Psihologijske teme, 31 (2022), 2, 455-475

Izvorni znanstveni rad UDK: 159.953.2.072

https://doi.org/10.31820/pt.31.2.12

The Short-Term Storage as a Buffer Memory between Long-Term Storage and the Motor System? Not Exactly

Ratko Đokić

University of Sarajevo, Faculty of Philosophy, Sarajevo, Bosnia and Herzegovina

Abstract

This study reports a replication and extension of Ulrich and Dietz (1985) on the possible role of the short-term storage (STS) as a buffer memory between long-term storage (LTS) and the motor system in free recall tasks. This STS buffer would allow for the parallel unfolding of the search of LTS and the motor output of responses by temporarily storing retrieved items until their final production. In the study, this view was examined against the alternative hypothesis whereby the fast initial production of items in free recall relies on a controlled or "strategic" search of LTS rather than on the accrual of retrieved items in STS. In two independent replications, 54 native German language speakers and 39 native Bosnian language speakers performed letter verbal fluency task under the manipulation of STS preload (no preload vs. preload) and delay interval (1 sec vs. 10 sec between the presentation of the target letter and the signal for the start of recall). Contrary to the assumption of STS as a buffer memory, the results revealed that STS preload did not cancel out the effect of the delay. Accordingly, the results were interpreted in the context of controlled LTS search models.

Keywords: short-term storage, working memory, long-term storage, motor system, controlled search

Introduction

Short-term storage (STS) serves to maintain a limited number of separate representations active for ongoing processing. Although originally postulated as a unitary and relatively passive memory domain (Atkinson & Shiffrin, 1968; Baddeley, 2012), the concept evolved into a hierarchical, dynamic model of working memory (WM; Baddeley & Hitch, 1974). According to Baddeley and Hitch model, WM employs two information-specific short-term buffers: the phonological loop for

✉ Ratko Đokić, current affiliation: Institute for Social Research in Zagreb, Frankopanska 22/I, 10000 Zagreb, Croatia. E-mail: *ratko@idi.hr*

The author would like to thank Prof. Rolf Ulrich for his inspiration and support for the study. Also, many thanks to the anonymous reviewers for their useful comments on the earlier version of this manuscript and Tana Morić for English proofreading.

storage of verbal information, and the visuospatial sketchpad for upholding visual and/or spatial representations. Furthermore, representations held in these short-term buffers are subjected to a collection of processes jointly termed the central executive (Jonides et al., 2008). In this way, WM concept differentiates simple information storage from executive processes that act upon this information (Baddeley, 2012; Jonides et al., 2008).

Thanks to the central executive, that is to the “crucial importance of its attentional capacity” (Baddeley, 2012, p. 6), WM is implicated in several higher and lower cognitive tasks such as language learning and processing, fluid reasoning, focused or divided attention, to name just a few (see Unsworth & Engle, 2007, for a review). Accordingly, much research has been devoted to exploring the role of the central executive in a broad spectrum of cognitive phenomena, whereas the short-term buffers have been studied mainly in the context of transferring information to long-term storage (LTS). Apparently, there is a lack of insight about the role of STS when information flows in the opposite direction: from LTS toward the output systems. While the central executive functions in the LTS search process are rather well delineated (e.g., Unsworth, 2016), there are not many attempts to account for the role of STS (i.e., phonological loop, specifically) when permanently stored information has to be recalled and reported overtly.

One of the exceptions is the stochastic framework of Ulrich and Dietz (1985; for other proposals on the STS as a buffer in the retrieval process, see Cowan, 2001). Ulrich and Dietz proposed the Simultaneous-Processing Model of LTS search in which STS serves as a buffer memory between LTS and the motor system. The present study aims to replicate and extend Ulrich and Dietz’s work on a possible contribution of the STS when subjects continually retrieve information from LTS, as in the case of free recall. Before presenting the details of the current study, Ulrich and Dietz’s model will be introduced within the context of the basic principles of LTS search models.

LTS search models typically imply both controlled and random components (e.g., Shiffrin & Atkinson, 1969; Troyer et al., 1997). Operations such as setting up a retrieval plan, utilizing appropriate retrieval cues to guide the search, monitoring retrieved information, and deciding whether to continue the search comprise the controlled component. Once the retrieval cue activates the limited area of LTS (or the search set), the representations are subsequently sampled and recovered from this search set. This sampling-with-replacement process constitutes the random component of the search process. It is assumed that the controlled component relies upon the central executive, while the random component exploits associative relations between the cue and the representations in the search set (Gruenewald & Lockhead, 1980; Unsworth & Engle, 2007). While the majority of search models are agnostic on the role of STS in the retrieval process or discuss this role in terms of the inherent STS limitations (see Raaijmakers & Shiffrin, 1980), Ulrich and Dietz (1985) propose that STS contributes to the search process as a waiting room in which responses retrieved from LTS reside until their final motor output. In this way, STS acts as a memory buffer that allows simultaneous search of LTS and generation of motor output.

For example, suppose the participant is instructed to list as many members of a specified taxonomic category (e.g., "animals") as can be recalled. In that case, probably, the appropriate retrieval cues (e.g., "domestic animals", "forest animals", "birds", etc.) will be successively used to delimit the LTS search area to the search set that consists of relevant items (e.g., "cat", "dog", "cow", etc.) and some irrelevant items. In the next retrieval step, items are randomly sampled from within the activated search set one by one. Following each sampling, the monitoring process examines whether the corresponding item is relevant (i.e., belongs to the specified category) and whether it has not been recalled before. Once the item is approved for the final output, the motor channel starts preparing, initializing, and executing motor programs of retrieved information. This retrieval model implies a serial alternation between the retrieval of the item in LTS and its word production. However, since the components of the verbal output (speech preparation, muscle activation, and the pronunciation of the recalled word) all take time, such a time-consuming switching between retrieval and output generation should be reflected in a slow responding rate at the beginning of the recall. Yet, this is not the case: responses at the beginning of the recall are usually produced in a rapid burst (e.g., Bousfield & Sedgewick, 1944). Ulrich and Dietz (1985, p. 245) explain this by assuming that retrieval of new units in LTS and word production of previously retrieved items can go in parallel and "almost independently" from each other. This simultaneity would be possible by introducing the STS as a buffer memory between the retrieval and motor stages of recall. Within this model, the principal role of such a buffer is to receive units from LTS and make them available for the final motor output. In that way, "the retrieval process can put the next unit into the buffer immediately after it finishes putting the current one there irrespective of whether the motor process runs it immediately or not" (Ulrich & Dietz, 1985, p. 245).

