Psychon Bull Rev
DOI 10.3758/s13423-015-0853-0

BRIEF REPORT

Forward scanning in verbal working memory updating

Yoav Kessler[1] · **Klaus Oberauer**[2]

© Psychonomic Society, Inc. 2015

Abstract Effective use of working memory (WM) for high-level cognitive tasks requires coordinating two conflicting requirements: robust maintenance and rapid updating. Models of WM suggest that these demands are coordinated by a gate between perceptual input and WM. Previous work with a letter-updating paradigm (Kessler & Oberauer, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*, 738–754, 2014) supported a scanning and gate-switching (SGS) model of WM updating. The present work provides further evidence for the SGS model. Participants were required to keep track of the last letter that appeared in each of a row of frames on the screen. On each updating step, a variable subset of letters in varying positions in the row had to be updated. The SGS model assumes that on each updating step, participants scan through the memory set sequentially, opening the gate when a letter requires updating, and closing the gate when the next letter needs to be maintained. As is predicted by the SGS model, the reaction times for each updating step increased with the number of updated items and with the number of gate switches. In addition, the present experiment provides direct evidence supporting the scanning assumption of the model. Hebrew-speaking participants performed the task with either Hebrew or English letter stimuli, in different blocks. As was predicted, the scanning direction of the stimulus set was from left to right in English and from right to left in Hebrew. The SGS model fit the data only when the scanning direction was taken into account, establishing the role of item-based forward scanning during WM updating.

Keywords Working memory · Updating · Forward scanning · Gating

Working memory (WM) provides the ability to robustly maintain task-relevant information in mind, as well as to update it with new information when needed. These two functions, maintenance and updating, are needed in a host of high-level cognitive tasks, such as mental arithmetic and reasoning. Recent empirical and modeling work has emphasized the conflict between the maintenance and updating functions of WM and focused on the mechanisms by which this conflict is resolved. Specifically, WM enables robust maintenance of information even in the face of interference from the inner or outer environment. Computational models, based on neurophysiological data, suggest that shielding WM representations from interference is achieved by controlled gating of the WM input (e.g., Braver & Cohen, 2000; Frank, Loughry, & O'Reilly, 2001; O'Reilly & Frank, 2006). This is implemented by a gate that separates the perceptual input from WM representations and by recurrent activation within WM. The gate to WM is closed by default, and updating is achieved by a transient opening of this gate when relevant information is available.

The idea of WM gating is supported by a large body of neurophysiological evidence and computational modeling. However, the behavioral costs of gate opening and closing have only recently been observed. In a previous work (Kessler & Oberauer, 2014), we examined the time course of updating verbal information in WM. We used a letter-updating paradigm devised by Kessler and Meiran (2008; see also

✉ Yoav Kessler
kesslery@bgu.ac.il

[1] Department of Psychology and Zlotowski Center for Neuroscience, Ben-Gurion University of the Negev, Beer-Sheva, Israel 84105

[2] Department of Psychology, University of Zurich, Zurich, Switzerland

Published online: 12 May 2015

Ecker, Lewandowsky, & Oberauer, 2014). Participants had to keep track of the last letter that appeared inside each of a row of frames on the screen. The number of frames (and, hence, the memory set size) varied between two and four in separate runs. Each run of trials began with the presentation of the frames, with a letter inside each of them. Then, a sequence of trials began. In each trial, new letters appeared inside some, all, or none of the frames. The participants were instructed to update their memory with the new letters, so that they only remembered the last letter that had appeared in each frame. A keypress was required in each trial in order to move to the next trial. The reaction time (RT) for this response was the dependent measure, reflecting the time taken to update WM in each trial. The number of updated items in each trial was manipulated, as well as the sequence of updated and repeated items. For example, updating two out of four letters could involve updating the two leftmost letters (Condition UURR; U = update, R = repeat), the two rightmost letters (RRUU), the letters in Positions 1 and 3 (URUR), and several other patterns. Thus, most trials would involve the updating of some items in the current memory set while maintaining the others, and the different patterns of U and R would impose different demands on the control of updating and maintenance by opening and closing the gate into WM, as we detail below.

