

RESEARCH ARTICLE

Registered Replication Report: Testing Disruptive Effects of Irrelevant Speech on Visual-Spatial Working Memory

A Partial Replication of “Functional Equivalence of Verbal and Spatial Information in Serial Short-Term Memory (Jones, Farrand, Stuart, & Morris, 1995; Experiment 4)”

Tatiana Kvetnaya

The irrelevant speech effect (ISE)—the phenomenon that background speech impairs serial recall of visually presented material—has been widely used for examining the structure of short-term memory. In Experiment 4, Jones, Farrand, Stuart, and Morris (1995) employed the ISE to demonstrate that impairment of performance is determined by the changing-state characteristics of the material, rather than its modality of origin. The present study directly replicated the spatial condition of Experiment 4 with $N = 40$ German participants. In contrast to the original findings, no main effect of sound type was observed, $F(2, 78) = 0.81$, $p = .450$, $\eta^2_p = .02$. The absence of an ISE in the spatial domain does not support the changing state hypothesis.

Keywords: irrelevant speech effect; short-term memory; visual-spatial memory; changing state hypothesis; replication study

Original citation

Jones, D. M., Farrand, P., Stuart, G., & Morris, N. (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4), 1008–1018. DOI: <https://doi.org/10.1037/0278-7393.21.4.1008>

Introduction

The irrelevant speech effect (ISE)—the well-established phenomenon that background speech interferes with serial recall of visually presented material—is considered a fruitful paradigm for examining the structure of short-term memory (Banbury, Macken, Tremblay, & Jones, 2001). Yet there is an ongoing debate about which working memory models are most suitable to explain it (e.g., Baddeley, 2000; Jones & Tremblay, 2000; Neath, 2000). Salamé and Baddeley (1982) proposed an explanation based on a modular theory of working memory (Baddeley & Hitch, 1974). This model posits that verbal and spatial information is processed in modality-specific subsystems, namely the phonological loop and the visuospatial sketchpad, which are coordinated by the central executive (Baddeley, 1996; Guérard & Tremblay, 2008). According to this account, interference arises when two concurrent activities are similar in content (Schendel, 2006) and thus rely on the same subsystem: For example, when participants hear irrelevant

speech during verbal serial recall tasks, their recall performance is poorer compared to a silent condition because both activities are processed in the phonological loop (Baddeley, 1996; Guérard & Tremblay, 2008).

By contrast, Jones, Farrand, Stuart, and Morris (1995) suggested that interference results from a modality-independent similarity of process rather than content (Beaman & Jones, 1997; Schendel, 2006). They proposed a unitary model of working memory, the *object-oriented episodic record (O-OER) model*, to explain the ISE. In this framework, concurrent activities are represented as streams of abstract, amodal basic units—so-called objects—in a single unitary representational space (Jones, Beaman, & Macken, 1996). The organization of the to-be-remembered objects, rather than their modality of origin, is the pivotal factor in determining the pattern of recall. More specifically, the degree of disruption is determined by the amount of serial order information that concurrent activities contain.

To demonstrate this, Jones et al. (1995) conducted a series of four experiments in which they equated verbal and spatial memory tasks. In the verbal condition, participants recalled a randomly generated sequence of seven letters, while in the visual-spatial condition, the same task was performed with a sequence of seven dots presented in random locations on the screen. In line with their assumptions, serial order recall in both domains was equally susceptible to interference from both a secondary spatial task (rote tapping in Experiment 2), a secondary verbal task (mouthed articulatory suppression in Experiment 3), as well as from irrelevant speech in Experiment 4. Moreover, disruption was more

marked if the interference conditions involved a changing sequence of actions or materials, but not if a single event (tap, mouthed utterance, or sound) was repeated. The use of the irrelevant speech paradigm in Experiment 4 was crucial for ruling out the alternative interpretation that the interference found in Experiments 2 and 3 may stem from a central-executive involvement (Jones et al., 1995).