Ulrich and Dietz (1985) tested the assumptions of their model by fitting it to the temporal course of the participants' recollection of items in a letter verbal fluency task (fluency task). In this task, participants were recalling 10 words starting with the target letter (e.g., if the participant is presented with the letter "Z", she could cite responses such as "zebra", "zoo", "zombie", etc.). Participants performed the task in two conditions. In one condition, they had a chance to preload STS with items retrieved from LTS before the appearance of the actual signal to start recall (i.e., participants had a 10-sec delay between the presentation of the target letter and the start signal for the overt recall). In the other condition, there was no such preload opportunity (there was a 1-sec delay between the presentation of the target letter and the start signal). The study showed that mean intercompletion times – intervals between the completions of any two consecutive recalls – were shorter for the first four recalled items for the 10-sec delay condition than for the 1-sec delay condition.

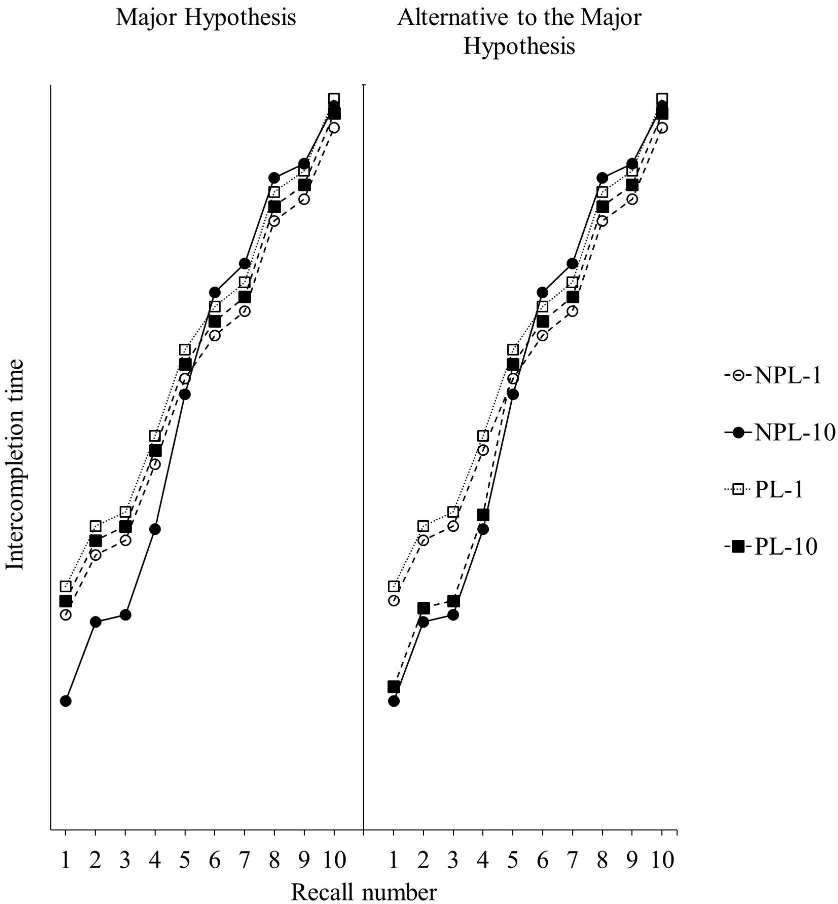
The authors interpreted these results as evidence that participants use the longer delay to begin the retrieval process immediately and to store recalled items in the STS buffer from where they are one by one charged into the motor channel for the final output.

However, an alternative interpretation of this result pattern is possible. For example, during the 10-sec delay, participants may develop a retrieval plan (e.g., how to search for clusters of phonemically related words, such as “zoo”-“zoom”-“Zumba”-“zucchini”) rather than immediately engage in LTS search and storage of retrieved items in STS for the final motor output. Once the overt recall starts, this controlled search component could also lead to the fast initial production of individual items, possibly in the form of a cluster of responses associated with the retrieval cue predefined in a preparatory phase (Gruenewald & Lockhead, 1980). According to this scenario, STS preload would be irrelevant to the velocity of the recall process since the controlled search component resides in the central executive, not in STS. Thus, to strengthen the notion that STS can serve as a buffer memory between LTS and final output, a direct manipulation of STS preload is deemed necessary. In order to evaluate this possibility, the current study employed not only a manipulation of the delay interval (like in Ulrich & Dietz, 1985) but also of the STS preload level. Thus, in half of all trials in each delay condition (1-sec vs. 10-sec), participants generated their recalls while remembering strings of four random digits (STS-preload condition); in the other half of the trials, participants were exempt from such taxation of their STS (no-STS-preload condition). By such factorially combining STS preload and delay, four experimental conditions were created: no-STS-preload – 1-sec delay (NPL-1), no-STS-preload – 10-sec delay (NPL-10), STS-preload – 1-sec delay (PL-1), and STS-preload – 10-sec delay (PL-10) condition.

Therefore, according to our major hypothesis, if STS serves as a buffer memory during LTS search, this should be reflected in the NPL-10 condition (for illustration purposes, please see the graph on the left in Figure 1). In this condition, participants would use the 10-sec delay interval to store as much as possible of items retrieved from LTS in their (initially empty) STS buffer before the overt recall starts. Once the overt production of responses begins, items already waiting in STS would be unloaded into the motor channel one by one in fast succession. This should be observed in short intercompletion times at the beginning of the recall process, as Ulrich and Dietz (1985) did for the first four recalled items (i.e. at the first four recall numbers) in their experiment. In contrast, if there is no time to preload STS with items retrieved from LTS, like in NPL-1 and PL-1 conditions, or the STS buffer is already preloaded, like in PL-10 condition, the LTS search has to be interrupted as soon as the first item is retrieved. Consequently, the observed intercompletion times should not differ among these three conditions.

Figure 1

Mean Intercompletion Time as a Function of Preload, Delay, and Recall Number Expected According to the Major Hypothesis and the Alternative to the Major Hypothesis



Note. Results expected according to the major hypothesis in the NPL-1 and NPL-10 conditions are modeled based on empirical data reported in Ulrich and Dietz (1985, Figure 4); results in the remaining conditions are modeled relative to these two. NPL-1 = no-STS-preload – 1-sec delay; NPL-10 = no-STS-preload – 10-sec delay; PL-1 = STS-preload – 1-sec delay; PL-10 = STS-preload – 10-sec delay; Recall number = position of the recalled item in an array of 10 recalled items.