Two versions of the paradigm were examined in separate experiments. In the *partial-display* paradigm, only updated letters were presented in each trial; an asterisk was instead presented inside frames in which the maintained letter was not updated. For example, assume that the maintained letters are L, G, P, and Z, and that the last letter had to be updated to B. In this case, asterisks would appear inside the three leftmost frames, to denote that their corresponding letters did not change, and B would appear in the rightmost frame. In the *full-display* paradigm, all of the letters appeared in each trial. Continuing the example above, in this case the letters L, G, P, and B would be presented within Frames 1–4, respectively.

The results of both paradigms showed that RTs depended on two factors. First, the RT increased with the number of updated items. For example, condition URRR was faster than UUUR. On the basis of the idea that WM consists of associations between items and their serial positions (e.g., Lewandowsky & Farrell, 2008), we suggested that updating is carried out serially, with a cost for establishing each association between a new (updated) item and its position (see also Kessler & Meiran, 2008, for a similar assumption). The second factor was the exact order of updated and repeated items. For example, condition URUR was slower than RRUU, although having the same set-size and the same number of updated items.

We developed a "scanning and gate-switching" (SGS) model in order to explain the complex pattern of RTs over the entire set of U–R patterns. According to the model, participants process each updating trial by scanning the stimuli from left to right, one item at a time, consistent with their habitual reading direction of reading the letters, while scanning their representation of the current list in WM in parallel. When encountering a stimulus that does not require updating (i.e., an asterisk or a letter identical to the letter held in the same frame in WM), they maintain the memory item in that frame. When encountering a stimulus that requires updating (i.e., a new letter not matching the one held in the same frame in WM), they update that item in memory. Updating involves removing the association of the old item to its context (viz. the specific frame it had appeared in) and establishing an association between the new stimulus and that context. The time for this updating process is reflected in the item-specific updating cost (captured by the *item–position* parameter of the model). Because scanning is serial, RTs are predicted to increase linearly with the number of item–position associations that need to be updated. In addition, when shifting the focus of attention from one item to the next while scanning the frames, participants switch between updated and repeated items. For example, in condition RRUU, Position 3 ("U") involves switching between repetition and updating, whereas Positions 2 and 4 involve repeating the same condition that was applied to the previous letter (i.e., "R" after "R" or "U" after "U," assuming scanning from left to right). We found that switching from updated to repeated items, or vice versa, was associated with a large RT cost, relative to repeating the same condition: The RT for a trial increased with the number of update switches to be carried out during scanning the list from left to right.

We interpreted the cost of switching between U and R positions as the time taken to open or close the gate to WM. In other words, this is the time required to switch between two states of the WM system—maintenance, as is required by repetition steps, and updating. Because the gate is closed by default, WM is in a maintenance state at the beginning of each trial. Accordingly, for example, Condition UURR involves two switches: from maintenance to updating in Position 1, and from updating to maintenance in Position 3. Condition URUR involves four switches: from maintenance to updating in Positions 1 and 3, and from updating to maintenance in Positions 2 and 4.

Kessler and Oberauer's (2014) SGS model provided the best fit to the data of four experiments, as well as to a reanalysis of Kessler and Meiran's (2008) Experiment 3 results, among several theoretical models that were tested. However, a good fit is a necessary but not a sufficient condition for confirming a theory (Roberts & Pashler, 2000). Whereas no other known theory of WM updating can explain the dependence of RTs on the specific order of updated and repeated items, it is still possible—although unlikely—that other processes than SGS might govern WM updating, and the fit of the model's predictions to the pattern of RTs across conditions could be merely accidental. To address this concern, more

direct evidence for forward scanning during WM updating is needed. Obtaining such evidence was the goal of the present study.

Participants performed the partial-display paradigm with a set size of four items. Table 1 presents the conditions and parameter values. Critically, the stimuli were either English or Hebrew letters, presented in different blocks. We hypothesized that participants would scan the stimulus sets according to the typical reading direction in the relevant language. Accordingly, English letters would be scanned from left to right, and Hebrew from right to left. If so, then the number of gate switches to be carried out for a given U–R pattern should depend on the language. For example, condition UURR would require two gate switches in English (in Positions 1 and 3), but only one in Hebrew (switching from maintenance to updating at the third item from the right). If, however, our model fit the data for some reason other than forward scanning, the pattern of RTs over the different U–R patterns should not differ between the languages.

Method

Participants

Twenty undergraduate students (11 females, nine males; age: $M = 25.00$ years, $SD = 1.92$) from Ben-Gurion University of the Negev participated in the experiment in return for monetary compensation. The participants reported having no learning disabilities or neuropsychological dysfunction.