This *changing state hypothesis* was called into question by at least three conceptual replications (Guérard & Tremblay, 2008; Klatte & Hellbrück, 1997; Meiser & Klauer, 1999) and one direct replication (Guitard & Saint-Aubin, 2015) of the experiments in Jones et al. (1995). Both Guérard and Tremblay (2008) as well as Meiser and Klauer (1999) observed some detrimental effect of tapping as well as articulatory suppression on verbal and spatial memory tasks. While the pattern of errors was found to be similar for both tasks, in contrast to Jones et al.'s (1995) findings, interference across domains was smaller than interference within domains. Likewise, a direct replication of Experiments 2 and 3 (Guitard & Saint-Aubin, 2015) reproduced the main findings of Jones et al. (1995) only partially: Changing state was more detrimental than steady state in a rote tapping secondary task, but no such difference between changing and steady state interference conditions occurred in the case of an articulatory suppression task. Again, contrary to the original results and in line with a modular theory of working memory, interference within the same domain was more marked than across domains. Finally, in their conceptual replication of Experiment 4, Klatte and Hellbrück (1997) compared the effect of a single speaker (changing state) vs. several speakers (steady state) on the serial recall of digits vs. Corsi block sequences. While the ISE emerged in the case of the verbal digit task, no differential effect of irrelevant sound on performance in the spatial Corsi block task was observed. A follow-up study employing a spatial task nearly identical to the one used by Jones et al. (1995) found, once again, no ISE in the spatial domain (Klatte & Hellbrück, 1997).

Despite the crucial role of Experiment 4 in the original authors' argument for the changing state hypothesis, however, the unpublished studies by Klatte and Hellbrück (1997) remain its only known replications up to this date. Moreover, while the ISE was extensively investigated in regards to verbal serial recall tasks, this paradigm was rarely employed in a spatial domain (Tremblay, Saint-Aubin, & Jalbert, 2006). Yet, as stressed by Jones et al. (1995), any effect of irrelevant speech on serial spatial memory is strong support for a theory of interference based on functional equivalence of material originating from different modalities.

Therefore, the present study aimed to replicate the spatial condition of Experiment 4. The main objectives were (1) to establish whether the ISE would occur in a spatial domain and (2) to examine whether the changing state hypothesis would be supported in this domain.

Target of Replication

A successful replication would find a main effect of sound type on performance in the spatial memory task. In addition, the changing state condition should be significantly more disruptive than the steady state condition in regards to serial order errors.

Method

Power Analysis and Sample

To determine the size of the planned sample, a power simulation based on the key statistics provided in the original paper and an alpha of .05 was conducted. In each simulation run, a virtual data set was generated based on the means shown in Figure 5 of the original study, and the reported mean square errors. The analysis yielded that a target sample size of $N = 40$ ensures over 95% power to detect the main effect. An executable R script of the power analysis can be found in the Supplementary Material.

In total, 41 participants were recruited through an online social networking site and personal contact from May through June 2017. Prerequisites for participation included normal or corrected vision and hearing ability, as in the original study. An additional exclusion criterion was put in place, which posits that participants who responded incorrectly on all serial positions of a trial sequence in 50% of all trials or more would be excluded from data analysis. This criterion was chosen to rule out possible bias of the results due to a lack of attentiveness or incorrect understanding of instructions by the participants. None of the participants met this exclusion criterion.

One participant was excluded from data analysis due to reporting a hearing impairment. Thus, the final sample consisted of $N = 40$ participants (33 women and 7 men) with an age range of 18–53 years, $M = 22.27$ years ($SD = 5.98$ years). They were predominantly psychology students from the University of Tuebingen who received course credit in turn for their participation. Unlike in the original study, participants were additionally rewarded with an honorarium of maximum €10 depending on their performance. For example, if a participant correctly recalled the position in the respective sequences for 75% of all dots displayed, they were rewarded with €7.50, etc. This procedure intended to increase the participants' attentiveness and motivation during the completion of the task, and therefore, the quality of the obtained data. A preliminary small-scale pilot study ($N = 3$) was conducted to check whether the use of financial reimbursement might cause floor or ceiling effects on performance. The resulting error rates (range: 4–10 out of possible 16 errors per serial position and sound type condition, $M = 6.56$ errors, $SD = 2.69$ errors) closely resembled those in the original study, indicating that no floor or ceiling effects are likely to occur. No prior knowledge from the participants is assumed.

