Irrelevant Background Speech Disrupts Serial Short-Term Memory for Verbal but not for Spatial Information: A Pre-Registered Replication Study

Tatiana Kvetnaya[1], Kristina Schopf[2], Florian Wickelmaier[2]

[1] *Department of Psychology, Goethe-University Frankfurt am Main, Germany, e-mail: tkvet@web.de*
[2] *Department of Psychology, University of Tübingen, Germany*

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 [1]. Yet there is an ongoing debate about which working memory models are most suitable to explain it [2][3][4]. In 1982, Salamé and Baddeley [5] proposed an explanation based on a modular theory of working memory [6]. 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 [7][8]. According to this account, interference arises when two concurrent activities are similar in content [9] 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 [7][8].

By contrast, Jones, Farrand, Stuart, and Morris [10] suggested that interference results from a modality-independent similarity of process rather than content [9][11]. 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 [12]. Within this framework, the degree of disruption is determined by the amount of serial order information that concurrent activities contain.

To demonstrate this, Jones et al. [10] 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 steadily repeated. The use of the ISE 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 [10].

However, while the ISE was extensively investigated in regards to verbal serial recall tasks, it has rarely been shown in a spatial domain [13]. Despite the crucial role of Experiment 4 in the original authors' argument for the O-OER model, to the best of our knowledge, an unpublished series of studies by Klatte and Hellbrück [14] remains its only known conceptual replication up to this date. Contrary to the original results and in line with a modular theory of working memory, Klatte and Hellbrück observed an ISE in a verbal, but not in a spatial serial recall task.

Because an effect of irrelevant speech on serial spatial memory constitutes strong support for a unitary model of working memory, the present pre-registered study aimed to directly replicate Experiment 4. A successful replication would find a main effect of sound type across task domains, with changing state sound being markedly more disruptive than steady state sound. Additionally, the absence of an interaction effect between sound type and task domain is expected, that is, performance should be disrupted equivalently in both verbal and spatial domains.

Method

The present cumulative replication study consisted of two bachelor's thesis projects conducted at the University of Tübingen between 2017 and 2018 [15][16]. Each thesis pre-registered and directly replicated one task condition of Experiment 4, respectively. Both manuscripts alongside the primary data, materials, and executable R scripts of all statistical analyses, are available on www.osf.io/hba2p.

Participants

To determine the size of the planned sample, power simulations based on the key statistics provided in the original paper were conducted. The analyses demonstrated that a sample size of n = 40 per task condition ensured over 95% power to detect a main effect of sound type, and approximately 80% power to detect an interaction of sound and task type.

Normal hearing and vision were set as prerequisites for participation. Additionally, 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. None of the participants met the latter exclusion criterion, and one participant was excluded due to impaired hearing.

The final sample consisted of 80 participants (62 women and 18 men, with an age range of 18–53 years, M_{age} = 23.14,

SD_{age} = 4.77) who were predominantly psychology students from the University of Tübingen. They received course credit in turn for their participation. In order to increase attentiveness and motivation during task completion, participants were additionally rewarded with an honorarium of maximum €10 depending on their performance. For example, if a participant correctly recalled the item position in the respective sequences for 75% of all items displayed, they were rewarded with €7.50, etc.

Material

Verbal 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 the target stimuli F, K, L, M, Q, R, and Y was presented in bold uppercase letters in 30-point white sans serif font against a gray background. During each trial, the sequence was displayed at a rate of one letter per two seconds (one second “on” and one second “off”) in a randomized order. After a 10-second retention interval, all seven letters were displayed simultaneously in a row, each within a drawn box. Participants were asked to reproduce the letter sequence using a mouse by clicking on each letter in the order of its original presentation. Shading of the box changed to blue to signify it has been selected. The selection could not be reversed or altered once it was done. After a participant selected all letters, the green button to begin the next trial was displayed again. Participants were given practice for three trials before the start of the 48 experimental trials.

