2023 ■ Vol. 19 ■ no. 3

Capacity and duration of iconic memory from partial reporting of brief stimuli: A replication of Sperling's experiment (1960)

Meenakshie Bradley-Garcia [a], Salma Ben Messaoud [a], Ashley Bossilkov [a], Chloé Cateaux [a], Carmel Tshila N'Sa [a], Annabelle Potvin-Pilon [a], Hajira Abdi-Nassir [a], Nardeen Assad [a], Fiona Balcom [a], Salma Belfaiz [a], Cassandra Blackburn [a], Julia Bolduc [a], Idriss Boutarbouch [a], Natasha Béchamp [a], Mireyva Tatiane Calderon Robles [a], Véronique Charbonneau [a], Isabelle Chatillon [a], Lauren Cooper-Ballinger [a], Maria Falconi [a], Vincent Francoeur [a], Asterie Gashema [a], Victoria Healey [a], Victoria Johnston [a], Cloé Lafleur [a], Josiane Laliberté [a], Jadzia Lareau [a], Camille Leblanc [a], Erin Miess [a], Victoria Méthot [a], Habib Najjar [a], Jamila Palanee [a], Suzanne Pradelle [a], Deevesh Raj Ramphul [a], Mahika Sarrazin [a], Nich Savage [a] & Monica Tremblay Jimenez [a]

[a]Department of Psychology, University of Ottawa

Abstract ■ It has been widely accepted that iconic memory has a capacity of minimally 9 elements and a duration of approximately 0.3 seconds. However, Sperling's (1960) partial report methodology influenced the study of iconic memory by demonstrating its larger capacity than previously considered. Due to the limited number of participants in the original study, a replication study was necessary to corroborate the results of Sperling (1960) to validate the scientific merit of results, thereby strengthening the validity of the study. The present study aimed to replicate Sperling's (1960) partial report experiment with modern technology amongst a larger and demographically heterogeneous sample. Male and female participants (n = 64) aged 18-59 years old (M = 30.61, SD = 13.211) were recruited to complete four online tasks via Qualtrics. Tasks 1 and 4 involved the recall of briefly presented sequences of 3, 4, 5, and 6 letters. Tasks 2 and 3 required participants to recall an array of 3 and 4 letters and numbers presented in 2 and 3 rows respectively. In Tasks 2 and 3, an auditory cue was presented for 0.05 s at two (high and low) and three different frequencies (high, medium, low), respectively, which indicated the row to be reported at different interstimulus intervals (-0.10, 0.0, 0.15, 0.30, 0.50, 1.0 s) relative to the presented stimuli. Analyses revealed that the interaction between the number of letters and the interstimulus interval, the number of rows and the interstimulus interval, as well as the number of letters and number of rows was statistically significant. The findings of this replication study support the results of Sperling (1960) suggesting that partial report accuracy is influenced by the number of briefly presented characters. Future studies should explore the influence of a controlled environment to explain the effects of the variables on recall abilities.

Keywords ■ Iconic memory, sensory memory, memory, Sperling, replication, online study.

Acting Editor ■ Denis Cousineau (Université d'Ottawa)

Reviewers
■ None of the author(s) of the original article reviewed the methodology

✉ mbrad048@uottawa.ca

doi 10.20982/tqmp.19.3.r011

Introduction

Remember sparkling stick fireworks when we were kids? We used to have fun drawing figures in the air, using the persistent line of light. The sensory register ensures that the perception quickly decays, as it only serves to give our brain time to process the information and maintain a fluid, continuous perception of the environment (Cappiello & Zhang, 2016; Coltheart, 1980; Di Lollo & Dixon, 1988; Irwin & Thomas, 2008; Lemaire & Didierjean, 2018; Neisser, 1967; Rensink, 2014).

The visual repository, also called iconic memory

(Neisser, 1967) has been the object of much research given the importance of visual perception in human functioning. Iconic memory is characterized by a high encoding capacity and a short lifetime span (less than 500 ms), making its capacities greater than those of working memory (Cappiello & Zhang, 2016; Cowan, 2001; O'Regan & Noë, 2001; Sperling, 1960). Alternatively, iconic memory (Neisser, 1967) could retain the visual stimulus for up to 100 ms before transferring it to working memory (Bradley & Pearson, 2012). To quantify the amount of visual information retained during a brief exposure, Miller (1956) used the full report that is characterized by exposing a limited number of symbols (letter, number, etc.) to a participant for a brief time, and then asking them to report all symbols remembered (Fortin & Rousseau, 2016).

George Sperling's research, among others, focused on conceptualizing and estimating fundamental characteristics of iconic memory. Prior to Sperling's research, early studies had accounted for an immediate memory span, for example a limit to the amount of information that participants could report from an array of letters briefly presented through a tachistoscope (Miller, 1956). Such limit was equal to no more than 4.5 letters on average (Coltheart, 1980; Sperling, 1960). Sperling (1960) believed that the duration of information in sensory memory was different than that of short-term memory. This theory-based assumption led him to quantitatively investigate the decay of available information within iconic memory. Through a series of ground-breaking experiments, Sperling (1960) provided insight into both the capacity and duration of iconic memory. He used the full ratio and developed another quantification approach, the partial report. By using the partial report and different cues at various intervals, the capacity and duration of iconic memory could be inferred. In this study, participants were briefly presented with an array of letters and/or numbers and were then asked to either report the entire array (whole report), or a single row of the array (partial report). During the partial report task, the row that participants were asked to report was determined by the frequency of a signal tone that was sounded before, during, or after the presentation of the array of letters and/or numbers. Variations of the time chosen to display the signal cue allowed estimating the duration of information in iconic memory. Indeed, Sperling believed that unlike full reports, which he thought were constrained by the immediate memory span, partial reports would unveil all the accessible information in iconic memory (1960). Sperling found that 80-90% of the stimulus was still available, but that this availability diminished rapidly after 150 ms.