Material

Spatial memory task. Before the experimental procedure participants received written instructions explaining the task. They were asked to respond as quickly and accurately as possible, and specifically instructed not to use verbal strategies for holding the items in mind. The items were presented on a 48 cm (19") computer screen (Eizo FlexScan S1921). Participants clicked on a green "start" button on the screen to begin each trial. A sequence of seven dots was presented in a quasi-random position generated within a 500×500 matrix. The randomly generated dots had a radius of 12 units and could not be located closer

than 85 units on either axis of the matrix. In such a case, another set of coordinates would be generated by the software. During each trial, the sequence was displayed at a rate of one dot per two seconds (one second “on” and one second “off”). After a 10-second retention interval, all dots in the sequence were displayed simultaneously in their original position. Participants were asked to reproduce the item sequence using a mouse by clicking on each dot in the order of its original presentation. Once a dot was selected, its shading changed to indicate that the response was registered by the computer. The selection could not be reversed or altered once it was done. After a participant selected all dots, the green button to begin the next trial was displayed again. They were given practice for three trials before the start of the 48 experimental trials.

All characteristics of the task in the present replication study were identical with the original study except for the following changes: First, matrix array size was adjusted from the original 300×300 units and minimum distance of 60 units to 500×500 units and a minimum distance of 85 units on either axis of the matrix. With a given resolution of 1280×1024 pixels on the 48 cm (19”) screen, this adjustment ensured that the size of the resulting stimulus square corresponded roughly to the original measures of 16.5×16.5 cm (Jones et al., 1995: p. 1010). Second, in addition to the original task, participants were informed that they had completed one and two thirds of the session after 16 and 32 trials, respectively, and that they may take a short break. This served to maintain the participants’ attentiveness and motivation throughout the duration of the experiment.

Irrelevant speech. Speech in a female voice spoken at two consonants per second was recorded digitally in the present study. Two experimental conditions were created: The steady state condition consisted of the repeated syllable “ah” and the changing state condition used the seven constants in the alphabetic sequence “a” to “g”. In the latter condition, the order of the consonants was fixed, but when played back to the participants, the recording started from a random point in the sequence. Both recordings were played throughout the presentation and retention phases of the task, and stopped at the beginning of the recall phase. The stimuli were played back to the participants via headphones (Beyerdynamic DT 990 Pro). Participants were instructed to ignore any sound they heard and were reassured that the contents of the sound would not be tested for throughout the procedure. The order of the sound type conditions (changing state vs. steady state vs. silent control condition) was randomized from trial to trial.

All characteristics of this task in the present replication study were identical with the original study, with the exception that the auditory stimuli were spoken in a female voice instead of a male voice, because no other recording of comparably high quality could be obtained.

Design and Procedure

Since the present replication study regarded only the spatial memory task condition of Experiment 4, participants were tested in a within-subjects design. Exactly following the original study, the independent variable was

the sound type (changing state vs. steady state vs. silent control condition). The dependent variable was the number of serial order errors, measured for each of the seven serial positions in the spatial memory task.

The experiment and the instructions were carried out by the author of the present replication report. Participants were tested individually in a sound-insulated booth in the Psychoacoustics Lab at the University of Tuebingen (see Wickelmaier, 2013, for details). Stimulus presentation and data collection were controlled by software written in Python/Psychopy (Peirce, 2007) running on a Linux operating system. Forty-eight trials were run in total, with a total of 16 trials per each of the three sound type conditions. The sound type condition was assigned randomly by the software from trial to trial. In addition, information about participants’ age, gender, and reaction time between all clicks was collected. However, reaction times were not used in the data analysis of the present study. After completion of the experiment, participants received their respective financial reward. This amount was calculated directly after completion of the procedure. Participants were verbally informed about the purpose of the study upon request.

All statistical analyses were conducted with R (R Core Team, 2017). All experimental material, including the Python code of the experiment, the auditory stimuli, instructions for participants, and R code can be found in the Supplementary Material.

Results

Across all sound conditions, participants achieved an overall error rate of $M = 6.54$ ($SD = 2.82$) errors out of a maximum of 16 errors per sound type and serial position. They achieved a mean error rate of $M = 6.68$ ($SD = 1.39$) errors in the changing state condition, $M = 6.35$ ($SD = 1.25$) errors in the steady state condition, and $M = 6.58$ ($SD = 1.33$) errors in the silent control condition, respectively.