However, if STS does not serve as a buffer memory, an alternative pattern of results is expected (please see the graph on the right in Figure 1). In this case, only the delay factor but not the STS preload level should affect the intercompletion times. The reason for the dominance of the delay factor in this scenario is that the 10-sec delay interval would be necessary to prepare an elaborated search of LTS. Once initialized, this "strategic" search process, based on the executive functions, could also result in an initial burst of responses, but this time irrespective of STS preload.

The current experiment examined these predictions. It employed the fluency task procedure from Ulrich and Dietz (1985) and extended it by an STS preload factor. Consistent with prior work, the STS preload was set at the upper bound of four items (e.g., Cowan, 2001). Thus, participants performed the fluency task in the STS-preload condition while remembering strings of four random digits, and in the no-STS-preload condition, they always remembered the same digit string: 1234. This procedure, along with the results of Ulrich and Dietz (1985) described above, as they observed significant effects in the first four recall numbers, put the special focus on four initial recall positions in the succeeding analyses.

In addition, theoretical assumptions were tested in parallel in two samples of participants – one consisting of native German language speakers and the other consisting of native Bosnian language speakers. This approach aimed at strengthening the conclusions of the study by comparing the patterns of results observed in two independent, cross-cultural replications (e.g., Ardila, 1995). Pertinent to this, decision to run separate analyses for two samples was also motivated by only scarce evidence of comparable cross-linguistic (letter) verbal fluency norms (especially for the Bosnian language), both at the level of word production (e.g. Oberg & Ramírez, 2006) and at the level of the time-course of retrieval (e.g. Luo et al., 2010).

Method

Participants

Sample of 54 native German language speakers (German sample) participated in the experiment (34 females; age range: 19 – 48, M_{age} = 28.31, SD_{age} = 7.63). All participants received either course credit (at the University of Tübingen) or 6 euros as compensation for their participation in the experiment.

The sample of native Bosnian language speakers (Bosnian sample) consisted of 39 participants (37 females; age range: 20 – 33, M_{age} = 23.90, SD_{age} = 2.93), students at the University of Sarajevo. All participants received course credit for their participation in the experiment.

All participants gave written informed consent that their data can be used for research purposes and published in an anonymous form. The research protocol was approved by the Department of Psychology at the University of Tübingen and the Department of Psychology at the University of Sarajevo.

Materials and Design

The experiment was written in E-Prime 2.0 software (Schneider et al., 2007) and administered by a laptop.

In the experiment, participants tried to memorize strings of digits they saw on the screen while recalling words beginning with a given letter. Participants performed the fluency task in conditions defined by crossing two levels of STS preload (no-STS-preload, STS-preload) with two delay intervals (1-sec, 10-sec). The order of the STS preload conditions was block-randomized. Two delay conditions were presented in random order within each block. The target letters were randomly sampled without replacement from the following 20 capital letters: A, B, D, E, F, G, H, I, J, K, L, M, N, O, P, R, S, T, U, V. Therefore, there were five letters randomly assigned to each of the four experimental conditions.

Procedure

An individual trial proceeded as follows: after the participant initialized the trial by pressing any key on the keyboard, four digits were presented one by one on the screen. The rate of digit presentation was 1 sec per digit. Depending on the STS preload condition, the sequence comprised four distinct random digits (STS-preload condition) or always the same digit order: 1234 (no-STS-preload condition). Immediately after presenting the last digit in the sequence, a target letter appeared on the screen. Participants had to withhold an overt recall until the recall signal ("START") replaced the initial letter on the screen. Depending on the delay condition, there could be a 1-sec or 10-sec delay between the presentation of the target letter and the recall signal.

Participants' task was to recall as many nouns starting with the target letter as possible. They were instructed not to repeat responses as well as not to cite personal names or geographical toponyms. Participants pressed the spacebar with every noun they pronounced; the computer registered the response time for every spacebar press. The intercompletion times were measured as time intervals between two consecutive spacebar presses. After the participant recalled 10 nouns, the signal "???" was presented on the screen. Participants should stop recalling words on this signal and were asked to reproduce the digits presented at the beginning of the trial. Participants' task was to recall all the digits from the sequence in the correct order of their presentation. They reproduced digits by typing them on the keyboard.

After the reproduction of the digit sequence, the computer provided feedback (lasting 4 sec) about the number of digits correctly recalled in the current trial. After the feedback, there was a 14-sec pause before the start of the subsequent trial. Participants were told that both parts of the experiment – digit-reproduction task and fluency task – were equally important. At the beginning of each block of trials, participants were informed which type of digit sequence (random or "1234") they would be presented with.

There were four practice trials, one for each STS preload x delay condition, at the beginning of the experiment using the initial letters C, W, Q, and Z (C, Š, Z, and Ž, in the Bosnian sample). The session length was about 35 min and all sessions were conducted in a sound-attenuated chamber.

Results

All statistical analyses were performed in parallel in both German and Bosnian samples, and results were compared between samples at the descriptive level.

Digit-Reproduction Task

Digit-reproduction performance was almost identical in two samples (Table 1). In no-STS-preload conditions, where participants had to recall the same 1234 digit sequence, the digit-reproduction was perfect. In the STS-preload conditions, the accuracy decreased. Post-hoc comparisons (at the Bonferroni-corrected alpha level $\alpha = .05/6 = .008$) between two levels of STS preload confirmed this, both at 1-sec ($F(1, 53) = 132.46$, $\eta_p^2 = .71$, in the German sample; $F(1, 38) = 92.54$, $\eta_p^2 = .71$, in the Bosnian sample) and 10-sec delay ($F(1, 53) = 81.39$, $\eta_p^2 = .61$, in the German sample; $F(1, 38) = 58.38$, $\eta_p^2 = .61$, in the Bosnian sample), all ps < .001. This decrement in the digit-reproduction accuracy between two STS preload levels indicates that presenting the participants with the random four-digit sequences taxed their STS.

Table 1

Mean Proportions (and Standard Deviations) of Correctly Recalled Digits

		Native German Language Speakers (N = 54)		Native Bosnian Language Speakers (N = 39)	
STS preload	Delay	1-sec	10-sec	1-sec	10-sec
no-STS-preload		1 (0)	1 (0)	1 (0)	1 (0)
STS-preload		.67 (.21)	.73 (.22)	.67 (.21)	.74 (.21)

Furthermore, accuracy levels between the two conditions with STS preload (1-sec vs. 10-sec delay) did not differ due to the Bonferroni correction ($F(1, 53) = 4.31$, $\eta_p^2 = .08$, in the German sample; $F(1, 38) = 4.65$, $\eta_p^2 = .11$, in the Bosnian sample; both ps = .04). The absence of this difference indicates that participants were not adjusting their strategy in the digit-reproduction task to the current delay interval, in order to increase their performance in the fluency task. Due to this, the following analyses of the fluency task results were performed in a complete sample, that is without filtering out any results based on the criteria of participants' accuracy in the digit-reproduction part of the experiment (see also Đokić et al., 2018)[1].

