving performance (i.e., a car-following task) when listening to music at loud or moderate volumes. In addition, numerous studies identified volume as a means for drivers to counter fatigue and main-

[2] Navarro et al. (2018) controlled for tempo without composing original music, but by manipulating the pre-existing tracks to increase or decrease tempi by 30%.

[3] Later work by Brodsky and colleagues draws on earlier findings, arguing that high music complexity is the consequence, among other things, of faster tempo, and that "the higher the complexity the greater the cost on attention resources" (Brodsky & Kizner, 2012, p. 165).

[4] Assuming however that HR, arousal, and to some extent musical complexity result from fast tempo should not lead us to conclude that increased HR, arousal, or musical complexity results in increased speeds, inaccuracies, violations, etc.

tain alertness (Dibben & Williamson, 2007) and to decrease boredom (Hargreaves & North, 2010). A more nuanced perspective is offered by Turner et al. (1996), who investigated the effect of volume on the response to unexpected visual events and reported that music at a moderate volume (70 dBA) facilitates performance (i.e., faster reaction to signals) as compared to conditions with music at low volume (60 dBA) or high volume (80 dBA). Consistent with these results, Consiglio et al. (2003) reported that music at medium volumes leads to performance comparable to a silent condition (in braking and vigilance tasks). The results of this meta-analysis suggest that music at low volumes have a detrimental effect on longitudinal control; this aligns, in part, with the results reported by Turner et al. (1996).

Instrumentation. Through several studies, instrumentation emerges as a relevant variable, particularly when contemplating the inclusion or exclusion of vocals (and lyrics). Within contexts indirectly related to driving, Crawford and Strapp (1994) showed that vocal music disrupts performance (i.e., scanning speed and logical reasoning) significantly more than instrumental music; similarly, Avila, Furnham, and McClelland (2012) showed that lyrics interfere with the processing of verbal information. Within the context of music and driving, therefore, we may expect that vocal music (lyrics) may disrupt performance and may interfere with the proper reading of verbal driving signs. Within the reviewed studies, Brodsky (2001), Brodsky and Kizner (2012) and Brodsky and Slor (2013) attended to the inclusion/exclusion of vocals, arguing that lyrics add a "level of perceptual complexity that [may] divert mental resources while driving" (Brodsky & Slor, 2013, p. 385). These notions notwithstanding, the number of studies that included selections without vocals is relatively small, precluding us from running analysis that would single out this variable.

Style/genre. Drawing on survey data, some studies report that young drivers (around 16–30 y/o) listen to pop, rock, dance, hip-hop, house, and rap styles when driving (Brodsky & Kizner, 2012; Brodsky & Slor, 2013; Brodsky, 2001), and that some particular styles (i.e., rap and hip-hop) "prompt aggressive conduct" and "have adverse effects on at least 50% of drivers" (Brodsky & Kizner, 2012, p. 165). The reporting of style in most studies, however, conforms to a generalized attempt to increase the ecological validity of the experiments by exposing drivers to music they would normally listen to while driving; the results of individual studies, nevertheless, can only be generalized to the style(s) of music used within each study. Santoso et al. (2013) investigated the effect of music genre on driving speed, albeit with only two conditions each with a single music selection drawn from Indonesian popular music; the study reported no significant effect. While investigating the effect of volume on driving-related tasks, Dalton and Behm (2007) found that hard rock music affects heart rate and reaction time, which in turn may affect tasks involving concentration and attention; the study, however, only included hard rock music and panpipe music (labeled as 'classical' in the study). Due to the relatively small number of studies coinciding in selected styles, and to the fluid nature of the concept of style, generalization is not warranted, and no analysis was conducted. Further research is needed to understand how independent variables interact with different musical styles to influence driving performance.

