Attentional control and metacognitive monitoring of the effects of different types of task-irrelevant sound on serial recall

Florian Kattner[1,2*] and Donna Bryce[3]

[1]*Health and Medical University*

[2]*Technical University of Darmstadt*

[3]*University of Tübingen*

Author Note

**Corresponding author's address:*

Florian Kattner
Health and Medical University
Olympischer Weg 1
14471 Potsdam
Germany
florian.kattner@health-and-medical-university.de

Abstract

The presence of task-irrelevant sound disrupts short-term memory for serial information. Recent studies found that enhanced perceptual task-encoding load (static visual noise added to target items) reduces the disruptive effect of an auditory deviant, but does not affect the task-specific interference by changing-state sound, indicating that the deviation effect may be more susceptible to attentional control. This study aimed to further specify the role of attentional control in shielding against different types of auditory distraction, examining speech and non-speech distractors presented in laboratory and web-based experiments. To further elucidate the role of controlled processes, we tested whether the detrimental effects of distractor sounds – and their modulation by attentional control – reach participants' awareness. We found that changing-state sound and auditory deviants in steady-state sound equally affected both objective recall performance and metacognitive confidence judgments, but did not affect the accuracy of confidence judgments. Most importantly, across four experiments, an increase of task load (visual degradation of the to-be-remembered items) did not reduce either type of auditory distraction. A close replication of the original modulation of the deviation effect by perceptual task load (in an online environment) even revealed a stronger deviation effect at high task load, suggesting that the manipulation may have influenced cognitive load and the ability to control distractor interference in memory. In line with a unitary account of auditory distraction, the results suggest that, while both types of distraction reach metacognitive awareness, they may be equally unrelated to perceptual load and the availability of attentional resources.

Keywords: irrelevant sound effect; auditory distraction; metacognitive monitoring; serial recall; attentional capture; changing-state effect

Public Significance Statement

Our ability to hold information in short-term memory suffers in the presence of background sound, but it is unclear to what extent auditory distraction depends on attentional control and metacognitive monitoring. This study reassessed a finding, whereby the diversion of attention by deviant sounds is reduced when the focal task becomes more difficult to process (via perceptual degradation). A series of experiments showed that both the effect of auditory deviants and the interference by changing-state sound is largely resistant to a manipulation of task load, indicating that distraction is not susceptible to attentional control. Nevertheless, participants appeared to be well aware of the detrimental sound effects on performance, as reflected in metacognitive confidence judgments. The findings have important implications for theoretical accounts of auditory distraction, indicating that disruption is due to automatic attentional capture, which cannot be controlled despite us being aware of it.

Attentional control and metacognitive monitoring of the effects of different types of task-irrelevant sound on serial recall

As most readers of this article will have experienced, mental activities such as reading, writing, listening, or memorizing become more difficult when extraneous sound is present in the background. Think, for example, of trying to keep a phone number in mind while two strangers next to you start having an argument, or while your favorite song is being played on the radio. These disruptive effects of irrelevant sound on cognitive performance may be the downside of the general openness of the auditory system (i.e., the ears cannot be closed) enabling continuous monitoring and detection of potentially relevant information (e.g., Hughes & Jones, 2001; Miller, 1974). The aim of the present study is to investigate to what extent the detrimental effects of different types of irrelevant sound on cognitive performance (a) depend on attentional control and (b) reach participants' conscious awareness as reflected in metacognitive monitoring judgments. Therefore, the degree of perceptual task-encoding load was manipulated to study the role of attentional control as a possible shield against two different forms of auditory distraction, resulting from speech and non-speech distractor sounds: the changing-state effect and the auditory deviation effect.

Interference-by-process

Previous studies have shown that performance in short-term memory tasks that require the maintenance of serial-order information is particularly sensitive to the presence of task-irrelevant speech (e.g., Colle & Welsh, 1976; Salamé & Baddeley, 1982). Specifically, the presence of background speech impairs the memorization of visually presented lists of verbal items (e.g., letter or digits), even when the speech is totally irrelevant to the memorization tasks or presented in a foreign language (for a short review, see Ellermeier & Zimmer, 2014).

Originally, this disruptive effect of speech was explained in terms of an automatic access of spoken information to a phonological short-term memory store, producing interference with the active articulatory rehearsal process that is used to refresh information in the store (Baddeley & Hitch, 1974). However, it was later demonstrated that various temporally changing non-speech sounds (e.g., random tone sequences or instrumental music) also produce distraction in verbal serial short-term memory tasks, rendering a phonological interference-by-content account at best incomplete (Jones et al., 1999; Jones & Macken, 1993; Kattner & Meinhardt, 2020; Marsh et al., 2020; Nittono, 1997; Salamé & Baddeley, 1989; Schlittmeier et al., 2008).

According to the interference-by-process (or changing-state) account of auditory distraction, the disruptive effects of irrelevant sound are not due to phonological interference, but due to the interference between automatically processed order cues in the acoustical background and the maintenance of serial order in short-term memory (Hughes & Jones, 2001; Jones et al., 1996). More specifically, pre-attentively processed changes between successive auditory events are assumed to provide order cues to the formation of an auditory stream (enabling segregation and grouping in auditory scene analysis; Bregman, 1990), which then compete with the order of to-be-remembered items for inclusion and maintenance in the articulatory rehearsal sequence. In line with this account, several studies demonstrated that changing-state sounds such as sequences of changing syllables or tones randomly varying in frequency (which provide many order cues) produce more interference in serial recall tasks than repeated syllables or tones (e.g., Elliott, 2002; Jones et al., 1992; Jones & Macken, 1993; Tremblay & Jones, 1999), whereas non-serial memory tasks were found to be largely immune to a changing-state effect (e.g., Beaman & Jones, 1997; Jones & Macken, 1993).

Attentional Capture

In contrast, attentional accounts of working memory (e.g., the embedded-processes model; Cowan, 1995, 1999) generally explain auditory distraction with a reorientation of attentional resources from the relevant task items to the irrelevant auditory information (Bell et al., 2012; Elliott, 2002). Such attentional capture may occur either as the result of violations of expectations or acoustical regularities (e.g., an unexpected beep or a sudden voice change), or due to the specific meaning or goal-relevance of the contents in irrelevant sound (e.g., taboo words or one's own name; Röer et al., 2013; Röer, Körner, et al., 2017). The most frequently studied example of attentional capture is the disruptive effect of a single auditory deviant in an otherwise regular (and thus predictive) sequence of sounds, such as a sudden change in frequency. This auditory deviation effect has been studied extensively using the oddball paradigm (for a review see Parmentier, 2014), demonstrating not only disruption to memory performance, but also delayed response times in speeded classification tasks or enhanced electrophysiological indicators of distraction (Escera et al., 1998; Getzmann et al., 2013; Leiva et al., 2016; Parmentier, 2016; Parmentier et al., 2008, 2018; Wessel et al., 2016; Wessel & Aron, 2013). Importantly, acoustical deviants in both steady-state and changing-state sequences of irrelevant sound were also shown to disrupt performance in a serial recall task (Hughes et al., 2005, 2007, 2013; Marois & Vachon, 2018; Vachon et al., 2012), suggesting independent and additive effects of changing-state sound and auditory deviants (with the deviation effect typically being smaller in magnitude than the changing-state effect).

According to a unitary attentional account of auditory distraction (e.g., Bell et al., 2012), any change in a sound should capture more attention than a repetition of the same sound or continuous noise. Therefore, the above-mentioned changing-state effect can be explained with the same mechanisms as the deviation effect, i.e., a diversion of attention from the focal task, assuming that changing-state sound captures more attention than steady-state sound (i.e., due to habituation of the attentional orienting response to repeated sounds; Bell et al., 2012, 2019; but see Jones et al., 1997).

Duplex-mechanism account

On the other hand, according to the duplex-mechanism account there are two functionally distinct forms of auditory distraction resulting from either (a) specific interference between properties of the irrelevant sound and psychological processes required for the focal task or (b) a more general (task-unspecific) diversion of attentional resources from the focal task to the irrelevant sound (Hughes, 2014; Hughes et al., 2005, 2007). More specifically, the account posits that the changing-state effect should be restricted to tasks that require serial-order processing, whereas sounds that are meaningful (e.g., one's own name; Röer et al., 2013) or contain a violation of the predictions of the neural model (e.g., an unexpected change in voice; Hughes et al., 2005) should lead to an attentional orienting response that disrupts performance on any cognitive task requiring central attentional resources. Therefore, the changing-state effect on serial recall is supposed to be due to interference-by-process only, whereas the deviation effect is supposed to be the result of attentional capture. In addition, it has been argued that attentional capture (e.g., the deviation effect) should be a form of auditory distraction that is susceptible to top-down cognitive control (Hughes, 2014) and awareness (Hughes & Marsh, 2019), whereas an individual should not be aware of and not be able to control the interference produced by changing-state sound (see also Bell et al., 2021).

In line with this account, several studies observed that changing-state sound does not disrupt performance when serial rehearsal was either prevented (through articulatory suppression; Jones et al., 2004) or unlikely to be the memorization strategy used to perform the task (e.g., in the missing-item task; Beaman & Jones, 1997; Hughes & Marsh, 2020; Jones & Macken, 1993; Kattner & Ellermeier, 2018). In contrast, attentional capture by semantic properties or auditory deviants in irrelevant sound seems to affect performance also in non-serial short-term memory tasks (Hughes et al., 2007; Kattner & Ellermeier, 2018; Marsh et al., 2018). There is also evidence that high working memory capacity (thought to be a crucial feature of cognitive top-down control) is associated with a smaller deviation effect, but unrelated to the changing-state effect (Ellermeier & Zimmer, 1997; Hughes et al., 2013; Sörqvist, 2010; Sörqvist et al., 2013). Others, however, reported the changing-state effect and the deviation effect to be equally unrelated to working memory capacity (Körner et al., 2017). Moreover, a simple warning signal or specific foreknowledge of the distractors was found to eliminate the deviation effect (Hughes et al., 2013; Sussman et al., 2003), whereas foreknowledge and expectancies did not affect the changing-state effect on serial recall (Bell et al., 2017; Röer et al., 2015). These dissociations suggest that distraction due to changing-state sound may be less susceptible to top-down control than attentional capture.

Task load and attentional control

The influence of peripheral attentional control can also be studied by manipulating the difficulty of the focal task. In line with the perceptual load theory (Lavie, 2005, 2010), enhanced task-encoding load is expected to draw on the same perceptual or attentional resources that are required for the processing of irrelevant sound. Therefore, any increase in task difficulty should make it less likely that irrelevant sound will be processed (due to a lack of perceptual resources), thus reducing auditory distraction. Interestingly, it was found that an increase of perceptual task load in the serial recall task reduces the degree of distraction produced by semantic properties and auditory deviants in task-irrelevant speech – but it did not affect the changing-state effect – suggesting that the enhanced task load may have shielded against attentional disengagement from the focal task (Hughes et al., 2013; Hughes & Marsh, 2019; Marsh et al., 2015, 2020). Hughes et al. (2013), for example, presented task-irrelevant sequences consisting of ten different letters as changing-state sound, either with all letters spoken by the same voice or with a single letter being spoken by a different voice (the deviant), while participants were asked to remember the serial order of visually presented digits. To manipulate task-encoding load, the digits were either clearly visible or degraded by adding static Gaussian visual noise. In the first two experiments, the authors showed that the voice deviant reduced serial recall accuracy with clearly visible digits, but not with visually degraded digits, suggesting that enhanced perceptual task-encoding load prevented the processing of the voice deviant in a stream of changing-state speech (in line with perceptual load theory; Lavie, 2005), thus eliminating the deviation effect. However, providing a warning signal prior to the deviant eliminated its disruptive effect on serial recall also with low task-encoding load, suggesting that the disruptive effect is due to a violation of expectations regarding the acoustical environment, which could be eliminated through top-down cognitive control by providing foreknowledge. In a third experiment, the authors tested the changing-state effect and found that changing letters were more distracting than steady-state letters (i.e., ten repetitions of a single letter), regardless of the task-encoding load and foreknowledge condition. These results clearly indicate that enhanced perceptual task load and cognitive top-down control reduces the distraction produced by a single auditory deviant (in an acoustical changing-state background), whereas it did not affect the changing-state effect. Surprisingly, the authors did not investigate the influence of task load on the effect of a deviant in a steady-state sequence of irrelevant sound (i.e., the deviation effect in the absence of a changing-state effect). Therefore, the effects of task load and cognitive control could be restricted to distraction produced by an interaction of continuous changing-state sound and an auditory deviant. More recent studies have shown that the auditory deviants in steady-state sound are also less disruptive when there are enhanced demands on visual attention (comparing a global vs. local focus of attention; Marsh et al., 2020) and cognitive control (i.e., inhibitory control demands; Hughes & Marsh, 2019). Others, however reported that a manipulation of task engagement using monetary incentives did not modulate the deviation effect (Bell et al., 2020).

One aim of the present study is to further specify the role of peripheral attentional control in shielding against isolated types of auditory distraction. Therefore, it was tested whether an increase of perceptual task-encoding load selectively reduces the auditory deviation effect also in the absence of changing-state sound (i.e., when the deviant is presented in a steady-state sequence of irrelevant sounds), while not affecting the interference produced by changing-state sound. In addition, we investigate whether the attentional control of the deviation effect is restricted to speech distractors or generalizes to non-speech sound (i.e., sequences of steady-state and changing-state tones; see Elliott, 2002; Jones et al., 1999; Jones & Macken, 1993; Marsh et al., 2020; Tremblay & Jones, 1998). As such, we aim to establish the boundary conditions of such a shielding effect of attentional control and thus further constrain the duplex-mechanism account.

Metacognitive monitoring

Another goal of the present study is to determine whether participants are consciously aware of the differential effects of auditory distractors, and their modulation by attentional control, on memory performance. Under the duplex-mechanism account it has been suggested that participants should be aware of the impact of an auditory deviant on their memory performance but not of the impact of changing-state sound (Hughes & Marsh, 2019). Further, one could speculate that if people were aware of the detrimental effect of auditory deviants, they may be more inclined to increase top-down cognitive control to reduce their impact. To investigate conscious awareness, we collected trial-by-trial metacognitive monitoring judgments and measured their accuracy (how well participants' subjective perception of their performance matched their objective performance) across different experimental conditions.