Researchers who worked on similar projects (Averbach & Coriell, 1961; Coltheart, 1980; Loftus et al., 1992) and more recent studies (Graziano & Sigman, 2008; Rensink, 2014; Sligte, 2010) seem to confirm his theory of visual persistence. However, results vary greatly depending on the method, for example depending on the type of visual task, iconic memory may last as little as 120 ms or as long as 240 ms (Rensink, 2014). Several years after Sperling's article, it was theorized that iconic memory is the expression of an information process and not a simple sensory store in which information starts decaying at the end of visual perception (Di Lollo, 1977). Other research focused on whether attention is required for iconic memory to form, and that question may still be unresolved (Bachmann & Aru, 2015; Botta et al., 2023; Mack et al., 2016). Recent studies found that iconic memory is related to neuronal activity in the visual cortex and may be improved through training (Gong et al., 2022; Teeuwen et al., 2021).

The objective of this study is to replicate the fourth experiment conducted by Sperling in 1960 which examined the duration of information in brief visual presentations. Replicating a study involves repeating the research procedure under similar conditions to assess the reliability and validity of the initial findings (National Academies of Sciences, Engineering, and Medicine, 2019). We also strived to address some of the limitations identified in the original research such as recruiting a larger and more heterogeneous sample that will allow the generalization of our results to a broader sample of the general population. Indeed, Sperling's original sample consisted of five highly trained participants, four of which were students at Harvard University and one of which was a faculty member of this university. This limited the statistical power of the conducted analyses and increased the likelihood of overlap in participant characteristics, potentially deviating from the diversity found in the general population. Furthermore, many early studies on memory capacity, such as Sperling's (1960) study, used tachistoscopes to briefly present visual stimuli to participants and hand-written response grids to code their answers. However, these methods of testing have become outdated as they may introduce additional variance and have since been computerized. Thus, the use of modern technology will allow for more precise and accurate measurements of the decay of available information in iconic memory with time.

Investigating iconic memory has significant implications for the field of cognitive psychology and the broader community. All visual information perceived by an individual is initially processed by and temporarily held in their iconic memory system (Sperling, 1960). This memory system acts as the primary receiver of visual information within the cognitive system, preceding the involvement of short-term and long-term memories. Understanding the parameters of iconic memory is therefore fundamental to our understanding of the early stages of visual sensory processing and subsequent human cognitive and perceptual abilities. Additionally, the study of this memory system has practical implications, such as its potential clinical use in assisting patients with memory disorders, as well as providing insight into common perceptual phenomena such as partial visual memories.

Based on the results of Sperling (1960), a gradual decline in reported correct responses as the signalling delay for the various sound tones increases (0.0, +0.15, +0.30, +0.5, +1.0 sec.) should be observed, confirming that sensory information deteriorates over time. This would be observable by the reduced performance on the partial report given by the participant. Additionally, presenting the auditory instruction before the visual stimulus (-0.10 sec.) should increase the amount of information available in the participant's immediate memory, making the report more accurate.

Method

Recruitment

Undergraduate students enrolled in a cognitive psychology course at the University of Ottawa during the 2023 winter semester collaborated on this project for course credit. Each student recruited friends and/or family members by sending a standard email containing an invitation to the study including a Qualtrics link to access the experiment. Participants had two weeks to complete the study and were instructed to reserve a 45-to-50-minute period, free from any distraction, to complete the study. This method ensured a consistent recruitment process and standardized instructions for all participants.

Participants

Data were obtained from 119 participants. Participants with a history of epilepsy should have been excluded (n = 2), due to the rapid presentation of visual stimuli as they could provoke a seizure; however, their data were kept as they did not significantly differ from other participants. Participants below the age of 18 years old (n = 0) would have been excluded and those who did not complete all four tasks (n = 55) were excluded. Data was retained from 64 participants (cisgender woman, n = 41; cisgender man, n = 20; preferred not to respond, n = 1; other, n = 2) between 18 and 59 years old (M = 30.61, SD = 13.211). This sample included participants who had completed various levels of education (high school, college, bachelor, Masters, or Ph.D.), diverse ethnic backgrounds (White, Black, Non-white Latin American, or East Asian) and had a sufficient level of English to complete the experiment. See Table 1 at the end for a detailed list of the demographic information obtained by each participant.

Materials

The study was administered through an online questionnaire using Qualtrics as the programming software. The program was set up to present the experiment on full screen mode to minimize potential distractions. Participants were asked to use their personal electronic devices, such as desktop, laptop, or tablet and to ensure they had functional speakers or headphones on said devices. Participant were asked to not use their smartphones to complete the study as the stimulus may appear very small making it more difficult to complete the tasks. We also recommended that participants with visual impairments wear their visual aid and those with hearing impairments to also wear their hearing aid.