Letter Verbal Fluency Task

Prior to the final analyses, trials with the intercompletion time shorter than 200 msec and longer than $M + 3SD$ for a particular recall number in a particular STS preload x delay condition were discarded from further analysis. The percentage of such trials was 2.10% in the German and 2.38% in the Bosnian sample.

Mean intercompletion time as a function of STS preload, delay, and recall number for both German and Bosnian samples is presented in Figure 2; Panel A compares current results to the estimates of the Ulrich and Dietz (1985) model; Panel B emphasizes the effects observed at recall numbers 1-4.

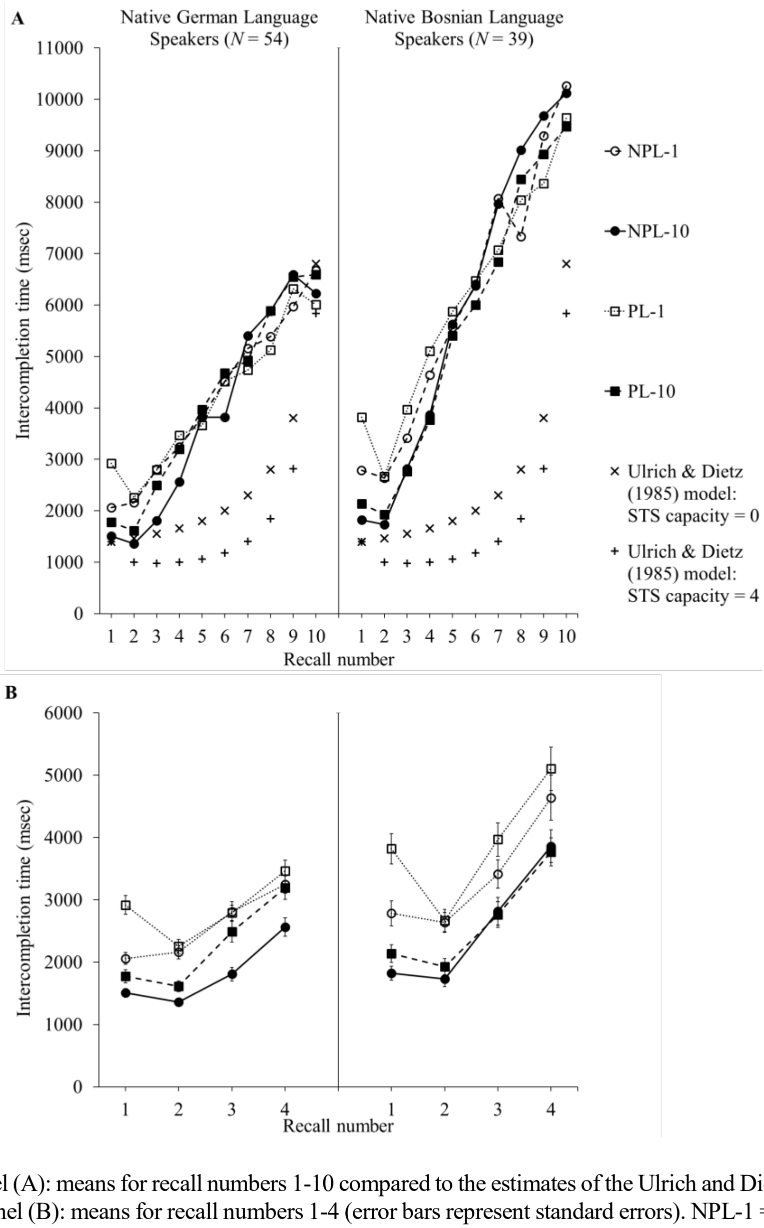
At a descriptive level, the expected pattern of results emerged in the German sample. At the first four recall positions, intercompletion times in the NPL-10 condition were shorter than in the three remaining conditions (Figure 2: Panel B). As for these three later conditions, mean intercompletion times in the PL-10 condition were followed by those in two 1-sec delay conditions, with or without STS preload. From the recall number 5 on, this arrangement of mean intercompletion times was lost, with the four experimental conditions constantly interchanging their positions (Figure 2: Panel A). However, in the Bosnian sample, results inclined more toward the view alternative to the major hypothesis, as a more pronounced convergence of NPL-10 and PL-10 conditions at recall numbers 1-4 led to a dissociation between 1-sec and 10-sec delay conditions.

Based on the research aims and hypothesis stated in the Introduction, the rest of the result section is organized around the following two points: a) assessing the replicability of Ulrich and Dietz (1985) and b) testing the major hypothesis of the current study.

[1] However, a crosscheck analysis performed in a subsample of 28 native German language speaking participants, who exceeded 65% accuracy criteria in the digit-reproduction task in both STS preload conditions, confirmed the results reported here.

Figure 2

Mean Intercompletion Times as a Function of STS Preload, Delay, and Recall Number



Note. Panel (A): means for recall numbers 1-10 compared to the estimates of the Ulrich and Dietz (1985) model. Panel (B): means for recall numbers 1-4 (error bars represent standard errors). NPL-1 = no-STS-preload – 1-sec delay; NPL-10 = no-STS-preload – 10-sec delay; PL-1 = STS-preload – 1-sec delay; PL-10 = STS-preload – 10-sec delay; Recall number = position of the recalled item in an array of 10 recalled items.

Replicability of Ulrich and Dietz (1985) Results

Replicability of Ulrich and Dietz (1985) empirical results was assessed by mirroring their original analysis of the delay (1-sec, 10-sec) x recall number (1 to 10) interaction within the no-STS-preload condition. As in the original study, this interaction was significant both in German, $F(9, 477) = 4.12$, $\eta_p^2 = .07$, $p < .001$, and Bosnian samples, $F(9, 342) = 2.46$, $\eta_p^2 = .06$, $p = .01$. Further comparisons of mean intercompletion times for conditions with 1-sec and 10-sec delay revealed a significant difference ($\alpha = .05$)† up to the recall number 4‡. Therefore, as expected, the overall pattern of the Ulrich and Dietz's results is generally replicated: when there was no STS preload, mean intercompletion times at the beginning of the recall process were shorter for the condition with the longer delay interval. However, there are two noticeable differences between theirs and the current results.

Thus, in the Ulrich and Dietz (1985) study, the first intercompletion time (i.e., the time from onset of the actual recall signal up to the completion of the first recall) in both delay conditions was much shorter than the second one. However, the only significant difference (at the Bonferroni-corrected alpha level of $\alpha = .05/4 = .0125$) in the current experiment had the opposite direction: in PL-1 condition mean intercompletion time was longer at recall number 1 than at recall number 2 ($F(1, 53) = 17.72$, $\eta_p^2 = .25$, in the German sample; $F(1, 53) = 15.32$, $\eta_p^2 = .29$, in the Bosnian sample; both *p*s < .001). In three other conditions, intercompletion times for the first and the second item were equal.

The second difference between the results of the two experiments was a deviation of the current mean intercompletion times from the ones estimated by the stochastic model of Ulrich and Dietz (1985). While predictions of the model closely fitted empirical results in the original study, this was not the case in the current experiment (Figure 2: Panel A). More precisely, compared to the model estimates for STS capacity of zero items (which correspond to the present NPL-1, PL-1, and PL-10 conditions) and four items (which correspond to the NPL-10 condition), current mean intercompletion times were longer and demonstrated a much steeper increase across consecutive recall numbers.

Test of the Major Hypothesis

As stated above, in accordance with the major hypothesis of the study, mean intercompletion times for the first four recalled items should be shorter in the NPL-10 condition than in the three remaining experimental conditions, which are supposed to be homogenous in this regard. Operationally defined, the test of the

† α-level adopted from Ulrich and Dietz (1985).

‡ In addition to these, differences in the recall number 6 in the German and recall number 8 in the Bosnian sample were also significant in the present study (results are not presented).

major hypothesis comprised of the assessment of two planned contrasts, examined for each of 10 recall numbers. The first contrast tested whether the mean intercompletion time in the NPL-10 condition was shorter compared to the pooled mean of three other conditions (NPL-1, PL-1, PL-10). If this was the case, the second contrast tested whether these three later conditions were homogeneous regarding the mean intercompletion times; for this purpose one–way repeated measures analysis of variance (RM ANOVA across NPL-1, PL-1, and PL-10 condition) was employed, and non-significant F-ratio was expected. This pattern of results (significant first contrast and non-significant second contrast) was expected at recall numbers 1-4, whereas for the remaining six recall positions no significant differences between experimental conditions were assumed. While emulating the original analysis at the level of individual recall numbers applied in Ulrich and Dietz (1985; please see above), this approach to testing the major hypothesis was based on its exact assumptions and the intention toward maximal reduction of the number of required statistical tests in an attempt to protect against the Type I error. In that sense, the adopted alpha-level for all planned contrasts was $\alpha = .01$.

As already hinted at the descriptive level, the comprehensive statistical confirmation of the major hypothesis fell short (Table 2). Namely, in the German sample, mean intercompletion times for recall numbers 1-4 in the NPL-10 condition, as expected, were significantly shorter compared to the corresponding pooled mean intercompletion times for three other experimental conditions. However, contrary to the major hypothesis, mean intercompletion times in these three other experimental conditions were not homogeneous at recall numbers 1 and 2, indicated by significant second planned contrast. Further post hoc analysis (at the Bonferroni-corrected alpha-level $\alpha = .05/20 = .003$) revealed that for these two recall numbers mean intercompletion times in PL-10 condition were shorter than pooled mean for the two 1-sec delay conditions; in addition, at the recall number 1, mean intercompletion time in NPL-1 condition was shorter than in PL-1 condition (results are not reported).

Table 2

Two Planned Contrasts per each Recall Number Within the Test of the Major Hypothesis

		Native German Language Speakers (N = 54)			Native Bosnian Language Speakers (N = 39)		
Recall number	Contrast*	F	p	η_p^2	F	p	η_p^2
1	1	100.76	<.001	.66	64.45	<.001	.63
	2	54.10	<.001	.51	38.20	<.001	.50
2	1	99.42	<.001	.65	29.41	<.001	.44
	2	22.88	<.001	.30	13.29	<.001	.26
3	1	51.24	<.001	.49	6.19	.017	.14
	2	2.10	.128	.04	10.40	<.001	.21
4	1	23.36	<.001	.31	5.27	.027	.12
	2	1.13	.328	.02	6.29	.003	.14

Table 2 - *Continued*

		Native German Language Speakers ($N = 54$)			Native Bosnian Language Speakers ($N = 39$)		
Recall number	Contrast*	F	p	η_p^2	F	p	η_p^2
5	1	0.00	.993	.00	0.00	.982	.00
	2	0.82	.443	.02	0.49	.615	.01
6	1	11.81	.001	.18	0.05	.824	.00
	2	0.18	.834	.00	0.61	.544	.02
7	1	2.90	.094	.05	1.47	.233	.04
	2	0.97	.382	.02	2.27	.111	.06
8	1	2.11	.153	.04	3.09	.087	.08
	2	1.95	.148	.04	1.37	.261	.03
9	1	0.54	.464	.01	1.47	.232	.04
	2	0.68	.508	.01	0.77	.467	.02
10	1	0.26	.612	.00	0.22	.641	.01
	2	1.60	.206	.03	0.47	.626	.01

Note. *Contrast 1: Mean intercompletion time in no-STS-preload – 10-sec delay vs. pooled mean intercompletion time in no-STS-preload – 1-sec delay, STS-preload – 1-sec delay, and STS-preload – 10-sec delay. German sample: $df = 1, 53$. Bosnian sample: $df = 1, 38$.

Contrast 2: ANOVA with experimental condition (no-STS-preload – 1-sec delay, STS-preload – 1-sec delay, STS-preload – 10-sec delay) as within-subjects factor. German sample: $df = 2, 106$. Bosnian sample: $df = 2, 76$.