Some previous studies have assessed monitoring judgments and monitoring accuracy in memory tasks including either auditory distraction or perceptual degradation. Bell and colleagues (2021) recently reported two studies in which they assessed participants' metacognitive beliefs, prospective judgments, and retrospective reports about the impact of changing-state sound and auditory deviants on serial recall performance. Their findings offered no support for the hypothesis that people are aware of the effect of auditory deviants but not of changing-state sound. Instead, they observed that after experiencing the auditory stimuli, participants could report that changing-state sequences disrupted memory recall more than auditory deviants (via both trial-by-trial prospective judgments and global retrospective judgments). They interpreted this as counter evidence to the duplex-mechanism account. The present study complements theirs by collecting trial-by-trial retrospective judgments and aiming to assess whether participants are aware of the previously reported sparing of their performance when an auditory deviant occurs under high task load conditions. In line with the duplex-mechanism account, both the disruptive effect of an auditory deviant and its modulation by task load, but not the changing-state effect should be reflected in metacognitive confidence judgments. In contrast, according to a unitary attentional account of auditory distraction, the two effects should be equally reflected in confidence judgments.

Further, assessing whether monitoring accuracy differs between different auditory distraction conditions, as is done here, may offer a more fine-grained evaluation of the hypothesis that participants are 'more aware' of the impact of auditory deviants than changing-state sounds. Beaman et al. (2014), for instance, found that task-irrelevant speech impaired not only memory performance in a recognition paradigm but also monitoring accuracy. Röer and colleagues focused specifically on metacognitive processes during serial recall under conditions of auditory distraction and found that the subjective rating of the distractibility produced by changing-state sound depended on the participants' metacognitive beliefs (i.e., expectations), whereas their objective serial recall accuracy did not (Röer, Rummel, et al., 2017). This suggests that participants may not have very accurate introspections about the impact of auditory distraction on their performance. Here we test whether trial-by-trial monitoring accuracy will be worse in the changing-state and auditory deviant conditions as compared to the steady-state condition. With regards to perceptual degradation in recall tasks, it has been found that monitoring accuracy was compromised when stimuli were degraded in an effortful listening condition (Amichetti et al., 2013). Consistent with this observation, we also expect that using trial-by-trial monitoring judgments will uncover poorer monitoring accuracy in the high perceptual task-load than the low task-load condition, regardless of the irrelevant sound condition.

Experiment 1

Methods

Participants

A power analysis (using the wp.rmanova() function from the {WebPower} package for R; Zhang & Yuan, 2018) based on the previously reported effect size for the crucial modulation of the deviation effect by task load (corresponding to Cohen's $\hat{f} = .76$; see Hughes et al., 2013; Exp. 1; p. 542), revealed that a minimum sample size of $N = 26$ is required to demonstrate the interaction with a statistical power of $1\text{-}\beta = .95$ ($\alpha = .05$). In addition, the power analysis showed that a minimum sample size of $N = 22$ is required to demonstrate the main effect for disruption by the presence of a deviant ($\hat{f} = .83$), whereas a minimum effect size of $N = 9$ participants is required to demonstrate the main effect for a disruption by changing-state sound which is not expected to depend on task load ($\hat{f} = .83$; see Hughes et al., 2013; Exp. 3a; p. 546). Hence, we started with a sample size of $N = 26$, but data collection was continued in batches of 10 participants until the Bayes factors (see Data Analysis for further details) provided conclusive evidence either for ($BF_{Interaction}$ / $BF_{Two\ main\ effects} > 7$) or against ($BF_{Two\ main\ effects}$ / $BF_{Interaction} > 7$) a modulation of the deviation effect by task load (i.e., referring to the likelihood of a model containing the interaction between sound and task load).

With this Bayesian stopping rule, thirty-six participants (22 women and 14 men) were recruited collaboratively at the University of Kassel, the Technical University of Darmstadt, and the University of Tübingen (thus exceeding the sample size of $N = 27$ in Hughes et al., 2013; Exp. 1). Ages ranged between 19 and 45 years ($M = 24.7$; $SD = 6.0$). The data of two additional female participants with extremely poor serial recall performance (less than two digits recalled on average) were not included in the analyses. All participants reported normal hearing and normal or corrected-to-normal vision. Student participants were compensated with course credits or payment (10€ per hour).

Apparatus and Stimuli

The experiment was conducted in quiet, dimly-lit sound attenuated listening booths (at University of Tübingen and Technical University of Darmstadt, respectively). Visual stimuli were presented on a 17-inch LCD monitor, and participants were seated at approximately 60 cm viewing distance. Sounds were played via headphones (Beyerdynamics DT-990 and Sony MDR-XD200). The experimental routines were programmed in MATLAB (Mathworks, Natick, MA) utilizing the Psychophysics toolbox 3.0 extensions (Brainard, 1997; Kleiner et al., 2007; Pelli, 1997).

The visual *to-be-remembered items* were 284×256-pixel images of black digits (from the set of 1 to 9) on white background. Static Gaussian noise of either low (*Var* = 1) or high variance (*Var* = 18) was added to the image (using the function 'imnoise' in MATLAB) on trials with low or perceptual high task-encoding load, respectively. An illustration of a digit in the low and high task-encoding load condition is provided in the Appendix (Fig. A1).

Recordings of the letters A, E, G, I, K, L, T, and U, each spoken by a female voice, were presented via headphones as *task-irrelevant sound* at approximately 65 dB(A). Each recording (44.1 kHz) had a duration of 1 s with the utterance covering about 500 ms in the middle of the signal. On a steady-state trial, a single letter was drawn randomly from the set of eight letters and played thirteen times (1 letter / s). On a changing-state trial, thirteen letters were drawn randomly with replacement from the set of eight letters. Deviant trials were identical to a steady-state trial except that one letter at a randomly drawn position in the second half of the sequence (uniformly distributed between the positions 7 and 12) was replaced by a different letter. An additional analysis was conducted in which the position of the deviant was dichotomized into early deviants occurring during the presentation of the to-be-remembered digits (letter positions 7-9) or late deviants occurring during the retention interval (letter positions 10-12).

Design and Procedure

A 3 (sound: steady-state, changing-state, steady-state + deviant) × 2 (task load: low, high) within-subjects design was implemented with 20 repetitions in each experimental condition. After reading the instructions, participants started with two practice trials (one with low and one with high task load, and with a randomly assigned irrelevant sound). Data of the practice trials was not included in the analysis. Participants then completed 120 trials with the order of sound and load conditions randomized for each run of 6 successive trials. Short breaks could be made after every block of 25 trials (i.e., after trial 25, 50, 75, and 100). Each trial started with the 1-s presentation of a warning signal, i.e., a blue square decreasing in size. A random sequence of eight unique digits (from 1-9) was then presented on the screen at a rate of one digit/s, followed by a 5-s retention interval showing a blank grey screen. Irrelevant sound was presented during both the digit presentation and the retention interval (i.e., for 13 s). After the retention interval, a numeric pad was shown on the screen, and participants were asked to click the eight digits in correct serial order. Entered responses could not be corrected, and no feedback on the recall accuracy was provided. After all digits had been entered, participants were asked to judge their confidence in the accuracy of their response (“Please indicate how sure are you that the digits you entered are correct:”) by clicking on a 7-point Likert scale ranging from “very uncertain” to “very certain”.

Data Analysis

Analogue strategies of data analysis were pursued for all four experiments. *Metacognitive monitoring accuracy* was calculated as robust γ-rank correlation coefficients between the confidence judgments and recall accuracy for each participant and experimental condition (using the R package {RoCoCo}; Bodenhofer et al., 2013; Bodenhofer & Klawonn, 2008). Serial recall accuracy, confidence judgments and metacognitive monitoring accuracy were analyzed initially with a 3 (sound: steady-state, steady-state + deviant, changing-state) × 2 (task load: low, high) repeated-measures analyses of variance (ANOVA). In case of a significant main effect of sound or a significant interaction, the specific effects of auditory distraction were further analyzed with separate ANOVAs of the changing-state effect (steady-state vs. changing-state) and the deviation effect (steady-state vs. steady-state + deviant). For all effects, generalized eta squared (η_G^2) is reported as a measure of effect size (as recommended by Bakeman, 2005). Additional Bayesian ANOVAs were conducted to estimate the likelihood of different models with main effects and interaction terms (using the {BayesFactor} package for R; Rouder et al., 2012). For these analyses, standard Cauchy-distributed priors were specified with a width of $r = 0.5$ for fixed effects and $r = 1$ for random effects, respectively. The estimated Bayes factors BF_{10} indicate the likelihood of an initial model containing a main effect of sound relative to the null hypothesis (intercept only). For additional main effects and interaction terms, BF_{10} refers to the likelihood of a model including the additional effect and the previously reported effects (e.g., both main effects of sound and task load) relative to the null hypothesis. Bayes factors BF_{10} greater than 3 are typically considered as moderate evidence for the alternative hypothesis (i.e., the alternative hypothesis is more than 3 times as likely as the null hypothesis given the present data), whereas a BF_{10} lower than 1/3 may be considered as moderate evidence for the null hypothesis (see

Jeffreys, 1961; Morey et al., 2016; for more details on the quantification of evidence). Below, the resulting Bayes factors will be reported together with the respective effects of the frequentist ANOVA.

Results

Serial recall accuracy

Fig. 1a illustrates the average *recall accuracy* as a function of task-encoding load and irrelevant sound. Changing-state speech clearly impaired serial recall, as compared to steady-state speech. In contrast, a deviant in a steady-state stream of sound appeared to produce less distraction. A 3 (sound) × 2 (task load) repeated-measures ANOVA confirmed a significant main effect of sound, $F(2,70) = 16.99$; $MSE = 0.003$; $p < .001$; $\eta_G^2 = 0.023$; $BF_{10} = 6937.59 \pm 0.02\%$, with lower recall accuracy on trials with changing-state speech ($M = 0.62$; $SD = 0.17$), compared to trials with steady-state speech ($M = 0.68$; $SD = 0.13$) and steady-state speech containing a deviant ($M = 0.66$; $SD = 0.15$). Separate analyses confirmed a significant changing-state effect on serial recall accuracy (steady-state vs. changing-state), $F(1,35) = 26.07$; $MSE = 0.002$; $p < .001$; $\eta_G^2 = 0.034$; $BF_{10} = 44559.9$, but did not reveal a significant deviation effect when all deviants (early and late) were included in the analysis (steady-state vs. steady-state with deviant), $F(1,35) = 3.21$; $MSE = 0.001$; $p = .082$; $\eta_G^2 = 0.003$; $BF_{10} = 0.099$. There was also a significant main effect of task load, $F(1,35) = 13.30$; $MSE = 0.006$; $p = .001$; $\eta_G^2 = 0.015$; $BF_{10} = 4035640 \pm 2.47\%$ (likelihood of a model with two main effects), with higher recall accuracy at low load ($M = 0.68$; $SD = 0.15$) than at high load ($M = 0.64$; $SD = 0.15$). However, there was no interaction between sound and load, $F(2,70) = 1.48$; $MSE = 0.003$; $p = .234$; $\eta_G^2 = 0.002$; $BF_{10} = 69347.62 \pm 1.98\%$, providing strong evidence against a modulation of auditory distraction by task-encoding load (i.e., the stopping rule; note that the Bayes factors indicate that a model with main effects of sound and task load is about 58.2 times more likely than a model with the two main effects and the interaction). Hence, the degree of auditory distraction on serial recall was not affected by task-encoding load.

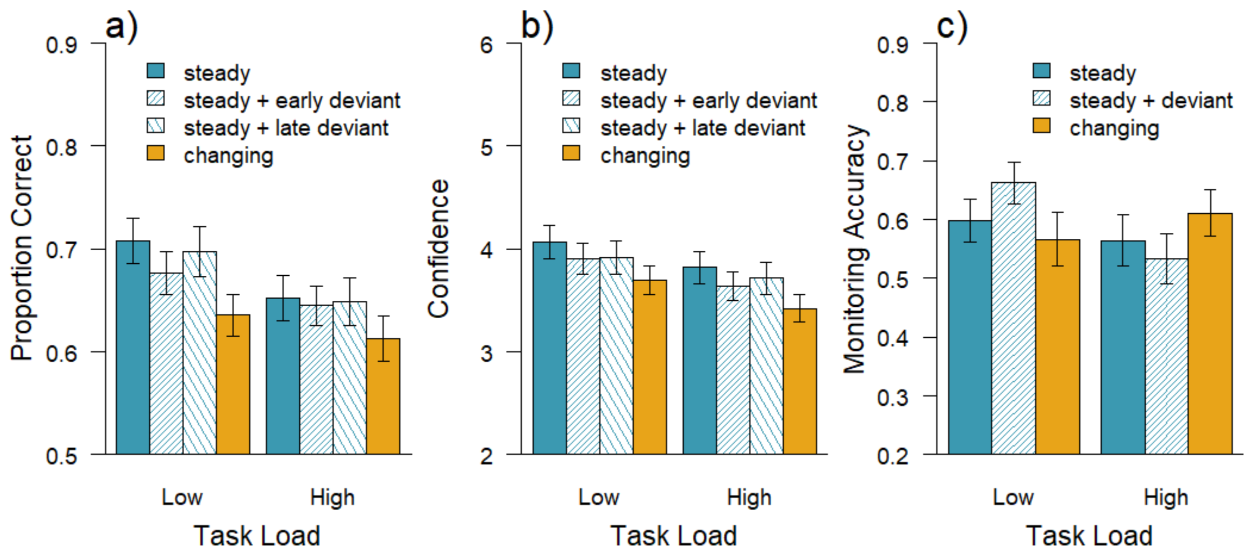


Figure 1. **(a)** Serial recall accuracy (proportion correct) in Experiment 1 as a function of the type of task-irrelevant speech played during maintenance of digits presented visually with low or high task-encoding load, **(b)** confidence judgments of serial recall accuracy, and **(c)** monitoring accuracy (gamma correlations between accuracy and confidence) in the same experimental conditions. Error bars depict standard errors of the mean.

As the separate analyses above revealed no evidence for disruption by the auditory deviant, the deviation effect was further analyzed as a function of whether the deviant occurred early in the sequence (during presentation of the to-be-remembered items) or late (during the retention interval; note that no Bayesian ANOVA was conducted due to the reduced number of observations). For early deviants, there was a significant deviation effect, $F(1,35) = 4.24$;

MSE = 0.003; *p* = .047; η_G^2 = 0.004 (with deviant: *M* = 0.66; *SD* = 0.15 vs. without deviant: *M* = 0.68; *SD* = 0.13), as well as a main effect of task load, *F*(1,35) = 12.48; *MSE* = 0.003; *p* = .001; η_G^2 = 0.021, but this ‘early’ deviation effect was not modulated by task load, *F*(1,35) = 0.99; *MSE* = 0.005; *p* = .327; η_G^2 = 0.002. In contrast, for the late deviants, there was only a main effect of task load, *F*(1,35) = 12.78; *MSE* = 0.008; *p* = .001; η_G^2 = 0.028, but the presence (vs. absence) of a deviant did not affect recall accuracy, *F*(1,35) = 0.40; *MSE* = 0.004; *p* = .531; η_G^2 < 0.001 (with deviant: *M* = 0.67; *SD* = 0.16 vs. without deviant: *M* = 0.68; *SD* = 0.13). There was also no interaction in case of late deviants, *F*(1,35) = 0.05; *MSE* = 0.008; *p* = .820; η_G^2 < 0.001.

Monitoring judgments and accuracy

Confidence judgments of serial recall are illustrated in Fig. 1b, and they also seem to be sensitive to both the type of irrelevant sound and the task load. A 3 (sound) × 2 (task load) repeated-measures ANOVA confirmed this observation with a significant main effect of sound, *F*(2,70) = 23.75; *MSE* = 0.112; *p* < .001, η_G^2 = 0.018; BF_{10} = 38933.48 ± 0.02%, demonstrating lower confidence on trials with changing-state speech (*M* = 3.56; *SD* = 1.09) than on trials with steady-state speech (*M* = 3.94; *SD* = 1.11), and steady-state speech with a deviant (*M* = 3.79; *SD* = 1.20). Follow-up analyses confirmed that confidence was sensitive to both changing-state sound, *F*(1,35) = 39.39; *MSE* = 0.066; *p* < .001; η_G^2 = 0.030; BF_{10} = 3565687 ± 2.24%, and the presence of an auditory deviant, *F*(1,35) = 8.92; *MSE* = 0.043; *p* = .005; η_G^2 = 0.004; BF_{10} = 7.48395 ± 5.38% (Bayes factor refers to the likelihoods of models with main effects of the respective sound conditions and task load, which was 15.44 and 17.87 times more likely than the interaction model for the changing-state and deviation effect, respectively). Confidence judgments were also subject to a main effect of task load, *F*(1,35) = 10.50; *MSE* = 0.307;

$p = .003$, $\eta_G^2 = 0.011$; $BF_{10} = 26367383 \pm 1.66\%$ (low load: $M = 3.89$; $SD = 1.14$; high load: $M = 3.64$; $SD = 1.13$). However, consistent with the objective measure of recall accuracy, there was no sound × task load interaction on confidence judgments, $F(2,70) = 0.19$; $MSE = 0.090$; $p = .825$; $\eta_G^2 < 0.001$; $BF_{10} = 159099.5 \pm 1.7\%$ (i.e., the model with an additional interaction term is about 165.73 times less likely than the model with only two main effects). This pattern suggests that subjective confidence in recall accuracy is sensitive to distraction by both changing-state sound and auditory deviants, regardless of task-encoding load.

The effect of an auditory deviant on confidence judgments was also analyzed as a function of the position of the deviant in the sequence of irrelevant sound. For early deviants, confidence differed significantly between trials with ($M = 3.77$; $SD = 1.19$) and without a deviant ($M = 3.94$; $SD = 1.11$), $F(1,35) = 9.77$; $MSE = 0.105$; $p = .004$, $\eta_G^2 = 0.005$, as well as between low ($M = 3.98$; $SD = 1.15$) and high task load ($M = 3.73$; $SD = 1.18$), $F(1,35) = 8.23$; $MSE = 0.284$; $p = .007$, $\eta_G^2 = 0.012$, but there was no interaction, $F(1,35) = 0.03$; $MSE = 0.111$; $p = .855$, $\eta_G^2 < 0.001$. In contrast, for the late deviants presented during the retention interval, there was only a main effect of task load, $F(1,35) = 6.47$; $MSE = 0.279$; $p = .016$, $\eta_G^2 = 0.009$, but no significant deviation effect, $F(1,35) = 3.46$; $MSE = 0.154$; $p = .071$, $\eta_G^2 = 0.003$, and no interaction, $F(1,35) = 0.12$; $MSE = 0.122$; $p = .728$, $\eta_G^2 < 0.001$. This pattern indicates that participants' monitoring is sensitive enough to detect that the disruptive impact of deviants is restricted to 'early' deviants presented during the encoding phase.

Monitoring accuracy was high, with a grand mean of 0.59 ($SD_\gamma = 0.16$) for the γ correlation, $p < .001$. Correlations in the different sound and load conditions are illustrated in Fig. 1c. A 3 (sound) × 2 (task load) repeated-measures ANOVA on monitoring accuracy (γ correlations) revealed no main effect of sound, $F(2,70) = 0.01$; $MSE = 0.051$; p = .909,

$\eta_G^2 = 0.001$; $BF_{10} = 0.02 \pm 0.03\%$, and no significant main effect of task load, $F(1,35) = 2.70$; $MSE = 0.032$; $p = .1097$, $\eta_G^2 = 0.007$; $BF_{10} = 1566.402$. However, there was a significant and highly likely interaction between task load and sound, $F(2,70) = 3.86$; $MSE = 0.036$; $p = .026$, $\eta_G^2 = 0.021$; $BF_{10} = 1.05 \cdot 10^{17} \pm 2.73\%$, suggesting that monitoring accuracy on trials with a deviant increased at low task-encoding load (see Fig. 1c).

Discussion

The presence of changing-state speech clearly impaired serial recall accuracy in Experiment 1, as compared to steady-state speech, and this effect did not depend on whether the to-be-remembered items were presented with low or high task-encoding load. Interestingly, the presence of an auditory deviant in a sequence of steady-state speech disrupted serial recall only when presented early during the encoding of the to-be-remembered items, but not when it was presented later during the retention interval. However, in contrast to previous findings (Hughes et al., 2013), this deviation effect did not attenuate with increased task-encoding load. Obviously, this pattern of results is not consistent with the duplex-mechanism account (Hughes, 2014; Hughes et al., 2005) which predicts that the deviation effect should be more sensitive to task load than the changing-state effect. A closer look at the data of Experiment 1 indicates that an increase in task load may have affected serial recall accuracy primarily in the steady-state speech condition (see Fig. 1a), which suggests that enhanced task load reduced the ability to ignore even repeated speech sounds. It could be speculated thus whether the previously reported small steady-state effect on serial recall (Bell et al., 2019) may be sensitive to attentional control, leading to more disruption by steady-state sound in case of high task load. As a consequence, both the changing-state effect (i.e., the difference between changing-state and steady-state sounds conditions) and the deviation effect (i.e., the difference between deviant and steady-state conditions) in the present experiment could have been reduced at high task load just because the level of performance in the steady-state control condition was lower. Hence, it is possible that an increase of the deviation effect (i.e., a greater difference in performance between deviant and steady-state at high task load) has been masked by the task-load effect on steady-state performance.

In addition to its effect on short-term memory, task-irrelevant changing-state speech also affected the participants' metacognitive confidence judgments regarding their performance in the serial recall task. Furthermore, confidence judgments reflected the disruptive effect of an auditory deviant presented during encoding, with confidence being higher on trials that did not contain a deviant. While task load was also reflected in confidence judgments (lower confidence with increased task load), neither effect of the presence of an auditory distractor on subjective confidence depended on task-encoding load, thus mirroring the data pattern in objective recall accuracy. Given that all effects in serial recall performance were reflected in confidence judgments, it is not surprising that there was a high gamma correlation between recall accuracy and confidence judgments in this experiment (i.e. high relative monitoring accuracy). Interestingly, we observed that metacognitive monitoring accuracy was higher on trials with an auditory deviant under low task load than on deviant trials under high task load. As this finding was rather unexpected, its reliability was assessed in Experiments 2 and 3.

As expected, the results suggest that enhanced attentional control did not shield against the disruptive effect of changing-state sound. While the overall deviation effect was not significant in Experiment 1, the disruptive effect of early deviants (those presented during encoding) was not alleviated by enhanced attentional control either. This suggests that the deviation effect may not depend as much on perceptual task load as previously suggested (Hughes et al., 2013; Exp. 1-2), and that both the changing-state effect and the deviation effect may be equally unrelated to cognitive control (and driven by automatic attentional capture; Bell et al., 2020; Körner et al., 2017). To test this interpretation, and assess the generalizability of the effects observed in Experiment 1, we conducted two additional experiments with increased statistical power: Experiment 2 employed non-speech sounds as auditory distractors, and Experiment 3 was a web-based close replication (with some necessary extensions) of Experiment 1.

Experiment 2

To investigate the generalizability of the disruptive effects of (a) changing-state sound and (b) steady-state sequences with deviants on serial recall and metacognitive monitoring, we presented non-speech sounds (i.e., tone sequences) as auditory distractors in Experiment 2. The same manipulations of visual task-encoding load as in Experiment 1 were used to test the susceptibility of distraction to attentional control.

Methods

Participants

Since no modulation of auditory distraction by task load was observed in Experiment 1 (there was Bayesian evidence against an interaction), the original effect size of the interaction ($\hat{f}$ = .76) may have been overestimated. In addition, as the deviation effect produced by non-speech distractors (as well as its modulation by task load) may be even smaller than for speech distractors, we reduced the hypothetical effect size for the interaction by at least a third. A sensitivity analysis of statistical power revealed that an increased sample size of $N = 53$ participants is required to demonstrate a reduced effect size of $\hat{f} = .50$ for the modulation of the deviation effect by task load with a statistical power of $1 - \beta = .95$ ($\alpha = .05$).

Fifty-three participants (41 women and 12 men) were recruited at the University of Kassel and at Technical University of Darmstadt. Ages raged between 18 and 40 years ($M = 22.4$; $SD = 4.0$). All participants reported normal hearing and normal or corrected-to-normal vision. Student participants received course credits as a compensation.

Apparatus and stimuli

The apparatus was the same as in Experiment 1. The to-be-remembered stimuli (digits) were the same as in Experiment 1. To manipulate task-encoding load, either low or high visual noise was added to the stimuli (see Fig. A1 in the Appendix). The only difference to Experiment 1 were the task-irrelevant stimuli. Instead of spoken letters, sequences consisting of 52 sinusoid tones varying in frequency from 261.63 Hz to 523.25 Hz in semitone steps were presented as task-irrelevant sound in Experiment 2. Each tone had a duration of 100 ms, including 10-ms rise and fall times with cosine-shaped envelopes. The tone sequences had a total duration of 13 s (i.e., covering the digit presentation and retention interval), with a new tone being presented every 250 ms. For the changing-state trials, 52 tones were drawn with replacement from the full set of 13 frequencies, whereas a single random tone was repeated 52 times on steady-state trials. The deviant trials were identical to the steady-state trials except for one randomly drawn tone in the second half of the sequence and before the end of the presentation of the to-be-remembered items (i.e., uniformly distributed between tone position 26 and 32) being replaced by a random tone with a different frequency. As such, in contrast to Experiment 1 all deviants can be considered ‘early’ deviants (i.e. presented during the encoding phase).

Design and Procedure

As in Experiment 1, a 3 (sound) × 2 (load) within-subjects design was implemented with 20 repetitions per experimental condition, resulting in a total of 120 trials. There were two additional practice trials at the beginning of the experiment. The participants' task was to recall the series of eight visually presented digits in correct serial order while ignoring the tones played via headphones. At the end of each trial, participants were asked to rate their confidence in serial recall accuracy using the same scale as in Experiment 1.

Results

Serial recall accuracy

Fig. 2a illustrates serial recall accuracy in Experiment 2. A 2 × 3 repeated-measures ANOVA revealed a significant and highly likely main effect of sound, $F(2,104) = 18.16$; $MSE = 0.003$; $p < .001$; $\eta_G^2 = 0.011$; $BF_{10} = 769.20$, but no main effect of load, $F(1,52) = 0.45$; $MSE = 0.002$; $p = .506$; $\eta_G^2 < 0.001$; $BF_{10} = 23.75 \pm 1.14\%$ (likelihood of a model with two main effects), and no interaction, $F(2,104) = 0.88$; $MSE = 0.003$; $p = .419$; $\eta_G^2 = 0.001$; $BF_{10} = 0.19 \pm 4.33\%$ (likelihood of a model with two main effects and an interaction). Follow-up analyses revealed a significant and very likely changing-state effect, $F(1,52) = 25.78$; $MSE = 0.002$; $p < .001$; $\eta_G^2 = 0.016$; $BF_{10} = 4509.79$, but no deviation effect, $F(1,52) = 3.55$; $MSE = 0.001$; $p = .065$; $\eta_G^2 = 0.001$; $BF_{10} = 0.08$.

Monitoring judgments and accuracy

Fig. 2b illustrates the confidence judgments in Experiment 2. As in Experiment 1, confidence was higher under conditions of steady-state sound ($M = 3.86$; $SD = 1.06$), than under conditions of changing-state sound ($M = 3.64$; $SD = 0.98$), with confidence on trials with a deviant tone being in between (*M* = 3.75; *SD* = 1.02). This difference in confidence judgments was confirmed by a significant main effect of sound, $F(2,104) = 7.44$; $MSE = 0.177$; $p = .001$; $\eta_G^2 = 0.008$; $BF_{10} = 8.55 \pm 0.01\%$, whereas there was no significant main effect of task-encoding load, $F(1,52) = 2.98$; $MSE = 0.131$; $p = .090$; $\eta_G^2 = 0.001$; $BF_{10} = 0.92 \pm 8.98\%$ (likelihood of a model with two main effects), and no interaction, $F(2,104) = 0.81$; $MSE = 0.145$; $p = .447$; $\eta_G^2 = 0.001$; $BF_{10} < 0.01 \pm 1.5\%$ (it is about 1350 times more likely that there is only a main effect of sound). Follow-up analyses again revealed that confidence judgments clearly reflected awareness of the impact of changing-state sound, $F(1,52) = 13.24$; $MSE = 0.100$; $p = .001$; $\eta_G^2 = 0.012$; $BF_{10} = 151.89$, and there is also some frequentist evidence for metacognitive awareness of the auditory deviant, $F(1,52) = 4.04$; $MSE = 0.08$; $p = .050$; $\eta_G^2 = 0.003$; $BF_{10} =$ 0.27.

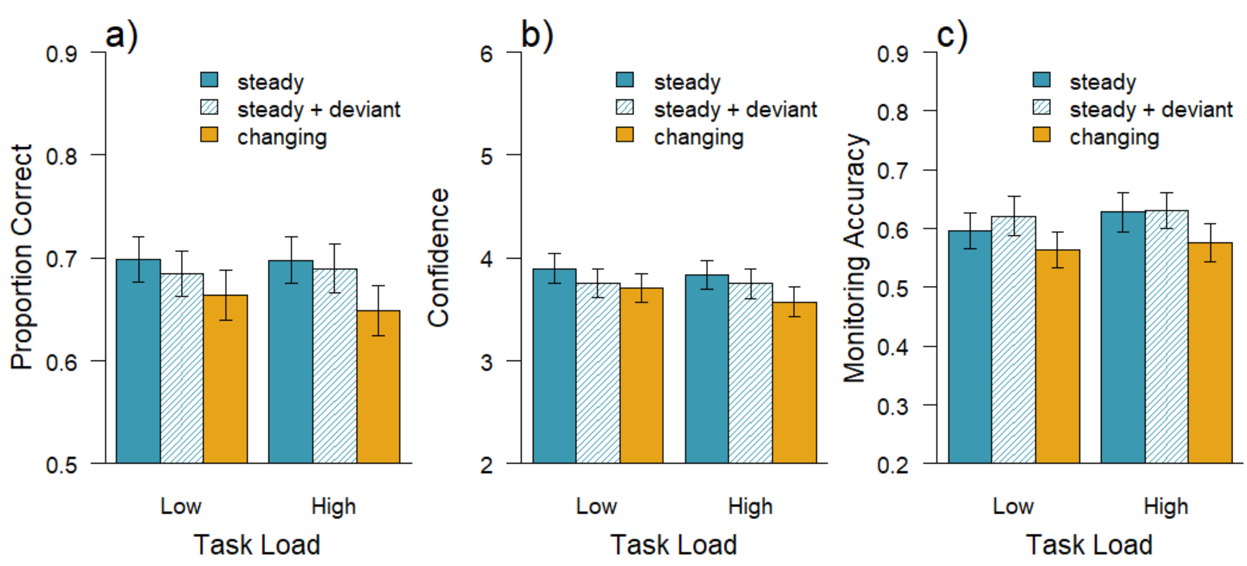


Figure 2. **(a)** Mean serial recall accuracy (proportion correct) in Experiment 2 as a function of the irrelevant tones played during maintenance of digits presented visually with low or high task-encoding load, **(b)** confidence judgments of serial recall accuracy, and **(c)** monitoring accuracy

(gamma correlations between accuracy and confidence) in the same experimental conditions. Error bars depict standard errors of the mean.

As in Experiment 1, confidence judgments correlated significantly with serial recall accuracy ($M_\gamma = .60$; $SD_\gamma = .16$; $p < .001$). As can be seen in Fig. 2c, the accuracy of confidence judgments varied much less between experimental conditions than in Experiment 1. A 3 × 2 repeated-measures ANOVA revealed no significant main effect of sound, $F(2,104) = 2.36$; $MSE = 0.037$; $p = .099$; $\eta_G^2 = 0.010$; $BF_{10} = 4.58 \cdot 10^{11} \pm 0.01\%$, no main effect of load, $F(1,52) = 0.73$; $MSE = 0.035$; $p = .397$; $\eta_G^2 = 0.002$; $BF_{10} = 1.58 \cdot 10^{12} \pm 1.44\%$ (i.e., a model with two main effects was 3.45 times more likely than a model with only a main effect of sound), and no interaction between sound and load conditions, $F(2,104) = 0.12$; $MSE = 0.034$; $p = .884$; $\eta_G^2 < 0.001$; $BF_{10} = 3.03 \cdot 10^{10} \pm 5.55\%$ (i.e., an interaction model is about 52.20 times less likely than a model with two main effects).

Discussion

Experiment 2 demonstrated that changing tone sequences clearly disrupted serial recall performance as compared to steady-state tone sequences (in line with previous studies; e.g., Jones et al., 1999; Jones & Macken, 1993) and that enhanced task load did not shield against the disruptive effect of changing-state tones. In contrast to Experiment 1, a single deviant tone presented 'early' during the encoding phase did not significantly disrupt serial recall in Experiment 2. However, also in contrast to Experiment 1, the visual degradation of to-be-remembered items also had no overall effect on serial recall performance, which might suggest that the manipulation of task-encoding load was not successful (it can only be speculated that the processing of visually degraded digits may be less demanding in the presence of non-speech sound than in the presence of speech sound). This of course limits our interpretation of the impact of attentional control on either effect of auditory distraction.

Interestingly, participants' subjective confidence in their performance was lower both when changing-state tone sequences or an auditory deviant was presented (though there was no Bayesian evidence for the latter effect), as compared to when steady-state tone sequences were presented. This may suggest that metacognitive judgments of performance are more sensitive to small disruptive effects (such as the deviation effect with non-speech stimuli) than objective recall accuracy. Consistent with objective performance, a manipulation of task-encoding load had no effect on confidence judgments, indicating that subjective confidence varies with task load only when there is an impact on recall accuracy. In addition, while the accuracy of confidence judgments was high across all conditions in Experiment 2 (i.e., it correlated well with recall performance), there was no evidence for an effect of either task-irrelevant sound or task-encoding load on metacognitive monitoring accuracy. Taken together with Experiment 1, there is little evidence of auditory distraction affecting the accuracy of metacognitive monitoring. It is very well possible that the selectively improved monitoring accuracy for deviants at low load, which was observed in Experiment 1, may have been due to imprecisions with a smaller sample size. Therefore, an aim of a third experiment was to replicate the possible effects of different types of irrelevant speech on serial recall and metacognition.

Experiment 3

The aim of Experiment 3 was to test the reliability of the results of Experiment 1 with enhanced statistical power. Due to the restrictions during the Covid-19 pandemic, it was not possible to run the experiment in the laboratory. Therefore, a web-based experiment was developed, with participants running the tasks online using their own computers at home. To reinvestigate the changing-state effect and the auditory deviation effect on serial recall and metacognitive monitoring judgments with irrelevant speech sounds, the presence of (a) changing-state speech and (b) a presumably more effective non-speech deviant (a pure tone) in steady-state speech was contrasted with steady-state speech background. In addition, a quiet control condition was added in Experiment 3 allowing us to control whether participants were listening to the irrelevant sounds, or whether they had removed the headphones (using catch trials). Moreover, the inclusion of a quiet condition allows us to additionally test for the small steady-state effect on serial recall and metacognitive monitoring (Bell et al., 2019).

Method

Participants

A sensitivity analysis of statistical power was conducted, assuming a further reduced effect size of $\hat{f} = 0.45$ for the modulation of the deviation effect by task load. This reduced effect size estimate was considered appropriate because (a) the effect size might be smaller in an online experiment than in the laboratory and (b) a higher number of to-be-excluded participants was expected in case online experimentation. The analysis revealed that a sample size of $N = 66$ is required to demonstrate the interaction with a statistical power of $1 - \beta = .95$ ($\alpha = .05$).

Sixty-six participants (51 women, 15 men) were recruited simultaneously at the University of Tübingen and at Technical University of Darmstadt to take part in an online experiment. Ages ranged between 18 and 54 years ($M = 21.3$; $SD = 4.7$). All participants reported normal hearing and normal or corrected-to-normal vision. Student participants who completed the entire task received partial course credit as compensation.

Before starting the experiment, all participants confirmed that the task will be completed alone in a dimly lit room using either a desktop or laptop computer or laptop. They also confirmed that they were not expecting interruption for about 60 min and that their cellphone was turned off during the task.

Apparatus and Stimuli

The experiment was programmed in PsychoPy (Peirce, 2007, 2009; Peirce et al., 2019) and converted to a PsychoJS (JavaScript) program, which could be run online on the Pavlovia server (https://pavlovia.org/). Participants were permitted to complete the experiment only with a keyboard using either a desktop or a laptop computer (not with a tablet or smartphone). In addition, participants were instructed not to use loudspeakers, but to wear headphones for the experiment. Based on self-report, 42.4% of all participants used over-ear or on-ear headphones (21.2% each), 43.9% used in-ear headphones, and the remaining 13.6% used ear buds.

The stimulus materials were similar to the ones used for Experiment 1. The same to-be-remembered visual displays of digits were presented in the online experiment. For the irrelevant sound conditions, twenty unique steady-state and twenty unique changing-state sequences were created, each consisting of twenty 700-ms recordings of letters spoken by a female voice (B, F, G, K, L, M, S, T). All sounds were converted from wav to mp3 format. Five deviant trials were created by replacing one randomly chosen letter between the fifth and eighth position of a steady-state sequence (i.e., between 3.5 s and 6.3 s after the onset of the to-be-remembered sequence) with a 659 Hz tone (700 ms) as the auditory deviant. As such, all deviants were presented during processing of the to-be-remembered digits (i.e. during encoding). In addition to the experimental conditions from Experiment 1, there were 10 quiet trials with no background sound.

Procedure

The experiment started with a headphone-screening test to ensure that participants were using headphones and had the volume set to an appropriate level (Woods et al., 2017). At the beginning, participants were asked to wear their headphones and to adjust the volume of their computer to a comfortable level while continuous pink noise (RMS level 0.10) was presented. The headphone-screening test was a 3-alternative forced choice task with three 200-Hz tones presented successively. Each tone was presented in stereo for a duration of 1000 ms together with an orange square showing the number “1”, “2”, or “3”. The first square was presented 4° to the left, the second in the center, and the third 4° to the right. A 500-ms inter-stimulus interval (and a blank screen) followed each tone. The level of one tone was 6 dB lower than the level of the two other tones. One of the two louder tones had the phase reversed between the left and right channels (in case of loudspeakers, this phase reversal was expected to reduce the sound pressure level in air, thus making it more difficult to detect the low-intensity tone; see Woods et al., 2017). After the third tone was presented, participants were asked to indicate which of the three tones was quieter than the other two by pressing the respective number key on their keyboard. The headphone test was completed if at least five correct responses were given within a 6-trial block. If less than five responses were correct, the test continued with the next 6-trial block until either a minimum of five correct responses were made within a block or the fifth block was completed. If the headphone test was not passed within five blocks, a message was shown on the screen, telling the participant that the study was terminated because the audio system was likely to be insufficient to proceed (they were allowed to restart the experiment though). Otherwise, the experiment continued with the serial recall task.

The procedure of the serial recall task was analogue to Experiment 1. On each trial, eight digits were drawn randomly without replacement from 1-9 and presented sequentially on the screen. After a fixation triangle (1000 ms), each digit was presented for 800 ms each and followed by a 200 ms inter-stimulus interval (i.e., one digit per second). For all but the quiet trials, irrelevant sound composed of twenty spoken letters was played during both the digit presentation and the subsequent 6-s retention interval. After the total trial duration of 14 s, a response matrix consisting of the numbers 1-9 was shown on the screen and participants were asked to click the order of digits (using a mouse). The sequence of entered digits appeared below the matrix, and participants could not correct their responses. After the eighth digit, participants were asked to indicate their recall confidence on a 7-point scale ranging from “very uncertain” to “very certain”. For motivational purposes during this web-based experiment, text feedback was then presented for 1 s reporting the number of correctly recalled digits before the next trial started.

The digits were presented with high and low task-encoding load (see Fig. A1 in the Appendix) on 45 trials each. Within each load condition, steady-state and changing-state sound were each presented on fifteen trials, steady-state sound with the auditory deviant was presented on five trials, and there were ten silent trials. The resulting 90 trials were presented in full random order. Participants could take a short break after the 22nd, 45th, and 68th trial. At the end of six randomly chosen catch trials, participants were asked to press the “j” (or “n”) key to indicate whether they had heard a sound during the trial (or not).

Results

Data processing

To make sure that the headphones were not taken off during the task, data were included only for participants who responded correctly in at least 5 of 6 catch trials (indicating whether they heard a sound). As six participants failed this selection criterion, the following results are based on the remaining $N = 59$ participants (45 women; $M_{age} = 21.5$ years; $SD_{age} = 4.9$ years).

Serial recall accuracy

Fig. 3a illustrates the accuracy during the online serial recall task of Experiment 3. It is obvious that changing-state speech clearly disrupted serial recall compared to steady-state sound, which in turn disrupted performance compared to the quiet condition. Moreover, steady-state speech with an auditory deviant (a single sine tone) impaired performance to a similar degree as changing-state speech did. However, while overall performance appears to be lower in the high-task load condition, neither type of auditory distraction seems to depend on task load.

A 4 (sound) × 2 (task load) repeated-measures ANOVA confirmed these observations with a significant main effect of sound, $F(3,141) = 36.13$; $MSE = 0.006$; $p < .001$; $\eta_G^2 = 0.071$; $BF_{10} = 2677.04$. As in Experiment 1, there was also a significant main effect of task load, $F(1,47) = 13.50$; $MSE = 0.010$; $p = .001$; $\eta_G^2 = 0.016$; $BF_{10} = 5520.34 \pm 2.74\%$ (likelihood of model with two main effects), with higher recall accuracy at low load ($M = 0.67$; $SD = 0.14$) than at high load ($M = 0.64$; $SD = 0.13$). The sound × load interaction was not significant and unlikely, $F(3,141) = 1.53$; $MSE = 0.007$; $p = .210$; $\eta_G^2 = 0.003$; $BF_{10} = 309.79 \pm 21.13\%$ (i.e., a model with two independent main effects is 17.82 times more likely than the interaction model), indicating that auditory distraction (changing-state, deviation, and steady-state effects) did not depend on task-encoding load.

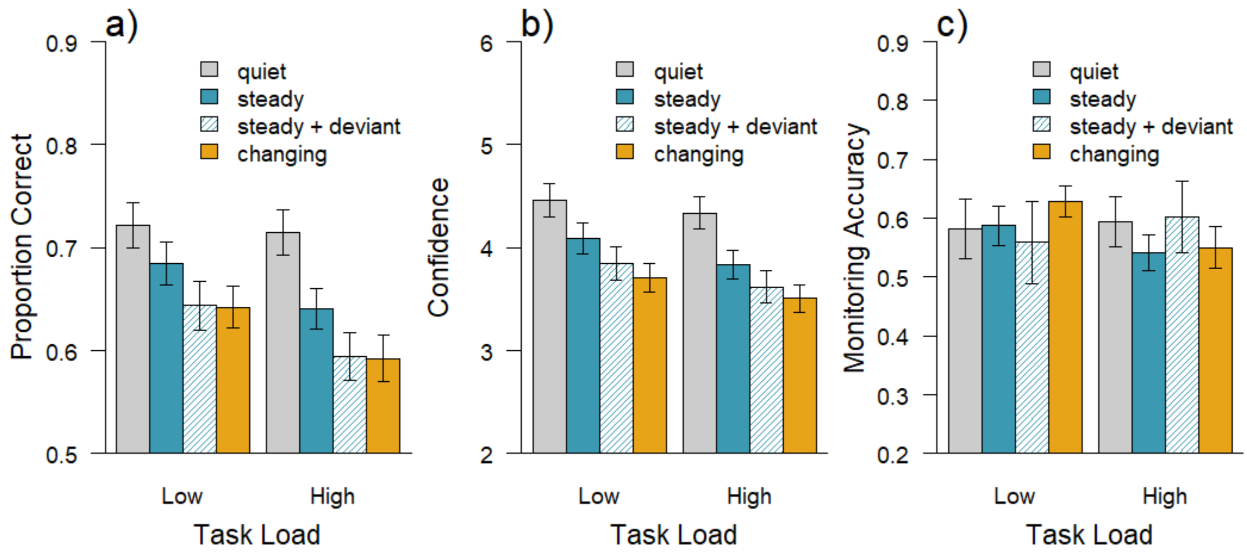


Figure 3. **(a)** Mean serial recall accuracy in Experiment 3 (online) as a function of the irrelevant sound played during maintenance of digits presented visually with low or high task-encoding load, **(b)** confidence judgments of serial recall accuracy, and **(c)** monitoring accuracy (gamma correlations between accuracy and confidence) in the same experimental conditions. Error bars depict standard errors of the mean.

Separate analyses revealed a significant changing-state effect, $F(1,47) = 20.08$; $MSE = 0.002$; $p < .001$; $\eta_G^2 = 0.028$; $BF_{10} = 1.39$ (inconclusive Bayesian evidence; a model containing also the main effect of load is only 1.65 times more likely), with impaired serial recall during changing-state speech ($M = 0.62$; $SD = 0.14$), compared to steady-state speech ($M = 0.66$; $SD = 0.13$). There was also a significant deviation effect, $F(1,47) = 20.08$; $MSE = 0.002$; $p < .001$; $\eta_G^2 = 0.028$; $BF_{10} = 0.97$ (inconclusive Bayesian evidence; the most likely model containing only the main effect of load is 1.34 times more likely), with disrupted performance in the presence of an auditory deviant in steady-state speech ($M = 0.62$; $SD = 0.15$). The disruptive effect of steady-state speech compared to the quiet control condition ($M = 0.72$; $SD = 0.14$; and see Fig. 3a) was also significant, $F(1,47) = 28.48$; $MSE = 0.003$; $p < .001$; $\eta_G^2 = 0.042$; $BF_{10} = 3.84$ (i.e., the most likely model with only a main effect of sound: steady vs. quiet).

Monitoring judgments and accuracy

Fig. 3b illustrates the confidence judgments in the different experimental conditions of Experiment 3. As in the previous experiments, 4 (sound) × 2 (task load) repeated-measures ANOVA revealed a significant main effect of sound, $F(3,141) = 41.44$; $MSE = 0.278$; $p < .001$, $\eta_G^2 = 0.078$; $BF_{10} = 11630.02 \pm 0.01\%$, with lower confidence on trials with changing-state speech ($M = 3.61$; $SD = 0.90$) than on trials with steady-state speech without ($M = 3.96$; $SD = 0.96$) or with an auditory deviant ($M = 3.73$; $SD = 1.01$). Confidence was highest in the quiet control trials ($M = 4.39$; $SD = 1.03$). The ANOVA also revealed a main effect of task load, $F(1,47) = 8.79$; $MSE = 0.442$; $p = .005$, $\eta_G^2 = 0.005$; $BF_{10} = 7297.62 \pm 1.65\%$ (low load: $M = 4.02$; $SD = 0.99$; high load: $M = 3.82$; $SD = 0.92$), but again no interaction between task load and sound, $F(3,141) = 0.35$; $MSE = 0.216$; $p = .787$, $\eta_G^2 = 0.001$; $BF_{10} = 214.55 \pm 4.11\%$ (i.e., a model with only a main effect of sound is 54.21 times more likely than the interaction model).

Further analyses revealed inconclusive Bayesian evidence, but they demonstrated a significant difference in confidence between changing-state and steady-state sound, $F(1,47) = 36.57$; $MSE = 0.083$; $p < .001$; $\eta_G^2 = 0.036$; $BF_{10} = 2.86$, between steady-state with and without an auditory deviant, $F(1,47) = 8.94$; $MSE = 0.139$; $p = .004$; $\eta_G^2 = 0.013$; $BF_{10} = 0.50$, and between steady-state and quiet trials, $F(1,47) = 30.93$; $MSE = 0.146$; $p < .001$; $\eta_G^2 = 0.046$; $BF_{10} = 6.59$.

The average robust γ-rank correlation between confidence judgments and recall accuracy in Experiment 3 was $M_\gamma = .58$ ($SD_\gamma = .14$) and differed significantly from zero, $p < .001$ (see Fig. 3c). As in Experiment 2, a 4 (sound) × 2 (task load) repeated-measures ANOVA on

γ correlations revealed no main effect of sound, $F(3,141) = 0.15$; $MSE = 0.086$; $p = .931$; $\eta_G^2 = 0.001$; $BF_{10} = 0.01$, no main effect of task load, $F(1,47) = 0.32$; $MSE = 0.0.091$; $p = .575$, $\eta_G^2 = 0.001$; $BF_{10} = 0.13$, and no interaction, $F(3,141) = 0.73$; $MSE = 0.099$; $p = .533$, $\eta_G^2 = 0.006$; $BF_{10} < 0.001$. Thus, the previous observation of enhanced monitoring accuracy in case of a deviant and low task-encoding load (see Experiment 1) could not be replicated in this online experiment with enhanced statistical power and using a different (presumably more effective) type of auditory deviant. Moreover, none of the auditory distractors used in Experiment 3 impaired monitoring accuracy compared to a quiet control condition[1].

Discussion

Experiment 3 was a web-based investigation of the disruptive effects of different types of irrelevant sound on serial recall, confidence judgments, and monitoring accuracy. The results show that the presence of both irrelevant changing-state speech and an acoustically distinct auditory deviant (a sine tone) in steady-state speech disrupted serial recall compared to plain steady-state speech and silence. Moreover, in line with previous observations (Bell et al., 2019), a significant steady-state effect could be observed in the present online experiment, with disrupted performance during steady-state speech compared to a quiet control condition (see Fig. 3a).

Consistent with our interpretations of the previous experiments, all three types of auditory distraction were found to be insensitive to manipulations of task-encoding load. Hence, the results suggest that the disruptions produced by both changing-state sound and auditory

[1] We note that the correlations in the quiet condition (and in the deviant condition) were based on a smaller number of trials than in the steady-state and changing-state conditions, which may have reduced reliability (see the larger error bars in Fig. 3c).

deviants (as well as the steady-state effect) cannot be modulated easily through peripheral attentional control (i.e., an increase of perceptual task load). In principle, this result might indicate that all types of auditory distraction are based on automatic attentional capture, and that an individual has very limited attentional control over the disruptive sound effects (see Bell et al., 2020; Körner et al., 2017; Röer, Körner, et al., 2017).

However, metacognitive judgments also indicate that participants were very well aware of the disruptive effects, thus casting doubt on the automaticity assumption of such an account. Again, the irrelevant sound effects on objective performance were mirrored in participants' metacognitive confidence judgments, which were lower in case of more disruptive sounds (changing-state and deviation effects) and in case of high task load. In addition, even the presence of steady-state speech was found to lead to lower confidence judgments as compared to quiet trials, indicating that participants are also aware of the relatively small disruptive effect of repeated speech sounds. In line with the previous experiment, there was no indication that either type of auditory distraction or task-encoding load affected the trial-by-trial precision of these metacognitive monitoring judgments.

Experiment 4

Given that the attenuation of the deviation effect by increased task load (Hughes et al., 2013; Exp. 1) had not been replicated across three experiments, a fourth experiment was conducted as a close replication of the original experiment. Therefore, we generated analogue stimulus materials and used exactly the same procedures except that – due to pandemic-related shutdown of the laboratory – the experiment was conducted in a web-based environment. In the terminology used in Experiments 1 to 3, this experiment contrasted the disruptive effects of changing-state sequences of spoken items with changing-state sequences that contain an auditory deviant (the deviant was a single speech item presented in a different voice). As this was intended to be a close replication of the original experiment, no confidence judgments were collected.

Method

Participants

Forty-two participants (32 women, 10 men) were recruited at the University of Tübingen. Ages ranged between 19 and 56 years (M = 22.1; SD = 5.6). This sample size would be sufficient to demonstrate an interaction effect of $\hat{f}$ = 0.57 (which would be only 75% of the effect size in the original experiment; Hughes et al., 2013; Exp. 1) in a simple 2 (deviant) × 2 (task load) design with a statistical power of 95% (α = .05). All participants reported either normal (n = 23) or corrected-to-normal vision (n = 19), and no participant reported hearing impairment. Before starting the task, all participants confirmed that they were working on a desktop computer or laptop alone in a dimly lit room without interruption for about 60 min and that their cellphone was turned off.

Apparatus and stimuli

The experimental routines for the different online tasks of Experiment 4 (pilot experiment, headphone screening task, serial recall task) were written in PsyToolkit syntax (Stoet, 2010, 2017).

As in the original study (Hughes et al., 2013; Exp. 1), the *to-be-remembered items* consisted of the digits 1-8, which were presented visually on the screen in random order (without replacement). On half of the trials, the digits were clearly visible and in half of the trials uniformly distributed visual noise was added (with an R script) to the JPEG images of the digits in black font on a white background (converting 50% of the 284 x 256 pixels to black or white), thus creating similar percepts as in the original study (see Fig. A2 in the Appendix). The digits were presented in an online pilot experiment to ten participants who did not participate in the main experiment (8 women, 2 men; age: $M = 30.7$; $SD = 7.1$ years). The pilot experiment consisted of 256 trials, with each digit being presented 16 times without and 16 times with visual noise. Participants were asked to classify the digits as even or odd digits as quickly and as accurately as possible by pressing the “A” or “L” key, respectively. Consistent with the original study, the results confirmed that the response times were significantly faster without visual noise ($M = 574$ ms; $SD = 55$ ms) than with visual noise ($M = 613$ ms; $SD = 51$ ms), $t(9) = -9.12$; $p < .001$. Accuracy of classification responses was high and did not differ as a function of task-encoding load (without noise: $M = 91.0\%$; $SD = 5.1\%$ vs. with noise: $M = 90.7\%$; $SD = 5.8\%$), $t(9) = 0.31$; $p = .76$.

The to-be-ignored sound consisted of the same list of ten spoken letters (A, B, C, G, J, K, L, M, Q, and S) as in the original study. The letters were pronounced as in the German alphabet and each letter was spoken by a male and a female German speaker using a monotone voice. Each letter recording had a duration of 500 ms, which was twice the duration that was reported for the English spoken letters in the original study (Hughes et al., 2013; p. 542) – note that some adjustments were made to the timing of the sequence to match the original procedure (see below). Changing-state sequences without a deviant were created by concatenating the ten letters spoken by the same voice in random order. There was a 150-ms gap of silence after each letter, resulting in a total sequence duration of 6.5 s. For the sequences with a deviant, the sixth letter was replaced by the same letter in the other voice.

Design and Procedure

Before starting the serial recall task, all participants had to pass the headphone-screening test which was identical to the one in Experiment 3 (but programmed in PsyToolkit).

The serial recall experiment had a 2 (deviant: absent, present) × 2 (task-encoding load: without noise, with noise) within-subjects design. As in the original study, there were two blocks of 45 trials, one with the letters spoken by a female voice and one with the letters spoken by the male voice. Each block contained 39 trials without a deviant and six trials with the sixth letter being presented in the deviant voice. For each block and sound condition, half of the trials (i.e., 19/20 trials without a deviant in the male/female voice block, and 3 deviant trials) were presented with low load (clearly visible digits) and half of the trials (i.e., 19/20 trials without a deviant in the female/male voice block, and 3 deviant trials) were presented with high load (digits with visual noise).

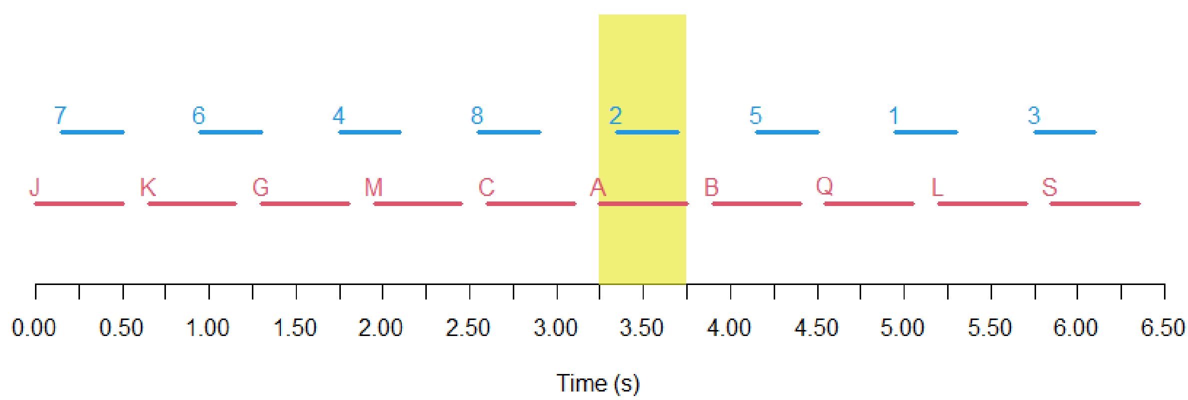


Figure 4. Illustration of the procedure in Experiment 4: The to-be-remembered digits (blue) were presented for 350 ms each, with a 450-ms inter-stimulus interval. The to-be-ignored sequence of spoken letters (red) started 150 ms before the onset of the first digit. Each letter had a duration of 500 ms and was followed by an inter-stimulus interval of 150 ms. On some trials, the sixth letter was presented in a different voice (yellow area).

Each digit was presented on the white screen for 350 ms, followed by a blank 450-ms inter-stimulus interval, resulting in a total duration of 6.4 s for the to-be-remembered sequence. The sequence of irrelevant letters started 150 ms before the first digit. With these timing parameters (which are identical the original study), the deviant letter would start 100 ms prior to the onset of the fifth to-be-remembered digit (see Fig. 4 for an illustration of the relevant and irrelevant sequence), which is very close to position of the deviant reported by Hughes et al. (2013, p. 542; “125 ms before the fifth to-be-remembered item”). Another 1500 ms after the offset of the last to-be-remembered digit, a response grid with the digits 1-8 was presented on the screen and participants were asked to click the eight digits in the order they were presented. The clicked sequence of digits was presented on the screen below the response grid. As in Experiments 1 and 2, no feedback on recall accuracy was presented to participants.

Results

Serial recall accuracy on trials with and without a voice deviant is illustrated in Fig. 5 as a function of task-encoding load. As can be seen, there was a deviation effect at high task load (i.e., with visually degraded digits), but not at low task load. This finding was confirmed by a 2 (deviant) × 2 (load) repeated-measures ANOVA, revealing inconclusive Bayesian evidence, but a significant main effect of sound (presence vs. absence of deviant), $F(1,41) = 8.62$; $MSE = 0.005$; $p = .005$; $\eta_G^2 = 0.011$; $BF_{10} = 0.37$, a significant main effect of task load, $F(1,41) = 5.83$; $MSE = 0.005$; $p = .020$; $\eta_G^2 = 0.008$; $BF_{10} = 0.10 \pm 1.31\%$, as well as a significant interaction between sound and task load, $F(1,41) = 6.01$; $MSE = 0.005$; $p = .019$; $\eta_G^2 = 0.007$; $BF_{10} = 0.04 \pm 1.81\%$ (we note that the Bayesian ANOVA suggests that a model with a main effect of sound only is about 9.48 times more likely than the interaction model, whereas neither model is more likely than the intercept-only model, thus rendering the Bayesian statistics inconclusive). The significant interaction reflects the fact that serial recall accuracy was affected by the presence of a deviant under high visual task-encoding load, $F(1,41) = 13.86$; $MSE = 0.005$; $p = .001$; $\eta_G^2 = 0.039$; $BF_{10} = 0.95$ (Bayes factor inconclusive), but not under low task-encoding load; $F(1,41) = 0.15$; $MSE = 0.005$; $p = .697$; $\eta_G^2 < 0.001$; $BF_{10} = 0.23$.

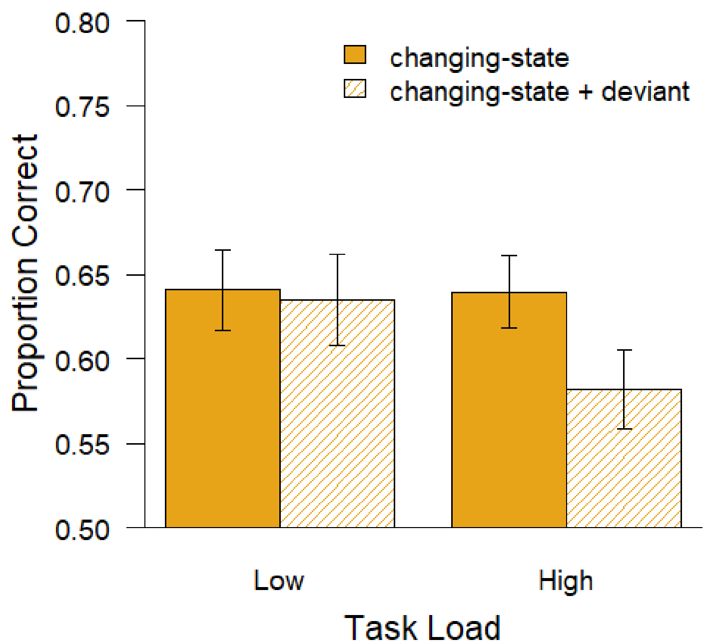


Figure 5. Serial recall accuracy as a function of task-encoding load and the presence of a deviant in Experiment 4. Error bars depict standard errors of the mean. Note, in the equivalent Figure in Hughes et al. (2013; Figure 2) the conditions ‘changing-state’ and ‘changing-state + deviant’ were labelled ‘No Deviant’ and ‘Deviant’, respectively.

Discussion

Experiment 4 was designed as a close replication of the modulation of the deviation effect by task load reported by Hughes et al. (2013; Exp. 1) with sufficient statistical power. Specifically, the disruptive effect of infrequent trials with changing-state sound containing a voice deviant was contrasted with changing-state sound not containing a deviant when the to-be-remembered visual digits were either clearly visible or degraded by visual noise. In the original study, an increase in task load eliminated the deviation effect, and this effect has been interpreted in terms of a reduction of perceptual load (Lavie, 1995, 2005, 2010; Murphy et al., 2016), which would be required for the perceptual-attentional system to process the deviant sound. However, the results of Experiment 4 did not replicate this original finding. That is, in contrast to the predictions of the duplex-mechanism account (Hughes, 2014), an increase in perceptual task-encoding load did not reduce or eliminate the auditory deviation effect on serial recall. This indicates that the previously reported elimination of the deviation effect by increased perceptual load may not be a replicable finding.

Surprisingly, the exact opposite pattern of results was observed in Experiment 4, as the deviation effect was present under high task load, but not at low task load. This might indicate that the cognitive resources available at low task load enabled participants to prevent the diversion of attention from the focal task to the auditory deviant at a late attentional selection stage (e.g., through enhanced inhibitory control; see Kattner, 2021), thus reducing the degree of distraction. It has been found in many previous studies that a manipulation of cognitive load has the opposite effect to perceptual load: With high load on cognitive control processes (e.g., working memory load), the inhibition of a distractor interference is more likely to fail because active cognitive control is needed to resolve the conflict between target and distractor information (Lavie, 2005; see also Lavie et al., 2004; Lavie & De Fockert, 2005). On the other hand, within the oddball paradigm it has been found also that increased working-memory load (postponing the response) reduced the distracting effect of an auditory deviant on response times in a visual categorization task (Berti & Schröger, 2003; but see Parmentier et al., 2008; Exp. 2).

Hence, it could be speculated whether the results of Experiment 4 can be explained in terms of cognitive rather than perceptual load. Similar to visual distraction effects (Lavie, 2010), it is plausible that the auditory deviation effect may be susceptible to both perceptual and cognitive control capacities. The exact conditions to observe a deviation effect may require (a) sufficient perceptual-attentional resources to process the deviant (which could be the case only in the low-load condition in Hughes et al., 2013, but at both load conditions of the present study), and (2) a lack of cognitive resources that would enable active control of the disruption produced by the task-irrelevant deviant on memory performance. It could be argued that the perceptual load was low enough to process the deviant in both conditions of the present study (but only in the low-load condition in Hughes et al., 2013), whereas the cognitive load was low enough to control the deviant only in the low-load condition of the present study (but not in the study by Hughes et al., 2013). In other words, while it is possible that modulation of deviation effect by perceptual task load is not replicable, it could be argued also that the manipulation used in the present online experiment (in contrast to Hughes et al., 2013) affected both perceptual and cognitive task load. Specifically, the visual degradation of digits may not only have delayed the perceptual processing of the digits (enhanced perceptual task-encoding load), but the perceptual constraints may have made the memorization task more difficult at a higher processing level (e.g., because less time is left for the rehearsal of slowly processed digits, or simply because attention needs to be shifted between the patches of visual noise containing the digit and the blank screen in the inter-stimulus interval), thus affecting the availability of cognitive resources. The exact influence of a task-load manipulation on perceptual and cognitive load may depend on factors such as the specific physical stimulus properties and the participants' motivation, task engagement, and working memory capacity. It is certainly possible that these factors differed between a well-controlled laboratory setting (in Hughes et al., 2013; Exp. 1) and the online setting of the present experiment. Specifically, the perceptual load imposed by the visually degraded digits may be more variable when participants run the task on their own computers and monitors than when the same stimuli are presented under controlled laboratory conditions. For instance, due to the specific screen contrast or ambient illumination conditions (which can be optimized at home), both visual task load conditions may have provided sufficient perceptual resources to still process the auditory deviant. However, the degradation of digits may still have affected the cognitive processing load available to memorize the serial order while ignoring the distractor. In particular, for individuals who completed the task at home, the baseline level of cognitive load may have been higher due to a plethora of additional uncontrolled distractions (a ringing phone, traffic from outside, other people speaking nearby etc.). While acknowledging these critiques, we do find it important to note that the web-based experimental context probably has higher ecological validity than laboratory experiments and it seems important to establish the impact of auditory distraction particularly under such conditions. Further research will be necessary to further specify these possible effects of task-load manipulations on the perceptual and cognitive control of auditory distractors.

General Discussion

Across four experiments we investigated the role of attentional control in mitigating the disruptive effects of (a) changing-state sound and (b) the presence of an auditory deviant in irrelevant background sound on objective serial recall performance and metacognitive monitoring judgments (and their accuracy). Therefore, both types of auditory distraction were contrasted between low and high task-encoding load conditions, which was expected to modulate the degree of peripheral attentional control available to process irrelevant sound. Table 1 summarizes the main findings of each experiment with regard to the statistical significance of the changing-state and deviation effect, the task load effect and the modulation of auditory distraction by task load according to the directional predictions of the duplex-mechanism account (i.e., a reduction of the deviation effect by task load). In line with many previous studies (e.g., Jones et al., 2004; Jones & Macken, 1993) it was found that both changing-state speech (Experiments 1 and 3) and changing-state tones (Experiment 2) consistently impaired serial short-term memory compared to the respective steady-state conditions. Moreover, it was shown for the first time that this acoustical interference is reflected also in trial-by-trial metacognitive confidence judgments (Experiments 1 – 3), suggesting that participants are aware of the disruptive effect of task-irrelevant sound on serial recall from short-term memory.

In addition, three experiments of the present study also demonstrated an auditory deviation effect on serial recall performance (for ‘early’ deviants in Experiment 1 as well as in Experiments 3 and 4), which has been labelled a benchmark finding in working memory research (Oberauer et al., 2018). However, the present data also indicate that the deviation effect may be smaller and less reliable than other forms of auditory distraction (e.g., the changing-state effect) unless much larger sample sizes are used. In addition, some aspects of our data suggest that disruption of serial recall may be restricted to deviants presented during the encoding of the to-be-remembered items (see Exp. 1). Moreover, it remains to be confirmed whether a deviant in purely non-speech sounds (e.g., a deviant frequency in a sequence of tones as in Exp. 2) is less disruptive than a deviant in speech (e.g., a voice shift). Interestingly, the auditory deviation effect was reflected also in participants’ lower confidence judgments on deviant trials (Experiments 1 – 3), suggesting that participants were aware of the potentially detrimental effect of an auditory deviant on their memory processing even when it did not affect objective performance (as in Experiment 2). This indicates that metacognitive judgments may be more susceptible than objective recall accuracy to the processing of relatively unobtrusive auditory distractors such as a single deviant frequency in a task-irrelevant sequence of repeated tones.

Table 1. *Summary of results from Experiments 1 to 4 based on the significance of effects (Yes: $p < .05$; No: $p \geq .05$) in the frequentist analyses.*

Measure	**Changing-state effect**	**Deviation effect**	**Task load effect**	**Deviation effect reduced at high task load**
Experiment 1. Lab-based, speech sounds				
Recall	Yes	Yes[1]	Yes	No
CJs	Yes	Yes[1]	Yes	No
Experiment 2. Lab-based, non-speech sounds				
Recall	Yes	No	No	No
CJs	Yes	Yes	No	No
Experiment 3. Web-based, speech sounds				
Recall	Yes	Yes	Yes	No
CJs	Yes	Yes	Yes	No
Experiment 4. Web-based, speech sounds with voice deviant[2]				
Recall	--	Yes	Yes	No[3]

Note.
CJs = Confidence judgments
[1] For ‘early’ deviants presented during encoding.
[2] Replication of Experiment 1 from Hughes et al. (2013) as a web-based experiment.
[3] The deviation effect was significantly larger at high load than at low load.

In contrast to several previous results (Hughes et al., 2013; Hughes & Marsh, 2019; Marsh et al., 2020), there was no indication in the present study of an attenuation of the deviation effect under conditions of enhanced task-encoding load (i.e., when visual noise was added to the to-be-remembered digits). In the first three experiments, an increase in perceptual task-encoding load affected neither the deviation effect nor the changing-state effect, suggesting that both types of distraction may be equally unrelated to attentional control. In particular, the observation that an increase in perceptual task-encoding load did not reduce the deviation effect is inconsistent with the directional predictions of the duplex-mechanism account (Hughes, 2014) and suggests that the previously reported modulation of the deviation effect by task load may not be replicable. More generally, the present data indicate that the cognitive disruptions produced by task-irrelevant sound could not be reduced through enhanced perceptual task load, thus avoiding processing of the irrelevant sound (Lavie, 2005, 2010). However, such findings are not unique. For instance, studies using the oddball paradigm requiring simple categorization responses (e.g., even/odd judgments of digits) rather than memorization have shown that the degree of attentional capture by a deviant sound does not depend on visual task-encoding load (i.e., visual degradation of the digits; Parmentier et al., 2008). More recently, similar results have been reported also for distraction in the serial recall task using monetary incentives, which are supposed to increase task engagement (Bell et al., 2020). In that study, providing monetary incentives did boost short-term memory performance in general, but it did not reduce the disruptive effect of changing-state sound (spoken words) and auditory deviants (a change in voice). In the present study, enhanced task-encoding load also affected short-term memory performance across all conditions in three out of four experiments, but it did not selectively reduce the effect of one type of auditory distraction. The present results may be consistent with an automatic attentional capture account, assuming that changing task-irrelevant sounds divert attention from the focal task as a result of an acoustically driven, automatic perceptual analysis (e.g., Körner et al., 2017; Parmentier, 2008). On the other hand, such an automaticity assumption is challenged by the fact that participants were clearly aware of the disruptive effects of irrelevant sound (see also Bell et al., 2021). Having set out to constrain the duplex-mechanism account proposed by (Hughes, 2014), our data in fact contradict it, suggesting an account whereby both the changing-state effect and the auditory deviation effect are explained with the same mechanism involving involuntary (though aware) attentional orienting responses to certain changes in a stream of irrelevant sound.

It could be argued, of course, that the sparing effects of task load for the deviation effect (as reported by Hughes et al., 2013) depends on the exact experimental context. The present study investigated the auditory deviation effect and the changing-state effect within the same experiments and with the same participants, whereas Hughes et al. (2013) tested the influence of task-encoding load on the deviation effect (Exp. 1 and 2) separately from the changing-state effect (Exp. 3a). It is possible that a deviant sound captures more attention in a between-subjects design when participants do not experience different types of sound without a deviant (i.e., steady-state and changing-state sequences). Moreover, as described in the introduction, Hughes et al. (2013) presented the deviants (a voice shift) in changing-state speech (varying spoken letters), whereas steady-state sequences (of speech or tones) were used as the control conditions in Experiments 1 – 3 of the present study (which may be a presentation procedure that is more similar to the oddball paradigm; e.g., Parmentier, 2008). It is less obvious why a deviant should capture more attention in changing-state background (Hughes et al., 2013) – and be more susceptible to manipulations of task load – than when it is embedded in steady-state background (current study). If anything, the occurrence of a deviant steady-state sequences of irrelevant sound may result in an even stronger violation of the listener's predictive model than a deviant in changing-state sequences (which should make the predictive model learn to expect greater acoustical variability). In any case, the present Experiments 1 and 3 demonstrate that deviants presented during encoding within steady-state sequences of speech produce statistically reliable disruption of serial recall, and that this disruption does not depend on task-encoding load. Further, the results of Experiment 4 indicate that our failure to replicate the previously reported elimination of the deviation effect with enhanced task-encoding load (Hughes et al., 2013) cannot be solely attributed to deviants occurring in a different context (e.g., a deviant within steady-state sound), as in this close replication of the original paradigm an auditory deviation effect was observed only under conditions of enhanced task difficulty, but not with low visual task-encoding load. While this finding is clearly inconsistent with an interpretation in terms of perceptual task load (and with the duplex-mechanism account), it could be speculated whether the visual degradation of to-be-remembered items enhanced cognitive task load (in particular in an online setting providing additional uncontrolled distraction), which may have prevented cognitive-inhibitory control of auditory distraction (compare Kattner, 2021; Marsh et al., 2012). Future research should endeavor to independently manipulate perceptual and cognitive load and examine their differential effects on the distraction produced by auditory deviants.

Another potential explanation for the results of Experiments 1 – 3 diverging from the findings of Hughes et al. (2013) is that the addition of confidence judgments altered task processing. In the field of metacognition such effects are referred to as reactivity effects and evidence for reactivity to monitoring judgments is mixed (see Double & Birney, 2019 for a review). Mitchum et al. (2016) identified three possible hypotheses regarding how collecting monitoring judgments may affect performance on the primary task: First, a positive reactivity effect may occur if the elicited monitoring judgment prompts participants to use more effective strategies. Second, the collection of monitoring judgments may lead to a change in participants' goals (to be less mastery-oriented). Third, a negative effect of reactivity could result from competition for resources between the primary task and providing the monitoring judgment (referred to as the dual-task hypothesis). Importantly, these authors, and much of the research in this field, focused on the impact of collecting Judgments of Learning during paired-associate learning. As Judgments of Learning are collected during study and before the memory test, they have considerable potential to interfere with memory recall. In our experimental design confidence judgments were collected after serial recall and as such could not directly influence strategy use on the same trial, making the impact of the first two hypotheses very unlikely. It is nevertheless conceivable that the requirement to provide confidence judgments draws cognitive resources away from memory processing. However, one would then expect overall poorer performance in our Experiments 1 – 3 as compared to a context where monitoring judgments were not collected, such as Experiment 4 and the other studies by Hughes et al. (2013). This was not the case and overall serial recall performance was very comparable across these studies, if not better in Experiments 1 – 3. Further, we find it unlikely that any reactivity effects of confidence judgments would be so selective as to specifically eliminate the effect of task load on the deviation effect. In conclusion, we do not consider reactivity effects a plausible alternative explanation for our divergent result pattern.

The collection of trial-by-trial confidence judgments provided novel insights into participants' metacognitive awareness of their cognitive performance as a function of different types of auditory distractors. Irrelevant sound effects and task-encoding load effects on participants' confidence judgments were largely consistent with their effects on objective memory performance, corroborating and extending recent findings of Bell and colleagues

(2021). Moreover, participants were sensitive to trial-by trial variations in their performance, but their metacognitive monitoring accuracy was largely unaffected by both the type of auditory distractor and task-encoding load. As such, in contrast to the predictions of the duplex-mechanism account (Hughes, 2014; Hughes & Marsh, 2019), there was no evidence that participants were ‘more aware’ of the impact of an auditory deviant (the diversion of attention) than of the impact of changing-state sound (interference-by-process) on serial recall from short-term memory.

Together with other recent findings (Bell et al., 2021), the equal sensitivity of metacognitive judgments to the changing-state and the deviation effect in the present series of experiments could be interpreted as further evidence against the dissociations assumed by the duplex-mechanism account of auditory distraction. Certainly, metacognitive awareness appears to be one more characteristic that does not dissociate between the different types of auditory distraction (in addition to the role of working memory capacity; Körner et al., 2017). Another property that did not dissociate between the two types of auditory distraction in the present study was attentional control, with both effects being equally unrelated to manipulations of task load (except for Experiment 4, which might indicate an involvement of both perceptual and cognitive load; see above). This observation is also consistent with previous findings showing that the two types of auditory distraction are equally unrelated to motivational task engagement (Bell et al., 2020). However, we argue that these findings alone cannot refute the duplex-mechanism account and we would caution against drawing strong conclusions regarding cognitive control and automaticity based solely on the present effects on metacognitive judgments. What remains unanswered by the approach taken in these experiments is to what extent participants are aware of *what* is causing these variations in performance. In principle, participants could be directly monitoring the success of their memory recall (i.e. via the strength of the memory trace, see Nelson & Narens, 1990) without being able to attribute this success to the presence or absence of an auditory deviant, for instance. In order to directly answer this question, the two types of metacognitive judgments – local (as collected here) and global (see Bell et al., 2021) – should be combined and contrasted within one study.

To sum up, three of the present experiments demonstrated reliable changing-state effects on both objective performances in a serial recall task and the participants' subjective confidence regarding task performance, regardless of the level of task-encoding load. In addition, the presence of an auditory deviant in steady-state speech was found to affect both objective performance and confidence judgments (in Experiments 1 and 3), and this effect was equally unrelated to task-encoding load. It is important to note that the changing-state and deviation effects observed with the purely web-based experiments (Experiments 3 and 4) are largely consistent not only with the in-person experiments of this study (at least in terms of the changing-state effect and the absence of a modulation by task load), but also with the large body of literature on auditory distraction, which is almost exclusively based on well-controlled laboratory experiments. However, in contrast to the literature, the auditory deviation effect observed in the web-based Experiment 4 was evident only in case of enhanced task load, suggesting that the low load condition may have enabled participants to alleviate the disruptive effect through cognitive control. Together, these findings suggest that although participants were aware of the impact of auditory distractors (reflected also in high monitoring accuracy), neither the changing-state effect nor the deviation effect can be attenuated through enhanced attentional control. It seems more likely that both types of auditory distraction can be explained with the automatic detection of novel sounds or acoustical changes leading to the diversion of attentional resources from the focal task (compare Bell et al., 2020, 2021; Körner et al., 2017).

Open Practices Statement

The data of all four experiments of this study (proportion correct, confidence judgments and monitoring accuracy aggregated for each participant and experimental condition) are openly available as csv-files in an Open Science Framework (OSF) repository at this view-only link: https://osf.io/s4ney. Additional information such as the scripts of the experimental routines or statistical analyses can be made available upon request.

References

Amichetti, N. M., Stanley, R. S., White, A. G., & Wingfield, A. (2013). Monitoring the capacity of working memory: Executive control and effects of listening effort. *Memory and Cognition*, *41*(6), 839–849. https://doi.org/10.3758/s13421-013-0302-0

Baddeley, A. D., & Hitch, G. J. (1974). Working Memory. In G. A. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (pp. 47–89). Academic Press. https://doi.org/10.1016/j.cub.2009.12.014

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*(3), 379–384. https://doi.org/10.3758/BF03192707

Beaman, C. P., Hanczakowski, M., & Jones, D. M. (2014). The effects of distraction on metacognition and metacognition on distraction: Evidence from recognition memory. *Frontiers in Psychology*, *5*, 439. https://doi.org/10.3389/fpsyg.2014.00439

Beaman, C. P., & Jones, D. M. (1997). Role of serial order in the irrelevant speech effect: Tests of the changing-state hypothesis. *Journal of Experimental Psychology: Learning Memory and Cognition*, *23*(2), 459–471. https://doi.org/10.1037/0278-7393.23.2.459

Bell, R., Mieth, L., Buchner, A., & Röer, J. P. (2020). Monetary incentives have only limited effects on auditory distraction: evidence for the automaticity of cross-modal attention capture. *Psychological Research*. https://doi.org/10.1007/s00426-020-01455-5

Bell, R., Mieth, L., Röer, J. P., & Buchner, A. (2021). The metacognition of auditory distraction: Judgments about the effects of deviating and changing auditory distractors on cognitive performance. *Memory & Cognition 2021*, 1–14. https://doi.org/10.3758/S13421-021-01200-2

Bell, R., Röer, J. P., Dentale, S., & Buchner, A. (2012). Habituation of the irrelevant sound effect:

Evidence for an attentional theory of short-term memory disruption. *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*(6), 1542–1557. https://doi.org/10.1037/a0028459

Bell, R., Röer, J. P., Lang, A. G., & Buchner, A. (2019). Distraction by steady-state sounds: Evidence for a graded attentional model of auditory distraction. *Journal of Experimental Psychology: Human Perception and Performance*, *45*(4), 500–512. https://doi.org/10.1037/xhp0000623

Bell, R., Röer, J. P., Marsh, J. E., Storch, D., & Buchner, A. (2017). The effect of cognitive control on different types of auditory distraction: A preregistered study. *Experimental Psychology*, *64*(5), 359–368. https://doi.org/10.1027/1618-3169/a000372

Berti, S., & Schröger, E. (2003). Working memory controls involuntary attention switching: Evidence from an auditory distraction paradigm. *European Journal of Neuroscience*, *17*(5), 1119–1122. https://doi.org/10.1046/j.1460-9568.2003.02527.x

Bodenhofer, U., & Klawonn, F. (2008). Robust rank correlation coefficients on the basis of fuzzy. *Mathware & Soft Computing*, *15*(1), 5–20.

Bodenhofer, U., Krone, M., & Klawonn, F. (2013). Testing noisy numerical data for monotonic association. *Information Sciences*, *245*, 21–37. https://doi.org/10.1016/j.ins.2012.11.026

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 443–446. https://doi.org/10.1163/156856897X00357

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound* (4th ed.). MIT Press.

Colle, H. A., & Welsh, A. (1976). Acoustic masking in primary memory. *Journal of Verbal Learning and Verbal Behavior*, *15*(1), 17–31. https://doi.org/10.1016/S0022-

5371(76)90003-7

Cowan, N. (1995). Attention and Memory: An Integrated Framework. In *Attention and Memory: An Integrated Framework*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780195119107.001.0001

Cowan, N. (1999). An Embedded-Processes Model of Working Memory. In *Models of Working Memory: Mechanisms of active maintenance and executive control* (pp. 62–101). Cambridge University Press. https://doi.org/10.1017/cbo9781139174909.006

Double, K. S., & Birney, D. P. (2019). Reactivity to measures of metacognition. *Frontiers in Psychology*, *10*, 2755. https://doi.org/10.3389/fpsyg.2019.02755

Ellermeier, W., & Zimmer, K. (1997). Individual differences in susceptibility to the “irrelevant speech” effect. *Journal of the Acoustical Society of America*, *102*, 2191–2199. https://doi.org/10.1121/1.419596

Ellermeier, W., & Zimmer, K. (2014). The psychoacoustics of the irrelevant sound effect. *Acoustical Science and Technology*, *35*(1), 10–16. https://doi.org/10.1250/ast.35.10

Elliott, E. M. (2002). The irrelevant-speech effect and children: Theoretical implications of developmental change. *Memory and Cognition*, *30*, 478–487.

Escera, C., Alho, K., Winkler, I., & Näätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience*, *10*(5), 590–604. https://doi.org/10.1162/089892998562997

Getzmann, S., Gajewski, P. D., Hengstler, J. G., Falkenstein, M., & Beste, C. (2013). BDNF Val66Met polymorphism and goal-directed behavior in healthy elderly — evidence from auditory distraction. *NeuroImage*, *64*(1), 290–298. https://doi.org/10.1016/J.NEUROIMAGE.2012.08.079

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

Hughes, R. W., Hurlstone, M. J., Marsh, J. E., Vachon, F., & Jones, D. M. (2013). Cognitive control of auditory distraction: impact of task difficulty, foreknowledge, and working memory capacity supports duplex-mechanism account. *Journal of Experimental Psychology. Human Perception and Performance*, *39*(2), 539–553. https://doi.org/10.1037/a0029064

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

Hughes, R. W., & Marsh, J. E. (2019). Dissociating two forms of auditory distraction in a novel Stroop serial recall experiment. *Auditory Perception and Cognition*, *2*(3), 129–142.

Hughes, R. W., & Marsh, J. E. (2020). When is forewarned forearmed? Predicting auditory distraction in short-term memory. *Journal of Experimental Psychology: Learning Memory and Cognition*, *46*(3), 427–442. https://doi.org/10.1037/xlm0000736

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. https://doi.org/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. https://doi.org/10.1037/0278-7393.33.6.1050

Jeffreys, H. (1961). *Theory of Probability*. Clarendon Press.

Jones, D. M., Alford, D., Bridges, A., Tremblay, S., & Macken, B. (1999). Organizational factors in selective attention: The interplay of acoustic distinctiveness and auditory streaming in the irrelevant sound effect. *Journal of Experimental Psychology: Learning Memory and Cognition*, *25*(2), 464–473. https://doi.org/10.1037/0278-7393.25.2.464

Jones, D. M., Beaman, C. P., & Macken, W. J. (1996). The object-oriented episodic record model. In S. E. Gathercole (Ed.), *Models of short-term memory* (pp. 209–238). Psychology Press.

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*, 369–381. https://doi.org/10.1037/0278-7393.19.2.369

Jones, D. M., Macken, W. J., & Mosdell, N. A. (1997). The role of habituation in the disruption of recall performance by irrelevant sound. *British Journal of Psychology*, *88*, 549–564. https://doi.org/10.1111/j.2044-8295.1997.tb02657.x

Jones, D. M., Macken, W. J., & Nicholls, A. P. (2004). The phonological store of working memory: Is it phonological and is it a store? *Journal of Experimental Psychology: Learning Memory and Cognition*, *30*, 656–674. https://doi.org/10.1037/0278-7393.30.3.656

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

Kattner, F. (2021). Transfer of working memory training to the inhibitory control of auditory distraction. *Psychological Research*.

Kattner, F., & Ellermeier, W. (2018). Emotional prosody of task-irrelevant speech interferes with the retention of serial order. *Journal of Experimental Psychology: Human Perception and Performance*, *44*(8), 1303–1312. https://doi.org/10.1037/xhp0000537

Kattner, F., & Meinhardt, H. (2020). Dissociating the Disruptive Effects of Irrelevant Music and Speech on Serial Recall of Tonal and Verbal Sequences. *Frontiers in Psychology*, *11*, 346. https://doi.org/10.3389/fpsyg.2020.00346

Kleiner, M., Brainard, D. H., Pelli, D. G., Broussard, C., Wolf, T., & Niehorster, D. (2007). What's new in Psychtoolbox-3? *Perception*, *36*(14), 1–16. https://doi.org/10.1068/v070821

Körner, U., Röer, J. P., Buchner, A., & Bell, R. (2017). Working memory capacity is equally unrelated to auditory distraction by changing-state and deviant sounds. *Journal of Memory and Language*, *96*, 122–137. https://doi.org/10.1016/j.jml.2017.05.005

Lavie, N. (1995). Perceptual Load as a Necessary Condition for Selective Attention. *Journal of Experimental Psychology: Human Perception and Performance*, *21*(3), 451–468. https://doi.org/10.1037/0096-1523.21.3.451

Lavie, N. (2005). Distracted and confused?: Selective attention under load. In *Trends in Cognitive Sciences* (Vol. 9, Issue 2, pp. 75–82). Elsevier. https://doi.org/10.1016/j.tics.2004.12.004

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, *19*(3), 143–148. https://doi.org/10.1177/0963721410370295

Lavie, N., & De Fockert, J. W. (2005). The role of working memory in attentional capture. *Psychonomic Bulletin & Review*, *12*(4), 669–674. https://doi.org/10.3758/BF03196756

Lavie, N., Hirst, A., De Fockert, J. W., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General*, *133*(3), 339–354. https://doi.org/10.1037/0096-3445.133.3.339

Leiva, A., Andrés, P., Servera, M., Verbruggen, F., & Parmentier, F. B. R. (2016). The role of age, working memory, and response inhibition in deviance distraction: A cross-sectional study. *Developmental Psychology*, *52*(9), 1381–1393. https://doi.org/10.1037/dev0000163

Marois, A., & Vachon, F. (2018). Can pupillometry index auditory attentional capture in contexts of active visual processing? *Journal of Cognitive Psychology*, *30*(4), 484–502. https://doi.org/10.1080/20445911.2018.1470518

Marsh, J. E., Beaman, C. P., Hughes, R. W., & Jones, D. M. (2012). Inhibitory control in memory: Evidence for negative priming in free recall. *Journal of Experimental Psychology: Learning Memory and Cognition*, *38*(5), 1377–1388. https://doi.org/10.1037/a0027849

Marsh, J. E., Campbell, T., Vachon, F., Taylor, P., & Hughes, R. W. (2020). How the deployment of visual attention modulates auditory distraction. *Attention, Perception, and Psychophysics*, *82*, 350–362. https://doi.org/10.3758/s13414-019-01800-w

Marsh, J. E., Sörqvist, P., & Hughes, R. W. (2015). Dynamic cognitive control of irrelevant sound: Increased task engagement attenuates semantic auditory distraction. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(5), 1462–1474. https://doi.org/10.1037/xhp0000060

Marsh, J. E., Yang, J., Qualter, P., Richardson, C., Perham, N., Vachon, F., & Hughes, R. W. (2018). Postcategorical auditory distraction in short-term memory: Insights from increased task load and task type. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(6), 882–897. https://doi.org/10.1037/xlm0000492

Miller, J. D. (1974). Effects of noise on people. *Journal of the Acoustical Society of America*, *56*(3), 729–764. https://doi.org/10.1121/1.1903322

Mitchum, A. L., Kelley, C. M., & Fox, M. C. (2016). When asking the question changes the ultimate answer: Metamemory judgments change memory. *Journal of Experimental Psychology: General*, *145*(2), 200–219. https://doi.org/10.1037/a0039923

Morey, R. D., Romeijn, J. W., & Rouder, J. N. (2016). The philosophy of Bayes factors and the quantification of statistical evidence. *Journal of Mathematical Psychology*, *72*, 6–18. https://doi.org/10.1016/j.jmp.2015.11.001

Murphy, G., Groeger, J. A., & Greene, C. M. (2016). Twenty years of load theory—Where are we now, and where should we go next? *Psychonomic Bulletin and Review*, *23*(5), 1316–1340. https://doi.org/10.3758/S13423-015-0982-5

Nelson, T. O., & Narens, L. (1990). Metamemory: A Theoretical Framework and New Findings. In G. H. Bower (Ed.), *Psychology of Learning and Motivation - Advances in Research and Theory* (Vol. 26, Issue C, pp. 125–173). Academic Press. https://doi.org/10.1016/S0079-7421(08)60053-5

Nittono, H. (1997). Background instrumental music and serial recall. *Perceptual and Motor Skills*, *84*(3), 1307–1313. https://doi.org/10.2466/pms.1997.84.3c.1307

Oberauer, K., Lewandowsky, S., Awh, E., Brown, G. D. A., Conway, A., Cowan, N., Donkin, C., Farrell, S., Hitch, G. J., Hurlstone, M. J., Ma, W. J., Morey, C. C., Nee, D. E., Schweppe, J., Vergauwe, E., & Ward, G. (2018). Benchmarks for models of short-term and working memory. *Psychological Bulletin*, *144*, 885–958. https://doi.org/10.1037/bul0000153

Parmentier, F. B. R. (2008). Towards a cognitive model of distraction by auditory novelty: The role of involuntary attention capture and semantic processing. *Cognition*, *109*(3), 345–362. https://doi.org/10.1016/j.cognition.2008.09.005

Parmentier, F. B. R. (2014). The cognitive determinants of behavioral distraction by deviant auditory stimuli: A review. *Psychological Research*, *78*, 321–338.

https://doi.org/10.1007/s00426-013-0534-4

Parmentier, F. B. R. (2016). Deviant sounds yield distraction irrespective of the sounds' informational value. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 837–846. https://doi.org/10.1037/XHP0000195

Parmentier, F. B. R., Elford, G., Escera, C., Andrés, P., & Miguel, I. S. (2008). The cognitive locus of distraction by acoustic novelty in the cross-modal oddball task. *Cognition*, *106*(1), 408–432. https://doi.org/10.1016/j.cognition.2007.03.008

Parmentier, F. B. R., Pacheco-Unguetti, A. P., & Valero, S. (2018). Food words distract the hungry: Evidence of involuntary semantic processing of task-irrelevant but biologically-relevant unexpected auditory words. *PLoS ONE*, *13*(1), e0190644. https://doi.org/10.1371/journal.pone.0190644

Peirce, J. W. (2007). PsychoPy-Psychophysics software in Python. *Journal of Neuroscience Methods*, *162*(1–2), 8–13. https://doi.org/10.1016/j.jneumeth.2006.11.017

Peirce, J. W. (2009). Generating stimuli for neuroscience using PsychoPy. *Frontiers in Neuroinformatics*, *2*(10), 1–8. https://doi.org/10.3389/neuro.11.010.2008

Peirce, J. W., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, *51*, 195–203. https://doi.org/10.3758/s13428-018-01193-y

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*(4), 437–442. https://doi.org/10.1163/156856897X00366

Röer, J. P., Bell, R., & Buchner, A. (2013). Self-relevance increases the irrelevant sound effect: Attentional disruption by one's own name. *Journal of Cognitive Psychology*, *25*(8), 925–931. https://doi.org/10.1080/20445911.2013.828063

Röer, J. P., Bell, R., & Buchner, A. (2015). Specific foreknowledge reduces auditory distraction by irrelevant speech. *Journal of Experimental Psychology: Human Perception and Performance*, *41*(3), 692–702. https://doi.org/10.1037/xhp0000028

Röer, J. P., Körner, U., Buchner, A., & Bell, R. (2017). Attentional capture by taboo words: A functional view of auditory distraction. *Emotion*, *17*(4), 740–750. https://doi.org/10.1037/emo0000274

Röer, J. P., Rummel, J., Bell, R., & Buchner, A. (2017). Metacognition in Auditory Distraction: How Expectations about Distractibility Influence the Irrelevant Sound Effect. *Journal of Cognition*, *1*(1), 2. https://doi.org/10.5334/joc.3

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*, 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

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

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

Schlittmeier, S. J., Hellbrück, J., & Klatte, M. (2008). Does irrelevant music cause an irrelevant sound effect for auditory items? *European Journal of Cognitive Psychology*, *20*, 252–271. https://doi.org/10.1080/09541440701427838

Sörqvist, P. (2010). 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 and Cognition*, *38*(5), 651–658. https://doi.org/10.3758/MC.38.5.651

Sörqvist, P., Marsh, J. E., & Nöstl, A. (2013). High working memory capacity does not always attenuate distraction: Bayesian evidence in support of the null hypothesis. *Psychonomic Bulletin and Review*, *20*(5), 897–904. https://doi.org/10.3758/s13423-013-0419-y

Stoet, G. (2010). PsyToolkit: A software package for programming psychological experiments using Linux. *Behavior Research Methods*, *42*(4), 1096–1104. https://doi.org/10.3758/BRM.42.4.1096

Stoet, G. (2017). PsyToolkit: A Novel Web-Based Method for Running Online Questionnaires and Reaction-Time Experiments. *Teaching of Psychology*, *44*(1), 24–31. https://doi.org/10.1177/0098628316677643

Sussman, E., Winkler, I., & Schröger, E. (2003). Top-down control over involuntary attention switching in the auditory modality. *Psychonomic Bulletin and Review*, *10*, 630–637. https://doi.org/10.3758/BF03196525

Tremblay, S., & Jones, D. M. (1998). Role of habituation in the irrelevant sound effect: Evidence from the effects of token set size and rate of transition. *Journal of Experimental Psychology: Learning Memory and Cognition*, *24*(3), 659–671. https://doi.org/10.1037/0278-7393.24.3.659

Tremblay, S., & Jones, D. M. (1999). Change of intensity fails to produce an irrelevant sound effect: Implications for the representation of unattended sound. *Journal of Experimental Psychology: Human Perception and Performance*, *25*(4), 1005–1015. https://doi.org/10.1037/0096-1523.25.4.1005

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*, 164–177. https://doi.org/10.1037/a0025054

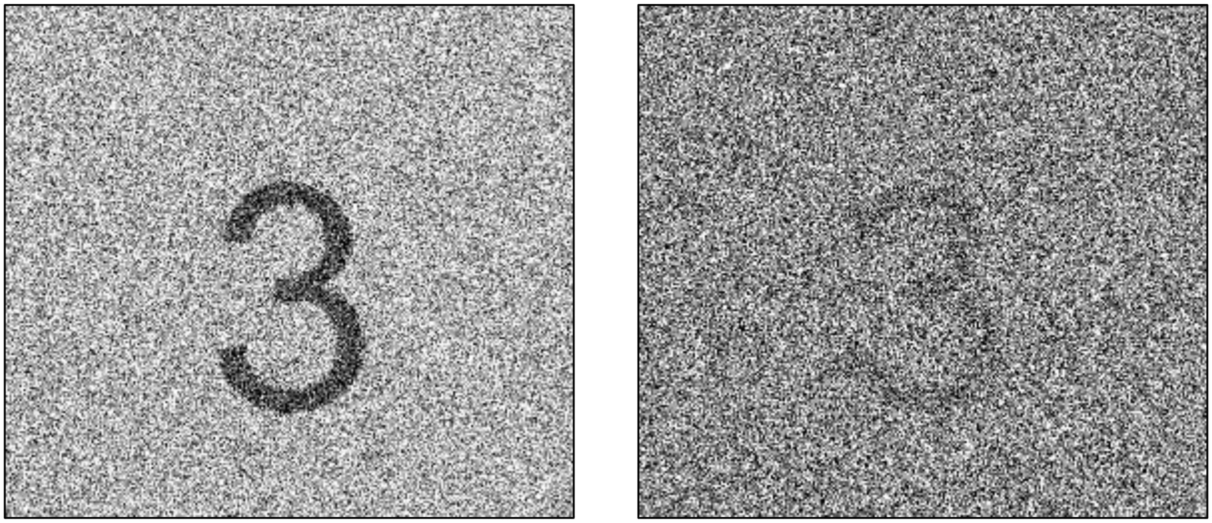
Wessel, J. R., & Aron, A. R. (2013). Unexpected events induce motor slowing via a brain mechanism for action-stopping with global suppressive effects. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, *33*(47), 18481–18491. https://doi.org/10.1523/JNEUROSCI.3456-13.2013

Wessel, J. R., Jenkinson, N., Brittain, J.-S., Voets, S. H. E. M., Aziz, T. Z., Aron, A. R., & Parmentier, F. B. R. (2016). Surprise disrupts cognition via a fronto-basal ganglia suppressive mechanism. *Nature Communications 2016 7:1*, *42*(1), 837–846. https://doi.org/10.1038/ncomms11195

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

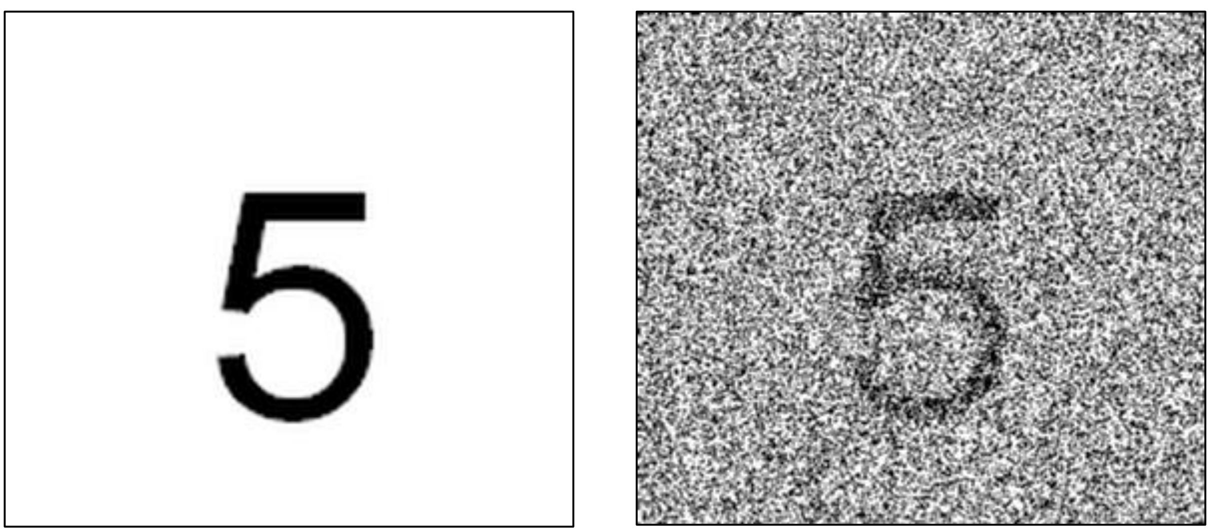
Zhang, Z., & Yuan, K.-H. (2018). *Practical Statistical Power Analysis Using Webpower and R*. ISDSA Press. https://webpower.psychstat.org/wiki/

Appendix



Low task load **High task load**

Figure A1. Illustration of a to-be-remembered digit as presented on the screen with static Gaussian noise of either low or high variance being added to the image on trials with low and high task load, respectively (Experiments 1 – 3).



Low task load **High task load**

Figure A2. Illustration of a to-be-remembered digit as it was presented on trials with low (left) and high (right) task load in the close replication Experiment 4.