4.2. Proposed conceptual model

Fig. 6 presents a conceptual model that introduces key factors identified in the current systematic review and meta-analysis that may influence the relationship between music-listening and vehicular performance. The key factors in the proposed conceptual model are (1) music-listening as an independent variable, (2) potential mediators including psychological and/or physiological response, (3) vehicular performance as an outcome, and (4) moderator variables including music features, driver characteristics, and other contextual factors. The proposed conceptual model delineates how music may directly or indirectly affect vehicular performance via psychological (e.g., mood, arousal, emotion, and anxiety) and physiological responses (e.g., heart rate, blood pressure, and cortisol levels); in the model, vehicular performance is operationalized as reaction time, longitudinal control, traffic signal violation, lateral control, and collisions. It is hypothesized that the relation between music-listening and driving performance can differ by other moderator variables, including (1) musical parameters such as volume, familiarity, tempo, instrumentation, and style/genre, (2) driver characteristics such as gender, driving experience, and age, and (3) contextual/environmental factors such as road type, time of day, and road visibility.

The model summarizes the music variables identified and manipulated in the studies included in the meta-analysis (i.e., volume, familiarity, tempo, instrumentation, and style/genre), yet it is not intended to reflect the findings exclusive of this meta-analysis. Meta-analyses seldom provide definitive conclusions; rather, meta-analyses are "interpreted as state-of-the-art empirical knowledge about a specific effect or research area" (Lakens, Hilgard, & Staaks, 2016). This model, therefore, is broadly designed to guide future research and to guide in the formation of testable hypotheses.

4.3. Limitations and future research

There are a number of methodological and theoretical limitations that need to be addressed for the current meta-analysis and systematic review. One of the most salient methodological limitations stems from the inability to validate the proposed conceptual model (shown in Fig. 6) as no single researcher or study has investigated all key factors included in the proposed conceptual model. Additional empirical research, therefore, is needed to validate the proposed model. In addition, due to the small number of studies included in the meta-analysis, generalization of the statistical findings in the current meta-analysis is limited to the population that the included studies were drawn from.

5. Conclusions

Much research explores the effects of music-listening while driving, but the results of these studies offer conflicting views at numerous levels. This meta-analysis integrated the findings of relevant studies, identified the potential moderators/mediators for variation in effects, and estimated the overall effects. The results of this meta-analysis resonate with empirical research and with anecdotal evidence alike, in that music-listening has an effect on driving performance; yet these results also reveal unexpected insights regarding the effect of particular musical parameters on driving performance: while loud volumes and fast tempi have long been suspected to have a detrimental effect on driving performance, the results of this meta-analysis indicate a detrimental effect of soft volumes and no significant difference in driving performance associated with tempo.

Naturally, at this point the reader may expect concrete recommendations related to music while driving. It is beyond the scope of this study, however, to speculate about (and suggest) behavioral changes, even when these suggestions are informed by the results of this meta-analysis. The concrete insights about tempo and volume notwithstanding, through this meta-analysis we identified shortcomings and limitations in existing studies; these limitations preclude us from speculating about the precise mechanisms whereby music and particular musical parameters support or hinder driving performance. Based on these shortcomings and limitations, the authors suspect that truly understanding the effect of music-listening on driving performance should be contextualized more broadly, investigating related mechanisms such as cortical entrainment (Doelling & Poeppel, 2015), stress regulation (Poock & Wiener, 1966), attention and memory improvement (Perham & Vizard, 2011), mental fatigue alleviation (Guo, Ren, Wang, & Zhu, 2015), and arousal and positive mood modulation (Easterbrook, 1959; Perham & Vizard, 2011).

Additionally, when investigating music-listening and driving performance within the broader context of psycho-physiological mechanisms, it is critical to attune the experimental design to evolving trends. For instance, existing research draws by-and-large on a cohort of participants who predate the widespread use of digital multimedia. The 'Generation Z' is for the most part excluded from this research; this demographic cohort (i.e., individuals born mid-1990s to mid-2000s) is composed of 'digital natives' (Prensky, 2001), whose perceptual and cognitive capacities have been shaped by a native engagement with new technologies, and who therefore may engage with music-listening and driving performance differently than earlier generations. Ultimately, because music-listening while driving is likely to continue to be a widespread behavior (Rentfrow & Gosling, 2003; Sloboda, 1999), traffic researchers, vehicle manufacturers, and insurance companies should attend more closely to the benefits and risks associated with it.