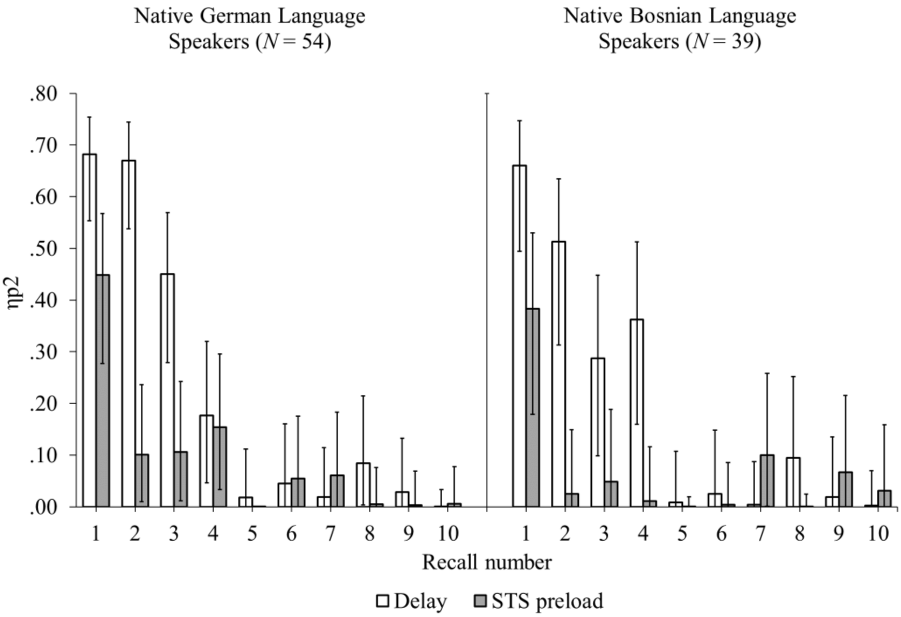
In comparison, results in the Bosnian sample moved even more from the major hypothesis and toward its alternative. While in the German sample first planned contrast was significant at all recall numbers 1-4, in the Bosnian sample this was the case only for recall numbers 1 and 2. The reason for this was a convergence of two 10-sec delay conditions (no-STS-preload and STS-preload) toward each other, which resulted in a more pronounced dissociation between 1-sec and 10-sec delay conditions. Consequently, the second planned contrast was significant at all recall numbers 1-4. Post hoc analysis ($\alpha = .003$) demonstrated that mean intercompletion times at recall numbers 1-4 in the PL-10 condition were shorter than corresponding pooled means for the two 1-sec delay conditions (results are not presented).

In order to reconsider the view alternative to the major hypothesis, namely the one by which mean intercompletion times at the first four recall numbers are more strongly influenced by delay than by STS preload, indicators of effect size (η_p^2) for these two factors were compared at the descriptive level (Figure 3). In accord with this view, effects of delay at first four recall positions were more prominent than the effects of STS preload, with no or relatively small overlap between respective confidence intervals (only exception being recall number 4 in the German sample, with both effects of about equal size)[2].

[2] All the analyses of intercompletion times were crosschecked in the complete sample of participants ($N = 93$) and the results aligned with the conclusions presented in the text.

Figure 3

Size of Effects (Error Bars Represent 90% Confidence Intervals) of Delay and STS Preload on Mean Intercompletion Time per Recall Number



Note. Due to the one-sided nature of *F*-test, 90% confidence intervals for η_p^2 correspond to the 95% confidence intervals normally calculated for two-sided statistics (Lakens, 2013). Recall number = position of the recalled item in an array of 10 recalled items.

Discussion

The aim of the study was the replication and extension of Ulrich and Dietz (1985) on the possible role of STS as a buffer memory between LTS and the motor system in situations of free recall. This STS buffer would allow for a simultaneous and almost independent unfolding of the search of LTS and the motor production of responses by temporary storage of retrieved items until their final verbal output. However, current results incline more toward the alternative view, by which the fast initial output of items in free recall tasks relies on the controlled search of LTS, and not on the preceding accrual of retrieved items in the STS buffer.

Thus, although the current experiment generally replicates the original results in Ulrich and Dietz (1985) – with no STS preload, mean intercompletion times at

Descriptives for intercompletion times and the results of the test of the major hypothesis (Contrast 1 and Contrast 2) are presented in the Appendix.

first four recall positions were shorter in 10-sec delay, than in 1-sec delay condition – there were two conceptually important deviations. First, Ulrich and Dietz observed much shorter intercompletion times for the first response than for the second one. As an explanation, they offered an assumption that the first item retrieved from LTS "immediately enters the motor channel and is prepared for the forthcoming recall before the actual recall signal appears" (p. 264). At the same time, the authors would expect the exact pattern that is observed in the current experiment – the first and second intercompletion times being about equal (or the second one shorter than the first) – under the assumption that only STS, and not motor channel is preloaded with copies from LTS. Namely, if at least two retrieved items are stored in STS before the actual recall signal appears, then intercompletion times for recall numbers 1 and 2 should simply indicate the rate of execution of motor programs for these two responses. Therefore, mean intercompletion times at recall numbers 1 and 2 should not differ (or the second one should be shorter, if the time to perceive the actual start signal is taken into account as well). However, this assumption in the current experiment is falsified by the observation of this pattern in all experimental conditions, irrespective of whether STS was preloaded or not even before the start of the LTS search process. Therefore, this finding advocates against the role of STS as a buffer from which retrieved items are one by one released into the motor channel for their final output.

At the same time, one possible explanation for longer intercompletion time at the first than at the second recall number in the PL-1 condition, and the lack of this difference in three other conditions, could be the interference caused by the parallel digit-reproduction task. Since in the former condition, participants had to remember the random string of four digits that was presented immediately before the start signal for the fluency task, it is probable that the initialization of overt production of responses in the fluency task was hampered by the subvocal rehearsal of the digits (e.g., Baddeley, 2012). In the three later conditions, this interference was absent either because there was no need for the rehearsal of digits (no-STS-preload conditions) or because participants had enough time to prepare for the appearance of the start signal and to distribute the rehearsal of digits over the delay interval (PL-10 condition).

The second deviation of current results from those in Ulrich and Dietz (1985) was that the current mean intercompletion times were observably longer and increased successively in a much steeper fashion than in the original experiment and compared to the predictions of the model. At this point, a reliable explanation for such a difference between two sets of empirical results could not be offered; one possible reason could be differences in the characteristics of the two samples or differences between the two cohorts from which samples were drawn. However, from the conceptual standpoint, more important is the deviation of the steepness (slope) of the increase of mean intercompletion times across 10 recall numbers. The more pronounced increase in current results compared to the one predicted by the model could be interpreted as another argument against the role of STS as a buffer memory in which already retrieved items await their transfer into the motor system and their final output. As already discussed above, if this was the case, initial intercompletion times would simply reflect the pace of motor processing of prepared responses, and a more flattened curve of mean intercompletion times would be expected, at least in its section from recall number 1 to 4.