Spatial memory task: The properties and setup of the spatial memory task exactly paralleled those of the verbal task, except for the following: The target stimuli consisted of a sequence of seven white dots, which were presented in a quasi-random position generated within a 500 × 500 matrix. The dots had a radius of 12 units and could not be located closer than 85 units on either axis of the matrix. In the retrieval phase, all dots in the sequence were displayed simultaneously in their original position. After a participant selected all dots, the green button to begin the next trial was displayed again.

Irrelevant speech: Speech in a female voice spoken at two syllables per second was recorded digitally for the present study. Exactly following the original study, two experimental conditions were created: The steady state condition consisted of the repeated syllable “ah”, the changing state condition used the seven syllables in the alphabetic sequence “a” to “g”. In the latter condition, the order of the syllables was fixed, but when played back to the participants, the recording started from a random point in the sequence. Recordings were delivered via headphones (Beyerdynamic DT 990 Pro).

Design and Procedure

Exactly following the original study, participants were tested in a mixed 2 (task type: verbal vs. spatial – between subjects) × 3 (sound type: changing state vs. steady state vs. silent control condition – within subjects) design. The dependent variable was the number of serial order errors measured for each of the seven serial positions in either memory task.

Participants were tested individually in a sound-insulated booth in the Psychoacoustics Lab at the University of Tübingen (see [17] for details). In total, participants undertook 48 trials, with 16 trials in each of the three sound conditions, including a silent control condition. For the changing and steady state conditions, speech recordings were played throughout the presentation and retention phases of the memory task and stopped at the beginning of the retrieval phase. The order of the sound conditions was randomized from trial to trial for each experimental session. After completion of the experiment, participants received their financial reward.

The feasibility of the procedure was verified in a small-scale ($N = 6$) pilot study. No prior knowledge from the participants is assumed. Stimulus presentation and data collection were controlled by software written in Python/PsychoPy [18] running on a Linux operating system. All statistical analyses were conducted with R [19].

Results

Figure 1 displays the mean number of errors for the spatial and verbal conditions alongside the original results of Experiment 4 by Jones et al. [10]. In the verbal condition, participants achieved an overall error rate of M = 3.08 (SD = 1.29) out of 16 possible errors per sound condition and serial position (changing state: M = 3.83, SD = 1.46; steady state: M = 3.14, SD = 1.16; silent control: M = 2.28, SD = 0.82). In the spatial condition, participants achieved an overall error rate of M = 6.54 (SD = 2.82) errors (changing state: M = 6.68, SD = 1.39; steady state: M = 6.35, SD = 1.25, silent control: M = 6.58, SD = 1.33).

Serial order errors were aggregated over serial position and subjected to an analysis of variance (ANOVA) with the factors sound type (changing state vs. steady state vs. silent control condition – within subjects) and task type (verbal vs. spatial memory task – between subjects). Contrary to the original study, there was a significant interaction between task domain and sound type, $F(2, 156) = 8.96$, $p < .001$, $\eta^2_p = .10$, reflecting a strong interference effect in the verbal domain that was absent in the spatial domain. Paired t-tests for each task domain revealed that changing state sound was significantly more disruptive than steady state sound in regards to performance in the verbal task ($t(39) = 2.60$, $p = .013$, $\eta^2_p = .15$.), but not in the spatial task ($t(39) = -1.22$, $p = .229$, $\eta^2_p = .04$). In addition to the interaction, the main effect of task type was significant, $F(1, 78) = 84.22$, $p < .001$, $\eta^2_p = .52$, reflecting the overall inferiority of the average performance in the spatial task. Taken together, these results indicate that irrelevant speech differentially impaired serial recall performance as a function of the task domain.