References

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.

Beh, H. C., & Hirst, R. (1999). Performance on driving-related tasks during music. *Ergonomics, 42*(8), 1087–1098. https://doi.org/10.1080/001401399185153.

Bellinger, D. B., Budde, B. M., Machida, M., Richardson, G. B., & Berg, W. P. (2009). The effect of cellular telephone conversation and music listening on response time in braking. *Transportation Research Part F: Traffic Psychology and Behaviour, 12*(6), 441–451. https://doi.org/10.1016/j.trf.2009.08.007.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Introduction to meta-analysis. *Wiley*. https://doi.org/10.1002/9780470743386.

Brodsky, W. (2001). The effects 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., & 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.

Caird, J. K., & Horrey, W. J. (2011). *Twelve practical and useful questions about driving simulation*. Raton, FL: CRC Press, Boca (pp. 5-1) http://doi.org/10.1201/b10836-6.

Caird, J. K., Johnston, K. A., Willness, C. R., Asbridge, M., & Steel, P. (2014). A meta-analysis of the effects of texting on driving. *Accident Analysis & Prevention, 71*, 311–318. https://doi.org/10.1016/j.aap.2014.06.005.

Cassidy, G., & MacDonald, R. (2009). The effects of music choice on task performance: A study of the impact of self-selected and experimenter-selected music on driving game performance and experience. *Musicae Scientiae, 13*(2), 357–386. https://doi.org/10.1177/102986490901300207.

Cassidy, G. G., & MacDonald, R. A. R. (2010). The effects of music on time perception and performance of a driving game. *Scandinavian Journal of Psychology, 51*(6), 455–464. https://doi.org/10.1111/j.1467-9450.2010.00830.x.

Consiglio, W., Driscoll, P., Witte, M., & Berg, W. P. (2003). Effect of cellular telephone conversations and other potential interference on reaction time in a braking response. *Accident Analysis & Prevention, 35*(4), 495–500. https://doi.org/10.1016/S0001-4575(02)00027-1.

Cooper, H. (2009). *Synthesizing research* (4th ed.). Thousand Oaks, CA: Sage.

Core Team, R. (2013). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing http://www.R-project.org/.

Crawford, H. J., & Strapp, C. M. (1994). Effects of vocal and instrumental music on visuospatial and verbal performance as moderated by studying preference and personality. *Personality and Individual Differences, 16*(2), 237–245. https://doi.org/10.1016/0191-8869(94)90162-7.

Cummings, P., Koepsell, T. D., Moffat, J. M., & Rivara, F. P. (2001). Drowsiness, counter-measures to drowsiness, and the risk of a motor vehicle crash. *Injury Prevention, 7*(3), 194–199. https://doi.org/10.1136/ip.7.3.194.

Dalton, B. H., & Behm, D. G. (2007). Effects of noise and music on human and task performance: A systematic review. *Occupational ergonomics, 7*(3), 143–152.

Dalton, B. H., Behm, D. G., & Kibele, A. (2007). Effects of sound types and volumes on simulated driving, vigilance tasks and heart rate. *Occupational Ergonomics, 7*(3), 153–168.

Dibben, N., & Williamson, V. J. (2007). An exploratory survey of in-vehicle music listening. *Psychology of Music, 35*(4), 571–589. https://doi.org/10.1177/0305735607079725.

Doelling, K. B., & Poeppel, D. (2015). Cortical entrainment to music and its modulation by expertise. *Proceedings of the National Academy of Sciences, 112*(45), E6233–E6242. https://doi.org/10.1073/pnas.1508431112.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological review, 66*(3), 183. https://doi.org/10.1037/h0047707.

Fairclough, S. H., van der Zwaag, M., Spiridon, E., & Westerink, J. (2014). Effects of mood induction via music on cardiovascular measures of negative emotion during simulated driving. *Physiology & Behavior, 129*, 173–180. https://doi.org/10.1016/j.physbeh.2014.02.049.