To test the hypothesis that sound type affects the serial recall error rate in the spatial memory task, a two-way repeated measures analysis of variance (ANOVA) with the factors sound type (changing state vs. steady state vs. silent control condition) and serial position (seven levels) was conducted. It yielded no main effect of sound type, $F(2, 78) = 0.807$, $p = .450$, and no interaction between sound type and serial position, $F(12, 468) = 0.595$, $p = .847$. Although the overall F -test was not significant, for the sake of comparability with the original study, a paired t -test was conducted to test the hypothesis that error rates in the changing state condition significantly differ from those in the steady state condition. The t -test did not yield a statistically significant difference between the changing state and steady state conditions, $t(39) = -1.222$, $p = .229$.

Figure 1 displays these results alongside Jones et al.’s (1995) original results. The respective effect sizes and test statistics are compared in **Table 1**.

Discussion

The present study constitutes the first pre-registered direct replication of an experiment by Jones et al. (1995), who employed the irrelevant speech paradigm in a visual-spa-

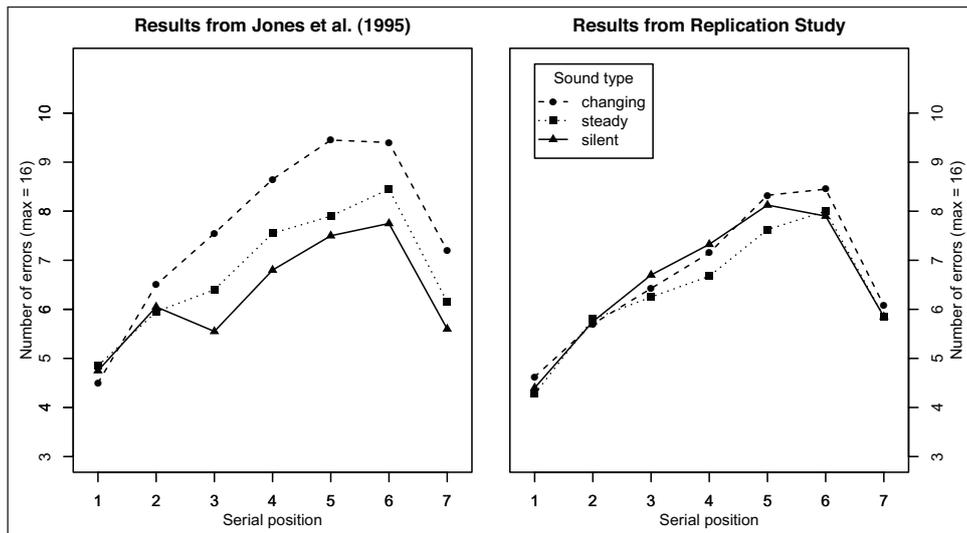


Figure 1: Effects of irrelevant speech on error rates per sound type condition (changing state vs. steady state vs. silent control condition) and serial position (1–7) in the spatial task condition of Experiment 4 by Jones et al. (1995; left panel) and the present replication study (right panel).

Table 1: Effect sizes and test statistics of Jones et al. (1995) and of the present replication study.

Effect or comparison	Jones et al. (1995)		Replication	
	Effect Size	Test statistic	Effect size	Test statistic
Main effect of sound type	$\eta^2_p = .27$	$F(2, 68) = 12.36, p < .001$	$\eta^2_p = .020$	$F(2, 78) = 0.807, p = .450$
Difference between changing and steady state conditions	$\eta^2_p = .16$	$F(1, 68) = 12.78, p < .001$	$\eta^2_p = .037$	$t(39) = -1.222, p = .229$

Note. Effect sizes from Jones et al. (1995) were converted from the test statistics provided in the original paper using methods described in Lakens (2013).

tial domain. In order for the present replication to be considered successful, the occurrence of (1) a main effect of sound type on performance in the spatial serial recall task, and (2) a higher error rate in the changing state condition compared to the steady state condition were expected. However, the results indicate that performance was not differentially impaired as a function of the three sound type conditions (changing state, steady state, and silent control condition). The difference in effect sizes between both studies was large, with sound type accounting for 27% of variance in the original study, while accounting for only 2% of variance in the present study. The results further indicate that performance did not differ between changing state and steady state conditions. Again, a considerable difference in effect sizes could be observed, with partial eta-squared values amounting to 16% in the original study and 3.7% in the present study, respectively. In summary, these results do not fit the predictions derived from Jones et al.'s (1995) changing state hypothesis in the context of the O-OER model, which supposes that irrelevant speech possessing changing state characteristics is more disruptive than both steady state speech and silence.

