**Auditory distraction during reading: A Bayesian meta-analysis of a continuing controversy**

Abstract

Everyday reading occurs in different settings, such as on the train to work, in a busy cafeteria, or at home, while listening to music. In all of these situations, readers are exposed to external auditory stimulation from nearby noise, speech, or music that may distract them from their task and reduce their comprehension. Although many studies have investigated auditory distraction effects during reading, the results have proved to be inconsistent and sometimes even contradictory. Additionally, the broader theoretical implications of the findings have not always been explicitly considered. In the present study, we report a Bayesian meta-analysis of 65 studies on auditory distraction effects during reading and use meta-regression models to test predictions derived from existing theories. The results showed that background noise, speech, and music all have a small, but reliably detrimental effect on reading performance. The degree of disruption in reading comprehension did not generally differ between adults and children. Intelligible speech and lyrical music resulted in the biggest distraction. While this last result is consistent with theories of semantic distraction, there was also reliable distraction by noise. It is argued that new theoretical models are needed that can account for distraction by both background speech and noise.

*Keywords*: reading, background noise, speech, music, meta-analysis

Reading is a critical skill that is indispensable in modern society. Although reading performance is best in silence when no distracting stimuli are present, such ideal conditions are rarely typical for daily life. Rather, much of everyday reading occurs in the presence of external auditory stimulation, such as noise from nearby traffic, music playing in the background, or a colleague talking on the phone. The interest in how auditory stimuli affect human performance is almost as old as modern psychology itself (e.g. Cassel & Dallenbach, 1918; Morgan, 1917). From to the widespread use of personal radios among students in the 1940s (Henderson, Crews, & Barlow, 1945; Miller, 1947) to the rise in popularity of the TV (Armstrong, Boiarsky, & Mares, 1991; Cool, Yarbrough, Patton, Runde, & Keith, 1994) and mobile devices (Kallinen, 2002), researchers and educators alike have been interested in whether background sounds can distract students from reading and other study-related tasks.

Over the past eight decades, many studies have examined how experimental exposure to speech, noise, and music affects the reading process. Although some interesting patterns of results have emerged, the research literature has been undermined by a fair number of inconsistent findings and the general lack of broader theoretical frameworks that can explain how auditory distraction during reading occurs. While a number of theoretical accounts have been developed in simpler tasks such as serial recall, it is currently not known how well they can account for all the findings from reading comprehension tasks that have been accumulated over the past several decades. Additionally, due to the mixed findings on some topics, it is currently not well understood what the magnitude of auditory distraction effects is, or even if they are reliably different from zero.

In the present paper, we address these issues in two ways. First, we present the first attempt to make a statistical synthesis of previous findings in a reading task in order to find out whether, and to what extent, auditory stimuli can interfere with reading performance. To do this, we adopted a Bayesian meta-analysis approach that makes it possible to quantify the degree of belief, given the data, that background sounds can disrupt reading. Second, we used Bayesian meta-regression models to test the predictions derived from existing theories on auditory distraction and to estimate how likely it is that they can explain the available data. The present paper starts with a brief overview of the literature that highlights the existing inconsistencies. Then, we consider theories that can explain auditory distraction effects during reading. Finally, the predictions from these theories are outlined and tested.

**The effect of background noise, speech, and music on reading: An overview**

**Background noise.** Although a number of epidemiological studies have suggested that chronic exposure to traffic noise is associated with lower reading ability in children (e.g., Haines, Stansfeld, Job, Berglund, & Head, 2001; Hygge, Evans, & Bullinger, 2002; Papanikolaou, Skenteris, & Piperakis, 2015; Stansfeld et al., 2005), only few studies have examined the effect of experimental exposure to noise. In one early study, Johansson (1983) found that 10-year-old children had the same reading comprehension and reading speed under quiet conditions, and conditions of continuous, or intermittent acoustical noise. More recently, Dockrell and Shield (2006) investigated the effect of typical classroom noise on reading comprehension in 8-year-old children. Participants completed the Suffolk Reading Scale in one of three conditions: silence, noise consisting of childrens’ babble, and the same babble combined with intermittent environmental noise. The results showed that children performed better in quiet than in the babble noise condition. Surprisingly, however, reading performance was best when the babble and the environmental noise were combined. Using similar sound stimuli, Ljung, Sörqvist, and Hygge, (2009) found that road traffic noise impaired the reading speed of 12-13-year-old children, but not their reading comprehension. However, a condition of children’s babble intermixed with irrelevant speech affected neither measure.

Studies of exposure to noise in adults have resulted in similarly mixed findings, sometimes even when done with the same materials (cf. Martin, Wogalter, & Forlano, 1988, Experiments 4 and 5). While most studies have failed to find an effect of acoustic or environmental noise on reading comprehension (Gawron, 1984; Jahncke, Hygge, Halin, Green, & Dimberg, 2011; Johansson, Holmqvist, Mossberg, & Lindgren, 2012; Veitch, 1990), others have found such an effect after examining the mediating role of personality characteristics (Furnham, Gunter, & Peterson, 1994; Ylias & Heaven, 2003). In summary, studies investigating the effect of background noise on reading comprehension have yielded inconsistent results, although some of them suggest that exposure to noise may be detrimental.

**Background speech.** Similar to noise, background speech is also often a nuisance to readers. However, speech has different acoustic properties than noise and it also carries semantic meaning if readers can understand it (which is very often the case in daily life). Perhaps owing to its semantic content, background speech is often rated as more distracting and more annoying than acoustical noise (Haka et al., 2009; Haapakangas et al., 2011; Landström, Söderberg, Kjellberg, & Nordström, 2002). Consistent with this subjective perception, intelligible background speech has been found to disrupt reading comprehension in a number of experiments (Armstrong et al., 1991; Baker & Madell, 1965; Martin et al., 1988; Sörqvist, Halin, & Hygge, 2010; however, see Venetjoki, Kaarlela-Tuomaala, Keskinen, & Hongisto, 2006). Additionally, there is some evidence to suggest that this disruption effect may be larger for participants who have poorer immediate suppression mechanism to ignore the background speech (Sörqvist, Halin, & Hygge, 2010; Sörqvist, Ljungberg, & Ljung, 2010).

Due to its implications for performance at the workplace, the effect of background speech on proofreading has also been investigated. Proofreading is a more cognitively demanding task than reading alone because it also requires allocating attention to look for mistakes, in addition to reading the text. There are generally two types of mistakes that have been considered in proofreading studies: contextual mistakes that require understanding the meaning of the text to detect, and non-contextual (i.e. spelling) mistakes that require only processing of the current word to detect. Due to the semantic content of intelligible speech, it can be hypothesized that background speech would disrupt the detection of contextual errors more than the detection of non-contextual errors.

Some support for this prediction was found by an early study by Weinstein (1977) who reported that background speech consisting of radio news significantly impaired the detection of contextual, but not the detection of non-contextual errors. However, Jones, Miles, and Page (1990) found exactly the opposite effect in another study. The authors manipulated both by the intelligibility of background speech (normal vs reversed) and the intensity of the sound (50 vs 70 dBA). They found that the intensity of the sound did not affect proofreading performance, but that normal (i.e. intelligible) speech reduced the number of non-contextual errors that were detected. Critically, however, the intelligibility of speech did not affect the detection of contextual errors (Jones et al., 1990). More recently, Venetjoki et al. (2006) found that background speech compared to continuous noise reduced the overall accuracy on a similar proofreading task. However, even though the task included both contextual and non-contextual errors, there was no significant effect of background speech on either error type in isolation. In a similar study, Landström et al. (2002) found that background speech, compared to broadband noise, did not affect proofreading performance for either contextual or non-contextual errors. The background stimuli were presented at comparable sound intensity level to Venetjoki et al. (2006), although the speech consisted of random spoken statements. Finally, Smith-Jackson and Klein (2009) also found no effect of background speech (intermittent or continuous) on overall proofreading accuracy.

Interestingly, a few studies have also suggested that the detrimental effect of background speech on reading and proofreading can be diminished by making the task harder and thus increasing participants’ engagement with it (Halin, 2016; Halin, Marsh, Haga, Holmgren, & Sörqvist, 2014a; Halin, Marsh, Hellman, Hellström, & Sörqvist, 2014b). In a few experiments, Halin et al. showed that performance on a reading/ proofreading task was disrupted by background speech only when the text was formatted in a familiar, but not in an unfamiliar (i.e. more difficult to read) font, and when the text was printed normally, but not when it was visually degraded (i.e. harder to read). Therefore, these results suggest that increasing task engagement may decrease the detrimental effect of background speech on reading comprehension and proofreading accuracy (see Sörqvist & Marsh, 2015 for a discussion).

Most studies that were considered so far have investigated only the end product of reading (i.e., comprehension accuracy or the overall time taken to read the text). However, these studies do not tell us how the reading process is influenced on a moment-to-moment basis. More recently, several eye-tracking studies have addressed this question by showing that the effect of background speech on reading can also be found at the level of fixation durations and fixations probabilities (Cauchard, Cane, & Weger, 2012; Hyönä & Ekholm, 2016; Vasilev, Liversedge, Rowan, Kirkby & Angele, n.d.; Yan, Meng, Liu, He, & Paterson, 2017). One key finding from these studies is that background speech leads to an increased number of re-reading fixations. While these studies have been successful in explaining where in the reading process disruption by background speech occurs, one puzzling aspect is that, unlike behavioural studies, none of the eye-tracking experiments have replicated the disruption effect in comprehension accuracy. It is currently not known why this inconsistency exists, but this raises the question of how reliable the effect of background speech on reading comprehension is.

In summary, background speech has been found to disrupt reading comprehension and proofreading accuracy in a number of experiments. In addition, the available evidence suggests that this disruption is due to processing the semantic meaning of the speech sound. These effects appear to be more reliable than the effect of background noise on reading, which has not been consistently replicated. Nevertheless, several studies in recent years have found no effect of background speech on reading comprehension, which casts doubt on its robustness and generalizability.

**Background music.** Unlike noise and speech, which are usually a nuisance, playing music in the background is often a personal choice or a habit. Interest in the potential effect of background music on reading started in the first half of the 20th century with the popularity of personal radios and their use by students. However, these early studies did not paint a clear picture of the relationship between background music and reading. While some of them found that music can negatively impact reading comprehension in children and university students (Henderson et al., 1945; Fendrick, 1937; Fogelson, 1973), others found that background music either does not affect reading (Freeburne & Fleischer, 1952; Miller, 1947; Mitchell, 1949) or that it actually improves it (Hall, 1952). Indeed, this controversy has persisted until the present day, and even the only two eye-tracking studies to address this question (Cauchard et al., 2012; Johansson et al., 2012) have failed to find any effect of background music on fixation durations or fixation probabilities.

To examine what conditions may give rise to distraction,somestudies have investigated whether the effect of background music on reading comprehension is modulated by personality traits (Avila, Furnham, & McClelland, 2011; Furnham & Allass, 1999; Furnham & Bradley, 1997; Furnham, Trew, & Sneade, 1999; Furnham & Stephenson, 2007; Furnham & Strbac, 2002). Based on Eysenck’s (1967) theory of personality, these studies have predicted that individuals high in extraversion will be distracted less by background music than individuals high in introversion due to the extroverts’ higher cortical arousal threshold. However, the results from these studies have been mixed. While some of them have found such an interaction between personality trait and background music (Daoussis & McKelvie, 1986; Furnham & Bradley, 1997; Furnham & Strbac, 2002), others have not (Avila et al., 2011; Furnham & Allass, 1999; Furnham et al., 1999; Furnham & Stephenson, 2007; Kou, McClelland, & Furnham, 2017). A number of factors may have led to these inconsistencies, such as the way in which participants were classified as introverts and extroverts, or the small sample size in some of the studies.

Another factor that has been considered is the genre of the music (Kallinen, 2002; Miller & Schyb, 1989; Mullikin & Henk, 1985; Tucker & Bushman, 1991). However, as the popularity of music genres changes with time, it is arguably better to investigate what aspects of the music may cause distraction. One factor that may play a role is participants’ preference for the music. For example, Etaugh and colleagues (Etaugh & Michals, 1975; Etaugh & Ptasnik, 1982) reported that preferred music decreased reading comprehension scores, but only for students who seldom study with music. In contrast to this, Johansson et al. (2012) found that participants had lower comprehension accuracy when listening to non-preferred music, but there was no effect of preferred music. Participants’ studying habits also did not modulate the results. Adding further to the confusion, Perham and Currie (2014) found that both liked and disliked lyrical music is equally disruptive to reading comprehension, although they did not report data on students’ studying habits.

The influence of background music on reading may also be modulated by the acoustic properties of the music. Some factors that have been considered are its informational load (Kiger, 1989), loudness and tempo (Thompson, Schellenberg, & Letnic, 2012), familiarity to participants (Hilliard & Tolin, 1979) and its capability to induce a startle response (Ravaja & Kallinen, 2004). These results are quite interesting in understanding what types of music may cause distraction, although they would benefit from further replication and extensions. In summary, previous studies suggest that certain types of music may be distracting, but a negative effect of background music on reading performance has not been consistently observed.

To summarize the discussion so far, the available evidence suggests that experimental exposure to background noise, speech, and music may disrupt reading performance. The effect of background noise and music appears to be less consistent, with many studies reporting non-significant effects on reading comprehension. Although the effect of background speech on reading appears to be more reliable, several experiments have also failed to find a distraction effect in reading comprehension and proofreading tasks. Therefore, considerable uncertainty exists with respect to the magnitude of these distraction effects and what aspects of background sounds may be responsible for them. One possibility is that only certain acoustical or linguistic properties of background sounds may account for the distraction. We now turn to this possibility by examining existing theories of auditory distraction.

**Theories of auditory distraction**

One of the earliest theoretical accounts of auditory distraction effects is the *phonological interference* hypothesis. This account is based on Baddeley and Hitch’s (1974; 1994) model of working memory, in which the phonological loop acts as an acoustic store where memories are registered and rehearsed through a process of sub-vocalization. Salamé and Baddeley (1982; 1987; 1989) reported a series of experiments in which they showed that memory for visually presented digits is impaired by unattended speech, but not by unattended acoustical noise. Additionally, a disruption effect was observed even if the speech sound was in a language that participants could not understand (Salamé and Baddeley, 1987). The authors argued that this is because speech sounds automatically gain access to the phonological loop and thus interfere with the encoding and rehearsal of visually presented items. Although this hypothesis is derived from a memory task, Salamé and Baddeley (1989) argued that a similar disruption may also be observed in more complex cognitive tasks such as reading.

Martin et al. (1988) were first to systematically test the phonological interference hypothesis in a reading comprehension task. In a series of experiments, they found that the disruptive effect of unattended speech was due to the semantic properties (i.e. meaning) of the speech, rather than its phonological features. More specifically, the authors found that English speech (intelligible to participants) was more distracting that Russian speech (unintelligible to participants). Similarly, a continuous speech stream of random words was found to disrupt comprehension more than a continuous speech stream of non-words. To account for these results, Martin et al. (1988) argued that, unlike serial recall tasks, reading comprehension requires understanding the meaning of the text. Therefore, the semantic properties of the irrelevant speech can interfere with building the semantic representations of the text that is being read. This prediction will be referred to as the *semantic interference* hypothesis.

The *changing-state* hypothesis (Hughes & Jones, 2001; Jones & Macken, 1993; Jones, Madden, & Miles, 1992) is another prediction that is also derived from serial recall tasks. According to this hypothesis, interference is caused by background sounds that exhibit considerable acoustic variation, but not by steady-state, aperiodic sounds that do not have such variation (Jones et al., 1992). For example, a sound consisting of different consonants (e.g., “B, F, P, S, N”) should cause more interference than a sound made up of the same consonant (e.g., “M, M, M, M, M”) because it exhibits more acoustic variation. The hypothesized mechanism through which interference occurs is that changing-state sounds contain information about the serial order of their constituent sound elements (Hughes & Jones, 2001). This information can then interfere with maintaining the serial order of items in a memory task.

Although reading is a more complex cognitive task, it also involves maintaining the order of words in the sentence, as well as their syntactic relations. For example, models of parallel word processing such as SWIFT (Engbert, Nuthmann, Richter, & Kliegl, 2005) assume, at least implicitly, that readers are somehow able to maintain word-order information while processing multiple words at the same time. Additionally, some models of reading comprehension (e.g., Kintsch, 1998) assume that word meanings are combined to form propositions or “idea units” according to their syntactic relationships (Kintsch & Rawson, 2005). Therefore, the forming of these units must also involve establishing the order of words in the sentence and how they relate syntactically to one another.

A final account that is relevant in a reading task is the *duplex theory* of auditory distraction (Hughes, 2014; Hughes, Vachon, & Jones, 2005; 2007; Sörqvist, 2010a). According to this theory, auditory distraction can occur from two different processes: *interference-by-process* and *attentional capture* (Hughes, 2014). Interference-by-process (Marsh, Hughes, & Jones, 2008; 2009; Marsh & Jones, 2010) occurs when the background sound interferes with a process that is important for the main task. For example, in a reading task, the semantic processing of meaningful speech would interfere with the task because reading also requires semantic processing to extract the meaning of the text. Alternatively, auditory distraction can also be caused by attentional capture (Hughes et al., 2005; Vachon, Hughes, & Jones, 2012) where attention is temporally directed away from the main task. For example, the sound “B” in the sequence “AAAAAA**B**A” would capture attention because another “A” is expected in the sequence (Hughes, 2014; see also Parmentier, 2014 for a review of similar effects due to deviant sounds).

In a reading task, the interference-by-process part of the duplex theory makes the same prediction as the semantic interference hypothesis by Martin et al. (1988) discussed earlier. In Marsh et al.’s (2008, 2009) account, distraction occurs because processing the meaning of the background speech depends on the same process used for extracting the meaning of the text that is being read. In contrast, Martin et al. (1988) assume that it is the semantic properties of the speech that cause the interference. These two very similar views are difficult to disentangle empirically, and since they make the same prediction, we will consider them together in the present analysis. The second part of the duplex theory- attentional capture- is a very interesting concept. However, because tasks such as reading typically involve longer exposure to sounds, it is more difficult to study and will not be considered further in this analysis.

**Present Study**

The review of the literature revealed that background noise, speech, and music may be detrimental to reading performance, but that considerable uncertainty exists as to whether such disruption effects can be consistently observed and what their magnitude is. This uncertainty does not make it possible to make firm conclusions about the experimental effects, nor does it tell us about their real-world significance. Are background sounds reliably disruptive to reading, and is this disruption large enough to be of any practical significance? Additionally, after 80 years of research on the topic, what theoretical conclusions can be made about the types of background sounds that are disruptive to reading?

The present study addressed these questions by performing a Bayesian random-effects meta-analysis of studies investigating experimental exposure to noise, speech, or music in the background. Both studies with adults and children were considered. Bayesian inference is especially suited to answer these questions because it uses the laws of probability to directly quantify the uncertainty of the estimate of auditory distraction effects, given the available evidence. This in turn makes it possible to derive the probability, given the data, that background noise, speech, and music can distract readers from their task. Bayesian meta-analytical models have traditionally been used in biology and medicine (e.g. Sutton & Abrams, 2001; Sutton et al. 2000), but more recently have also been introduced to psychology and linguistics (Vasishth, 2015; Vasishth, Chen, Li, & Guo, 2013; Jäger, Engelmann, & Vasishth, 2017). As such, they have been successfully used to address contentious research questions, such as the processing of relative clauses in Chinese (Vasishth et al., 2013), and the extent to which readers can pre-process words in parafoveal vision (Vasilev & Angele, 2017).

There are two available meta-analyses to date that have addressed how background noise and music affect a wide range of behavioural and cognitive tasks (Kämpfe, Sedlmeier, & Renkewitz, 2010; Szalma & Hancock, 2011). While the results from these meta-analyses are quite interesting, their more general focus on all types of cognitive tasks does not make it possible to make firm conclusions about reading in particular. Interestingly, Kämpfe et al. reported a separate analysis of reading-only studies and estimated the general effect of music to be r= -0.11 (d= -0.22). However, this estimate was based on only eight studies and thus does not include most of the currently available data. Therefore, one of the contributions of the present meta-analysis was to estimate the general effect of background noise, speech, and music on reading, and to calculate the probability, given all the available evidence, that these auditory stimuli are detrimental to reading performance.

The second and more important goal of the present analysis was to investigate what aspects of background sounds give rise to distraction. Although it can be informative to estimate the overall size of the effects, as previous meta-analyses have, this does not tell us what it is about these sounds that makes them distracting. As it was discussed previously, there are a few theories that make specific predictions about what type of auditory stimuli should be distracting. Therefore, the second aim of the study was to test the prediction of these theories using Bayesian meta-regression models (Welton, Sutton, & Cooper, 2012). As some of the theories outlined above were not originally developed in reading comprehension tasks, it is important to keep in mind that the present investigation is not a strict test of these theories. Rather, it aims to find out whether they can accommodate the existing evidence in reading tasks, and if not, to pave the way for the development of future theories.

**Predictions.** All of the predictions in the present analyses are summarised in Figure 1.The phonological interference hypothesis (Salamé, & Baddeley, 1982) makes the unique prediction that all types of speech should be equally distracting because they all gain access to the phonological store. Therefore, both intelligible speech (i.e., in participants’ native language) and unintelligible speech (i.e., in a foreign language) should be equally distracting. Additionally, the phonological interference hypothesis is not capable of explaining distraction by acoustical noise and non-lyrical music because both sounds do not gain access to the phonological store.

The semantic interference (Martin et al., 1988) and interference-by-process (Marsh et al., 2008) accounts both make the unique prediction that only intelligible speech that can be processed semantically by participants would cause distraction. Therefore, intelligible speech should be more distracting than unintelligible speech. Additionally, they also predict that: 1) lyrical music should be more distracting than non-lyrical music because the former contains lyrics that are intelligible to participants; and 2) intelligible speech should be more distracting than lyrical music because, on average, continuous speech has more semantic content than lyrical music. However, since lyrical music that is intelligible to participants contains not only semantic, but also phonological information, it is not possible to rule out any involvement of phonology in this effect. C:\Users\mvasilev\OneDrive\RSmeta\predictions.tif

*Figure 1*. A schematic summary of the predictions derived from theories on auditory distraction.

Finally, the changing-state hypothesis (Jones et al., 1992) predicts that sounds exhibiting considerable acoustic variation should be more distracting than steady-state sounds that do not exhibit such variation. This leads to two predictions. First, non-lyrical music should be more distracting than acoustical noise (e.g. white or pink noise). This is because the former exhibits more acoustical variation. Non-lyrical music is the strongest test of this prediction because it avoids any potential confounds from spoken language that would be present in lyrical music. Second, more complex environmental noise (e.g. traffic noise or office noise containing, phones ringing, indistinct chatter, etc.) should again be more distracting that steady-state acoustical noise because it also exhibits more acoustical variation.

**Method**

**Literature Search**

The search of the literature was conducted by following the PRISMA guidelines (Moher, Liberati, Tetzlaff, Altman, & Prisma Group, 2009). A flowchart of the process is presented in Figure 2. Google Scholar, Scopus, the Web of Science, and ProQuest Dissertations were searched with the following keywords: “background noise AND reading”, “background speech AND reading”, “background music AND reading”. The search for each of the three background sounds was done separately. The literature search covered articles published before the 25th of June, 2017. Additionally, the reference lists of all screened articles, as well as those of previous literature reviews and meta-analyses on similar topics (Beaman, 2005; Clark & Sörqvist, 2012; Dalton, & Behm, 2007; Kämpfe et al., 2010; Klatte, Bergström, & Lachmann, 2013; Shield & Dockrell, 2003; Szalma & Hancock, 2011), were also examined.

The identified articles were evaluated against the inclusion criteria presented in Appendix A. In short, the studies had to experimentally manipulate background noise, speech or music in a reading or a proofreading task, have a sound methodological design, and include reading in silence as a baseline condition. The inclusion criteria were developed prior to the meta-analysis with the help of a smaller, qualitative review of the literature. Epidemiological studies of chronic exposure to traffic noise in children were not included because they answer a qualitatively different question and are often confounded by other variables, such as social deprivation (Haines, Stansfeld, Head, & Job, 2002). Overall, 44 % of the experiments whose eligibility was assessed were included in the meta-analysis. Although the inclusion rate may appear to be low, it was necessary to ensure that only studies that are similar enough to be analyzed together are included. Information about the included studies and their effect sizes are presented in Appendix B.

C:\Users\Martin Vasilev\AppData\Local\Microsoft\Windows\INetCache\Content.Word\flowchart.tif

*Figure 2*. A flowchart illustrating the stages of the literature search process.

**Dependent Measures**

The main dependent variable was reading comprehension accuracy, which was available for 54 of the studies (83.1 %). Therefore, most of the reported analyses were based on reading comprehension. Moreover, effect sizes for reading speed were available for 13 studies (20 %), and these were analyzed separately. Finally, experiments reporting proofreading accuracy (N=7; 10.7 %) were also analyzed for completeness, but this was again done separately from the analysis on reading comprehension accuracy.

**Effect Size Calculation**

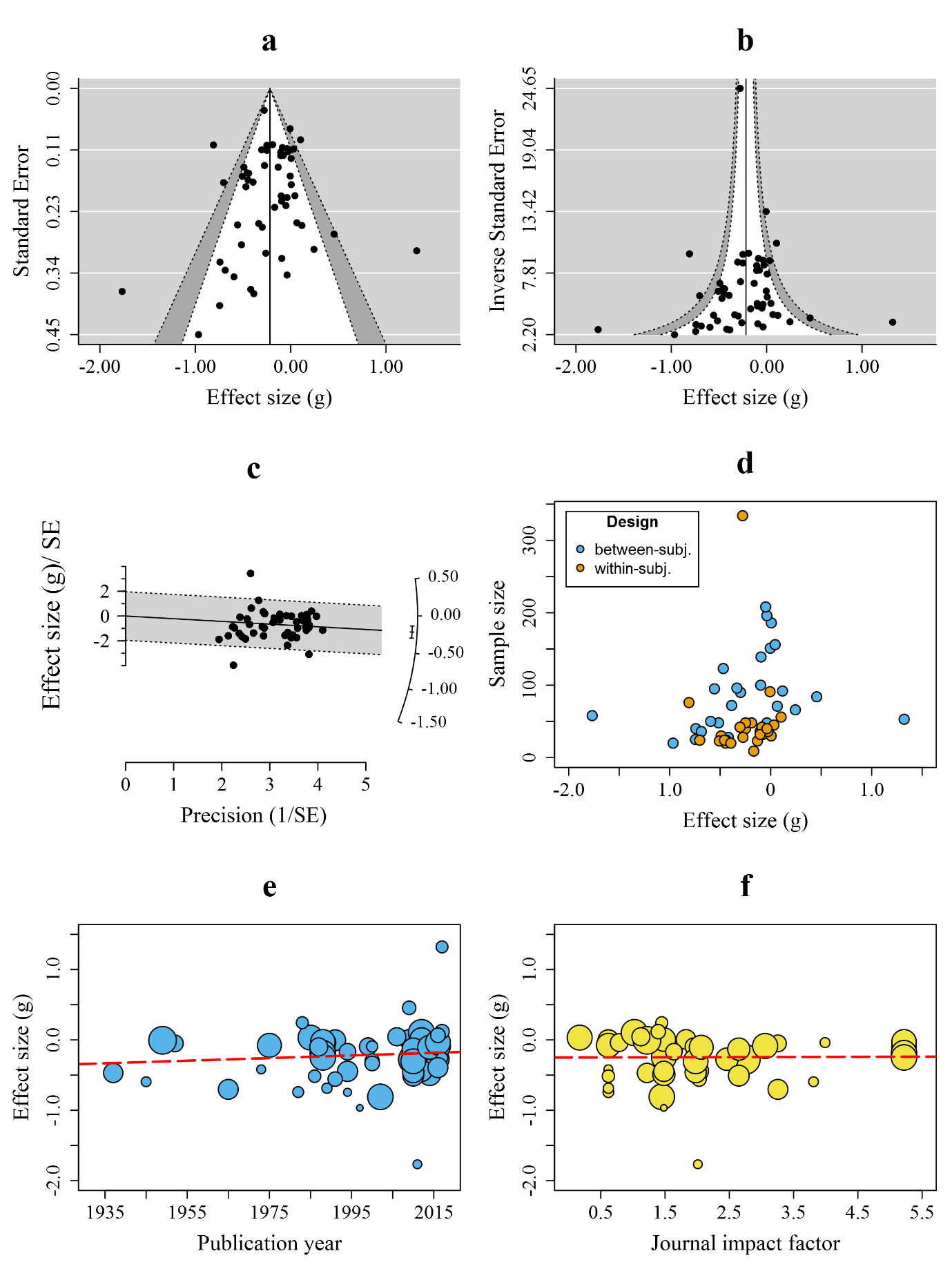
Standardized effect sizes of the mean difference (g) and their variances were calculated from the reported descriptive statistics. This was done by first calculating Cohen’s *d* for the respective design of the study and then applying Hedges’ *g* (Hedges & Olkin, 1985) correction for small sample bias. The effect sizes were calculated with formulas 12.11-12.22 from Borenstein (2009). In all effect sizes, silence was the control condition. Therefore, the effects represent the standardized mean difference between reading in the experimental sound condition and the control condition of reading in silence. If descriptive statistics were unavailable or incomplete, the effect sizes were calculated by digitalizing graphs (Rohatgi, 2015) or converted/approximated from the reported test statistics by using existing formulas (Borenstein, 2009; Lajeunesse, 2013)[[1]](#footnote-1). In the analysis of reading comprehension accuracy and proofreading accuracy, studies were coded so that negative effect sizes indicate lower comprehension/ proofreading accuracy in the experimental sound conditions. Similarly, in the analysis of reading speed, negative effect sizes also indicate slower reading speed in the experimental sound condition compared to silence.

Because 55.5% of the studies used a within-subject design, it was necessary to estimate the population correlation (ρ) between the control and experimental conditions. (Borenstein, 2009; Szalma & Hancock, 2011). Eight statistically-independent estimates were obtained from experiments for which the raw data were available, as well as from one study (Miller, 1947) that reported the required statistics. These represented a wide range of experimental sound types and included both reading comprehension and reading speed measures. We followed Szalma and Hancock’s (2011) approach to meta-analyze the obtained correlations and to obtain a weighted estimate of ρ. The resulting weighted value of 0.74 was used for calculating the effect sizes for all within-subject design studies.

Effect sizes from within- and between-subject studies are calculated with different standard deviation metrics and are thus not necessarily comparable (Morris & DeShon, 2002). Consistent with previous work (Kämpfe et al., 2010; Szalma & Hancock, 2011), the effect sizes from within-subject studies were transformed to make them comparable to the effect sizes of between-subject studies. This was done with formula 11 from Morris and DeShon (2002). Additionally, because some studies yielded more than one effect size, care was taken to avoid statistical non-independence in the analyses (see Noble, Lagisz, O'dea, & Nakagawa, 2017 for a recent overview). If a study contributed multiple effect sizes per analysis, these were averaged together to include only one effect size for that study (Lipsey & Wilson, 2001)[[2]](#footnote-2).

**Publication Bias**

In the present meta-analysis, 12.3% of all included studies were from the so-called grey literature (i.e. they were not formally published in a peer-reviewed journal or in an edited book at the time of analysis). Additionally, we performed visual and statistical investigation of publication and other related biases. This was done with the “meta” (Schwarzer, 2007) and “metafor” (Viechtbauer, 2010) R packages. The visualization of the results for reading comprehension is presented in Figure 3 (see the Supplemental Material for reading speed). The funnel plots (panels **a** and **b**) indicated that there was some heterogeneity in the data, but there was no clear evidence of asymmetry that could indicate publication bias. This was confirmed by a funnel plot test of asymmetry based on a weighted linear regression of the effect estimates on their standard errors (Sterne et al., 2011), which revealed no statistically significant evidence for asymmetry for either reading comprehension (t(52) = -0.42, p= 0.67) or reading speed (t(11)= 0.08, p= 0.93; proofreading accuracy was not considered here because funnel plot tests of asymmetry are not recommended when there are less than 10 studies; Sterne et al., 2011). Additionally, meta-regression analyses (Figure 3**e**-**f**) indicated that the size of auditory distraction effects was not predicted by the impact factor of the journal or the year of publication. In summary, there was no evidence to suggest that publication bias may have influenced the conclusions from the meta-analysis.

****

*Figure 3*. Visual assessment of publication and other related biases for reading comprehension accuracy (presentation format adapted from Nakagawa, Noble, Senior, & Lagisz, 2017, Figure 6). **a**: Funnel plot of effect sizes against their standard error.White and dark grey bounds indicate the 95% and 99% pseudo-confidence intervals, respectively. Studies with smaller standard error should appear at the top, while studies with smaller standard error should scatter at the bottom of the plot. **b**:The same funnel plot of effect sizes against the inverse of their standard error. **c**: Radial (Galbraith) plot of the z-statistic of each study (y axis) against the inverse of the standard error (x axis). Shading shows z-value bounds of ± 2. **d**: Plot of effect sizes against their sample size, broken down by study design type. **e** and **f**: Meta-regression models examining whether the size of effects is predicted by publication year (**e**) or impact factor of the journal where the study was published (**f**). Both models show that this was generally not the case. Red dotted line shows the meta-regression slope.