Febriandirza, A., Chaozhong, W., Zhong, M., Hu, Z., & Zhang, H. (2017). The effect of natural sounds and music on driving performance and physiological. *Engineering Letters, 25*(4).

Gleser, L. J., & Olkin, I. (2009). Stochastically dependent effect sizes. In H. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 357–376). New York, NY, US: Russell Sage Foundation.

Groene, R., & Barrett, S. (2012). The effect of music and suggestion on defensive driving responses of high school students: Implications for music therapy. *Music Therapy Perspectives, 30*(1), 56–64. https://doi.org/10.1093/mtp/30.1.56.

Guo, W., Ren, J., Wang, B., & Zhu, Q. (2015). Effects of relaxing music on mental fatigue induced by a continuous performance task: Behavioral and ERPs evidence. *PloS one, 10*(8), e0136446. https://doi.org/10.1371/journal.pone.0136446.

Hargreaves, D. J., & North, A. C. (2010). Experimental aesthetics and liking for music. In P. N. Justin & J. A. Sloboda (Eds.), *Handbook of music and emotion: Theory, research, applications* (pp. 515–546). Oxford: Oxford University Press.

Harvey, A. R., & Carden, R. L. (2009). Driving error and anxiety related to iPod mp3 player use in a simulated driving experience. *Perceptual and Motor Skills, 109*(1), 159–167. https://doi.org/10.2466/pms.109.1.159-167.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, C.A: Academic Press.

Ho, C., & Spence, C. (2017). *The multisensory driver: Implications for ergonomic car interface design*. CRC Press.

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology, 45*(1–3), 73–93. https://doi.org/10.1016/S0301-0511(96)05223-4.

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.

Husain, G., Thompson, W. F., & Schellenberg, E. G. (2002). Effects of musical tempo and mode on arousal, mood, and spatial abilities. *Music Perception: An Interdisciplinary Journal, 20*(2), 151–171. https://doi.org/10.1525/mp.2002.20.2.151.

Jimison, Z. N. (2014). The effect of music familiarity on driving: A simulated study of the impact of music familiarity under different driving conditions.

Konz, S., & Mcdougal, D. (1968). The effect of background music on the control activity of an automobile driver. *Human Factors, 10*(3), 233–243. https://doi.org/10.1177/001872086801000305.

Lakens, D., Hilgard, J., & Staaks, J. (2016). On the reproducibility of meta-analyses: Six practical recommendations. *BMC Psychology, 4*(1), 24. https://doi.org/10.1186/s40359-016-0126-3.

Lee, J. D., Roberts, S. C., Hoffman, J. D., & Angell, L. S. (2012). Scrolling and driving: How an MP3 player and its aftermarket controller affect driving performance and visual behavior. *Human Factors, 54*(2), 250–263. https://doi.org/10.1177/0018720811429562.

Lerner, N., Baldwin, C., Higgins, J. S., Lee, J., & Schooler, J. (2015). Mind wandering while driving: What does it mean and what do we do about it? *Proceedings of the human factors and ergonomics society annual meeting* (Vol. 59(1), pp. 1686–1690). Sage CA: Los Angeles, CA: SAGE Publications. https://doi.org/10.1177/1541931215591364.

London, J. (2012). *Hearing in time: Psychological aspects of musical meter*. Oxford University Press http://doi.org/10.1093/acprof:oso/9780199744374.001.0001.

Madison, G., & Paulin, J. (2010). Ratings of speed in real music as a function of both original and manipulated beat tempo. *The Journal of the Acoustical Society of America, 128*(5), 3032–3040. https://doi.org/10.1121/1.3493462.

Mizoguchi, K., & Tsugawa, S. (2012). Influence of in-vehicle music on driving: Experimental results with a driving simulator. In *2012 IEEE International Conference on Vehicular Electronics and Safety (ICVES)* (pp. 117–121). IEEE. https://doi.org/10.1109/ICVES.2012.6294296.

Navarro, J., Osiurak, F., & Reynaud, E. (2018). Does the tempo of music impact human behavior behind the wheel? *Human Factors, 60*(4), 556–574. https://doi.org/10.1177/0018720818760901.

North, A. C., & Hargreaves, D. J. (1995). Subjective complexity, familiarity, and liking for popular music. *Psychomusicology: A Journal of Research in Music Cognition, 14*(1–2), 77. https://doi.org/10.1037/h0094090.

North, A. C., & Hargreaves, D. J. (1999). Music and driving game performance. *Scandinavian Journal of Psychology, 40*(4), 285–292. https://doi.org/10.1111/1467-9450.404128.

Oron-Gilad, T., Ronen, A., & Shinar, D. (2008). Alertness maintaining tasks (AMTs) while driving. *Accident Analysis & Prevention, 40*(3), 851–860. https://doi.org/10.1016/j.aap.2007.09.026.

Perham, N., & Vizard, J. (2011). Can preference for background music mediate the irrelevant sound effect? *Applied Cognitive Psychology, 25*(4), 625–631. https://doi.org/10.1002/acp.1731.

Poock, G. K., & Wiener, E. L. (1966). Music and other auditory backgrounds during visual monitoring. *Journal of Industrial Engineering, 17*(6), 318–323.

Prensky, M. (2001). Digital natives, digital immigrants. *On the Horizon, 9*, 1–6.

Rentfrow, P. J., & Gosling, S. D. (2003). The do re mi's of everyday life: The structure and personality correlates of music preferences. *Journal of personality and social psychology, 84*(6), 1236. https://doi.org/10.1037/0022-3514.84.6.1236.

Santoso, G. A., Maulina, D., Adystia, C., & Oei, T. P. (2013). The influence of number of passengers and music genre on driving speed of young adult angkot drivers. *Transportation Research Part F: Traffic Psychology and Behaviour, 18*, 1–10. https://doi.org/10.1016/j.trf.2012.12.001.

Sloboda, J. A. (1999). Everyday uses of music listening: A preliminary study. *Music, Mind and Science*, 354–369.

Spinney, L. (1997). Pump down the volume. *New Scientist, 155*, 22.

Staum, M. J., & Brotons, M. (2000). The effect of music amplitude on the relaxation response. *Journal of Music Therapy, 37*, 22–39. https://doi.org/10.1093/jmt/37.1.22.

Turner, M. L., Fernandez, J. E., & Nelson, K. (1996). The effect of music amplitude on the reaction to unexpected visual events. *The Journal of General Psychology, 123*(1), 51–62. https://doi.org/10.1080/00221309.1996.9921259.

Ünal, A. B., de Waard, D., Epstude, K., & Steg, L. (2013a). 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. (2013b). Blocking-out auditory distracters 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., Dijksterhuis, C., de Waard, D., Mulder, B. L., Westerink, J. H., & 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.

Veroniki, A. A., Jackson, D., Viechtbauer, W., Bender, R., Bowden, J., Knapp, G., ... Salanti, G. (2016). Methods to estimate the between-study variance and its uncertainty in meta-analysis. *Research Synthesis Methods, 7*, 55–79. https://doi.org/10.1002/jrsm.1164. doi: 10.1002/jrsm.1164.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software, 36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03.

Wang, D. Y. D., Jimison, Z., Richard, D., & Chuan, C. H. (2015). Effect of listening to music as a function of driving complexity: A simulator study on the differing effects of music on different driving tasks.

Westergaard, P. (1975). *An introduction to tonal theory*. New York: Norton.

Wiesenthal, D. L., Hennessy, D. A., & Totten, B. (2000). The influence of music on driver stress. *Journal of Applied Social Psychology, 30*(8), 1709–1719.

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.

Young, K. L., Mitsopoulos-Rubens, E., Rudin-Brown, C. M., & Lenné, M. G. (2012). The effects of using a portable music player on simulated driving performance and task-sharing strategies. *Applied Ergonomics, 43*(4), 738–746. https://doi.org/10.1016/j.apergo.2011.11.007.