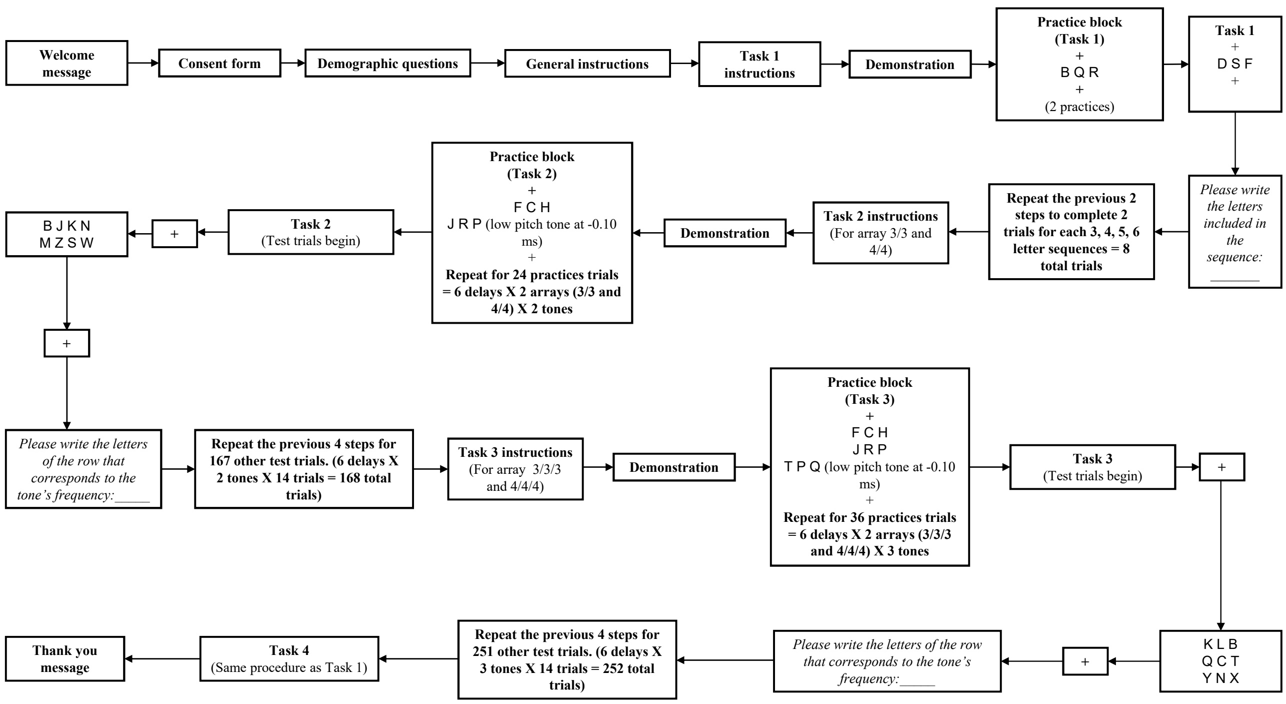
Stimuli

The stimuli used in this replication consisted of a randomly generated array of consonants only ('B', 'C', 'D', 'F', 'G', 'H', 'J', 'K', 'L', 'M', 'N', 'P', 'Q', 'R', 'S', 'T', 'V', 'W', 'X', 'Y', 'Z'), to prevent word recognition, or a combination of consonants and single-digit numbers (0-9). All stimuli were displayed in the middle of the screen, in white capital letters in Arial font, size 12, over a black background, for 50 ms. Before and after each array was presented, a plus sign (+) was shown in the middle of the screen for 33,34 ms as a fixation point. For the first and fourth task, the arrays consisted of one row of either 3, 4, 5 or 6 consonants. For the second task, two rows of 3 or 4 consonants (3/3 or 4/4) were presented. As for the third task, the arrays consisted of three rows of 3 letters (3/3/3) or three rows of 2 letters and 2 numbers each (4/4/4). For the second and third task, signal tones were used to indicate which row had to be reported, depending on its pitch (high, medium, or low). For a high-pitch tone (2,500 CPS), the participants had to report the top row of the grid and for a low-pitch tone (500 CPS), they had to report the bottom row of the grid. For the third task, a medium-pitch tone (650 CPS) was added to indicate that participants had to report the middle row of the grid. The signal tones were presented with an increasing or decreasing time delay (-0.10, 0.0, 0.15, 0.30, 0.50, 1.0 sec.) from the arrays and played for 0.5 sec.

Procedure

The procedure was identical for all participants, for the exception of counterbalancing and randomization procedures. Figure 1 provides a schematic. The experiment began with a welcome message, followed by a consent form and a demographic questionnaire (Table 1). Participants were then presented with general instructions to minimize technical issue and were given the chance to test their audio. Demonstrations and practice trials, where the correct answers were indicated to participants, preceded each task.

Figure 1 ■ Flowchart of the experimental protocol for this replication study of Sperling (1960). The fixation point (+) was presented before and after each array for 33,34 ms. Each array was presented for 50 ms. To illustrate the experimental protocol, typical arrays are used as example. As per counterbalancing, half of the participants started with 3/3 arrays, while the other half started with 4/4 arrays. The same counterbalancing method was applied to 3/3/3 and 4/4/4 arrays.

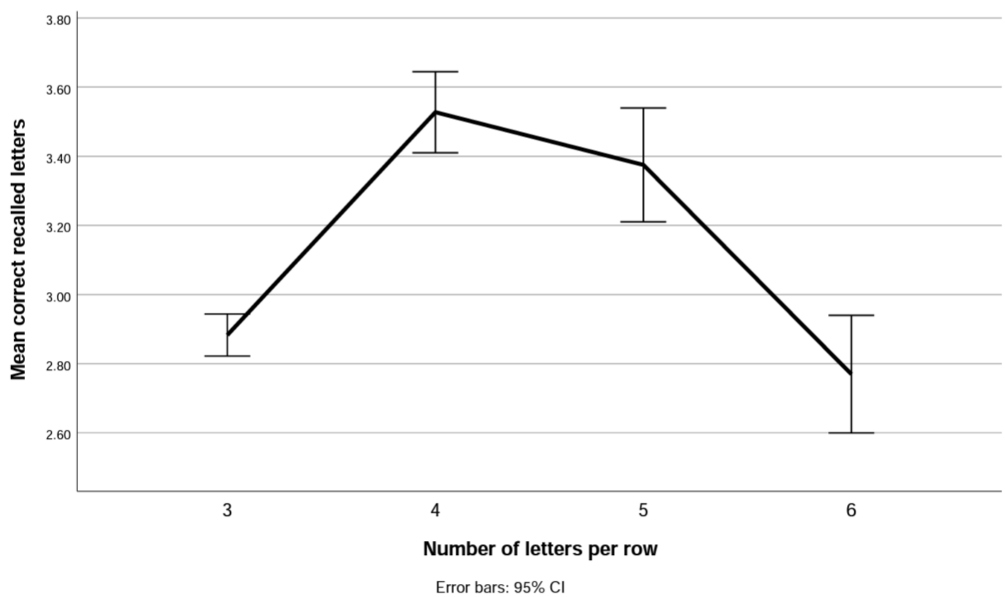


Participants were asked to type an answer, and indicate a "0" as their answer if they couldn't recall the response. For the first task, participants were given a short demonstration and two practice trials with two random sequences of either 3, 4, 5 or 6 letters. They were then asked to complete a total of eight trials, or two trials for each 3, 4, 5, or 6 letter sequences on a single row.

For the second task, instructions were presented on the screen, followed by a sound test, two demonstrations (one for each tone) and a practice block. The practice block consisted of 24 randomized trials, or two practices trials for every tone delay (-0.10, 0.0, 0.15, 0.30, 0.50, 1.0 sec) applied to each array type (3/3 and 4/4) and for the low and high pitched tones (6 delays × 2 tones × 2 practices). After the practice block, participants completed 168 trials (6 delays × 2 tones × 14 trials) to complete the second task with the order of presentation counterbalanced as described below. New instructions, indicating the addition of a medium-pitched tone, were presented for the third task, followed by a sound test, three demonstrations (one for each tone) and a practice block. Participants were asked to complete the practice block, which consisted of 36 randomized trials (6 delays × 3 tones × 2 practices). Participants were then instructed to complete a total of 252 trials (6 delays × 3 tones × 14 trials) to complete the third task with a similar counterbalanced order of presentation as the second task.

For the second and third task, there were counterbalancing variations in the order of presentation of the stimuli. As for the second task, half of the participants started with the 3/3 arrays while the other half started with the 4/4 arrays. In the same fashion, half of the participants started with the 3/3/3 arrays while the other half started with the 4/4/4 arrays in the third task. The viewing order was reversed for each participant between both tasks. In addition, half the participants heard the sound tone in increasing order of time delay (-0.10, 0.0, 0.15, 0.30, 0.50, 1.0 sec.), while

Figure 2 ■ Mean correct recalled letters as a function of the number of letters per row for tasks 1 and 4. Results of the mean correct number of letters recalled according to the number of letters per row. The mean of letters correctly reported by participants (n = 64) for each condition (3, 4, 5, and 6 letters). Error bars represent 95% confidence intervals adjusted for correlations and differences (Cousineau, 2017).



the other half heard the sound tone in decreasing order of time delay (1,0, 0.50, 0.30, 0.15, 0.0, -0.10 sec.). The listening order was reversed for each participant between both tasks.

For the fourth task, participants followed the same procedure as the first task. It was repeated to account for fatigue and practice effects.

The experiment ended with a thank you message, and participants were asked to provide feedback on technical or performance-related difficulties experienced during the experiment.

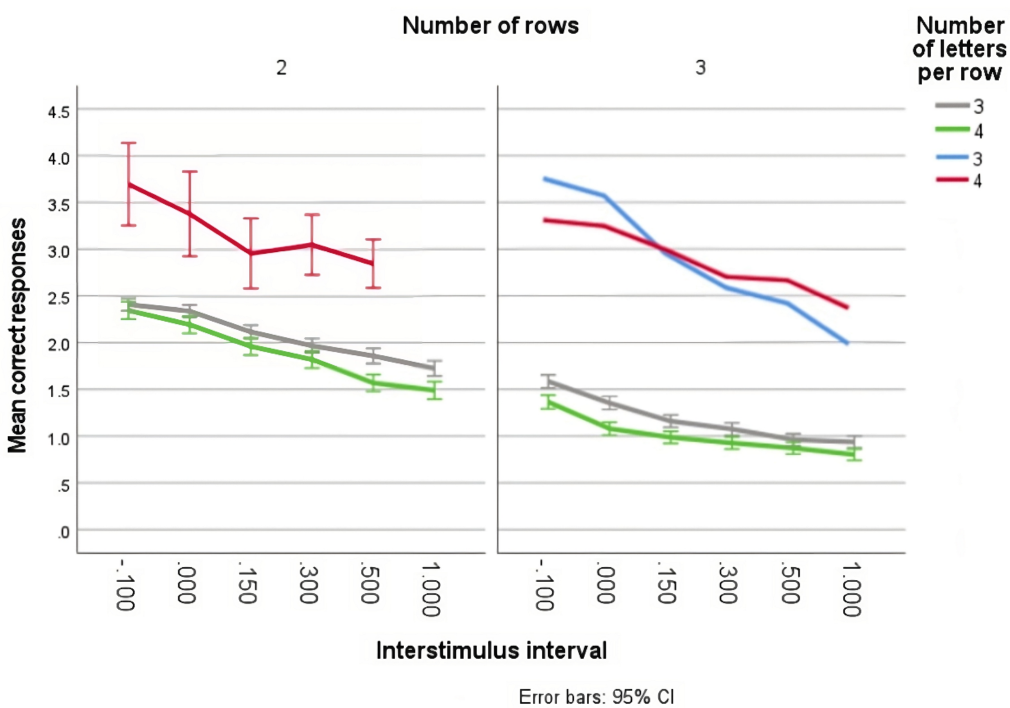
Coding

Recall of letters presented in each task was scored as follows: a score of 1 was given for each correct letter in the correct order, and a score of 0 was given to any missed letters or correct letters in the incorrect order. Given these parameters the score for tasks 1 and 4 could vary between 0 and 6 depending on the sequence presented, while tasks 2 and 3 could vary between 0 and 4 depending on the condition. The average of the scores for each condition (-0.10, 0.0, 0.15, 0.30, 0.50, 1.0 second delays) was then calculated by dividing the sum of the correct answers by the number of trials (number of correct letters by trial) or by the total number of presented letters (percentage of correct answers).

Statistical analysis

Statistical analyses were performed using SPSS 28. A 95% confidence interval was used with an alpha level set to $p < .05$ for all analyses of means. A repeated measures ANOVA (RMANOVA) was conducted on the scores from the first and fourth task to determine the effect of the number of letters on report accuracy. A Greenhouse-Geisser correction was implemented when the assumption for Mauchly's sphericity was violated. For the second and third task, another RMANOVA was conducted using number of rows and number of letters as within-subject factors and indicated stimuli interval (-0.10, 0.0, 0.15, 0.30, 0.50 and 1.0 sec) as between-subject factor to determine their respective main effects and joint interactions on report accuracy. A Greenhouse-Geiser correction was also implemented on those main effects and interactions when Mauchly's sphericity was violated. A paired-sample t-test was also conducted from the second and third tasks to verify if the order of presentation (increasing vs. decreasing) of the tone delays had a significant effect on the results. The stimuli and the raw data for this experiment have been made available online (osf.io/cynv8/).

Figure 3 ■ Comparison of the means of correct responses between Sperling's (1960) study and the current study. The grey and green lines represent the results of the present study while the results in red and blue are estimated from Sperling (1960) figures. On the left, the results of the correct number of letters reported by participants according to the number of letters per row and the interstimulus interval for two rows. On the right, the results of the correct number of letters reported by participants according to the number of letters per row and the interstimulus interval when there are three rows. The mean of correct letters recalled by participants (n = 64) when there are 3 and 4 letters in either 2 or 3 rows for each interstimulus interval (-100 ms, 0 ms, 150 ms, 300 ms, 500 ms, and 1000 ms). Error bars represent 95% confidence intervals adjusted for correlations and differences (Cousineau, 2017). The right panel with the 3 rows reflects the results of one participant from Sperling's study, thus the error bars could not be computed.



Results

First, a RMANOVA using the within-subject variable was performed to determine whether the number of letters observed (3, 4, 5, or 6) influenced the number of letters correctly reported for tasks 1 and 4. The assumption for Mauchly's sphericity test was violated [$\chi^2(5)$ = 13.212, p = .021]. After applying a Greenhouse Geisser correction (ε = .873), significant differences were observed [$F(2.62, 165.085) = 225.600$, $p < .001$, partial η^2 = .782]. As a result, the number of letters presented accounted for 78% of the variability in performance. Figure 2 indicates that most participants reported almost all the letters when 3 or 4 letters were presented, but recall became more difficult when 5 or 6 letters were shown. As found by Sperling (1960), there were no practice or fatigue effects.

Second, a paired sample t-test was executed to determine whether an increasing or decreasing order of interstimulus interval affected participants' performance for tasks 2 and 3. The test revealed that the order of presentation of the different tones had no significant impact on performance (p = .202). This suggests that participants retained a similar amount of information despite increasing or decreasing order delays between visual stimulus and tone cues. This differs from Sperling's (1960) study, for which there seems to be an interaction between order type and interstimulus interval parameters.

Third, a factorial ANOVA using the within-subjects variables was carried out to determine if the number of letters presented (3 or 4), the number of rows (2 or 3), the interstimulus interval (-0.10, 0.0, 0.15, 0.30, 0.50 and 1.0 sec) and their interactions affected performance for tasks 2 and

3. The analysis revealed that the main effect of the number of letters was statistically significant [$F(1,63) = 418.163$, $p < .001$, partial $\eta^2 = .869$]. The number of letters explained 86% of the variability in the correct number of letters reported. Participants performed better when asked to recall 3 letters rather than 4, as shown in Figure 3. Moreover, the main effect of the number of rows was statistically significant [$F(1,63) = 425.033$, $p < .001$, partial $\eta^2 = .871$]. It suggests that the number of rows explains 87% of the variability in performance. The number of letters correctly reported is better when there are 2 rows rather than 3 (see Figure 3). These results corroborate those of Sperling (1960), although he does not explicitly report this main effect (see Figure 3). For the interstimulus interval, Mauchly's sphericity test was violated [$\chi^2(14) = 103.272$, $p < .001$]. After a Greenhouse-Geisser correction ($\varepsilon = .529$), this main effect was also statistically significant [$F(5,315) = 109.998$, $p < .001$, partial $\eta^2 = .636$], accounting for 64% of the variability in performance. As the auditory cue became more distant from the visual stimulus presented (from time 0 to 1.0 sec), performance decreased. Moreover, the number of letters correctly reported was better when the auditory cue was before the visual stimulus (see Figure 3). These results corroborate those of Sperling (1960), who considers it crucial to give the auditory cue within the one-second timeframe.

Although Sperling (1960) provides little information on interactions, several interactions in this current analysis were discovered. The interaction between the number of letters and the number of rows was statistically significant [$F(1,63) = 19.451$, $p < .001$, partial $\eta^2 = .236$]. The number of letters by the number of rows interaction accounts for 24% of the variability in performance. According to Figure 3, performance peaks when there are 3 letters for 2 rows, while performance drops when there are 4 letters for 2 and 3 rows, thereby corroborating Sperling's findings. A Greenhouse-Geisser correction ($\varepsilon = .878$) was applied to the interaction between the number of letters and the interstimulus interval, considering the violation of Mauchly's sphericity test [$\chi^2(14) = 23.439$, $p = .054$]. Following correction, this interaction was also statistically significant [$F(5,315) = 3.272$, $p = .007$, partial $\eta^2 = .049$]. The number of letters combined with the interstimulus interval represents around 5% of the variability in the participants' performance. As shown in Figure 3, the peak performance occurs when 3 letters are presented with a sound cue at -0.10 sec, irrespective of the number of rows. However, the performance drops the most when 4 letters were presented with an auditory cue at 1.0 sec, regardless of the number of rows. Furthermore, the factorial ANOVA also revealed that the interaction between the number of rows and the interstimulus interval was statistically significant [$F(5,315) = 6.961$, $p < .001$, partial $\eta^2 = .099$]. The interaction accounted for 10% variability in performance. Specifically, the interaction between 2 rows and -0.10 sec gives the best performance, while the combination of 3 rows and 1.0 sec is the one for which performance decreases the most (see Figure 3). Lastly, Mauchly's sphericity test indicated that the sphericity assumption is not respected when examining the interaction between the number of letters, the number of rows, and the interstimulus interval [$\chi^2(14) = 24.619$, $p = .039$]. After applying the Greenhouse-Geisser correction ($\varepsilon = .864$), the latter interaction was not statistically significant [$F(5,315) = 1.957$, $p = .085$, partial $\eta^2 = .030$]. Indeed, this interaction would explain only 3% of the variability observed in participants' performance.

Discussion

The aim of the current study was to replicate the experimental design put forth by Sperling in 1960 to verify the validity and fidelity of his findings by utilizing a larger and more diverse sample of participants, as well as modern technology. As in Sperling (1960), participants were asked to report the letters and/or numbers of briefly exposed stimuli. Two methods of reporting were explored, first, using whole reporting tasks during which participants were required to report all letters presented on a single row (tasks 1 and 4); and second, using partial reporting tasks where participants were required to report letters from one of two or three rows cued by a sound tone (tasks 2 and 3).

The findings from this study demonstrate that participants never came close to the 4.4 letters reported by Sperling (1960) in his partial-report tasks, even at the extremity of a 95% confidence interval level (Figure 3). Moreover, the average number of letters reported by participants declined in the form of an inverted U-shape once the number of presented letters exceeded four. We can tell from Figure 3 that these results are not impacted by the moment of presentation of the auditory stimulus in the experiment when the participants executed the task. Both these issues are significant differences from Sperling (1960). A potential explanation would be the fact that Sperling's participants underwent much more training during his experiment, including whole reports with up to 12 letters. This lack of training may justify the drop in the performance once we reached a higher number of letters in a single row, which may have caused stress for our participants. Another explanation may rely on the size and demographic characteristics of our sample compared to Sperling's (1960). Consistent with the underperformance observed in the tasks 1 and 4, the partial reporting in our study resulted in a lower number of correct answers than Sperling's (1960) for each task and each condition, as shown in Figure 3. Based on our results, the capacity of iconic memory would be sharply lower than Sperling's at 4.4 (2 × 2.2) letters following task 2 and only 3.6 (3 × 1.2) letters following task 3. This suggest that a practice effect exists for one, but not the other. The findings of this study suggest that a one-second delay in the signalling tone will lead the partial report accuracy to be comparable to that of the whole report.

As also seen in Figure 3, the accuracy of partial reports decreased as the delay between the presentation of the visual stimulus and the sound tone increased. This is illustrated by a lower accuracy at a delay of 1.0 second than was found at the lowest average of the whole reporting tasks. Therefore, our results are consistent with Sperling (1960), especially for arrays with three rows. Past the 0.3 seconds delay, the number of correctly reported letters decreases more slowly and levels off, reflecting the passage of residual information into another type of memory. In other terms, this study demonstrated that the timing of the sound instruction has an impact on recall performance, suggesting that sound tones may serve as retrieval cues for the corresponding row in the array. This idea is in line with previous studies which claim that performance on partial report tasks is negatively influenced as the gap between stimulus presentation and auditory cue increases (Averbach & Coriell, 1961; Di Lollo & Dixon, 1988; Sperling, 1960). Our results show that Sperling's original theory, that human beings can retain more information in iconic memory than originally thought, has not yet been disproved.

In contrast, one of our results differs from what Sperling observed: the ascending or descending order in which the sound instruction is presented does not significantly affect participants' performance. This difference may be explained by the fact that Sperling only examined the effect of order when the stimuli were presented as two rows of letters. In contrast, our study showed that the number of letters and the number of rows also influenced recall accuracy. Finally, compared to Sperling, the present study also examined the presence of interactions (the number of letters and rows; the number of rows and indicated stimuli interval; the number of letters and indicated stimuli interval) likely to improve or deteriorate an individual's performance.

Our research adds to modern studies in the field of cognitive psychology that have expanded on Sperling's (1960) initial research regarding the duration, capacity, and other parameters of iconic memory. Some authors proposed encoding techniques that would enhance the ability to recall information, such as the specificity principle where Thomson and Tulving (1970) suggested that the probability of successful memory retrieval increased when the retrieval cues aligned with the contextual factors that were present during the original encoding of information. By applying the encoding specificity principle, it is reasonable to suggest that the incorporation of practice trials during the encoding phase could potentially enhance the partial recall of participants. Other authors conducted investigations into the potential impact of various variables on iconic memory, such as colour, on the accuracy of information recall (Phillips, 2011).

Our study demonstrated several methodological strengths which ultimately allowed for a more valid assessment of the capacity and duration of stimuli in iconic memory. One of these strengths was the inclusion of a diverse and inclusive sample of participants. In contrast to Sperling's (1960) study, the population used in the replication varied in age, gender, education level and had no specific pre-established training in psychology. Thus, allowing for the generalizability of the results to a wider sample of the general population. Despite efforts to attain a more extensive and heterogenous sample of the population, the sample was still primarily comprised of younger women with some degree of higher education. In addition, the use of modern technology facilitated the recruitment of this larger and more diverse sample by allowing us to overcome potential barriers resulting from geographical or scheduling constraints. This online format also allowed for greater standardized administration and control of the experimental conditions, thereby enhancing the study's internal validity. While this contrasted with Sperling's initial use of a Gerbrand's tachistoscope, it allowed the increase of the precision and accuracy of the experimental design, ultimately resulting in more reliable findings. We also provided breaks in-between experimental trials to mitigate potential eye strain, fatigue or boredom resulting from prolonged engagement with the task whereas Sperling conducted his study across several days. Moreover, the randomization of participants across the different experimental conditions was also a significant strength, as it helped reduce the carry-over effects resulting from the presentation of conditions in the same order to all participants. This is particularly relevant due to our utilization of a repeated-measures design.