Furthermore, the statistical confirmation of the current study's major hypothesis did not succeed. Initial (four) mean intercompletion times in the NPL-10 condition, in which participants had enough time to preload their STS with items retrieved from LTS before the overt start of recall, were shorter than in three other experimental conditions in the German sample, while in the Bosnian sample this was the case only for recall numbers 1-2. At the same time, mean intercompletion times in the PL-10 condition were shorter than in two 1-sec delay conditions, and this tendency was statistically confirmed for recall numbers 1 and 2 in the German sample and for all recall numbers 1-4 in the Bosnian sample. In concordance to this, it seems that mean intercompletion times at recall numbers 1-4 were more affected by delay than by STS preload. Taken all together, these results incline toward the view alternative to the major hypothesis, by which mean intercompletion times are the manifestation of controlled search of LTS, and not of the mechanical process of discharging items retrieved from LTS and stored in the STS into the motor system.

A controlled search of LTS consists of two basic phases. The first phase involves strategic selection and/or generation of retrieval cues, by which the search is limited to the so-called search sets: narrower, well-defined subcategories of relevant information. After defining the search set, another, randomized search phase follows (e.g., Unsworth & Engle, 2007). Within this phase, individual information is sampled one by one from the search set based on their associative links to the applied retrieval cue(s).

Thus, for example, in performing verbal fluency tasks, participants typically produce responses in "bursts" or clusters – groups of semantically related responses (e.g., Bousfield & Sedgewick, 1944; Troyer et al., 1997). Responses within the cluster are produced in a fast procession, corresponding to the randomized aspect of the search in which the items associated with the activated retrieval cue(s) are emitted. The clusters themselves are separated by longer time intervals, reflecting the strategic aspect of the search in which a new retrieval cue(s) is defined, and the search is directed into the new cluster. Accordingly, Gruenewald and Lockhead (1980) suggest that the search is not for individual items per se, but for search sets or, in their terminology, semantic fields. The semantic field represents "an organization of concepts or meanings" that, once accessed, allows the production of a cluster – a set of words associated in terms of common meaning, and not in terms of item-to-item relations (Gruenewald & Lockhead, 1980, p. 239). Numerous empirical findings suggest that executive, and not storage component of WM is responsible for managing the controlled search of LTS (for a review, see Unsworth & Engle, 2007; Unsworth, 2016).

The results of the current experiment fit into this framework. Disassociation of NPL-10 condition, followed by PL-10 condition, from the 1-sec delay ones at recall numbers 1-4 (seen in the German, but especially in the Bosnian sample) suggest that participants used the prolonged delay interval for setting up the controlled search of LTS, and not for retrieval of discrete items and their storage in STS. Once the overt recall started, participants were able to rapidly produce initial responses in the form of a cluster of items. Pronounced effect of delay, compared to the effect of STS preload, further corroborates this notion.

However, one could expect an even clearer picture if the "STS preload" aspect of the current experiment is modified. Namely, the STS preload in this experiment was fixed and was hence the same for all participants: it was set at four items, the average upper STS capacity bound defined in the literature (Cowan, 2001). However, this STS capacity seems to vary among individuals from one to about four (see Jonides et al., 2008, for a review). If the number of items to store in STS exceeds this capacity, successful retention of items demands the engagement of executive processes (Baddeley & Hitch, 1974; Unsworth & Engle, 2006). This subtle distinction between simple STS storage and manipulation of units in WM points to the possibility that, if some participants had STS capacity lesser than four items, their intercompletion times in the STS-preload conditions could be affected not only by STS preload but also by increased demands on attention (see also Baddeley et al., 1984). Therefore, further studies should assure that the STS preload procedure burdens only storage and not the executive component of WM. Specifically, the STS preload should be customized to the STS capacity of each individual participant. This could result in further convergence of mean intercompletion times between PL-10 and NPL-10 conditions.

In conclusion, current results by no means indicate that there is no role for STS in search of LTS. On the contrary, STS is probably a domain where various processes of controlled search take place. First such a candidate could be a monitoring process by which individual retrieved items are compared to the task requirements and/or to the already produced responses and cleared for overt report accordingly (Shiffrin & Atkinson, 1969; Unsworth, 2016). Another role could be upholding the count of failed retrievals (irrelevant items or already recalled items) that is compared to the implicitly defined criteria for stopping the search (Raaijmakers & Shiffrin, 1980). However, current results do indicate that the retrieval processes accommodated by STS probably could not be reduced to the temporary storage of items recalled from LTS while they passively wait to be released into the motor channel one by one, for their final output.

References

Ardila, A. (1995). Directions of research in cross-cultural neuropsychology. *Journal of Clinical and Experimental Neuropsychology, 17*(1), 143–150. https://doi.org/10.1080/13803399508406589

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The psychology of learning and motivation* (Vol. 2, pp. 89–195). Academic Press. https://doi.org/10.1016/S0079-7421(08)60422-3

Baddeley, A. D. (2012). Working memory: Theories, models, and controversies. *Annual Review of Psychology, 63*, 1–29. https://doi.org/10.1146/annurev-psych-120710-100422

Baddeley, A. D., & Hitch, G. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 8, pp. 47–89). Academic Press. https://doi.org/10.1016/S0079-7421(08)60452-1

Baddeley, A. D., Lewis, V. J., Eldridge, M., & Thomson, N. (1984). Attention and retrieval from long-term memory. *Journal of Experimental Psychology: General, 113*(4), 518–540. https://doi.org/10.1037/0096-3445.113.4.518

Bousfield, W. A., & Sedgewick, C. H. W. (1944). Analysis of sequences of restricted associative responses. *Journal of General Psychology, 30*(2), 149–165.

https://doi.org/10.1080/00221309.1944.10544467

Cowan, N. (2001). The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences, 24*(1), 97–185.

https://doi.org/10.1017/S0140525X01003922

Đokić, R., Koso-Drljević, M., & Đapo, N. (2018). Working memory span tasks: Group administration and omitting accuracy criterion do not change metric characteristics. *PLoS ONE*, *13*(10), Article e0205169. https://doi.org/10.1371/journal.pone.0205169

Gruenewald, P. J., & Lockhead, G. R. (1980). The free recall of category examples. *Journal of Experimental Psychology: Human Learning and Memory, 6*(3), 225–240. https://doi.org/10.1037/0278-7393.6.3.225

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology, 59*, 193–224. https://doi.org/10.1146/annurev.psych.59.103006.093615

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology, 4*, Article 863. https://doi.org/10.3389/fpsyg.2013.00863

Luo, L., Luk, G., & Bialystok, E. (2010). Effect of language proficiency and executive control on verbal fluency performance in bilinguals. *Cognition, 114*(1), 29–41.

https://doi.org/10.1016/j.cognition.2009.08.014

Oberg, G., & Ramírez, M. (2006). Cross-linguistic meta-analysis of phonological fluency: Normal performance across cultures. *International Journal of Psychology, 41*(5), 342–347. https://doi.org/10.1080/00207590500345872

Raaijmakers, J. G. W., & Shiffrin, R. M. (1980). SAM: A theory of probabilistic search in associative memory. In G. H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory.* (Vol. 14, pp. 207–262). Academic Press.

Schneider, W., Eschman, A., & Zuccolotto, A. (2007). *E-Prime user's guide*. Psychology Software Tools Inc.

Shiffrin, R. M., & Atkinson, R. C. (1969). Storage and retrieval processes in long-term memory. *Psychological Review, 76*(2), 179–193. https://doi.org/10.1037/h0027277

Troyer, A. K., Moscovitch, M., & Winocur, G. (1997). Clustering and switching as two components of verbal fluency: Evidence from younger and older healthy adults. *Neuropsychology, 11*(1), 138–146. https://doi.org/10.1037/0894-4105.11.1.138

Ulrich, R., & Dietz, K. (1985). The short-term storage as a buffer memory between long-term storage and the motor system: A simultaneous-processing model. *Journal of Mathematical Psychology, 29*(3), 243–270. https://doi.org/10.1016/0022-2496(85)90008-2

Unsworth, N. (2016). Working memory capacity and recall from long-term memory: Examining the influences of encoding strategies, study time allocation, search efficiency, and monitoring abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*(1), 50–61. https://doi.org/10.1037/xlm0000148

Unsworth, N., & Engle, R. W. (2006). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language, 54*(1), 68−80. https://doi.org/10.1016/j.jml.2005.06.003

Unsworth, N., & Engle, R. W. (2007). The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review, 114*(1), 104–132. https://doi.org/10.1037/0033-295X.114.1.104

Kratkoročno skladište kao međuspremnik između dugotrajnog skladišta i motoričkoga sustava? Zapravo ne

Sažetak

Istraživanje je replikacija i proširenje studije Ulricha i Dietza (1985) o mogućoj ulozi kratkoročnoga skladišta (KS) kao međumemorije između dugoročnoga skladišta (DS) i motoričkoga sustava u zadacima slobodnoga dosjećanja. Takva međumemorija omogućila bi paralelno odvijanje pretrage DS-a i motoričke produkcije odgovora privremenim pohranjivanjem pronađenih čestica do njihove verbalizacije. U studiji je to stajalište suprotstavljeno alternativnoj hipotezi prema kojoj se brza inicijalna produkcija odgovora u slobodnome dosjećanju temelji na kontroliranoj ili „strateškoj" pretrazi DS-a, a ne na akumulaciji odgovora u KS-u. U dvjema neovisnim replikacijama 54 izvorna govornika njemačkoga jezika i 39 izvornih govornika bosanskoga jezika izvršavali su zadatak fonemske verbalne fluentnosti u uvjetima manipulacije opterećenja KS-a (bez opterećenja naspram s opterećenjem) i intervala odgode (1 s naspram 10 s odgode između prezentacije ciljnoga slova i signala za početak produkcije odgovora). Suprotno pretpostavci o KS-u kao međumemoriji, rezultati nisu pokazali da je efekt odgode poništen opterećenjem KS-a. U skladu s time rezultati su interpretirani u kontekstu modela kontrolirane pretrage DS-a.

Ključne riječi: kratkoročno skladište, radno pamćenje, dugoročno skladište, motorički sustav, kontrolirana pretraga

Primljeno: 10. 11. 2021.

Appendix

Table 1A

Means and Standard Deviations for Intercompletion Times (msec) per Preload, Delay, and Recall Number (N = 93)

		Delay			
		1-sec		10-sec	
STS preload	Recall number	*M*	*SD*	*M*	*SD*
no-STS-preload	1	2363.40	1028.57	1641.65	612.67
	2	2360.65	907.60	1517.55	603.07
	3	3064.46	1296.82	2231.50	1180.00
	4	3830.35	1781.72	3107.91	1471.83
	5	4554.56	1900.89	4576.91	1985.61
	6	5308.02	2260.08	4892.14	2473.38
	7	6377.98	3159.94	6478.83	3095.04
	8	6200.78	2708.39	7194.92	3302.13
	9	7364.77	3747.19	7886.03	4082.43
	10	8179.93	3811.03	7858.90	4104.48
STS-preload	1	3296.68	1358.67	1927.40	840.95
	2	2428.05	982.41	1745.83	725.55
	3	3285.45	1397.55	2605.92	1264.06
	4	4152.29	1871.12	3435.48	1412.30
	5	4590.53	2408.94	4571.77	2018.91
	6	5331.60	2409.04	5231.83	2273.77
	7	5715.41	2820.97	5727.33	2448.98
	8	6352.14	3207.18	6961.60	3246.92
	9	7176.39	3037.81	7548.86	3624.23
	10	7534.10	4167.88	7799.64	3364.24

Table 2A

Two Planned Contrasts per each Recall Number Within the Test of the Major Hypothesis (N = 93)

Recall number	Contrast*	*F*	*p*	η_p^2
1	1	147.43	<.001	.62
	2	86.50	<.001	.48
2	1	105.75	<.001	.53
	2	34.96	<.001	.28
3	1	39.38	<.001	.30
	2	10.26	<.001	.10
4	1	22.77	<.001	.20
	2	6.58	.002	.07
5	1	0.00	.981	.00
	2	0.01	.989	.00
6	1	3.63	.060	.04
	2	0.08	.924	.00
7	1	3.95	.050	.04
	2	2.96	.054	.03
8	1	5.13	.026	.05
	2	2.43	.091	.03
9	1	1.98	.163	.02
	2	0.38	.686	.00
10	1	0.00	.955	.00
	2	1.14	.322	.01

Note. *Contrast 1: Mean intercompletion time in no-STS-preload – 10-sec delay vs. pooled mean intercompletion time in no-STS-preload – 1-sec delay, STS-preload – 1-sec delay, and STS-preload – 10-sec delay; *df* = 1, 92.

Contrast 2: ANOVA with experimental condition (no-STS-preload – 1-sec delay, STS-preload – 1-sec delay, STS-preload – 10-sec delay) as within-subjects factor; *df* = 2, 184.