Stimuli

The stimuli were 18 uppercase English consonants and 18 Hebrew consonants. The letters were presented in white against a black background, within four frames that were positioned horizontally at the center of the screen. Letters were not allowed to repeat in different positions within the same set.

Table 1 Parameter values by conditions

Condition	Item–Position	Gate Switch (English)	Gate Switch (Hebrew)
RUUU	3	1	2
RRUU	2	1	2
RRRU	1	1	2
UUUR	3	2	1
UURR	2	2	1
URRR	1	2	1
RURU	2	3	4
URUR	2	4	3
UUUU	4	1	1
RRRR	0	0	0

The gate switch predictor values are presented for English and Hebrew separately. The number of new item–position associations is not affected by the scanning direction.

Procedure

The experiment consisted of two blocks, one for each language. The order of blocks was counterbalanced across participants. Each block included 60 runs of trials, preceded by five practice runs. Each run started with the presentation of four frames with letters inside (see Fig. 1). The participants were instructed to memorize the letters and to press the spacebar to continue. Then, a sequence of trials began. The number of trials was varied randomly between one and five, so that the end of the run was unexpected. Each trial involved updating all, some, or none of the items in WM. The condition of each trial was selected randomly from the following conditions: RRRR (full repetition), UUUU (full update), RUUU, RRUU, RRRU, UUUR, UURR, URRR, RURU, and URUR. In each trial, only the updated letters were presented on the screen. An asterisk was presented in nonupdated frames, to denote that the letters in those frames did not change. The participants were instructed to remember the last letter that appeared in each frame. Accordingly, they had to update their WM whenever new letters appeared. They were then required to press the spacebar after updating the items to continue to the

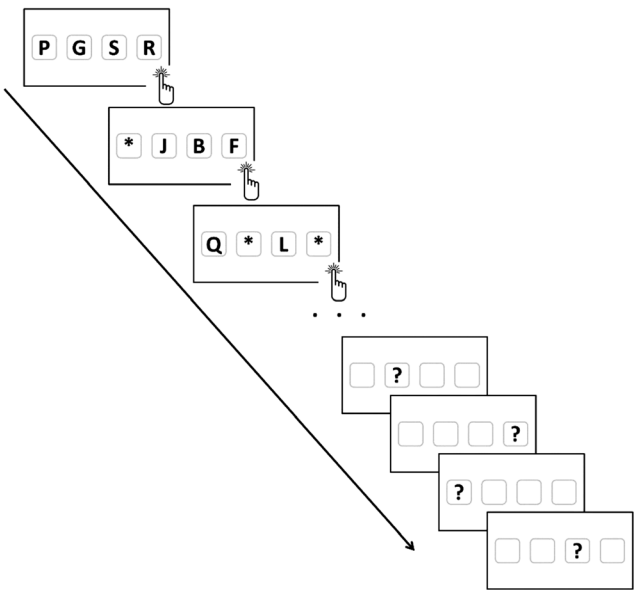

Fig. 1 Schematic depiction of the paradigm. Each run of trials begins with the presentation of four letters inside frames. The participants are required to memorize the letters and to press a key to continue. Then a sequence of one to five trials begins, each involving updating of none, some, or all of the letters (see the conditions in Table 1). An asterisk inside a frame indicates that the letter that corresponds to that frame does not update. A keypress is required to move to the next trial. This response serves to measure the RT for the updating trials. At the end of the run, the participants are asked to recall the last letter that had appeared in each frame; the frames are cued for recall in a random order

next trial. RTs for these keypress responses were measured. An intertrial interval of 1,000 ms followed the response, during which empty frames were presented. After one to five updating trials, the participants were required to recall the last letter in each frame. To probe the recall, a question mark appeared in each frame, one frame at a time and in a random order, and the participants had to type in the last letter that had appeared in that frame.

Results

Recall accuracies—namely the proportion of runs in which the participants recalled all of the letters correctly—were 83 % in the English block and 85 % in the Hebrew block [$t(38) = 0.60$, $p = .55$]. Only the trials from runs in which recall was correct were included in the RT analysis. The RT trimming was done in two steps. First, RTs longer than 15 s were removed. Second, RTs that deviated by more than three standard deviations from the mean for each participant in each condition were considered outliers and removed. In total, 1.1 % of the trials were removed as outliers. Figure 2 presents the mean RTs by condition in each language.

We started by performing an analysis of variance (ANOVA) on the mean RTs, with language and condition as within-subjects variables. We predicted that language and condition would be additive when the conditions were coded according to the specific reading direction of the language used (e.g., $UURR_{English} = RRUU_{Hebrew}$). An interaction was predicted when the reading direction was not taken into account (e.g., $UURR_{English} = UURR_{Hebrew}$). The results confirmed our hypothesis (see Fig. 2). The main effect of language was significant, $F(1, 19) = 6.34$, $p = .02$, $\eta_p^2 = .25$, as well as the main effect of condition, $F(9, 171) = 63.93$, $p < .001$, $\eta_p^2 = .77$, when recoded according to the reading direction, and $F(9, 171) = 61.36$, $p < .001$, $\eta_p^2 = .76$, when not recoded. As predicted, the two-way interaction was significant only when the conditions were not recoded, $F(9, 171) = 14.70$, $p < .001$, $\eta_p^2 = .44$, but not when the reading direction was taken into account, $F(9, 171) = 0.98$, $p = .46$, $\eta_p^2 = .05$. This finding confirms our hypothesis that the scanning direction plays an important role in WM updating.

We next moved to implementing the SGS model. We implemented the model as a multilevel linear regression model (Pinheiro & Bates, 2000) for RTs as the dependent variable, as was done by Kessler and Oberauer (2014). In this analysis, the number of gate switches and the number of item–position associations served as within-subjects predictors for RT in the various conditions. For example, Condition $UURR_{English}$ involves updating two item–position associations and two gate switches. The model assumes that RT is a linear function of the number of item–position associations and of gate switches, each with a separate regression weight that served to estimates the duration of each process. For instance, a gate-switching slope of 200 implies that the average gate switch cost was 200 ms.

The advantages of multilevel linear regression is that it allows the regression coefficients to vary between subjects, hence taking into account between-subjects variability in the durations of these processes, and enables a direct comparison between nonnested models, due to its reliance on maximum likelihood estimators. To this end, the model included subjects as a random effect on the intercept and on all slopes (i.e., on the effects of all predictors). Model comparison and selection were based on the Bayesian information criterion (*BIC*) statistic, which penalizes for free parameters and thereby rewards parsimony. In addition, Bayes factors (*BF*) were calculated on the basis of the BIC differences between models, using the formula $BF = \exp(\Delta BIC/2)$ (Wagenmakers, 2007). The BF value is interpreted as the factor by which the evidence favors the selected model as compared to the nonselected one. The models were implemented using the lme function of the nlme package in R (Pinheiro, Bates, DebRoy, Sarkar, & the R Development Core Team, 2012). A diagonal random-effect variance–covariance matrix was used, assuming that the random effects were independent.

Analogous to the ANOVA, we fitted the model to the data in two ways. First, we modeled the data in which the conditions were recoded according to the reading condition. On the basis of the ANOVA results, no interactions between the parameter values and language were expected. We started by fitting the full model, including item–position, gate switch, language, and all of the interactions among these predictors (see Table 2 for the parameter values and model fit statistics for all models). As predicted, none of the interactions with language was significant. We continued by removing nonsignificant predictors until we converged on the best-fitting model (for which further omitting any of the predictors led to a decrease in fit). This model included all of the main effects, plus the interaction between gate switch and item–position. Second, we fit the model to the data, this time without recoding the conditions according to the reading direction. The best-fitting model only included the interaction between gate switch and language, with a larger gate-switching cost in English than in Hebrew. Note that such a model is implausible theoretically, even before testing for model fit, because there is no reason to assume that the duration of gate switching would be language-specific or stimulus-specific. When comparing the fits of the best-fitting models, compelling evidence was observed favoring the model fit to the reading-direction-recoded data, $BF = 1.14 \times 10^{35}$. In other words, the SGS model fits the data significantly better when taking into account the reading direction (see Fig. 2 for the model predictions). As can be seen in Fig. 2b, the best-fitting model of the nonrecoded data fails to fit the Hebrew conditions.

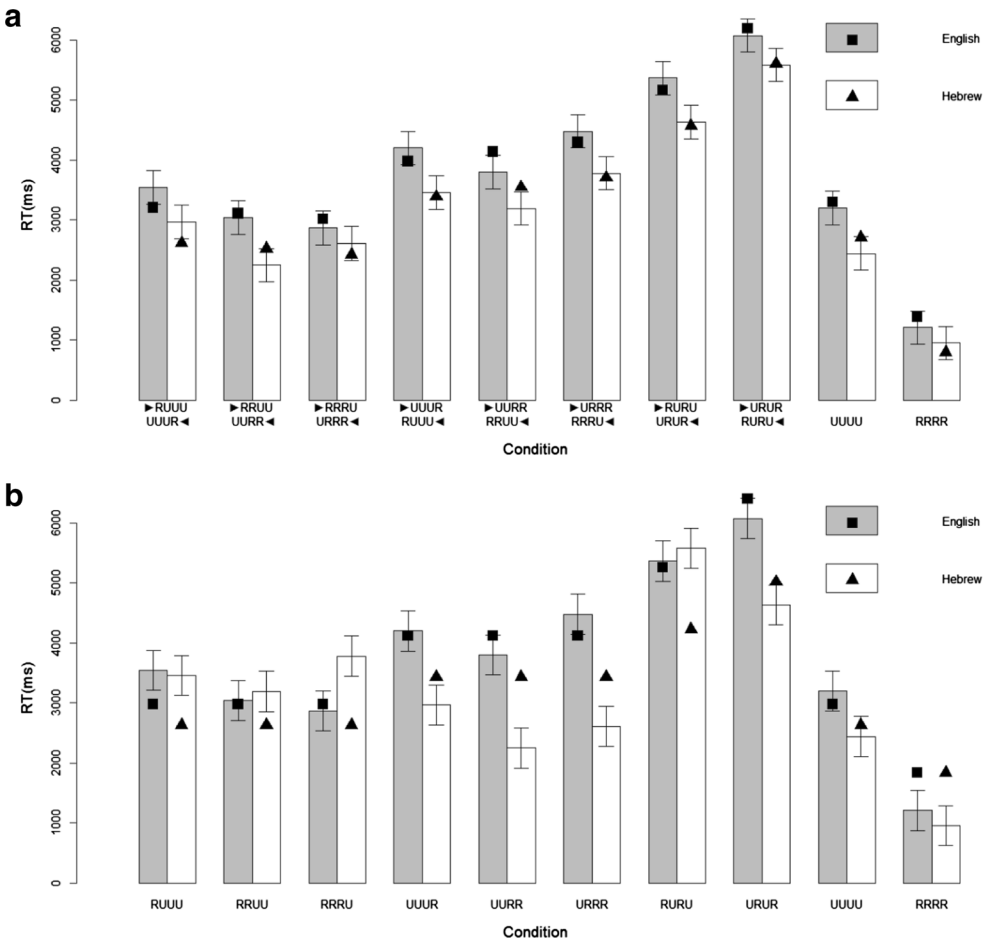

Fig. 2 Bars represent the mean RT (and 95 % confidence intervals) for each condition in each language. The square and triangle symbols represent the model predictions for English and Hebrew, respectively. (**a**) Observed data and model predictions when the conditions are recoded according to the reading (scanning) direction (e.g., $URUR_{English} = RURU_{Hebrew}$). Additive effects of language and condition are observed. The model predictions fit the data in both languages. (**b**) Observed data and model predictions when the reading direction is not taken into account (e.g., $URUR_{English} = URUR_{Hebrew}$). An interaction is observed between language and condition, reflecting the fact that the patterns of RTs across the conditions differ in the two languages. The model fails to fit this interaction, as can be observed in the lack of fit to the Hebrew conditions

Discussion

Our modeling results replicate and extend our previous finding that the SGS model provides an adequate fit to WM-updating latencies. The multilevel modeling framework enables testing the contribution of each predictor to explaining the data. Accordingly, even before turning to analyzing the effects of language and scanning direction, the fact that the best-fitting model included gate switch and item–position as predictors provides strong evidence for their roles in understanding WM-updating latencies, in accordance with the SGS model. More generally, the SGS model was selected by Kessler and Oberauer (2014) from a large class of theoretically inspired models that differed in their assumptions regarding the architecture of WM. These included item–position associations (as reflected by the item–position parameter in SGS), interitem associations, chaining, gate switching, and all possible combinations of these models. The results ruled out a role for interitem associations (including chaining), and therefore here we did not test models that included that predictor.

Our findings further corroborate the scanning assumption of the SGS model. Although the participants were not instructed regarding the required reading direction in each language, the results demonstrate equivalent effects of condition on updating RTs if, and only if, the conditions are coded according to the language-specific reading direction. The effects of language and condition were additive when the latter was coded according to the scanning direction. This is also reflected in the modeling results, in which the parameter values did not vary across the languages. When the direction of scanning was assumed to be the same for both languages—namely left to right—an interaction between gate switch and language arose. Such an interaction could not be explained by

Table 2 Model fit statistics and predictor estimates

	Predictor	Estimate (ms)	Std. Error	*df*	*t*	*p*
Full model (conditions *recoded* according to reading direction)	(Intercept)	2,077	181	373	11.50	<.001
	ItPos	434	89	373	4.90	<.001
BIC = 6,680.66	GateSw	1,035	125	373	8.26	<.001
AIC = 6,612.80	Lang	–583	156	373	–3.73	<.001
Log likelihood = –3,289.40	ItPos × GateSw	–289	83	373	–3.49	<.001
	ItPos × Lang	–169	116	373	–1.45	.15
	GateSw × Lang	–13	129	373	–0.10	.92
	ItPos × GateSw × Lang	69	108	373	0.63	.53
Best-fitting model (conditions *recoded* according to reading direction)	(Intercept)	2,081	233	376	8.93	<.001
	ItPos	350	70	376	5.00	<.001
BIC = 6,670.70	GateSw	1,028	92	376	11.21	<.001
AIC = 6,626.78	Lang	–592	220	376	–2.69	.008
Log likelihood = –3,302.40	ItPos × GateSw	–255	63	376	–4.04	<.001
$BF_{compared\ to\ Full} = 145.47$						
Full model (conditions *not recoded* according to reading direction)	(Intercept)	2,077	213	373	9.74	<.001
	ItPos	434	115	373	3.78	<.001
BIC = 6,880.09	GateSw	1,035	125	373	8.25	<.001
AIC = 6,812.23	Lang	–35	208	373	–0.17	.87
Log likelihood = –3,389.12	ItPos × GateSw	–289	107	373	–2.71	.007
	ItPos × Lang	–70	159	373	–0.44	.66
	GateSw × Lang	–329	145	373	–2.27	.02
	ItPos × GateSw × Lang	14	149	373	0.09	.93
Best-fitting model (conditions *not recoded* according to reading direction)	(intercept)	1,836	188	378	9.79	<.001
	GateSw (English)	1,142	120	378	9.49	<.001
BIC = 6,832.15	GateSw (Hebrew)	796	84	378	9.49	<.001
AIC = 6,804.21						
Log likelihood = –3,395.11						
$BF_{compared\ to\ Full} = 2.57 \times 10^{10}$						

ItPos = item–position; GateSw = gate switch, Lang = language. Language is coded as the difference between Hebrew and English (i.e., negative values indicate faster RTs for Hebrew). The ItPos predictor was centered at zero in the analysis to ensure that its main effect was not generated through its interaction terms. This was done by subtracting the mean predictor value from the predictor value in each condition.

any model of WM updating, because the cost of gating should not depend on the specific stimuli. Furthermore, even when equipped with different parameter estimates for the two languages, this model fell short of fitting the data.

The best-fitting model explained RTs using four predictors. The item–position and gate switch predictors had been included in the best-fitting models in our previous work. The contributions of these parameters were additive to that of language, asserting that the processes that they reflect are independent of the specific stimuli. The negative interaction coefficient between item–position and gate switch, although unpredicted, indicates that the additional cost of updating an item–position association is not independent of gate switches. We suggest that some aspects of the item–position association updating take place in parallel to gate switching, so that the cost of updating the first item in a row is smaller than the cost of updating the next ones. Specifically, recent findings have demonstrated the role of removal in WM updating. According to this idea, old information must be removed from WM prior to its substitution with new input. Ecker and colleagues (Ecker, Lewandowsky, & Oberauer, 2014; Ecker, Oberauer, & Lewandowsky, 2014) have shown that removal can be done in advance, even before the new information is presented. Although our model does not estimate the duration of removal independently, this duration is part of the item–position parameter. Unlike the creation of new item–position associations, which is only possible after opening the gate, removal could be done while the gate was being opened, leading to a smaller item–position cost for updating in combination with gate switching than for updating without gate switching.

Our results provide direct evidence of two specific assumptions of the SGS model. The first, as we explained in the introduction, is that the stimulus set is scanned in each trial from start to end. As we showed here, the direction of scanning depends on the specific stimuli that are presented. The second assumption is that each trial starts with WM in maintenance mode. Two alternative assumptions could have been conceived. One possibility is that each trial starts in updating mode. However, this is incompatible with the theoretical gating model upon which the SGS is based, which assumes that

maintenance (rather than updating) is the default mode of WM operation (Frank et al., 2001). Furthermore, it is not supported by the data. For example, such an assumption would lead to the wrong prediction that the RTs for Condition URUR (three gate switches when starting in U mode) should be shorter than those for RURU (four switches when starting in U mode). The second alternative assumption is that the state in which each trial starts depends on the operation required for the first stimulus. This alternative is ruled out by the present results, because it cannot explain the differences between the two languages: The number of gate switches would be independent of reading direction. Hence, our results support the assumption of starting each trial in a default maintenance mode.

To conclude, whereas a forward-scan mechanism had been suggested to take place in various aspects of WM processing (starting from Sternberg, 1966), the SGS model is the first to incorporate this process in order to understand updating latencies. Our results demonstrate that forward scanning is involved in updating verbal lists in WM. Understanding WM updating is highly important due to its role in high-level cognitive functioning, as well as in individual differences in executive functioning and intelligence. Establishing the central role of forward scanning and gate switching in this ability is a necessary step toward detailed process models of other WM-updating tasks.

Author Note This research received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under Grant Agreement № PCIG09-GA-2011-293832, awarded to the first author, and from the Swiss National Science Foundation, Grant №s. 100014_126766 and 100014_143333, to the second author.

References

Braver, T. S., & Cohen, J. D. (2000). On the control of control: The role of dopamine in regulating prefrontal function and working memory. In S. Monsell & J. Driver (Eds.), *Control of cognitive processes: Attention and performance XVIII* (pp. 713–737). Cambridge, MA: MIT Press.

Ecker, U. K. H., Lewandowsky, S., & Oberauer, K. (2014). Removal of information from working memory: A specific updating process. *Journal of Memory and Language, 74,* 77–90. doi:10.1016/j.jml.2013.09.003

Ecker, U. K. H., Oberauer, K., & Lewandowsky, S. (2014). Working memory updating involves item-specific removal. *Journal of Memory and Language, 74,* 1–15. doi:10.1016/j.jml.2014.03.006

Frank, M. J., Loughry, B., & O'Reilly, R. C. (2001). Interactions between frontal cortex and basal ganglia in working memory: A computational model. *Cognitive, Affective, & Behavioral Neuroscience, 1,* 137–160. doi:10.3758/CABN.1.2.137

Kessler, Y., & Meiran, N. (2008). Two dissociable updating processes in working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34,* 1339–1348. doi:10.1037/a0013078

Kessler, Y., & Oberauer, K. (2014). Working memory updating latency reflects the cost of switching between maintenance and updating modes of operation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 738–754. doi:10.1037/a0035545

Lewandowsky, S., & Farrell, S. (2008). Short-term memory: New data and a model. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 49, pp. 1–48). San Diego, CA: Academic Press.

O'Reilly, R. C., & Frank, M. J. (2006). Making working memory work: A computational model of learning in the frontal cortex and basal ganglia. *Neural Computation, 18,* 283–328. doi:10.1162/089976606775093909

Pinheiro, J. C., & Bates, D. M. (2000). *Mixed-effects models in S and S-Plus*. New York, NY: Springer.

Pinheiro, J., Bates, D. M., DebRoy, S., Sarkar, D., & the R Development Core Team. (2012). *nlme: Linear and nonlinear mixed effects models (R package version 3.1) [Computer software]*. Vienna, Austria: R Foundation for Statistical Computing.

Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review, 107,* 358–367. doi:10.1037/0033-295X.107.2.358

Sternberg, S. (1966). High-speed scanning in human memory. *Science, 153,* 652–654. doi:10.1126/science.153.3736.652

Wagenmakers, E.-J. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review, 14,* 779–804. doi:10.3758/BF03194105