While our study demonstrated notable strengths, it is necessary to acknowledge certain limitations it was faced with, one of which was its significant length, weight, and complexity. These factors could have posed a challenge for some participants and provided a plausible explanation for the high attrition rate of our study (45%). As the current study was conducted online and thus unsupervised, the likelihood of participants becoming distracted by outside stimuli, such as mobile notifications, may have been higher (Dontre, 2020). Despite the user-friendly instructions displayed on participants' screens during the experiment, and the possibility of carrying out the study from anywhere, thus reaching a larger number of people, the use of technology in such a study has its drawbacks. The electronic distribution of the study via e-mail, removed both a test administrator and controlled test environment. Both components would have benefited the study for control purposes, potentially enabling a higher rate of test completion and ensuring similar testing conditions. Another potential study limitation was the two-week timeframe participants had to complete the task; a higher rate of completion might have been attained if participants were offered a longer timeframe to complete the study.

Future research endeavors should consider limiting the number of experimental trials and the duration of the study. This simplification would improve participant engagement and potentially mitigate attrition rates. Thus, to enhance the generalizability of results, future studies could benefit from recruiting an even broader and more heterogeneous sample of the population. Future research should also investigate the potential mediating effect of demographic variables, such as age and gender, on the relationship between visual stimuli characteristics and accurate iconic memory recall. This could provide a more comprehensive understanding of how different individuals may process and recall briefly presented stimuli differently, and whether age-related declines in cognitive and perceptual abilities may impact iconic memory.

Authors' note

Various authors contributed as part of an undergraduate course project while Annabelle Potvin-Pilon, Ashley Bossilkov, Carmel Tshila N'sa, Chloé Cateaux, Salma Ben Messaoud contributed an equal amount towards the preparation of this manuscript. Meenakshie Bradley-Garcia conceptualized, programmed, supervised, and assisted with the preparation of this manuscript. The raw data, the stimuli used for this experiment, and the exploratory analysis have been made available on osf.io/cynv8/.

References

Averbach, E., & Coriell, A. S. (1961). Short-term memory in vision. *Bell System Technical Journal*, *40*(1), 309–328. doi: 10.1002/j.1538-7305.1961.tb03987.x.

Bachmann, T., & Aru, J. (2015). Comments on how mack et al. (2015) (do not) see iconic memory. *Consciousness and Cognition*, *34*, 73–74. doi: 10.1016/j.concog.2015.03.014.

Botta, F., Arévalo, E. M., Bartolomeo, P., & Lupiáñez, J. (2023). Attentional distraction affects maintenance of information in visual sensory memory. *Consciousness and Cognition*, *107*, 103453–. doi: 10.1016/j.concog.2022.103453.

Bradley, C., & Pearson, J. (2012). The sensory components of high-capacity iconic memory and visual working memory. *Frontiers in psychology*, *3*, 355–359. doi: 10.3389/fpsyg.2012.00355.

Cappiello, M., & Zhang, W. (2016). A dual-trace model for visual sensory memory. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(11), 1903–1922. doi: 10.1037/xhp0000274.

Coltheart, M. (1980). Iconic memory and visible persistence. *Perception & psychophysics*, *27*(3), 183–228. doi: 10.3758/BF03204258.

Cousineau, D. (2017). Varieties of confidence intervals. *AdvAnces in cognitive Psychology*, *13*, 140–155. doi: 10.5709/acp-0214-z.

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114. doi: 10.1017/S0140525X01003922.

Di Lollo, V. (1977). Temporal characteristics of iconic memory. *Nature*, *267*(5608), 241–243. doi: 10.1038/267241a0.

Di Lollo, V., & Dixon, P. (1988). Two forms of persistence in visual information processing. *Journal of Experimental Psychology: Human Perception and Performance*, *14*(4), 671–681. doi: 10.1037/0096-1523.14.4.671.

Dontre, A. J. (2020). *The influence of technology on academic distraction: A review*. doi: 10.1002/hbe2.229.

Fortin, C., & Rousseau, R. (2016). *Psychologie cognitive : Une approche de traitement de l'information*. Presses de l'Université du Québec.

Gong, L., Zhao, J., Dai, Y., Wang, Z., Hou, F., Zhang, Y., Lu, Z.-L., & Zhou, J. (2022). Improving iconic memory through contrast detection training with hoa-corrected vision. *Fundamental Research*, *S2667325822002734*, 1–9. doi: 10.1016/j.fmre.2022.06.006.

Graziano, M., & Sigman, M. (2008). The dynamics of sensory buffers: Geometric, spatial, and experience-dependent shaping of iconic memory. *Journal of Vision*, *8*(5), 9–99. doi: 10.1167/8.5.9.

Irwin, D. E., & Thomas, L. E. (2008). Visual sensory memory. In S. J. Luck & A. Hollingworth (Eds.), *Visual memory* (pp. 9–42). Oxford University Press.

Lemaire, P., & Didierjean, A. (2018). *Introduction à la psychologie cognitive (3e éd., tout en couleurs)*. De Boeck supérieur.

Loftus, G. R., Duncan, J., & Gehrig, P. (1992). On the time course of perceptual information that results from a brief visual presentation. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(2), 530–549. doi: 10.1037/0096-1523.18.2.530.

Mack, A., Erol, M., Clarke, J., & Bert, J. (2016). No iconic memory without attention. *Consciousness and Cogni-*

tion: An International Journal, *40*, 1–8. doi: 10.1016/j.concog.2015.12.006.
Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97. doi: 10.1037/h0043158.
National Academies of Sciences, Engineering, and Medicine. (2019). *Reproducibility and replicability in science*. The National Academies Press. doi: 10.17226/25303.
Neisser, U. (1967). *Psychologie cognitive*. Appleton-Century-Crofts.
O'Regan, J. K., & Noë, A. (2001). A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences*, *24*(5), 939–973. doi: 10.1017/s0140525x01000115.
Phillips, I. B. (2011). Perception and iconic memory: What sperling doesn't show. *Mind and Language*, *26*(4), 381–411.
Rensink, R. A. (2014). Limits to the usability of iconic memory. *Frontiers in Psychology*, *5*, 1–99. doi: 10.3389/fpsyg.2014.00971.
Sligte, I. G. (2010). Detailed sensory memory, sloppy working memory. *Frontiers in Psychology*, *1*, 1–99. doi: 10.3389/fpsyg.2010.00175.
Sperling, G. (1960). The information available in brief visual presentations. *Psychological Monographs: General and Applied*, *74*(11), 1–29. doi: 10.1037/h0093759.
Teeuwen, R. R. M., Wacongne, C., Schnabel, U. H., Self, M. W., & Roelfsema, P. R. (2021). A neuronal basis of iconic memory in macaque primary visual cortex. *Current Biology*, *31*(24), 5401–5414. doi: 10.1016/j.cub.2021.09.052.
Thomson, D. M., & Tulving, E. (1970). Associative encoding and retrieval: Weak and strong cues. *Journal of Experimental Psychology*, *86*(2), 255–262. doi: 10.1037/h0029997.

Open practices

The *Open Data* badge was earned because the data of the experiment(s) are available on osf.io/cynv8/
The *Open Material* badge was earned because supplementary material(s) are available on osf.io/cynv8/

Citation

Bradley-Garcia, M., Ben Messaoud, S., Bossilkov, A., Cateaux, C., Tshila N'Sa, C., Potvin-Pilon, A., Abdi-Nassir, H., Assad, N., Balcom, F., Belfaiz, S., Blackburn, C., Bolduc, J., Boutarbouch, I., Béchamp, N., Calderon Robles, M. T., Charbonneau, V., Chatillon, I., Cooper-Ballinger, L., Falconi, M., … Tremblay Jimenez, M. (2023). Capacity and duration of iconic memory from partial reporting of brief stimuli: A replication of Sperling's experiment (1960). *The Quantitative Methods for Psychology*, *19*(3), r1–r11. doi: 10.20982/tqmp.19.3.r011.

Received: 09/09/2023 ~ Accepted: 09/10/2023

Table 1 follows.

Table 1 ■ Table 1 ■ Demographic questions asked to participants following the consent form Questions

Questions and Response options	Responses
What is your biological sex?	
Male	$n = 22$
Female	$n = 42$
What gender do you identify with?	
Cisgender woman	$n = 41$
Cisgender man	$n = 20$
Transgender man	$n = 0$
Transgender woman	$n = 0$
Nonbinary or gender fluid	$n = 0$
Other (specify)	$n = 2$
I prefer not to respond	$n = 1$
What is your age?	
I am _____ years	18-59
What is your highest level of completed education?	
Highschool	$n = 22$
College	$n = 11$
University - Bachelor	$n = 26$
University - Master	$n = 4$
University - Ph.D or Psy.D	$n = 1$
Postdoctorate Degree	$n = 0$
Other (specify)	$n = 0$
Are you currently studying?	
Yes	$n = 37$
No	$n = 27$
(if yes) What program are you currently in?	
High school	$n = 1$
College	$n = 7$
University - Bachelor	$n = 26$
University - Master	$n = 1$
University - Ph.D or Psy.D	$n = 1$
Postdoctorate Degree	$n = 0$
Other (specify)	$n = 1$
What is your native language?	
French	$n = 40$
English	$n = 16$
Spanish	$n = 4$
Other (specify)	$n = 4$
What is your level of proficiency in English?	
Advanced	$n = 53$
Moderate	$n = 10$
Beginner	$n = 1$
What is your ethnic background? Select all that apply.	
White	$n = 45$
Black	$n = 3$
Non-White Latin American	$n = 4$
East Asian	$n = 2$
South Asian/East Indian	$n = 2$
Southeast Asian	$n = 3$
Non-white West Asian, North African, Arab	$n = 4$
Persons of Mixed Origin	$n = 1$
First Nations, Métis, or Inuit	$n = 0$
Other (specify)	$n = 0$
As the stimulus for this study will appear very quickly, do you have a history of photosensitive epilepsy?	
Yes	$n = 2$
No	$n = 62$
Do you have any medical or psychological conditions that could impact your results in an experiment involving short term memory?	
Yes	$n = 5$
No	$n = 59$
In the past 24 hours, have you consumed any recreational drugs or alcohol?	
Yes	$n = 12$
No	$n = 52$
How many hours of sleep did you get last night?	
1-2 hours	$n = 0$
3-4 hours	$n = 1$
5-6 hours	$n = 11$
6-7 hours	$n = 19$
7-8 hours	$n = 26$
9-10 hours	$n = 7$
More than10 hours	$n = 0$
Does this reflect your typical sleep quantity?	
Yes	$n = 48$
No	$n = 16$
Do you have visual impairment?	
Yes	$n = 13$
No	$n = 51$
Do you use visual aids?	
Glasses	$n = 11$
Contact	$n = 0$
Other (specify)	$n = 0$
No	$n = 2$
Do you have hearing impairments?	
Yes	$n = 2$
No	$n = 62$
Do you use hearing aids?	
Yes	$n = 2$
No	$n = 0$

Note. Participants had the option to select "I prefer not to answer," for all questions. If the response "I prefer not to answer" was not selected by at least one participant for that specific question, this answer choice was omitted from the present table.