These diverging findings might be either due to methodological differences or due to a non-replicability of the original findings. Notable differences between the original study and the present replication included, firstly, the use of a female voice instead of a male voice to create the irrelevant speech conditions, and secondly,

the performance-related financial reward the participants received. While there appear to be phonetic differences between typical male and female speech, such as higher pitch, frequency and pitch range (e.g., Pépiot, 2014), to the best of my knowledge, there is no evidence in the literature indicating that the use of a female voice would yield a differentially disruptive effect on short-term memory. Likewise, no such evidence is available in the case of financial reimbursement. The presence of bow shaped serial position curves demonstrating primacy and recency effects, as well as overall error rates resembling the original findings indicate that the basic processing mechanisms during the experimental task are comparable with the original study.

The results from the present replication study rather correspond to the findings of Klatte and Hellbrück (1997). In their conceptual replication study of Experiment 4, an interaction between task domain and sound condition was observed: While changing state speech was significantly more detrimental than both steady state speech and silence in a verbal serial recall task, no differences between these three sound conditions were observed for the performance in a spatial memory task. Taken together, these findings do not support a theory of interference assuming functional equivalence of material originating from different modalities. The present findings rather align with an explanation of the ISE based on the assumption of modularity of short-term memory. In

this framework, irrelevant speech will detrimentally affect serial recall of verbal material because both are processed in a shared domain-specific subsystem, such as the phonological loop. This would not be the case for the spatial task, as it would be processed in a distinct subsystem, the visual-spatial sketchpad (Baddeley, 1996).

Limitations

Other factors might have been present that led to an absence of the ISE. A possible limitation is that the present study replicated only the spatial condition of Experiment 4. The absence of a verbal task condition poses a risk of arguing in favor of the null hypothesis. While the estimated power of 95% or more serves to ensure that the risk of a Type II error is low, it is worth noting that power was simulated on the basis of the effect sizes derived from Jones et al. (1995). Thus, it can only be considered valid if the original effect sizes accurately reflect the underlying effect.

In order to conclusively rule out any contextual effects, a comparison of the present findings with a verbal serial recall task condition using the same stimulus material would be warranted. If an ISE would occur in a verbal domain, resulting in an interaction between sound type and task domain—as in the case of Klatt and Hellbrück (1997)—this could allow for a more definite conclusion about the (non)modularity of short-term memory.

Supplementary Material

The Supplementary Material for this article can be found online at: <https://osf.io/hba2p/>.

Acknowledgement

I would like to express special thanks to my supervisor Florian Wickelmaier for his support throughout the completion of this thesis project.

Competing Interests

The author has no competing interests to declare.