**Data Analysis**

**Meta-analysis**. The common choice in meta-analysis is between a fixed-effect and a random-effects model. A fixed-effect model assumes that all effect sizes that are combined are estimating the same true underlying effect, which we will call *θ*. Therefore, the effect size of the *i*-th study, *Ti*, is assumed to come from a normal distribution with some mean *θ* and variance :

(1)

In this model, any variability in the estimate is due to sampling error alone. On the other hand, a random-effects model relaxes this assumption by explicitly allowing for variability in the effect sizes between studies (Welton et al., 2012). In this case, the effect size of the *i*-th study *Ti* is assumed to be generated by a unique underlying effect for that study, denoted here by . This unique underlying effect is in turn assumed to come from a normal distribution with some (unknown) mean *θ* and between-study variance *τ2*: (2)

Therefore, the effect sizes of individual studies in a random-effects meta-analysis can be informally thought of as random samples from a normal distribution of effect sizes (Welton et al., 2012).

In the present meta-analysis, a random-effects model was chosen *a priori* because some between-study heterogeneity was expected due to differences in design, sound intensity levels, participants, reading materials, and so forth. A random-effects model can naturally account for such sources of variability between studies and is often the model of choice in studies on language processing (e.g. Jäger et al., 2017; Vasishth et al., 2013; Vasilev & Angele, 2017). The full Bayesian model was defined as follows (Jäger et al., 2017; Schmid & Mengersen, 2013):   
 (3)

,

,

where: is the observed effect size (in Hedges’ g) in the *i*-th study

is the true auditory distraction effect in the *i*-th study

is the true sampling variance of the *i*-th study, estimated from the within-study variance of the sampling distribution of study *i*

is the unknown true auditory distraction effect estimated by the model

is the unknown between-study variance

In this model, precision was defined as the inverse of the within-study variance of the sampling distribution. The last two lines in equation 3 indicate the prior probability distributions used for *θ* and *τ.* In the present analysis, we used Uniform priors that assign equal probability to any value on these intervals. As such, they have very little to no influence on the results. This was confirmed by doing a sensitivity analysis of the main results with alternative priors: *Normal (0, 104)* for *θ* and *Normal (0, 104) I(0, )* for *τ* (normal distribution truncated at 0)*.* The sensitivity analysis indicated that the choice of priors did not influence the results(see the Supplemental Material).

**Meta-regression**. Although random-effects meta-analysis can account for heterogeneity between studies, it does not tell us what causes this heterogeneity in the first place (Welton et al., 2012). However, it is possible to use meta-regression models to investigate how different study characteristics (e.g. whether the background music was lyrical or non-lyrical) are associated with the observed effect sizes. Meta-regression models are similar to the ordinary least-squares regression, but with the crucial difference that the estimate is adjusted by the precision of the studies (i.e., the inverse of the within-study variance of the sampling distribution; Welton et al., 2012). The model from equation 3 was extended by adding a regression coefficient *β* for the underlying effect of the covariate (the added parameters are formatted in bold; Jäger et al., 2017; Welton et al., 2012):

(4)

,

,

where: *β* is the regression coefficient for the underlying effect of the covariate .

is the true auditory distraction effect in the *i*-th study, adjusted for the covariate effect

is the unknown true auditory distraction effect, also adjusted for the covariate effect

All remaining parameters have the same interpretation as in equation 3.

The contrasts used for the covariate are presented in Table 1. These contrasts were used to test the predictions outlined in the introduction.

Table 1

*Type of Meta-Regression Comparisons and the Contrast Coding of Covariates*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Comparison | Covariate levels | | Contrast coding | |
| Level 1 | Level 2 | Level 1 | Level 2 |
| Non-lyrical vs lyrical music | non-lyrical | lyrical | -1 | 1 |
| Lyrical music vs intelligible speech | music | speech | -1 | 1 |
| Unintelligible vs intelligible speech | unintelligible | intelligible | -1 | 1 |
| Acoustical vs environmental noise | acoustical | environmental | -1 | 1 |
| Acoustical noise vs instrumental music | noise | music | -1 | 1 |
| Child vs adult participants | child | adult | -1 | 1 |

**Posterior sampling.** The posterior probability distribution was sampled with JAGS (Plummer, 2003) using the R software, v. 3.31 (R Core Team, 2016). Five Markov Chain Monte Carlo (MCMC) chains were run with 75 000 iterations each. Checks were made to ensure that the starting values of the MCMC chains did not influence the results. The first 3000 iterations were discarded as burn-in. A thinning interval of 5 was used for the MCMC chains (i.e., every fifth sample was retained) to reduce the influence of auto-correlation. The summary of the posterior distribution was based on 15 000 samples per chain (excluding the burn-in period). Convergence was assessed with visual inspection of the trace plots and with Gelman and Rubin’s (1992) convergence diagnostic. The diagnostics suggested that convergence had been achieved in all models.

The results are presented as the estimate of the effect sizes of interest and their corresponding 95 % credible intervals. Unlike the classical confidence intervals, credible intervals have the intuitive interpretation that they contain the true auditory distraction effect with 95% probability because the values within this interval make up 95% of the posterior probability distribution (cf. Morey, Hoekstra, Rouder, Lee, & Wagenmakers, 2016). All probabilities reported in the paper are the posterior probability, given the data, that auditory distraction effects exist. A more detailed summary of Bayesian methods and their interpretation is beyond the scope of this paper. However, Nicenboim and Vasishth (2016) provide an accessible overview.

**Results**

**Meta-analysis**

The results from the meta-analysis are presented in Table 2. Additionally, forest plots are presented in Figure 4 for the main measure of comprehension accuracy. To interpret the magnitude of the effects, we will consider Cohen’s (1988) guidelines of 0.20 for small effects, 0.50 for medium effects, and 0.80 for large effects. Overall, there was a small negative effect for reading comprehension (g= -0.21), which indicates that background sounds generally impaired comprehension accuracy. Consistent with the review of the literature, background speech had a stronger negative impact on reading comprehension (g= -0.26) compared to both background noise and music (g= -0.17 and -0.19, respectively). Nevertheless, the effect for all three sound types was fairly small in size. Reading speed and proofreading accuracy were also impaired by background sounds. However, the effect sizes for these two measures were very small and the 95% credible intervals all included 0 as a plausible value for the effect. Interestingly, however, the probability that these effects are negative was very high in all analyses (more than 90%). This means that, although the size of the effects was small, there was very high probability that background speech, noise, and music are detrimental to reading comprehension, reading speed, and proofreading accuracy.

Table 2

*Posterior Effect Size Estimates of Auditory Distraction Effects and 95% Credible Intervals from the Meta-analysis*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Type of analysis | N | Mean ES (g) | 95% CrI | p(ES<0 | Data) |
| Reading comprehension |  | | | |
| All sounds | 54 | -0.21 | [-0.30, -0.13] | > 0.99 |
| Noise | 12 | -0.17 | [-0.33, 0.002] | 0.97 |
| Speech | 20 | -0.26 | [-0.36, -0.17] | > 0.99 |
| Music | 36 | -0.19 | [-0.34, -0.05] | > 0.99 |
| Reading speed |  | | | |
| All sounds | 13 | -0.06 | [-0.15, 0.02] | 0.92 |
| Speech | 6 | -0.08 | [-0.20, 0.03] | 0.92 |
| Proofreading accuracy |  | | | |
| Speech and Noise | 7 | -0.14 | [-0.42, 0.04] | 0.94 |
| Speech a | 6 | -0.09 | [-0.30, 0.07] | 0.90 |

N: number of studies in the analysis. p(ES<0 | Data): probability that background sounds are distracting, given the data (i.e., probability that the effect size is smaller than 0).

a intelligible speech only

Because both studies with adults and children were included in the analyses above, we carried out meta-regression models to test whether the effect sizes differed between adults and children. Only reading comprehension was considered in these analyses, as there were too few child studies to reliably estimate differences in reading speed, and all proofreading studies were done with adults. The results are presented in Table 3. They show the estimated mean difference between studies with children compared to studies with adults, after adjusting for their precision in the analysis. Overall, the difference between adults and children was very close to 0, thus showing that background sounds were equally detrimental to reading comprehension for both children and adults. One exception was that background noise impaired reading comprehension in children slightly more than it did in adults, but the mean difference was still quite small (g= 0.05). Additionally, the effect was not highly reliable as there was only 74% probability of a true mean difference. Taken together, these results suggest that effect sizes for reading comprehension generally did not differ between adults and children. For this reason, child and adult studies were analyzed together in all remaining analyses.

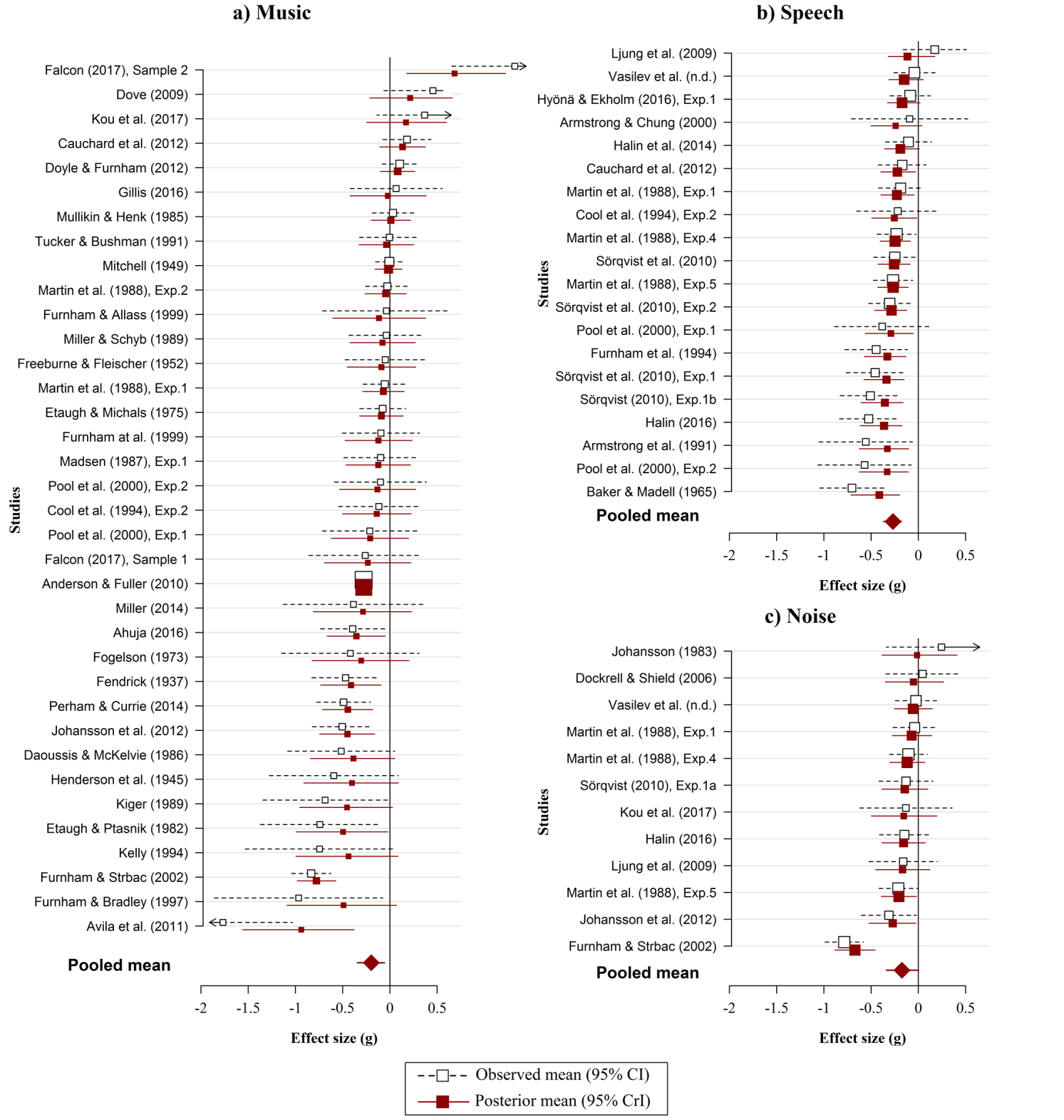


Table 3

*Mean Difference in the Effect Size Between Child and Adult Studies: Meta-regression Results*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Analysis | Number of studies | | Mean diff. (g) | 95 % CrI | p(ESCH> ESA | Data) |
| children | adults |
| Reading comprehension |  | | | | |
| All sounds | 18 | 36 | -0.01 | [-0.10, 0.08] | 0.43 |
| Noise | 5 | 7 | 0.05 | [-0.13, 0.22] | 0.74 |
| Speech | 5 | 15 | 0.00 | [-0.12, 0.12] | 0.51 |
| Music | 13 | 23 | 0.02 | [-0.12, 0.17] | 0.64 |

*Note*: Mean diff: Posterior estimate of the mean difference (in Hedges’ g) between adult and child participants. CrI: credible interval. p(ESCH> ESA): probability that the effect sizes for child participants are bigger than the effect sizes for adult participants, given the data.

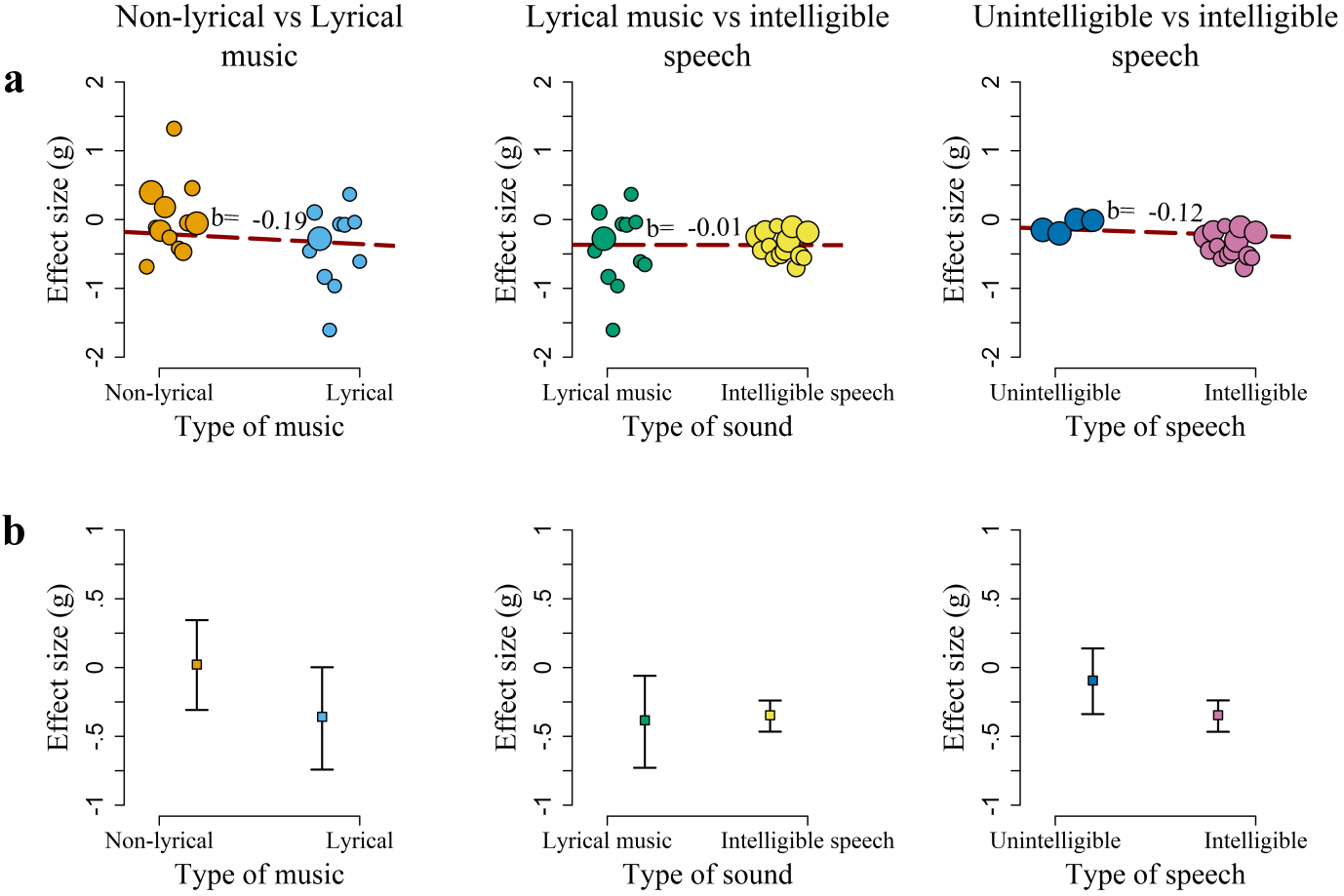
****

*Figure 4*. Forest plot for the main effect of background music (**a**), speech (**b**), and noise (**c**) on reading comprehension. Plotted are the observed (i.e. empirical) effect sizes with their 95% confidence intervals, and the posterior effect size estimate from the meta-analysis model and the corresponding 95% credible intervals. The size of squares is proportional to the weight of each study (i.e., the inverse of the within-study variance of the sampling distribution). The pooled estimate from the meta-analysis is shown by the dark red diamond at the bottom of each panel (with 95% credible intervals).

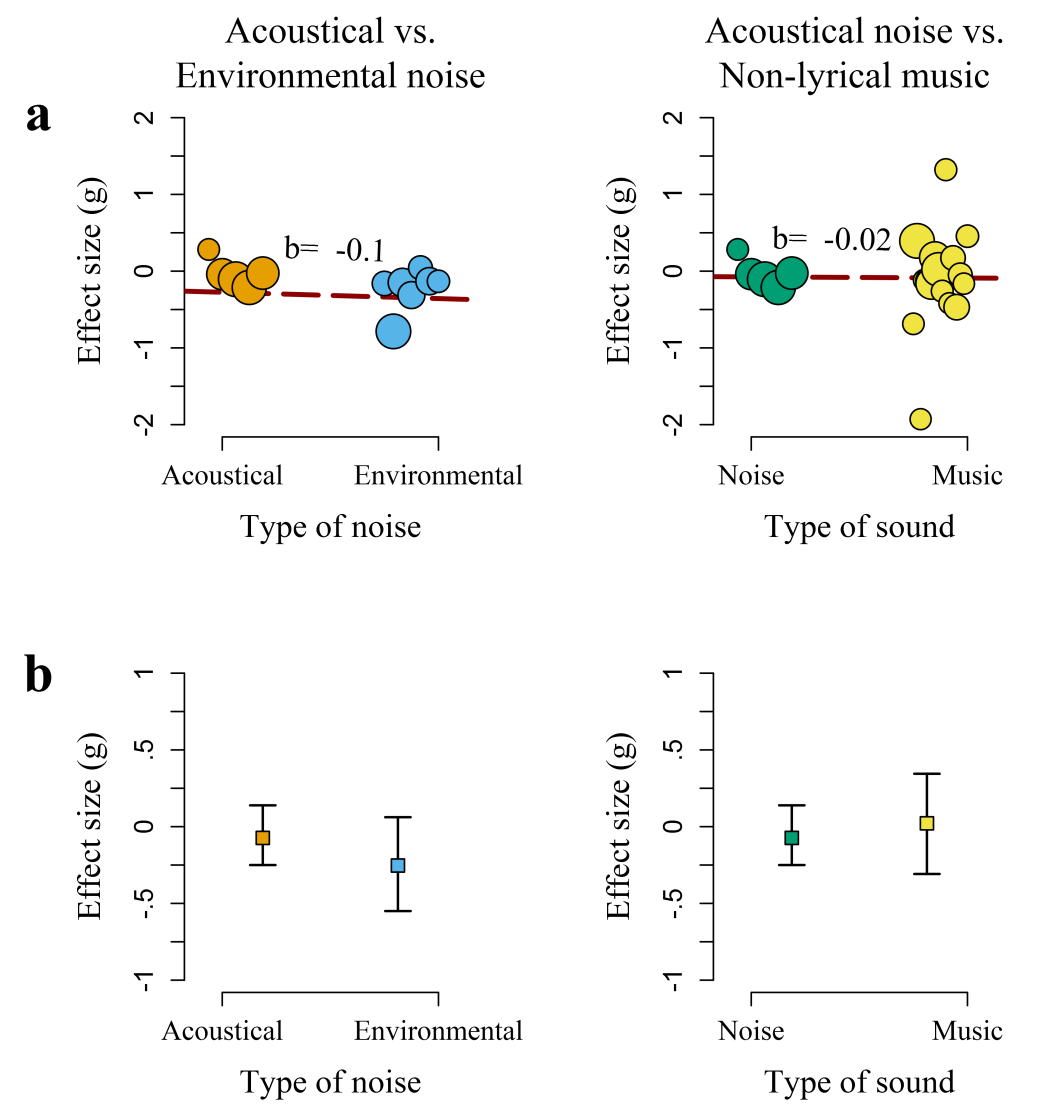
**Meta-regression**

The results from the meta-regression models testing the theoretical predictions outlined in the introduction are presented in Figures 5 and 6. Recall that the models yield a regression slope, which shows the estimated mean difference between the two groups, after adjusting for the precision of individual studies. Consistent with the semantic, but not with the phonological interference hypothesis, there was 99% probability that intelligible speech was more distracting than unintelligible speech (mean difference: g= -0.12). Additionally, in line with both the semantic and phonological interference hypotheses, there was 95% probability that lyrical music was more distracting than non-lyrical music (mean difference: g= -0.19). Surprisingly, however, intelligible speech and lyrical music did not differ between one another, and the estimated probability of a true difference was only 54% (with 50% being no difference, since the posterior probability density would lie evenly to the left and right side of 0).

Consistent with the changing-state hypothesis, there was 90% probability that environmental noise was more distracting than acoustical noise (mean difference: g= -0.10). However, there was only 55% probability of a difference between non-lyrical music and acoustical noise, thus suggesting that the two sound types did not generally differ. As Figure 6b shows, the size of both effects, as estimated by a random-effects meta-analysis, was very close to 0. This result is contrary to the predicted difference from the changing-state hypothesis.

****

*Figure 5*. Results of the meta-regression models testing the predictions of the semantic and phonological interference hypotheses. Panel **a** shows the regression slope and the observed effect size of the studies included in the analysis. The slope indicates the mean difference estimated by the meta-regression model (in terms of Hedges’s g) between the two groups. The size of circles is proportional to the weight of individual studies (inverse of the within-study variance of the sampling distribution). Panel **b** shows the posterior effect size for each group, as estimated by a random-effects meta-analysis of the simple effect. Error bars show the 95% credible intervals.



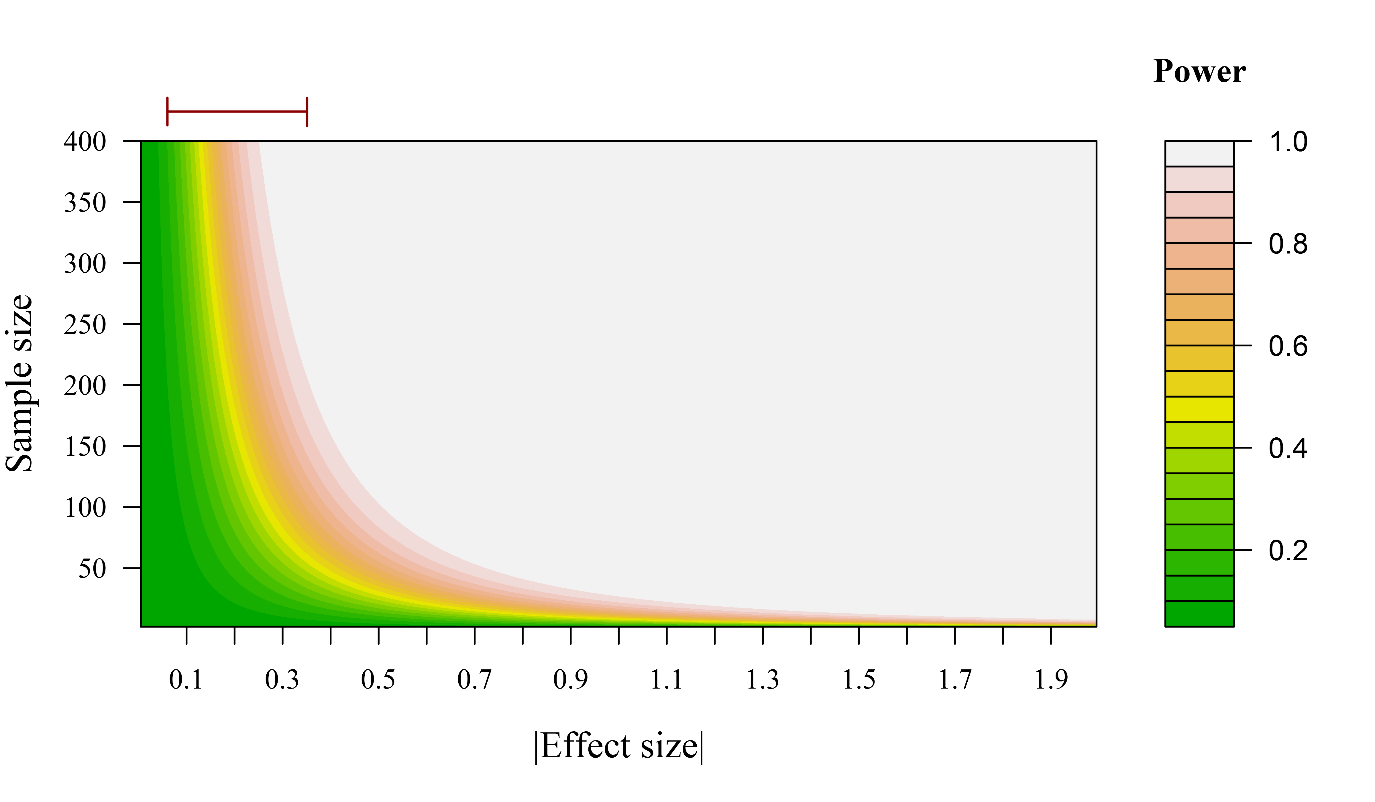
*Figure 6*. Results of the meta-regression models testing the predictions of the changing-state hypothesis. Panel **a** shows the regression slope and the observed effect size of the studies included in the analysis. The slope indicates the mean difference estimated by the meta-regression model (in terms of Hedges’s g) between the two groups. The size of circles is proportional to the weight of individual studies (inverse of the within-study variance of the sampling distribution). Panel **b** shows the posterior effect size for each group, as estimated by a random-effects meta-analysis of the simple effect. Error bars show the 95% credible intervals.

**Discussion**

The present study investigated the magnitude of auditory distraction effects during reading and how compatible these effects are with existing theories of distraction. We will first consider the overall size of the effects and then discuss their theoretical implications. The main findings from the meta-analysis can be summarized as follows. First, background speech, noise, and music all had a negative effect (indicating distraction) on reading comprehension accuracy. The magnitude of the effects was small, but highly reliable, meaning that there was very high probability that these sounds are detrimental to reading comprehension given the available evidence. Second, auditory distraction effects measured with reading comprehension did not generally differ between adults and children. Finally, background speech, noise, and music had a very small, negative effect on reading speed, and background speech and noise also had a small, negative effect on proofreading accuracy. Although both effects proved to be smaller than the ones observed in reading comprehension, there was still high probability that they were negative (>90%).

The present results provide the first comprehensive analysis of auditory distraction effects in a reading task. As the review of the literature showed, interest in this topic has a very long history that precedes the Cognitive Revolution, and indeed, most of the work on auditory distraction in other cognitive tasks. Traditionally, much of the interest in auditory distraction in reading tasks has been due to its practical implications for reading outside the psychological laboratory, such as revising for an exam or reading in the classroom. However, the inconclusive and sometimes contradictory evidence has made it difficult to derive clear conclusions until now. The present results advance our understanding of this topic by showing that external auditory input almost always comes at a cost for reading efficiency. Even though this cost was modest, especially for measures such as reading speed and proofreading accuracy, there was still relatively high probability that it exists. Therefore, the present study resolves some of the controversy highlighted in the introduction by showing that general auditory distraction effects by background noise, speech, and music almost certainly exist, but that their magnitude is small.

Given that there was very high probability that background speech, noise, and music are detrimental to reading comprehension, why have some of the previous findings been so inconsistent? One possibility is that some of the original studies may not have had sufficient power to detect the underlying effects. Figure 7 shows the relationship between sample size and statistical power for a range of effect sizes, including the ones observed in the present meta-analysis (cf. Wallisch, 2015). This is for illustrative purposes only, as statistical power is influenced not only by sample size and the magnitude of the true effect, but also by other factors, such as the reliability of the measure, missing data, sampling control and so on (Hansen & Collins, 1994). Nevertheless, as Figure 7 clearly shows, statistical power is related to sample size and generally a larger number of participants are required to achieve sufficient statistical power of detecting some of the auditory distraction effects observed in the present study. This suggests that, although most of the observed effects are negative in sign, statistical significance may not always be achieved if the underlying effect is small and the experiment is underpowered.



*Figure 7*. An illustration of the sample sizes needed to achieve different levels of statistical power for a range of realistic effect sizes. Dark red interval at the top shows the range of effect sizes observed in the present meta-analysis. Desirable levels of statistical power are depicted by warm colors. Statistical power was calculated with the “pwr” R package (Champely, 2012) and is based on an independent-samples t-test with equal groups, and an α level of 0.05 (two-tailed).

The second goal of the present study was to investigate what is it about background sounds that makes them distracting and to test what theoretical frameworks can explain the results. This is an important question as not all studies have explicitly considered the theoretical implications of their work, with some researchers taking a more applied approach of simply testing whether certain types of sounds are distracting to reading or not. More broadly, the present analysis provides a glimpse into how well readers can maintain focus on the main task (reading) while listening to a competing stream of auditory input that they try to ignore. The meta-regression results provided a few key insights into the nature of auditory distraction effects, as measured with reading comprehension accuracy.

First, lyrical music was found to be more distracting than non-lyrical music, but, surprisingly, equally as distracting as intelligible speech. Second, intelligible speech was in turn more distracting than non-intelligible speech. Finally, environmental noise was more distracting than acoustical noise, but there was no reliable difference between non-lyrical music and acoustical noise. These results provide strong support for the notion that the presence of language in background sounds is the strongest contributor to auditory distraction. Indeed, the two largest distraction effects were found for lyrical music (g= -0.35) and intelligible speech (g= -0.34).

This last finding is consistent with both the semantic interference (Martin et al., 1988) and interference-by-process (Marsh et al., 2008) accounts, which predict that either the semantic content of speech/ sung lyrics or the actual process of trying to extract their meaning can distract readers from their main task. The findings are generally not consistent with the phonological interference account for two reasons. First, it predicts that all speech sounds should be equally disruptive because they would all gain access to the phonological store; however, intelligible speech was reliably more distracting than non-intelligible speech. Additionally, background noise, which would not gain access to the phonological store, was also found to cause distraction. Finally, the results are only in part consistent with the changing-state account (Jones et al., 1992), which predicts that sounds with greater acoustical variation would cause greater distraction. This is because environmental noise was more distracting than acoustical noise (consistent with the theory), but non-lyrical music was not more distracting than acoustical noise (not consistent with the theory). In both cases, environmental noise and non-lyrical music exhibit more acoustic variation.

What type of theoretical framework could account for the present results? Clearly, none of the theories considered so far can account for all the findings. This is not necessarily a limitation because, as noted previously, not all theories were originally designed to account for distraction effects in a reading task. Therefore, it is more useful to consider a hypothetical model that can explain the data from reading tasks. One such framework could be a two-component model in which noise and speech influence reading through separate processes. In the first component, background noise would cause a small, but negative decrement in comprehension. The present data does not make it possible to explain why this disruption by noise occurs and more research is needed to understand this mechanism. The second component would cause quantitively greater decrements in comprehension from intelligible speech. Recent evidence suggests that the cognitive process of trying to analyze the meaning of the speech may be enough to cause distraction (Hyönä & Ekholm, 2016). Whether the semantic content and semantic representation of the speech sound are processed and cause additional distraction is an open question that needs to be explored in future research. This second component would also account for the effect of background music. This is because the present results suggest that distraction due to music is effectively reduced to distraction from the sung lyrics, since music without lyrics was not distracting.

**Limitations**

While meta-regression is a very useful tool for testing how auditory distraction differs between background sounds or age groups, the present results are only observational in nature (Thompson & Higgins, 2002). Therefore, direct evidence from laboratory experiments is required to verify these results. Nevertheless, as no attempt has been made so far to answer these questions with all the available evidence, we anticipate that our findings will prove to be very useful in guiding future experimental research and advancing our understanding of how auditory distraction during reading occurs.

Additionally, some of the meta-regression analyses were based on a smaller number of studies. However, this is not necessarily a limitation in the Bayesian approach that we have adopted here because the results reflect our best understanding of auditory distraction effects given the currently available data. Once more data is available, the present results can be easily updated via Bayes’ theorem, which will make it possible to arrive at an even more precise estimate of the effects.

**Future Directions**

The present study grouped background sounds into broad categories, such as noise, speech, or music. However, real-word sounds that readers are routinely exposed to do not always belong to only one of these categories. Rather, different sounds may be present at the same time, such as music playing from the TV and environmental noise from nearby traffic. Currently, there is limited understanding of how different types of sounds may interact to increase or decrease distraction. For example, there is some evidence that acoustical noise, when intermixed with background speech, can reduce the negative impact of the speech sound by reducing its intelligibility (Haapakangas et al., 2011; Hongisto, 2005; Venetjoki et al., 2006). Therefore, more research is needed to investigate sounds that are more complex and thus more realistic of auditory distraction in the real world.

Another question that deserves more attention is how auditory distraction may differ between age groups. Studies with adults and children have usually been done in isolation, which makes it challenging to assess how these groups differ under the same experimental conditions. The present meta-regression analyses are the closest (and arguably, the only possible way) of addressing this question with the currently available data. However, experiments directly comparing adults and children are needed to make firm conclusions. Traditionally, a lot of research has focused on large-scale epidemiological studies of chronic exposure to noise in schools such as the RANCH (Stansfeld et al., 2005) and west London studies (Haines et al., 2001a; 2001b). As such, it is surprising how little is known about the effect of experimental exposure to noise on reading in children. Eye-movement recordings may be particularly helpful in studying this topic as they can reveal subtle auditory distraction effects that may not show up in behavioral measures such as comprehension accuracy (Cauchard et al., 2012; Hyönä & Ekholm, 2016; Yan et al., 2017). Longitudinal studies of reading development have already made successful use of eye-tracking to study processes such as the development of the perceptual span (Sperlich, Meixner, & Laubrock, 2016), and this method also holds promise in understanding how children’s susceptibility to distraction may change during the school years and beyond.

Eye-tracking technology and ERP recordings (or a combination of both) are useful methods because they can provide rich data about the time course of auditory distraction effects during reading. We anticipate that this type of evidence will be crucial for gaining a better understanding of when and how these effects occur, and what their theoretical nature is. The field of eye-movements during reading (in silence) has already seen the successful development of advanced computational models such as the E-Z Reader (Reichle, Pollatsek, Fisher, & Rayner, 1998) and SWIFT (Engbert et al., 2005), which can simulate many empirical findings. Similarly, a more precise quantification of the time course of auditory distraction effects can move the field forward by making it possible to build computational models that can simulate these processes and to generate new predictions.

**Conclusion**

Auditory distraction during reading has been a topic of interest for the last 80 years and, as the surge of recent publications shows, it is likely to continue to be an active area of research in the future. The present study was the first attempt to make a comprehensive statistical synthesis of auditory distraction effects in a reading task. The results showed that background noise, speech and music are almost always distracting, even if the distraction effects are often small in size. Sounds that contain intelligible language (e.g., speech, lyrical music) were particularly distracting, most likely due to their semantic properties that interfere with processing the written text. The present findings also have some practical implications. For example, they suggest that listening to instrumental music while reading would not affect the comprehension of the text, whereas listening to lyrical music would do so. Additionally, readers exposed to background noise would likely incur a cost in terms of reduced comprehension, even if this cost is very small. Finally, the recent interest in measuring eye-movements during reading heralds the emergence of a new sub-field that may give an even more precise understanding of how and when auditory distraction occurs.

**Acknowledgments**

M.R.V was supported by a PhD studentship from Bournemouth University on auditory distraction effects during reading. The funding body had no role in the design, analysis, the writing of this paper, or decisions to publish. The authors declare no conflicts of interest with respect to this paper. All data analysis scripts and materials are available during the peer-review process at: [https://1drv.ms/f/s!AmDg\_dyhaECdyhtsDBqGRxsSSw83](https://1drv.ms/f/s!AmDg_dyhaECdyhtsDBqGRxsSSw83%20) and will be made publicly available at GitHub and the Open Science Framework upon the publication of this meta-analysis.

Appendix A

**Study Inclusion Criteria**

* The study investigated the effect of experimental exposure to background noise, speech, or music in a reading/ proofreading task.
* Only studies investigating the immediate effect of background sounds on reading/ proofreading were included. Experiments that studied the effect of long-term exposure to music as an intervention for reading were excluded. Studies that investigated the effects of chronic exposure to traffic noise were also excluded.
* The study contained a condition of reading in silence. This served as the baseline to which background sound manipulations were compared. Studies without a silence baseline were excluded.
* The study had appropriate randomization and counter-balancing of the sound conditions.
* Participants were native speakers of the language in which they were reading.
* The study was done with healthy, typically-developing participants (either children or adults).
* The external environment or any additional manipulations did not introduce confounds.
* Participants were not tested on the contents of the sound that they were listening to (e.g. speech).
* The assessment task emphasized comprehension of the text rather than reproducing the text from memory as accurately as possible.
* The comprehension assessment did not occur too long after the reading phase (usually within 10-15 minutes).
* The comprehension assessment was done in silence.

Appendix B

Table B1

*A Summary of the Studies and Their Effect Sizes That Were Included in the Meta-analysis*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | Study | NC | NE | Sample | Design | DV | Sound | Sound type | dB(A) | g | var |
| 1 | Sörqvist et al. 2010 | 40 | | adults | within | RC | speech | native | 72.5 | -0.24 | 0.01 |
| 1 | Sörqvist et al. 2010 | 40 | | adults | within | RS | speech | native | 72.5 | -0.05 | 0.01 |
| 2 | Ljung et al. 2009 | 70 | 50 | children | between | RC | noise | traffic | 62 | -0.16 | 0.03 |
| 2 | Ljung et al. 2009 | 70 | 50 | children | between | RS | noise | traffic | 62 | 0.71 | 0.04 |
| 2 | Ljung et al. 2009 | 70 | 66 | children | between | RC | speech | babble | 62 | 0.17 | 0.03 |
| 2 | Ljung et al. 2009 | 70 | 66 | children | between | RS | speech | babble | 62 | 0.21 | 0.03 |
| 3 | Fogelson 1973 | 14 | 14 | children | between | RC | music | pop | - | -0.42 | 0.14 |
| 4 | Tucker & Bushm. 1991 | 75 | 76 | adults | between | RC | music | rock & roll | 80 | 0.00 | 0.03 |
| 5 | Daoussis & McK. 1986 | 24 | 24 | adults | between | RC | music | rock | 50 | -0.52 | 0.08 |
| 6 | Etaugh & Michals 1975 | 32 | | adults | within | RC | music | preferred | - | -0.08 | 0.02 |
| 7 | Etaugh & Ptasnik 1982 | 20 | 20 | adults | between | RC | music | preferred | - | -0.74 | 0.10 |
| 8 | Kiger 1989 | 18 | 18 | children | between | RC | music | low load | - | 3.50 | 0.28 |
| 8 | Kiger 1989 | 18 | 18 | children | between | RC | music | high load | - | -0.69 | 0.11 |
| 9 | Miller & Schyb 1989 | 49 | 49 | adults | between | RC | music | classical | 47.5 | 0.11 | 0.04 |
| 9 | Miller & Schyb 1989 | 49 | 49 | adults | between | RC | music | pop | 47.5 | 0.23 | 0.04 |
| 9 | Miller & Schyb 1989 | 49 | 49 | adults | between | RC | music | vocal | 47.5 | -0.46 | 0.04 |
| 10 | Doyle & Furnham 2012 | 56 | | adults | within | RC | music | vocal | - | 0.10 | 0.01 |
| 11 | Anderson & Fuller 2010 | 334 | | children | within | RC | music | lyrical | 75 | -0.28 | 0.00 |
| 12 | Furnham & Strbac 2002 | 76 | | children | within | RC | noise | office | - | -0.78 | 0.01 |
| 12 | Furnham & Strbac 2002 | 76 | | children | within | RC | music | vocal/unfam. | - | -0.83 | 0.01 |
| 13 | Mullikin & Henk 1985 | 45 | | children | within | RC | music | classical | - | 0.39 | 0.01 |
| 13 | Mullikin & Henk 1985 | 45 | | children | within | RC | music | rock | - | -0.33 | 0.01 |
| 14 | Avila et al. 2011 | 19 | 20 | children | between | RC | music | vocal/ familiar | - | -1.61 | 0.13 |
| 14 | Avila et al. 2011 | 19 | 19 | children | between | RC | music | Instr./ familiar | - | -1.93 | 0.15 |
| 15 | Freeburne & Fleis. 1952 | 43 | 46 | adults | between | RC | music | classical | - | 0.02 | 0.04 |
| 15 | Freeburne & Fleis. 1952 | 43 | 46 | adults | between | RS | music | classical | - | -0.35 | 0.04 |
| 15 | Freeburne & Fleis. 1952 | 43 | 42 | adults | between | RC | music | pop | - | 0.04 | 0.05 |
| 15 | Freeburne & Fleis. 1952 | 43 | 42 | adults | between | RS | music | pop | - | -0.40 | 0.05 |
| 15 | Freeburne & Fleis. 1952 | 43 | 40 | adults | between | RC | music | semi-classical | - | -0.08 | 0.05 |
| 15 | Freeburne & Fleis. 1952 | 43 | 40 | adults | between | RS | music | semi-classical | - | -0.36 | 0.05 |
| 15 | Freeburne & Fleis. 1952 | 43 | 37 | adults | between | RC | music | jazz | - | -0.17 | 0.05 |
| 15 | Freeburne & Fleis. 1952 | 43 | 37 | adults | between | RS | music | jazz | - | -0.61 | 0.05 |
| 16 | Fendrick 1937 | 61 | 62 | adults | between | RC | music | semi-classical | - | -0.47 | 0.03 |
| 17 | Henderson et al. 1945 | 19 | 17 | adults | between | RC | music | classical | - | -0.12 | 0.11 |
| 17 | Henderson et al. 1945 | 19 | 14 | adults | between | RC | music | pop | - | -1.07 | 0.14 |
| 18 | Miller 2014 | 13 | 13 | adults | between | RC | music | classical lyrical | - | -0.84 | 0.16 |
| 18 | Miller 2014 | 13 | 17 | adults | between | RC | music | classical instr. | - | 0.13 | 0.13 |
| 18 | Miller 2014 | 13 | 11 | adults | between | RC | music | rock lyrical | - | -0.38 | 0.16 |
| 18 | Miller 2014 | 13 | 18 | adults | between | RC | music | rock instr. | - | -0.45 | 0.13 |
| 19 | Furnham & Allass 1999 | 16 | 16 | adults | between | RC | music | complex | - | -0.02 | 0.12 |

Table B1 (continued)

*A Summary of the Studies and Their Effect Sizes That Were Included in the Meta-analysis*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | Study | NC | NE | Sample | Design | DV | Sound | Sound type | dB(A) | g | var |
| 19 | Furnham & Allass 1999 | 16 | 16 | adults | between | RC | music | simple | - | -0.05 | 0.12 |
| 20 | Furnham & Bradl. 1997 | 10 | 10 | adults | between | RC | music | pop | - | -0.97 | 0.21 |
| 21 | Furnham at al. 1999 | 43 | 49 | children | between | RC | music | instrumental | - | -0.12 | 0.04 |
| 21 | Furnham at al. 1999 | 43 | 47 | children | between | RC | music | vocal | - | -0.07 | 0.04 |
| 22 | Perham & Currie 2014 | 30 | | adults | within | RC | music | disliked lyrical | 70 | -0.71 | 0.02 |
| 22 | Perham & Currie 2014 | 30 | | adults | within | RC | music | non-lyrical | 70 | -0.16 | 0.02 |
| 22 | Perham & Currie 2014 | 30 | | adults | within | RC | music | liked lyrical | 70 | -0.60 | 0.02 |
| 23 | Kelly 1994 | 13 | 12 | adults | between | RC | music | pop | 65 | -0.74 | 0.16 |
| 24 | Dove 2009 | 28 | 28 | adults | between | RC | music | sedat. classical | 62.5 | 0.10 | 0.07 |
| 24 | Dove 2009 | 28 | 28 | adults | between | RC | music | stimul. classical | 62.5 | 0.81 | 0.08 |
| 24 | Dove 2009 | 28 | 28 | adults | between | RS | music | sedat. classical | 62.5 | -0.07 | 0.07 |
| 24 | Dove 2009 | 28 | 28 | adults | between | RS | music | stimul. classical | 62.5 | -0.51 | 0.07 |
| 25 | Furnham et al. 1994 | 20 | | adults | within | RC | speech | TV drama | - | -0.45 | 0.03 |
| 26 | Johansson 1983 | 22 | 22 | children | between | RC | noise | continuous | 51 | 0.28 | 0.09 |
| 26 | Johansson 1983 | 22 | 22 | children | between | RC | noise | intermittent | 67.4 | 0.21 | 0.09 |
| 27 | Halin 2016 | 28 | | adults | within | RC | speech | native (easy) | 60 | -0.89 | 0.03 |
| 27 | Halin 2016 | 28 | | adults | within | RC | speech | native (diff) | 60 | -0.16 | 0.02 |
| 27 | Halin 2016 | 28 | | adults | within | RC | noise | traffic (easy) | 60 | -0.35 | 0.02 |
| 27 | Halin 2016 | 28 | | adults | within | RC | noise | traffic (diff) | 60 | -0.01 | 0.02 |
| 27 | Halin 2016 | 28 | | adults | within | RC | noise | aircraft (easy) | 60 | -0.23 | 0.02 |
| 27 | Halin 2016 | 28 | | adults | within | RC | noise | aircraft (diff) | 60 | -0.01 | 0.02 |
| 28 | Smith-J. & Klein 2009 | 54 | | adults | within | PR | speech | native | 65 | -0.04 | 0.01 |
| 29 | Cauchard et al. 2012 | 30 | | adults | within | RC | music | instrumental | 65 | 0.18 | 0.02 |
| 29 | Cauchard et al. 2012 | 30 | | adults | within | RC | speech | native | 65 | -0.17 | 0.02 |
| 29 | Cauchard et al. 2012 | 30 | | adults | within | RS | music | instrumental | 65 | 0.01 | 0.02 |
| 29 | Cauchard et al. 2012 | 30 | | adults | within | RS | speech | native | 65 | -0.20 | 0.02 |
| 30 | Johansson et al. 2012 | 24 | | adults | within | RC | music | preferred | 65 | -0.34 | 0.02 |
| 30 | Johansson et al. 2012 | 24 | | adults | within | RC | music | non-preferred | 65 | -0.67 | 0.03 |
| 30 | Johansson et al. 2012 | 24 | | adults | within | RC | noise | cafe | 65 | -0.31 | 0.02 |
| 30 | Johansson et al. 2012 | 24 | | adults | within | RS | music | preferred | 65 | -0.14 | 0.02 |
| 30 | Johansson et al. 2012 | 24 | | adults | within | RS | music | non-preferred | 65 | -0.10 | 0.02 |
| 30 | Johansson et al. 2012 | 24 | | adults | within | RS | noise | cafe | 65 | -0.07 | 0.02 |
| 31 | Weinstein 1974 | 15 | 18 | adults | between | PR† | noise | teletype | 70 | -0.56 | 0.12 |
| 31 | Weinstein 1974 | 15 | 18 | adults | between | PR‡ | noise | teletype | 70 | -1.26 | 0.14 |
| 32 | Weinstein 1977 | 29 | | adults | within | PR† | speech | native | 68 | -0.03 | 0.02 |
| 32 | Weinstein 1977 | 29 | | adults | within | PR‡ | speech | native | 68 | -0.29 | 0.02 |
| 33 | Martin et al. 1988, E1 | 36 | | adults | within | RC | speech | native | 82 | -0.20 | 0.01 |
| 33 | Martin et al. 1988, E1 | 36 | | adults | within | RC | speech | random | 82 | -0.18 | 0.01 |
| 33 | Martin et al. 1988, E1 | 36 | | adults | within | RC | music | instrumental | 82 | 0.00 | 0.01 |
| 33 | Martin et al. 1988, E1 | 36 | | adults | within | RC | music | random tones | 82 | -0.11 | 0.01 |
| 33 | Martin et al. 1988, E1 | 36 | | adults | within | RC | noise | white | 82 | -0.04 | 0.01 |
| 34 | Martin et al. 1988, E2 | 36 | | adults | within | RC | music | instrumental | 82 | 0.02 | 0.01 |
| 34 | Martin et al. 1988, E2 | 36 | | adults | within | RC | music | lyrical | 82 | -0.08 | 0.01 |
| 35 | Martin et al. 1988, E4 | 48 | | adults | within | RC | noise | white | 82 | -0.11 | 0.01 |

Table B1 (continued)

*A Summary of the Studies and Their Effect Sizes That Were Included in the Meta-analysis*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | Study | NC | NE | Sample | Design | DV | Sound | Sound type | dB(A) | g | var |
| 35 | Martin et al. 1988, E4 | 48 | | adults | within | RC | speech | native | 82 | -0.31 | 0.01 |
| 35 | Martin et al. 1988, E4 | 48 | | adults | within | RC | speech | foreign | 82 | -0.15 | 0.01 |
| 36 | Martin et al. 1988, E5 | 48 | | adults | within | RC | noise | white | 82 | -0.21 | 0.01 |
| 36 | Martin et al. 1988, E5 | 48 | | adults | within | RC | speech | non-word | 82 | -0.20 | 0.01 |
| 36 | Martin et al. 1988, E5 | 48 | | adults | within | RC | speech | random words | 82 | -0.33 | 0.01 |
| 37 | Cool et al. 1994, E2 | 9 | | children | within | RS | music | radio/ generic | - | 0.13 | 0.05 |
| 37 | Cool et al. 1994, E2 | 9 | | children | within | RS | speech | movies | - | 0.20 | 0.05 |
| 37 | Cool et al. 1994, E2 | 9 | | children | within | RC | music | radio/ generic | - | -0.12 | 0.05 |
| 37 | Cool et al. 1994, E2 | 9 | | children | within | RC | speech | movies | - | -0.22 | 0.05 |
| 38 | Mitchell 1949 | 91 | | children | within | RTS | music | radio/ generic | - | -0.01 | 0.01 |
| 39 | Armstrong et al. 1991 | 33 | 30 | adults | between | RTS | speech | TV ads | - | -0.63 | 0.07 |
| 39 | Armstrong et al. 1991 | 33 | 32 | adults | between | RTS | speech | TV drama | - | -0.48 | 0.06 |
| 40 | Pool et al. 2000, E1 | 30 | 30 | children | between | RC | speech | TV soap opera | 60 | -0.38 | 0.07 |
| 40 | Pool et al. 2000, E1 | 30 | 30 | children | between | RC | music | TV music | 60 | -0.21 | 0.07 |
| 41 | Pool et al. 2000, E2 | 48 | 24 | children | between | RC | speech | TV soap opera | 60 | -0.57 | 0.06 |
| 41 | Pool et al. 2000, E2 | 48 | 24 | children | between | RC | music | TV music | 60 | -0.10 | 0.06 |
| 42 | Dockrell & Shield 2006 | 52 | 52 | children | between | RTS | noise | babble | 65 | -0.49 | 0.04 |
| 42 | Dockrell & Shield 2006 | 52 | 52 | children | between | RTS | noise | babble+environ. | 65 | 0.58 | 0.04 |
| 43 | Hyönä & Ekh. 2016, E1 | 42 | | adults | within | RC | speech | native | 82.5 | -0.17 | 0.01 |
| 43 | Hyönä & Ekh. 2016, E1 | 42 | | adults | within | RC | speech | foreign | 82.5 | 0.00 | 0.01 |
| 43 | Hyönä & Ekh. 2016, E1 | 42 | | adults | within | RS | speech | native | 82.5 | -0.02 | 0.01 |
| 43 | Hyönä & Ekh. 2016, E1 | 42 | | adults | within | RS | speech | foreign | 82.5 | 0.06 | 0.01 |
| 44 | Hyönä & Ekh. 2016, E2 | 36 | | adults | within | RS | speech | scrambl.-differ. | 82.5 | -0.15 | 0.01 |
| 44 | Hyönä & Ekh. 2016, E2 | 36 | | adults | within | RS | speech | scrambl.-same | 82.5 | -0.18 | 0.01 |
| 45 | Hyönä & Ekh. 2016, E3 | 35 | | adults | within | RS | speech | native | 82.5 | -0.13 | 0.01 |
| 45 | Hyönä & Ekh. 2016, E3 | 35 | | adults | within | RS | speech | scrambled | 82.5 | -0.20 | 0.01 |
| 46 | Hyönä & Ekh. 2016, E4 | 36 | | adults | within | RS | speech | scrambled-sem. | 82.5 | -0.11 | 0.01 |
| 46 | Hyönä & Ekh. 2016, E4 | 36 | | adults | within | RS | speech | scrm-syn+sem | 82.5 | -0.14 | 0.01 |
| 47 | Armstrong & Chng 2000 | 19 | 20 | adults | between | RC | speech | native | - | -0.09 | 0.10 |
| 48 | Madsen 1987, E1 | 50 | 50 | adults | between | RC | music | various | 75 | -0.10 | 0.04 |
| 49 | Sörqvist 2010, E1a | 23 | | children | within | RC | noise | aircraft | 57.5 | -0.13 | 0.02 |
| 50 | Sörqvist 2010, E1b | 23 | | children | within | RC | speech | native | 57.5 | -0.51 | 0.03 |
| 51 | Sörqvist et al. 2010, E1 | 24 | | adults | within | RC | speech | native | 65 | -0.46 | 0.02 |
| 52 | Sörqvist et al. 2010, E2 | 42 | | adults | within | RC | speech | native | 65 | -0.30 | 0.01 |
| 53 | Halin et al. 2014 | 32 | | adults | within | RC | speech | native | 65 | -0.10 | 0.02 |
| 54 | Halin et al. 2014, E1 | 31 | | adults | within | PR‡ | speech | native | 65 | -0.09 | 0.02 |
| 54 | Halin et al. 2014, E1 | 31 | | adults | within | PR† | speech | native | 65 | 0.20 | 0.02 |
| 55 | Halin et al. 2014, E2 | 29 | | adults | within | PR‡ | speech | native | 65 | -0.13 | 0.02 |
| 55 | Halin et al. 2014, E2 | 29 | | adults | within | PR† | speech | native | 65 | 0.11 | 0.02 |
| 56 | Haapakangas et al. 2011 | 54 | | adults | within | PR‡ | speech | native | 48 | -0.09 | 0.01 |
| 56 | Haapakangas et al. 2011 | 54 | | adults | within | PR† | speech | native | 48 | -0.11 | 0.01 |
| 57 | Baker & Madell 1965 | 24 | | adults | within | RC | speech | native | - | -0.70 | 0.03 |
| 58 | Vasilev et al. n.d. | 40 | | adults | within | RC | noise | speech-spectr. | 60 | -0.03 | 0.01 |
| 58 | Vasilev et al. n.d. | 40 | | adults | within | RC | speech | foreign | 60 | -0.01 | 0.01 |

Table B1 (continued)

*A Summary of the Studies and Their Effect Sizes That Were Included in the Meta-analysis*

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ID | Study | NC | NE | Sample | Design | DV | Sound | Sound type | dB(A) | g | var |
| 58 | Vasilev et al. n.d. | 40 | | adults | within | RC | speech | native | 60 | -0.07 | 0.01 |
| 58 | Vasilev et al. n.d. | 40 | | adults | within | RS | noise | speech-spectr. | 60 | 0.04 | 0.01 |
| 58 | Vasilev et al. n.d. | 40 | | adults | within | RS | speech | foreign | 60 | -0.06 | 0.01 |
| 58 | Vasilev et al. n.d. | 40 | | adults | within | RS | speech | native | 60 | -0.15 | 0.01 |
| 59 | Falcon 2017, Sample 1 | 22 | 20 | children | between | RC | music | classical | 55 | -0.26 | 0.09 |
| 60 | Falcon 2017, Sample 2 | 25 | 28 | children | between | RC | music | classical | 55 | 1.32 | 0.09 |
| 61 | Ahuja 2016 | 20 | | adults | within | RC | music | liked | 60 | -0.71 | 0.04 |
| 61 | Ahuja 2016 | 20 | | adults | within | RC | music | disliked | 60 | -0.08 | 0.02 |
| 62 | Kou et al. 2017 | 31 | 29 | adults | between | RC | music | pop (vocal) | 65 | 0.37 | 0.07 |
| 62 | Kou et al. 2017 | 31 | 32 | adults | between | RC | noise | office | 65 | -0.13 | 0.06 |
| 63 | Sukowski et al. 2016 | 12 | | adults | within | PR | speech | native | 59.5 | -0.62 | 0.05 |
| 64 | Yan et al. 2017 | 42 | | adults | within | RS | speech | native | 62 | -0.16 | 0.01 |
| 64 | Yan et al. 2017 | 42 | | adults | within | RS | speech | meaningless | 62 | 0.06 | 0.01 |
| 65 | Gillis 2016 | 24 | 47 | adults | between | RC | music | various | - | 0.07 | 0.06 |

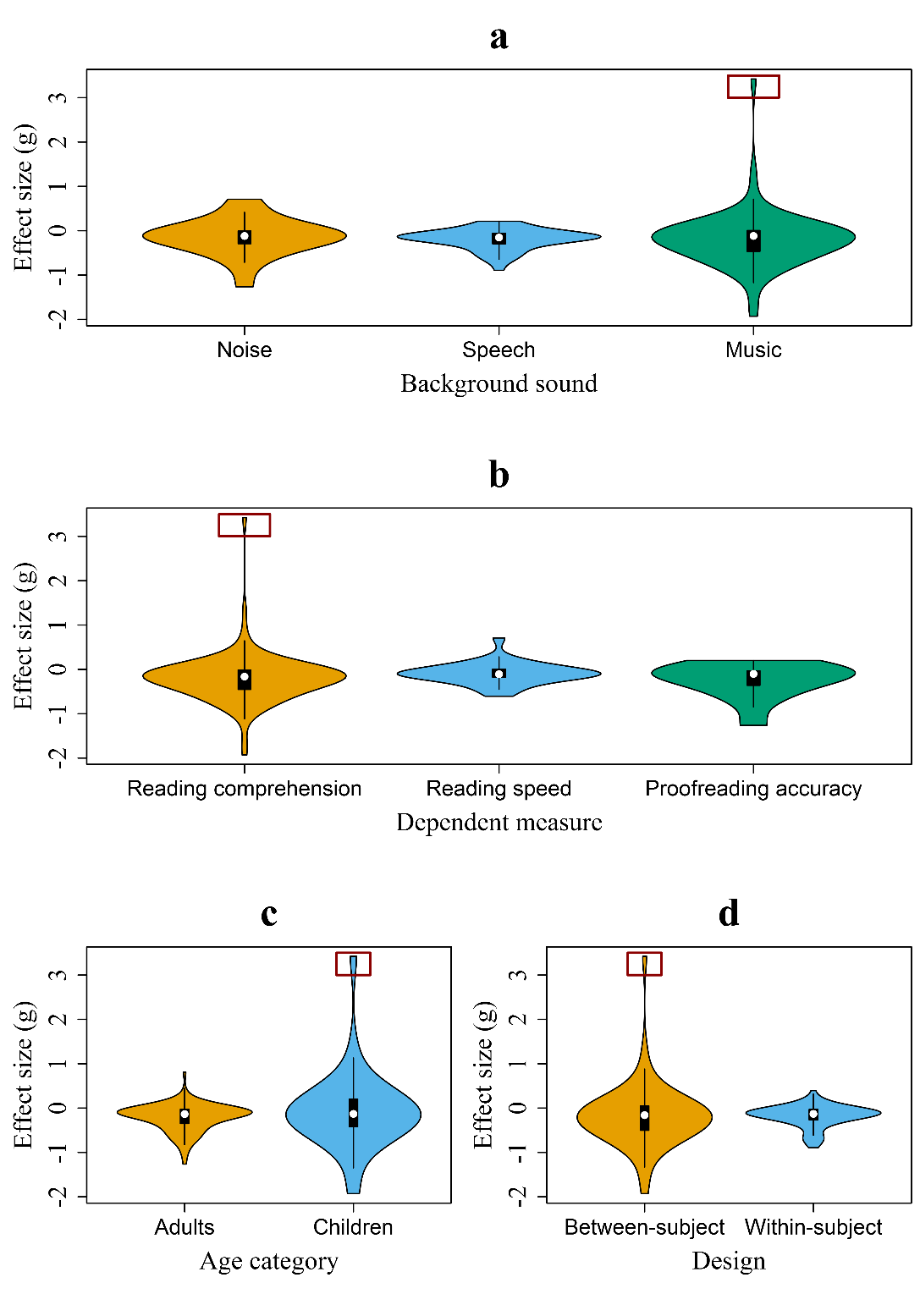
*Note*: NC: number of participants in the control (silence) condition. NE: number of participants in the experimental (sound) condition. RC: Reading comprehension. RS: reading speed. RTS: Reading test score. PR: Proofreading accuracy. ES: Effect size in Hedges’ g.

† Non-contextual errors (proofreading accuracy)

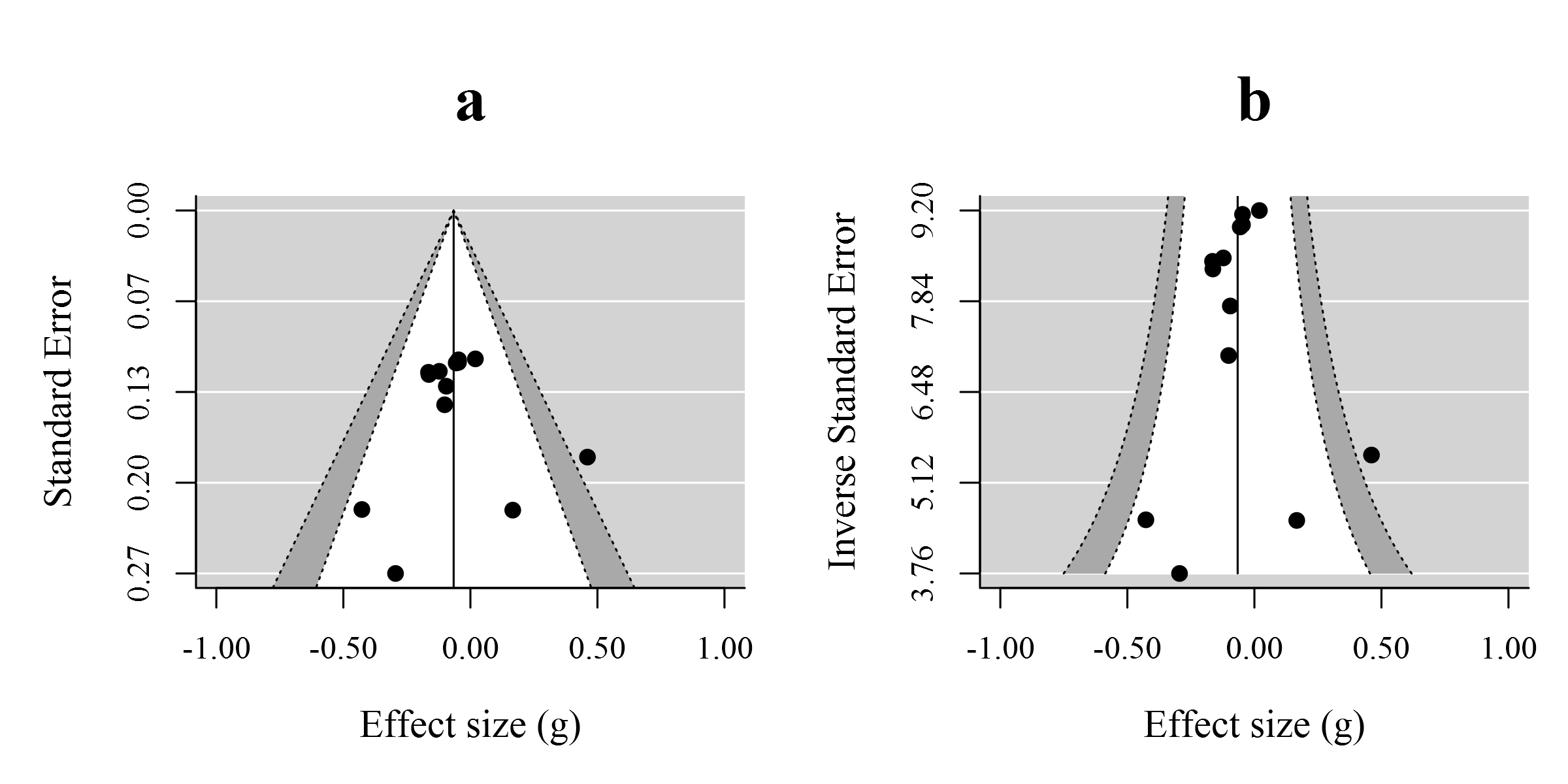
‡Contextual errors (proofreading accuracy)

Supplemental Material

**Visualization of Effect Sizes**

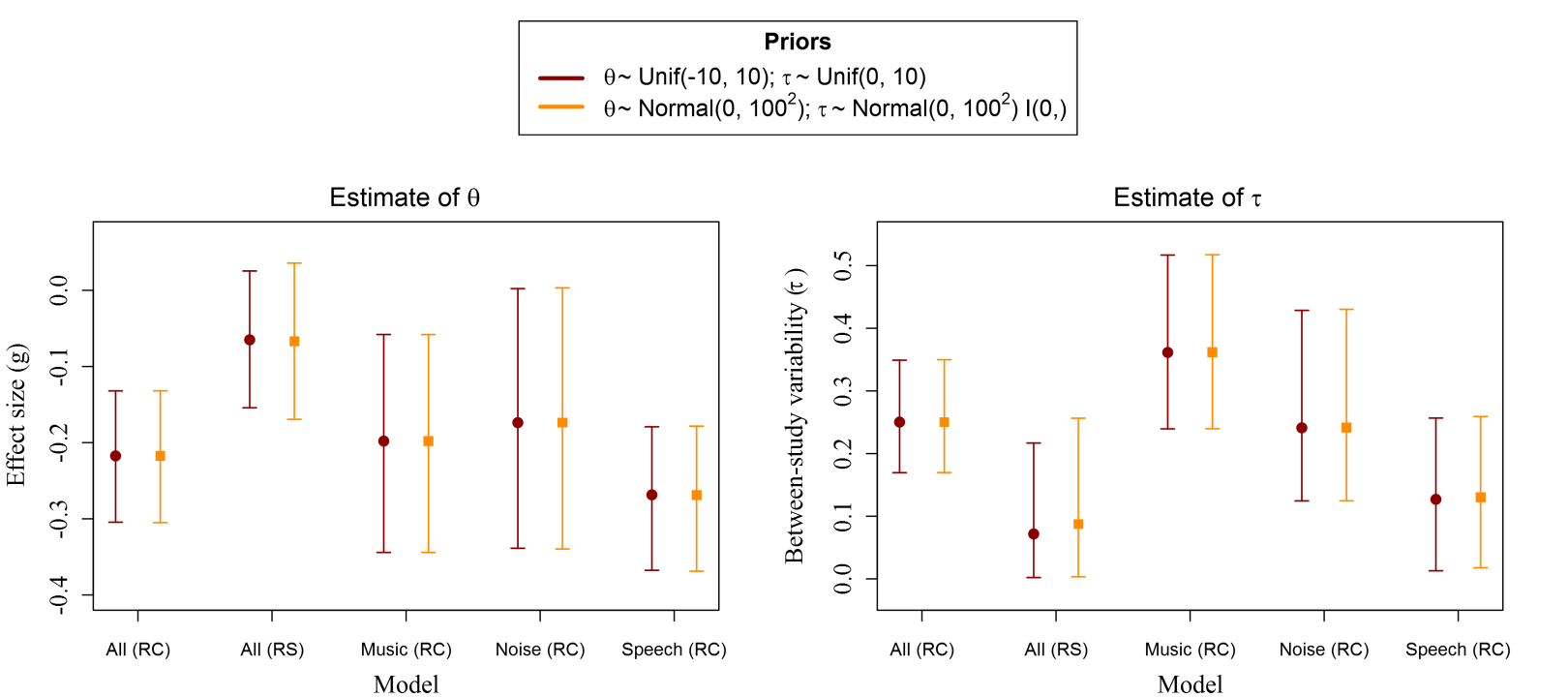
****

*Figure S1*. Box plots and probability densities of the effect sizes included in the meta-analysis. Breakdown shown by: background sound type (panel **a**), dependent measure (panel **b**), age of participants (panel **c**), and study design (panel **d**; computed after transforming within-subject effect sizes with Morris & DeShon’s, 2002, formula 11). Red rectangle shows one effect size that was excluded as an outlier.

****

*Figure S2*. Funnel plot of reading speed effect sizes plotted against their standard error (**a**) and the inverse of their standard error (**b**).

**Prior Sensitivity Analysis**



*Figure S3*. Sensitivity analysis with different priors on the θ and τ parameters for the main meta-analysis results. Uniform priors (dark red) were used in the analysis reported in the main paper. The results show that using diffuse normal priors (orange) did not change the main results reported in the paper. All: all studies. RC: reading comprehension. RS: reading speed.

**Robustness Check (Leave-one-out Method)**

Robustness analyses were carried out by using the leave-one-out method (see Greenhouse & Iyengar, 2009) to ensure that individual studies did not have very large influence on the effect size estimate. In this method, the meta-analysis is repeated by omitting one different study each time. The summary statistics of the results are reported in Table S1. Overall, the effect sizes changed little by omitting each one of the studies. The effect size range for proofreading accuracy was slightly bigger, but this was likely due to the small number of studies in this analysis (7). This greater variability is not unusual for random-effects meta-analysis with few studies because there is more uncertainty in estimating the between-study variance in the model (cf. Welton et al., 2012).

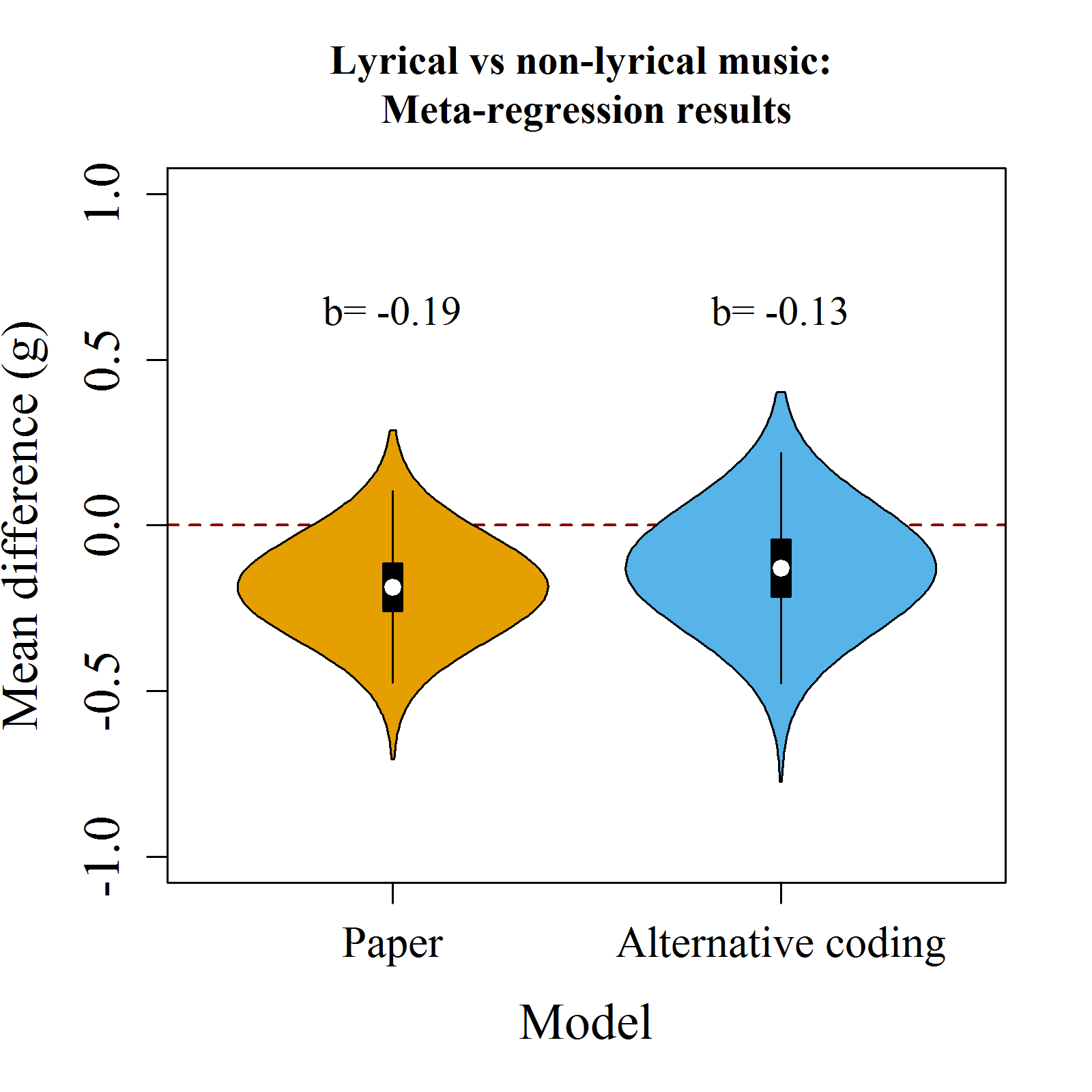
Table S1

*Summary of the Robustness Analysis Using the Leave-one-out Method*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Analysis | ES reported in main paper | Leave-one-out | | | |
| Mean ES | SD of ES | Min ES | Max ES |
| Reading comprehension |  | | | | |
| All sounds | -0.21 | -0.21 | 0.006 | -0.23 | -0.19 |
| Noise | -0.17 | -0.17 | 0.02 | -0.19 | -0.11 |
| Speech | -0.26 | -0.26 | 0.01 | -0.28 | -0.24 |
| Music | -0.19 | -0.19 | 0.01 | -0.22 | -0.16 |
| Reading speed | -0.06 | -0.06 | 0.01 | -0.09 | -0.05 |
| Proofreading accuracy | -0.14 | -0.15 | 0.04 | -0.19 | -0.08 |

**Lyrical vs Non-Lyrical Meta-regression: Robustness Check**

Some of the included studies had effect sizes for both lyrical and non-lyrical music. In order to avoid stochastical dependency among the effect sizes included in this meta-regression analysis, it was necessary to ensure that each study contributed one and only one effect size to either the “lyrical” or “non-lyrical” group. In the paper, the effect sizes were divided into the two groups in a way that maximized the number of effect sizes per group. This is because meta-regressions with larger and more balanced number of observations per group would generally yield more informative results. However, to check for subjectivity in this decision, we did the opposite division of the effect size to compare the results (this will be referred to as the “alternative coding”). The resulting posterior distributions of the mean difference are plotted in Figure S4. As it can be seen, the estimated mean difference was slightly smaller. In the model reported in the paper, there was 95% probability that lyrical music was more distracting than non-lyrical music. For the model with alternative coding, this probability was 83%. Therefore, even though there was slightly more uncertainty and the mean difference was slightly smaller with the alternative coding, our conclusions remain unchanged.

**

*Figure S4*. A plot of the posterior distributions of the estimated mean difference in effect sizes between lyrical and non-lyrical music. Plotted are the model reported in the paper (orange) and the model done with the alternative coding of the effect sizes (blue). The results indicate that the decision of which coding to use did not affect the conclusions in the paper.

**Unavailable Data**

Due to that fact that four studies did not contain enough information to compute effect sizes and to include them in the meta-analysis, statistical simulations were carried out to explore the consequences of this. The relevant information about these studies is presented in Table S2. For each study, a realistic interval was computed that should contain the effect size of interest given the available information. The simulations were done by taking 10 000 random draws from a Uniform distribution using the effect size bounds in Table S2. For the variance component, a random draw was also taken from a Uniform distribution with bounds corresponding to the range of variance values in the dataset. The random draws were taken from Uniform distributions to denote ignorance about where on the interval the real value may lie. Each randomly generated effect size was added to the dataset that was analyzed in the paper and the meta-analysis was then repeated. The results from the simulations are presented in Table S3 and compared to the effect sizes reported in the main paper. As the simulations show, the results change very little or not at all when the missing effect sizes were simulated and added to the analyses. Therefore, the lack of access to the effect sizes of these fours studies did not bias the conclusions from the meta-analysis.

Table S2

*Information about Studies with Unavailable Data and Their Anticipated Effect Size*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Study | N | Measure | Sound | Available information | Anticipated effect size |
| Hall (1952) | 245 | RC | Music | 2.37% increase in reading score in the music condition | 0< g <0.5 |
| Gawron (1984) | 32† | RC | Noise | Effect size known, but not the direction of the difference | |g|= 0.048 |
| Slater (1968) | 263 | RC | Noise | No sign. differences and “no trends indicative of… [an] effect” (p. 242) | -0.2< g <0.2 |
| Jones et al. (1990), E2 | 16 | PR | Speech | F-value <1; effect size is negative based on the means in Table 2 | -0.13< g <0 ‡ |

RC: reading comprehension accuracy. PR: proofreading accuracy. N: (combined) sample size. All effect sizes are with Morris and DeShon’s (2002) correction (where applicable).

† Only two schedules (2x16 participants) are relevant to the analysis

‡ -0.13 is the lowest bound since this would correspond to the effect size when the F-value is 1.

Table S3

*Results from the Statistical Simulations with Missing Data (SDs in parenthesis)*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Analysis | ES (paper) | Results from 10 000 simulations | | | |
| Mean ES | Range | Mean distribution | Variance distribution |
| PR | -0.14 | -0.13 (0.01) | [-0.15, -0.10] | Uniform(-0.13, 0) | Uniform(0.01, 0.13) |
| RC: Music | -0.19 | -0.19 (0.004) | [-0.20, -0.17] | Uniform(0, 0.5) | Uniform(0.01, 0.20) |
| RC: Noise | -0.17 | -0.16 (0.01) | [-0.17, -0.13] | Uniform(-0.2, 0.2)† | Uniform(0.01, 0.08) |

RC: reading comprehension. PR: proofreading.

† Used for Slater’s (1968) study. For Gawron’s (1984) study, the effect size was positive for half of the simulations (g= 0.048), and negative (g= -0.048) for the remaining half.

References

References marked with an asterisk indicate studies included in the meta-analysis.

\* Ahuja, J. K. (2016). *The effects of background music on cognitive performance* (Unpublished dissertation). The University of Arizona, Tucson, AZ, USA.

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

\* Armstrong, G. B., Boiarsky, G. A., & Mares, M. L. (1991). Background television and reading performance. *Communications Monographs*, *58*(3), 235-253. doi:10.1080/03637759109376228

\* Armstrong, G. B., & Chung, L. (2000). Background television and reading memory in context assessing TV interference and facilitative context effects on encoding versus retrieval processes. *Communication Research*, *27*(3), 327-352. doi:10.1177/009365000027003003

\* Avila, C., Furnham, A., & McClelland, A. (2011). The influence of distracting familiar vocal music on cognitive performance of introverts and extraverts. *Psychology of Music*, *40*(1), 84-93. doi:10.1177/0305735611422672

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47-89. doi:10.1016/S0079-7421(08)60452-1

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, *8*(4), 485-493. doi:10.1037/0894-4105.8.4.485

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

Beaman, C. P. (2005). Auditory distraction from low-intensity noise: A review of the consequences for learning and workplace environments. *Applied Cognitive Psychology*, *19*(8), 1041-1064. doi:10.1002/acp.1134

Borenstein, M. (2009). Effect sizes for continuous data. In H. M. Cooper, L. V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (2nd ed., pp. 221-235). New York, USA: Russel Sage Foundation.

Cassel, E. E., & Dallenbach, K. M. (1918). The effect of auditory distraction upon the sensory reaction. *The American Journal of Psychology*, *29*(2), 129-143.

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

Champely, S. (2012). *pwr: Basic functions for power analysis*. R package version 1.2-1. Retrieved from https://CRAN.R-project.org/package=pwr

Clark, C., & Sörqvist, P. (2012). A 3 year update on the influence of noise on performance and behavior. *Noise and Health*, *14*(61), 292-296. doi:10.4103/1463-1741.104896

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

\* Cool, V. A., Yarbrough, D. B., Patton, J. E., Runde, R., & Keith, T. Z. (1994). Experimental effects of radio and television distractors on children's performance on mathematics and reading assignments. *The Journal of Experimental Education*, *62*(3), 181-194. doi:10.1080/00220973.1994.9943839

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.

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

\* Dockrell, J. E., & Shield, B. M. (2006). Acoustical barriers in classrooms: The impact of noise on performance in the classroom. *British Educational Research Journal*, *32*(3), 509-525. doi:10.1080/01411920600635494

\* Dove, M. K. (2009). *The relationship of rhythmic and melodic perception with background music distraction in college level students* (Unpublished doctoral dissertation). University of Kansas, Kansas city, USA.

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

Engbert, R., Nuthmann, A., Richter, E. M., & Kliegl, R. (2005). SWIFT: A dynamical model of saccade generation during reading. *Psychological Review*, *112*(4), 777-813. doi:10.1037/0033-295X.112.4.777

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

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

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

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

\* Fendrick, P. (1937). The influence of music distraction upon reading efficiency. *The Journal of Educational Research*, *31*(4), 264-271. doi:10.1080/00220671.1937.10880749

\* Fogelson, S. (1973). Music as a distractor on reading-test performance of eighth grade students. *Perceptual and Motor Skills*, *36*(3c), 1265-1266. doi:10.2466/pms.1973.36.3c.1265

\* Freeburne, C. M., & Fleischer, M. S. (1952). The effect of music distraction upon reading rate and comprehension. *Journal of Educational Psychology*, *43*(2), 101-109. doi:10.1037/h0054219.

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

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

\* Furnham, A., Gunter, B., & Peterson, E. (1994). Television distraction and the performance of introverts and extroverts. *Applied Cognitive Psychology*, *8*(7), 705-711. doi:10.1002/acp.2350080708

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

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

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

Gawron, V. J. (1984). Noise: Effect and aftereffect. *Ergonomics*, *27*(1), 5-18. doi:10.1080/00140138408963460

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical Science*, *7*(4), 457-472.

Greenhouse, J. B., & Iyengar, S. (2009). Sensitivity analysis and diagnostics. In H. M. Cooper, L.V. Hedges, & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 418–434). New York, USA: Russell Sage Foundation.

\* Haapakangas, A., Kankkunen, E., Hongisto, V., Virjonen, P., Oliva, D., & Keskinen, E. (2011). Effects of five speech masking sounds on performance and acoustic satisfaction. Implications for open-plan offices. *Acta Acustica united with Acustica*, *97*(4), 641-655.

Haines, M. M., Stansfeld, S. A., Head, J., & Job, R. F. S. (2002). Multilevel modelling of aircraft noise on performance tests in schools around Heathrow Airport London. *Journal of Epidemiology & Community Health*, *56*(2), 139-144. doi:10.1136/jech.56.2.139

Haines, M. M., Stansfeld, S. A., Brentnall, S., Head, J., Berry, B., Jiggins, M., & Hygge, S. (2001a). The West London Schools Study: The effects of chronic aircraft noise exposure on child health. *Psychological Medicine*, *31*(8), 1385-1396. doi:10.1017/S003329170100469X

Haines, M. M., Stansfeld, S. A., Job, R. S., Berglund, B., & Head, J. (2001b). Chronic aircraft noise exposure, stress responses, mental health and cognitive performance in school children. *Psychological Medicine*, *31*(2), 265-277. doi:10.1017/S0033291701003282

Haka, M., Haapakangas, A., Keränen, J., Hakala, J., Keskinen, E., & Hongisto, V. (2009). Performance effects and subjective disturbance of speech in acoustically different office types– A laboratory experiment. *Indoor Air*, *19*(6), 454-467. doi:10.1111/j.1600-0668.2009.00608.x

\* Halin, N. (2016). Distracted while reading? Changing to a hard-to-read font shields against the effects of environmental noise and speech on text memory. *Frontiers in Psychology*, *7*, 1196. doi:10.3389/fpsyg.2016.01196

\* Halin, N., Marsh, J. E., Haga, A., Holmgren, M., & Sörqvist, P. (2014a). Effects of speech on proofreading: Can task-engagement manipulations shield against distraction? *Journal of Experimental Psychology: Applied*, *20*(1), 69-80. doi:10.1037/xap0000002

\* Halin, N., Marsh, J. E., Hellman, A., Hellström, I., & Sörqvist, P. (2014b). A shield against distraction. *Journal of Applied Research in Memory and Cognition*, *3*(1), 31-36. doi:10.1016/j.jarmac.2014.01.003

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

Hansen, W. B., & Collins, L. M. (1994). Seven ways to increase power without increasing *N. NIDA research monograph*, 142, 184-184.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

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

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

Hongisto, V. (2005). A model predicting the effect of speech of varying intelligibility on work performance. *Indoor Air*, *15*(6), 458-468. doi:10.1111/j.1600-0668.2005.00391.x

Hughes, R. W. (2014). Auditory distraction: A duplex-mechanism account. *PsyCh Journal*, *3*(1), 30-41. doi:10.1002/pchj.44

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

Hughes, R. W., Vachon, F., & Jones, D. M. (2005). Auditory attentional capture during serial recall: Violations at encoding of an algorithm-based neural model? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(4), 736-749. doi:10.1037/0278-7393.31.4.736.

Hughes, R. W., Vachon, F., & Jones, D. M. (2007). Disruption of short-term memory by changing and deviant sounds: Support for a duplex-mechanism account of auditory distraction. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*(6), 1050-1061. doi:10.1037/0278-7393.33.6.1050.

Hygge, S., Evans, G. W., & Bullinger, M. (2002). A prospective study of some effects of aircraft noise on cognitive performance in schoolchildren. *Psychological Science*, *13*(5), 469-474. doi:10.1111/1467-9280.00483

\* Hyönä, J., & Ekholm, M. (2016). Background speech effects on sentence processing during reading: An eye movement study. *PLoS One*, *11*(3), e0152133.

Jäger, L. A., Engelmann, F., & Vasishth, S. (2017). Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language*, *94*, 316-339. doi:10.1016/j.jml.2017.01.004

Jahncke, H., Hygge, S., Halin, N., Green, A. M., & Dimberg, K. (2011). Open-plan office noise: Cognitive performance and restoration. *Journal of Environmental Psychology*, *31*(4), 373-382. doi:10.1016/j.jenvp.2011.07.002

\* Johansson, C. R. (1983). Effects of low intensity, continuous and intermittent noise on mental performance and writing pressure of children with different intelligence and personality characteristics. *Ergonomics*, *26*(3), 275-288. doi:10.1080/00140138308963341

\* Johansson, R., Holmqvist, K., Mossberg, F., & Lindgren, M. (2012). Eye movements and reading comprehension while listening to preferred and non-preferred study music. *Psychology of Music*, *40*(3), 339-356. doi:10.1177/0305735610387777

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

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

Jones, D. M., Miles, C., & Page, J. (1990). Disruption of proofreading by irrelevant speech: Effects of attention, arousal or memory? *Applied Cognitive Psychology*, *4*(2), 89-108. doi:10.1002/acp.2350040203

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

Kämpfe, J., Sedlmeier, P., & Renkewitz, F. (2010). The impact of background music on adult listeners: A meta-analysis. *Psychology of Music*, *39*(4), 424-448. doi:10.1177/0305735610376261

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

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

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. USA: Cambridge University Press.

Kintsch, W., & Rawson, K. A. (2005). Comprehension. In M. J. Snowling and C. Hulme (Eds.), *The science of reading: A handbook* (pp. 209-226). United Kingdom: Blackwell Publishing.

Klatte, M., Bergström, K., & Lachmann, T. (2013). Does noise affect learning? A short review on noise effects on cognitive performance in children. *Frontiers in Psychology*, 4, 578. doi:10.3389/fpsyg.2013.00578

\* Kou, S., McClelland, A., & Furnham, A. (2017). The effect of background music and noise on the cognitive test performance of Chinese introverts and extraverts. *Psychology of Music*, Advanced online publication. doi:10.1177/0305735617704300

Lajeunesse, M. J. (2013). Recovering missing or partial data from studies: A survey of conversions and imputations for meta-analysis. In J. Koricheva, J. Gurevitch, & K. Mengersen (Eds.), *The handbook of meta-analysis in ecology and evolution* (pp. 195–206). Princeton: Princeton University Press.

Landström, U., Söderberg, L., Kjellberg, A., & Nordström, B. (2002). Annoyance and performance effects of nearby speech. *Acta Acustica United with Acustica*, *88*(4), 549-553.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, Calif: SAGE Publications, Inc.

\* Ljung, R., Sorqvist, P., & Hygge, S. (2009). Effects of road traffic noise and irrelevant speech on children's reading and mathematical performance. *Noise and Health*, *11*(45), 194-198.

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

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

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

Marsh, J. E., & Jones, D. M. (2010). Cross-modal distraction by background speech: What role for meaning? *Noise and Health*, *12*(49), 210-216.

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

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

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

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

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

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Prisma Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. doi:10.1371/journal.pmed.1000097

Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E. J. (2016). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review*, *23*, 103-123. doi:10.3758/s13423-015-0947-8

Morgan, J. B. (1917). The effect of sound distraction upon memory. *The American Journal of Psychology*, *28*(2), 191-208.

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, *7*(1), 105-125. doi:10.1037/1082-989X.7.1.105

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

Nakagawa, S., Noble, D. W., Senior, A. M., & Lagisz, M. (2017). Meta-evaluation of meta-analysis: Ten appraisal questions for biologists*. BMC Biology*, *15*(1), 18. doi:10.1186/s12915-017-0357-7

Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational Ideas-Part II*. Language and Linguistics Compass*, *10*(11), 591-613. doi:10.1111/lnc3.12207

Noble, D. W., Lagisz, M., O'dea, R. E., & Nakagawa, S. (2017). Nonindependence and sensitivity analyses in ecological and evolutionary meta-analyses. *Molecular Ecology*, *26*(9), 2410-2425. doi:10.1111/mec.14031

Papanikolaou, M., Skenteris, N., & Piperakis, S. M. (2015). Effect of external classroom noise on schoolchildren’s reading and mathematics performance: Correlation of noise levels and gender. *International Journal of Adolescent Medicine and Health*, *27*(1), 25-29. doi:10.1515/ijamh-2014-0006

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

Parmentier, F.B. R. (2014). The cognitive determinants of behavioral distraction by deviant auditory stimuli: A review. *Psychological Research*, *78*(3), 321-338. doi:10.1007/s00426-013-0534-4

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd international workshop on distributed statistical computing* (Vol. 124, p. 125). Technische Universität at Wien.

\* Pool, M. M., Van der Voort, T. H., Beentjes, J. W., & Koolstra, C. M. (2000). Background television as an inhibitor of performance on easy and difficult homework assignments. *Communication Research*, *27*(3), 293-326. doi:10.1177/009365000027003002

R Core Team (2016). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. Retrieved from <http://www.R-project.org/>.

Ravaja, N., & Kallinen, K. (2004). Emotional effects of startling background music during reading news reports: The moderating influence of dispositional BIS and BAS sensitivities. *Scandinavian Journal of Psychology*, *45*(3), 231-238. doi:10.1111/j.1467-9450.2004.00399.x

Reichle, E. D., Pollatsek, A., Fisher, D. L., & Rayner, K. (1998). Toward a model of eye movement control in reading. *Psychological Review*, *105*(1), 125-157. doi:10.1037/0033-295X.105.1.125

Rohatgi, A. (2015). digitizeR: Tool to extract numerical data from images of plots, maps, etc.. *R package version 1.0*. Retrieved from http://github.com/ankitrohatgi/digitizeR

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

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

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

Schmid, C. H., & Mengersen, K. (2013). Bayesian meta-analysis. In J. Koricheva, J. Gurevitch, & K. Mengersen (Eds.), *The handbook of meta-analysis in ecology and evolution* (pp. 145–173). Princeton: Princeton University Press.

Schwarzer, G. (2007). Meta: An R package for meta-analysis. *R News*, *7*(3), 40-45.

Shield, B. M., & Dockrell, J. E. (2003). The effects of noise on children at school: A review. *Building Acoustics*, *10*(2), 97-116. doi:10.1260/135101003768965960

Slater, B. R. (1968). Effects of noise on pupil performance. *Journal of Educational Psychology*, *59*(4), 239-243. doi:10.1037/h0026025

\* Smith-Jackson, T. L., & Klein, K. W. (2009). Open-plan offices: Task performance and mental workload. *Journal of Environmental Psychology*, *29*(2), 279-289. doi:10.1016/j.jenvp.2008.09.002

Sörqvist, P. (2010a). High working memory capacity attenuates the deviation effect but not the changing-state effect: Further support for the duplex-mechanism account of auditory distraction. *Memory & Cognition*, *38*(5), 651-658. doi:10.3758/MC.38.5.651

\* Sörqvist, P. (2010b). Effects of aircraft noise and speech on prose memory: What role for working memory capacity? *Journal of Environmental Psychology*, *30*(1), 112-118. doi:10.1016/j.jenvp.2009.11.004

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

\* Sörqvist, P., Ljungberg, J. K., & Ljung, R. (2010). A sub-process view of working memory capacity: Evidence from effects of speech on prose memory. *Memory*, *18*(3), 310-326. doi: 10.1080/09658211003601530

Sörqvist, P., & Marsh, J. E. (2015). How concentration shields against distraction. *Current Directions in Psychological Science*, *24*(4), 267-272. doi:10.1177/0963721415577356

Sperlich, A., Meixner, J., & Laubrock, J. (2016). Development of the perceptual span in reading: A longitudinal study. *Journal of Experimental Child Psychology*, *146*, 181-201. doi:10.1016/j.jecp.2016.02.007

Stansfeld, S. A., Berglund, B., Clark, C., Lopez-Barrio, I., Fischer, P., Öhrström, E., ... & RANCH Study Team. (2005). Aircraft and road traffic noise and children's cognition and health: A cross-national study. *The Lancet*, *365*(9475), 1942-1949. doi:10.1016/S0140-6736(05)66660-3

\* Sukowski, H., & Romanus, E. (2016). Effects of background speech on reading performance in adults. *Proceedings of Meetings on Acoustics 22ICA,* *28*(1), 050002. doi:10.1121/2.0000360

Sterne, J. A., Sutton, A. J., Ioannidis, J. P., Terrin, N., Jones, D. R., Lau, J., ... & Tetzlaff, J. (2011). Recommendations for examining and interpreting funnel plot asymmetry in meta-analyses of randomised controlled trials. *BMJ*, *343*, d4002. doi:10.1136/bmj.d4002

Sutton, A. J., & Abrams, K. R. (2001). Bayesian methods in meta-analysis and evidence synthesis. *Statistical Methods in Medical Research*, *10*(4), 277-303. doi:10.1177/096228020101000404

Sutton, A. J., Abrams, K. R., Jones, D. R., Jones, D. R., Sheldon, T. A., & Song, F. (2000). *Methods for meta-analysis in medical research*. Chichester, UK: John Wiley & Sons.

Szalma, J. L., & Hancock, P. A. (2011). Noise effects on human performance: A meta-analytic synthesis. *Psychological Bulletin*, *137*(4), 682-707. doi:10.1037/a0023987

Thompson, S. G., & Higgins, J. (2002). How should meta-regression analyses be undertaken and interpreted? *Statistics in Medicine*, *21*,1559-1573. doi:10.1002/sim.1187

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

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

Vachon, F., Hughes, R. W., & Jones, D. M. (2012). Broken expectations: Violation of expectancies, not novelty, captures auditory attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *38*(1), 164 -177. doi:10.1037/a0025054

Vasilev, M. R., & Angele, B. (2017). Parafoveal preview effects from word N+ 1 and word N+ 2 during reading: A critical review and Bayesian meta-analysis. Psychonomic bulletin & review, 24(3), 666-689. doi:10.3758/s13423-016-1147-x

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

Vasishth, S. (2015). *A meta-analysis of relative clause processing in Mandarin Chinese using bias modelling* (Unpublished master’s thesis). The University of Sheffield, United Kingdom.

Vasishth, S., Chen, Z., Li, Q., & Guo, G. (2013). Processing Chinese relative clauses: Evidence for the subject-relative advantage. *PLoS One*, *8*(10), e77006. doi:10.1371/journal.pone.0077006

Veitch, J. A. (1990). Office noise and illumination effects on reading comprehension. *Journal of Environmental Psychology,* *10*(3), 209-217. doi:10.1016/S0272-4944(05)80096-9

Venetjoki, N., Kaarlela-Tuomaala, A., Keskinen, E., & Hongisto, V. (2006). The effect of speech and speech intelligibility on task performance. *Ergonomics*, *49*(11), 1068-1091. doi:10.1080/00140130600679142

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1-48.

Wallisch, P. (2015). Brighter than the sun: Powerscape visualizations illustrate power needs in neuroscience and psychology. *arXiv preprint arXiv:1512.09368*.

\* Weinstein, N. D. (1974). Effect of noise on intellectual performance. *Journal of Applied Psychology*, *59*(5), 548-554. doi:10.1037/h0037338

\* Weinstein, N. D. (1977). Noise and intellectual performance: A confirmation and extension. *Journal of Applied Psychology*, *62*(1), 104-107. doi:10.1037/0021-9010.62.1.104

Welton, N. J., Sutton, A. J., & Cooper, N. (2012). *Statistics in practice: Evidence synthesis for decision making in healthcare*. Somerset: Wiley.

\* Yan, G., Meng, Z., Liu, N., He, L., & Paterson, K. B. (2017). Effects of Irrelevant Background Speech on Eye Movements during Reading. *The Quarterly Journal of Experimental Psychology*, 1-20. Advance online publication.

Ylias, G., & Heaven, P. C. (2003). The influence of distraction on reading comprehension: A Big Five analysis. *Personality and Individual Differences*, *34*(6), 1069-1079. doi:10.1016/S0191-8869(02)00096-X

1. Four studies did not contain any information that made it possible to calculate the effect sizes. As all of the studies were more than 25 years old, it was not possible to obtain the data from the authors. Therefore, these studies were discarded (they did not count towards the number of included studies). We explored the implications of this through statistical simulations and found no evidence that failing to include these studies biased the results (see the Supplementary Material). [↑](#footnote-ref-1)
2. One exception was the meta-regression model comparing lyrical vs non-lyrical music. We show in the Supplemental Material that the way the effect sizes were chosen did not influence the conclusions from this analysis. [↑](#footnote-ref-2)