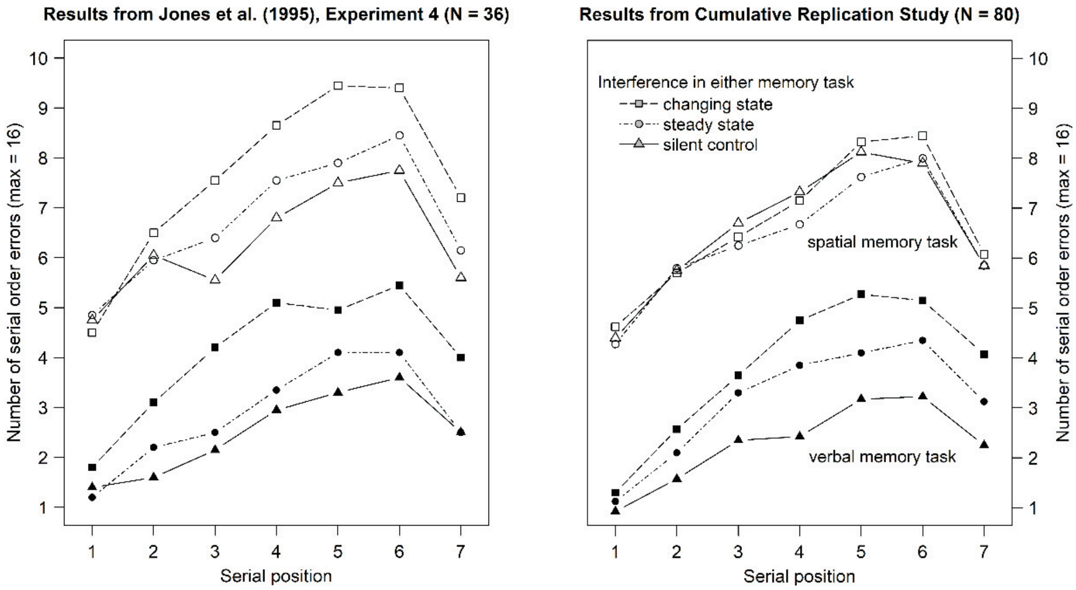

Figure 1: Effects of irrelevant speech on serial order errors in verbal and spatial memory tasks in Experiment 4 by Jones et al. (1995; left panel) and the present cumulative replication study (right panel). Serial order errors are displayed as a function of serial position (1−7) and interference condition (changing state condition: the alphabetic sequence "a" to "g" repeated vs. steady state condition: repeated "ah" vs. silent control condition: no speech). Maximum (max.) error = 16 for each serial position.

Discussion

The present study constitutes the first pre-registered direct replication of an experiment by Jones et al. [10], who employed the irrelevant speech paradigm in a visual-spatial domain. In order for the present replication to be considered successful, a main effect of sound type across both task domains (and therefore, the absence of an interaction between sound type and task domain) was expected, with changing state being markedly more disruptive than steady state sound. Contrary to this expectation, we found a strong interaction between task domain and sound type.

Indeed, in the verbal task, changing state speech was more disruptive of recall performance than steady state speech. Within the verbal domain, sound type accounted for 15 % of variance in the present study, comparable to 16 % in the original study [15]. However, no such pattern emerged within the spatial domain: No significant differences of recall performance between any three sound conditions were observed. The discrepancy between effect sizes in 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 [16]. Consequently, the observed interaction between task domain and sound type accounted for only 0.8 % of variance in the original study, while accounting for 10 % of variance in the present study [15].

In summary, these results do not fit the predictions derived from Jones et al.'s changing state hypothesis in the context of the O-OER model. 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, such as the visual-spatial sketchpad [7].

The results from the present replication study closely resemble the findings of Klatte and Hellbrück [14], who also observed an interaction between task domain and sound condition. In their conceptual replication of Experiment 4, changing state speech was significantly more detrimental than both steady state speech and silence in a verbal serial recall task. However, no differences between these three sound conditions were observed for the performance in a spatial memory task.

Limitations

These diverging findings might be either due to methodological differences or due to a non-replicability of the original findings. 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 use of performance-related financial reimbursement.

A further possible limitation is that data collection for the verbal and spatial memory tasks occurred in two separate studies. Thus, participants were not randomly assigned to one of the two memory task conditions. Therefore, the presence of contextual factors which might have contributed to an absence of the ISE in the spatial task cannot be ruled out with absolute certainty. However, since overall error patterns in both memory tasks closely resembled those in the original study, these methodological differences as well as the two-part composition of this replication study may be of minor relevance for explaining the present results.

Nonetheless, the changing state hypothesis appears to successfully predict interference based on the properties of the irrelevant sound and the verbal cognitive tasks, provided that both are processed in a shared cognitive subsystem.

In conclusion, our results indicate that the involvement of such a shared domain-specific subsystem may be necessary, but not sufficient for two concurrent activities to interfere with each other. In order to comprehensively account for memory effects such as the ISE, further elaboration on both modality-based specific and amodal general-purpose mechanisms is warranted. Future efforts to model the ISE should take integrative models encompassing multiple perceptual modalities and interactions between them into account.

The present work further highlights the importance of registered replications and the need for sufficiently powered experiments in order to increase the replicability of psychological findings.

Author's Note

Portions of this work will be presented at the "Tagung experimentell arbeitender Psychologen" (Conference of Experimental Psychologists, TeaP) in London in April 2019.

References

[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, 12–29

[2] Baddeley, A. D. (2000): The phonological loop and the irrelevant speech effect: Some comments on Neath (2000), *Psychonomic Bulletin & Review* 7, 544–549

[3] Jones, D. M., Tremblay, S. (2000): Interference in memory by process or content? A reply to Neath (2000), *Psychonomic Bulletin & Review* 7, 550–558

[4] Neath, I. (2000): Modeling the effects of irrelevant speech on memory, *Psychonomic Bulletin & Review* 7, 403–423

[5] Salamé, P., Baddeley, A. (1982): Disruption of short-term memory by unattended speech: Implications for the structure of working memory, *Journal of Verbal Learning and Verbal Behavior* 21, 150–164

[6] Baddeley, A. D., Hitch, G. (1974): Working memory, in: G.H. Bower (Ed.), Psychology of learning and motivation, Academic Press, New York, pp. 47–89

[7] Baddeley, A. D. (1996): The concept of working memory, in: S.E. Gathervole (Ed.), Models of short-term memory, Psychology Press, Hove, UK

[8] Guérard, K., Tremblay, S. (2008): Revisiting evidence for modularity and functional equivalence across verbal and spatial domains in memory, *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34, 556–569

[9] Schendel, Z. A. (2006): The irrelevant sound effect: Similarity of content or similarity of process?, Doctoral Dissertation, The Ohio State University

[10] 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, 1008–1018

[11] 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, 459–471

[12] Jones, D. M., Beaman, C. P., Macken, W. J. (1996): The object-oriented episodic record model, in: S.E. Gathervole (Ed.), Models of short-term memory, Psychology Press, Hove, UK, pp. 209–237

[13] Tremblay, S., Saint-Aubin, J., Jalbert, A. (2006): Rehearsal in serial memory for visual-spatial information: Evidence from eye movements, *Psychonomic Bulletin & Review* 13, 452–457

[14] Klatte, M., Hellbrück, J. (1997): Effects of irrelevant speech on serial recall of verbal and spatial materials, in: A. Schick, M. Klatte (Eds.), Contributions to psychological acoustics. Results of the seventh Oldenburg symposium on psychological acoustics, Oldenburg, 531-538

[15] Schopf, K. (2018): Effects of irrelevant background speech on verbal serial recall: A partial replication of "Functional equivalence of verbal and spatial information in serial short-term memory" (Jones, Farrand, Stuart, & Morris, 1995; Experiment 4), Bachelor's thesis, University of Tübingen, URL: https://osf.io/f73a4/

[16] Kvetnaya, T. (2018): 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), *Journal of European Psychology Students* 9, 10–15

[17] Wickelmaier, F.: Psychoacoustics laboratory at the Department of Psychology — acoulab, URL: http://www.mathpsy.uni-tuebingen.de/acoulab/index.html

[18] Peirce, J. W. (2007): PsychoPy—Psychophysics software in Python, *Journal of Neuroscience Methods* 162, 8–13

[19] R Core Team: R: A language and environment for statistical computing. URL: https://www.R-project.org/