References

- Baddeley, A. D.** (1996). The concept of working memory. In: Gathercole, S. E. (Ed.), *Models of short-term memory*, 1–27. Hove, UK: Psychology Press.
- Baddeley, A. D.** (2000). The phonological loop and the irrelevant speech effect: Some comments on Neath (2000). *Psychonomic Bulletin & Review*, 7(3), 544–549. DOI: <https://doi.org/10.3758/BF03214369>
- Baddeley, A. D., & Hitch, G.** (1974). Working memory. In: Bower, G. H. (Ed.), *Psychology of learning and motivation*, 8, 47–89. New York: Academic Press. DOI: [https://doi.org/10.1016/S0079-7421\(08\)60452-1](https://doi.org/10.1016/S0079-7421(08)60452-1)
- Banbury, S. P., Macken, W. J., Tremblay, S., & Jones, D. M.** (2001). Auditory distraction and short-term memory: Phenomena and practical implications. *Human Factors*, 43(1), 12–29. DOI: <https://doi.org/10.1518/001872001775992462>
- 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. DOI: <https://doi.org/10.1037/0278-7393.23.2.459>
- Guérard, K., & Tremblay, S.** (2008). Revisiting evidence for modularity and functional equivalence across verbal and spatial domains in memory. *Journal of Experimental Social Psychology: Learning, Memory, and Cognition*, 34(3), 556–569. DOI: <https://doi.org/10.1037/0278-7393.34.3.556>
- Guitard, D., & Saint-Aubin, J.** (2015). A replication of “Functional equivalence of verbal and spatial information in serial short-term memory (1995; Experiments 2 and 3)”. *The Quantitative Methods for Psychology*, 11(2), r4–r7. DOI: <https://doi.org/10.20982/tqmp.11.2.r004>
- Jones, D. M., Beaman, C. P., & Macken, W. J.** (1996). The object-oriented episodic record model. In: Gathercole, S. E. (Ed.), *Models of short-term memory*, 209–237. Hove, UK: Psychology Press.
- Jones, D. M., Farrand, P., Stuart, G., & Morris, N.** (1995). Functional equivalence of verbal and spatial information in serial short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 21(4), 1008–1018. DOI: <https://doi.org/10.1037/0278-7393.21.4.1008>
- Jones, D. M., & Tremblay, S.** (2000). Interference in memory by process or content? A reply to Neath. *Psychonomic Bulletin & Review*, 7(3), 550–558. DOI: <https://doi.org/10.3758/BF03214370>
- Klatte, M., & Hellbrück, J.** (1997). Effects of irrelevant speech on serial recall of verbal and spatial material. In: Schick, A., & Klatt, M. (Eds.), *Contributions to psychological acoustics. Results of the seventh Oldenburg symposium on psychological acoustics*, 531–538. Oldenburg: Bibliotheks- und Informationssystem der Universität Oldenburg.
- Lakens, D.** (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, 4. DOI: <https://doi.org/10.3389/fpsyg.2013.00863>
- Meiser, T., & Klauer, K. C.** (1999). Working memory and changing-state hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 25(5), 1272–1299. DOI: <https://doi.org/10.1037/0278-7393.25.5.1272>
- Neath, I.** (2000). Modeling the effects of irrelevant speech on memory. *Psychonomic Bulletin & Review*, 7(3), 403–423. DOI: <https://doi.org/10.3758/BF03214356>
- Pearce, J. W.** (2007). PsychoPy—psychophysics software in Python. *Journal of neuroscience methods*, 162(1), 8–13. DOI: <https://doi.org/10.1016/j.jneumeth.2006.11.017>
- Pépiot, E.** (2014). Male and female speech: a study of mean f0, f0 range, phonation type and speech rate in Parisian French and American English speakers. *Speech Prosody*, 7, 305–309. Retrieved from <https://halshs.archives-ouvertes.fr/halshs-00999332>
- R Core Team.** (2017). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from: <http://www.R-project.org/>.
- Salamé, P., & Baddeley, A. D.** (1982). Disruption of short-term memory by unattended speech: Impli-

cations for the structure of working memory. *Journal of Verbal Learning and Verbal Behavior*, 21, 150–164. DOI: [https://doi.org/10.1016/S0022-5371\(82\)90521-7](https://doi.org/10.1016/S0022-5371(82)90521-7)

Schendel, Z. A. (2006). *The irrelevant sound effect: Similarity of content or similarity of process?* (Doctoral dissertation), The Ohio State University, Retrieved from: https://etd.ohiolink.edu/rws_etd/document/get/osu1148590088/inline.

Tremblay, S., Saint-Aubin, J., & Jalbert, A. (2006). Rehearsal in serial memory for visual-spatial information: Evidence from eye movements. *Psychonomic Bulletin & Review*, 13(3), 452–457. DOI: <https://doi.org/10.3758/BF03193869>

Wickelmaier, F. (2013). Psychoacoustics Laboratory at the Department of Psychology. Retrieved from: <http://www.empg2014.uni-tuebingen.de/acoulab/index.html>.

How to cite this article: Kvetnaya, T. (2018). Registered Replication Report: Testing Disruptive Effects of Irrelevant Speech on Visual-Spatial Working Memory. *Journal of European Psychology Students*, 9(1), 10–15, DOI: <https://doi.org/10.5334/jeps.450>

Published: 06 April 2018

Copyright: © 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 Unported License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

]u[

Journal of European Psychology Students is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS 