Published in final edited form as:

J Cogn Psychol (Hove). 2011 ; 23(2): 204–219. doi:10.1080/20445911.2011.481621.

Three layers of working memory: Focus-switch costs and retrieval dynamics as revealed by the *N*-count task

Chandramallika Basak¹ and Paul Verhaeghen²

¹Beckman Institute, Urbana/Champaign, IL, USA

²Georgia Institute of Technology, Atlanta, GA, USA

Abstract

Two experiments explored the process of switching items in and out of the focus of attention using a new paradigm, the *N* count task (adapted from Garavan, 1998; *N* varied from 1 to 4). This task yielded a focus size of one, indicated by a substantial focus switch cost for 2 count. Additionally, the focus switch costs in response time increased with working memory load, indicating an effortful search process occurring at a speed of about 240 ms/item. Maintaining and switching to and from a passive load did not increase the focus switch costs or decrease memory accuracy, indicating that there is no crosstalk between passive and active items. The results support a concentric theory of working memory: a small focus at its core, a surrounding area of (at least) three readily available items referred to as the outer store, and a still wider region of passive storage, possibly more long term memory than working memory.

Keywords

Focus switching; Ageing; Working memory; Focus of attention; Retrieval

Not all items stored in working memory are equally accessible for processing. The distinction can be made intuitively—for instance, while having a conversation, one might suddenly think of a point to make and, rather than interrupt the flow of the conversation, one might store that point away in what is colloquially described as “the back of the mind”, for retrieval during the first pause in the conversation. Likewise, in traditional working memory tasks such as operation span, where to-be-remembered items are interleaved with bouts of processing, the items-to-be-remembered likely leave the focus of one’s attention while processing is ongoing; at the end of the trial, they need to be retrieved from a buffer outside immediate awareness, as testified by high correlations between working memory capacity measures in such tasks and tests of “secondary” memory (Unsworth & Engle, 2007).

More formally, working memory researchers such as Cowan (2001) make a distinction between a zone of immediate access, labelled the focus of attention, from a larger, activated portion of long-term memory in which items are stored in a readily available but not

immediately accessible state (we will label this zone immediately outside the focus of attention, i.e., the second tier, the “outer store”, after Verhaeghen, Cerella, & Basak, 2004).¹ This two-tier system is nested: If the number of items stored in working memory exceeds the capacity limit of the focus, the supernumerary items will be accommodated into the outer store. The size of the focus of attention is subject to debate—typically, the focus can hold three to four items (e.g., Cowan, 2001), although under some circumstances (i.e., items presented serially to unpractised participants; e.g., Garavan, 1998; McElree, 2001; Oberauer, 2002; Verhaeghen et al., 2004) it seems to accommodate only a single item. Regardless of the buffer sizes, one logical consequence of a two-store structure is the existence of a focus-switching operation (McElree, 2001; as far as we can determine, the term “focus switch” originated with Voigt & Hagendorf, 2002). When the number of items to be retained in working memory is smaller than or equal to the capacity of the focus of attention, they will be contained within the focus; there they are immediately retrievable, and access times will be fast. When the number of items to be retained exceeds the capacity of the focus, the excess items will be stored outside the focus of attention. In that case, accessing items for processing will necessitate a retrieval operation; this will slow down response time.

Evidence of such slow-downs has been observed in at least three paradigms. A first piece of evidence comes from a modified *N*-back task as adapted by McElree (2001). Participants are presented with a running series of digits on a computer screen, shown one at a time. The task is to judge whether the current digit is identical to the digit presented *N* items earlier, or not. Using a speed-accuracy tradeoff procedure, McElree found that access times were distinctly faster for *N* = 1 than for either *N* = 2 or *N* = 3. This result has been replicated in a simpler latency version of the *N*-back (Vaughan, Basak, Hartman, & Verhaeghen, 2008; Verhaeghen & Basak, 2005; Verhaeghen et al., 2004; Zhang & Verhaeghen, 2009). Second, Oberauer (2002, 2006) utilised a working memory updating task in which a set of single-digit numbers stored in working memory had to be updated one by one using simple arithmetic operations. He found that updating occurred faster when an item had just been updated than when other items had intervened. This again suggests that working memory has privileged access only to the single item last updated. A third piece of evidence—and this is the paradigm we are using in the present set of two experiments—comes from the running count task devised by Garavan (1998; see also Gehring, Bryck, Jonides, Albin, & Badre, 2003; Li et al., 2006; and Voigt & Hagendorf, 2002). In this task, participants keep a separate running count of triangles and squares that appear in a random sequence on a computer screen. Presentation is self-paced, allowing for the recording of response times. Response times are about 500 ms slower when the stimulus shape is changed from the previous trial than when both successive stimuli have the same shape. The results suggest that participants keep separate mental counters for each stimulus shape, and that switching between counters in working memory is a time-consuming process. This, in turn, suggests that only one of the counters is immediately accessible for updating and hence resides in the focus of attention, while the other counter is held in temporary storage outside the focus of attention. Additionally, one recent fMRI study using an item-recognition paradigm (Nee & Jonides, 2008) shows that

¹This outer store would correspond with the region of direct access in Oberauer’s model, when the focus of attention holds only one item. We are referring to the region just outside the focus of attention, which is involved in the focus switching phenomenon.

retrieval of items stored in the focus of attention is subserved by increased activation of the inferior temporal cortex, whereas retrieval of items from the outer store is associated with increased activation of the left mid-ventrolateral prefrontal cortex (VLPFC) and the medial temporal lobe (MTL), suggesting that the two types of retrieval are indeed distinguishable.

Oberauer (2002, see also 2001) has proposed that the accessibility hierarchy of working memory has a third level, namely that of items that are not needed for ongoing processing but might need to be recalled at some later time. Entrance into the latter tier is not driven by overflow but by functionality: Items not needed in the foreseeable future will be stored in this third tier, regardless of whether the second and third tier are stored to capacity or not. Precisely for this reason, Oberauer posits a strict boundary between the third tier and the two other subsystems, characterised by a lack of crosstalk (i.e., interference) between the third tier and the two other systems. This, then, is the proposed defining characteristic of items stored inside the third tier: Although they are retrievable with high accuracy after the task proper has been completed, they will not interfere with ongoing processing. Oberauer's evidence for his claim is that speed of updating of stored items is solely a function of the number of active items (i.e., items that subjects know need to be updated during the course of a trial), not of the number of passive items (i.e., items that are only probed after the trial is completed). Similar results were obtained much earlier by Sternberg (1969) in the context of memory scanning—scanning is much faster when memory is loaded with a passive set (items that are not to be remembered) than with an active set (items that are to be remembered). We note here that Oberauer derives no structural claims from his data; neither does Sternberg. Although the model can graphically be represented as concentric regions, or tiers—a focus surrounded by a set of active items surrounded by a set of passive items—Oberauer explicitly acknowledges that these do not represent structurally (let alone anatomically) distinct subsystems. Rather, they represent functional states in working memory—items selected by attention versus items held outside attention in either a ready-to-be-selected (“active”) state or in a background (“passive”) state.

The present set of two experiments was designed to replicate and perhaps augment Oberauer's findings using a different paradigm—a variant of Garavan's (1998) running count task described earlier. Garavan used two counts; we use multiple, up to four. The advantage of this task over our customary *N*-back task is that it allows us to probe items in a random order rather than the left-to-right forward order. In our mind, Oberauer's study left two questions that deserved closer scrutiny with a different paradigm; both are concerned with how the outer store is accessed. The first question, addressed in Experiment 1, deals with access of active items in the outer store; the second, addressed in Experiment 2, with the lack of crosstalk between active and passive items.

First: How are active items accessed? We have previously argued that the active part of the outer store is content-addressable (Verhaeghen & Basak, 2005; Verhaeghen et al., 2004), a feature similar to that of long-term memory (LTM). Our evidence came from our *N*-back task, where we found no increase in response times (RT) for values of *N* between 2 and 5. This is further supported by results from Nee and Jonides (2008), where retrieval of information from the outer store is found to be associated with increases in MTL activation. Because patients with damage to the MTL show deficits in LTM retrieval, not short-term

memory (STM) retrieval, the results suggest that neural mechanisms responsible for retrieval of information from the outer store are similar to those used in LTM. Oberauer's (2002, 2003) updating paradigm, however, casts doubt on this assertion: This paradigm yields substantial RT by set-size slopes, about 150 ms/item for switch items. This suggests the presence of a relatively inefficient search process (compared to search slopes in short-term memory scanning of about 40 ms/item for Sternberg, 1966; and 33 ms/item for search within the expanded focus of attention after extensive practice; Verhaeghen et al., 2004; we note here that we remain agnostic as to what the actual search mechanism entails: serial or parallel, exhaustive or self-terminating). What might be the issue here is predictability: Predictable forward probing appears to yield a zero (or very fast) slope; random probing a nonzero (slow) slope. Another possibility, however, is that there is a tradeoff between search slope and task difficulty. If both the task at hand and the storage/maintenance processes of working memory compete for the same resources, it is possible that more complex tasks drive a normally content-addressable system into a less efficient mode, namely that of controlled search. Oberauer's arithmetic updating is relatively complex compared to the simple identity judgements required by the *N*-back task and to the update-by-adding-1 requirement in the running count task. We do know that stimulus complexity influences the focus-switch cost: Zhang and Verhaeghen (2009) found a clear RT by set-size slope in an identity-judgement task using Chinese characters, perhaps due to increased resource demands for storing such complex stimuli. Likewise, Voigt and Hagendorf (2002) showed that when the starting count in a running count task differs from zero by a large amount, the focus-switch costs increased. If task difficulty influences switch costs, perhaps it influences the search process as well. In the present experiments, we used a running-count paradigm, where the operation is to add 1 to a count; it is hard to imagine a simpler operation. This then allows us to test the tradeoff hypothesis against the sequential/predictability hypothesis.

The second outstanding question, tackled in Experiment 2, concerns the lack of crosstalk between passive items and active items. Compared to Oberauer's (2002) task, where passive items are never accessed, our *N*-count paradigm allows us to increase competition between active and passive items by requiring the subject to store both active (i.e., to-be-updated) and passive (i.e., not-to-be-updated) counts, each with a specific random start-up value, and then presenting the shapes associated with both active and passive counts over the course of a single trial. If the shape is associated with an active count, the count is incremented by one; if it is associated with a passive count, it is to be ignored and the initial count is left untouched. This makes for a very stringent test of the no-crosstalk hypothesis. First, in the *N*-count task, most research participants report silent rehearsal of all the counts, thereby accessing all items frequently and presumably consistently (Basak, 2006). This repeated access should maximise the possibility for crosstalk between subsystems. Second, and most interestingly, cueing of passive counts allows us to directly investigate whether cueing an intervening passive item influences either the switch or the nonswitch RTs for two active items—if it does not, this would be strong evidence for the existence of a firewall between the set of active and the set of passive items.

We note here that these findings may have consequences for research designs (notably in social psychology or instructional psychology) that purport to manipulate processing load by

providing participants with some information (usually a set of digits) to be retained while a concurrent task is being performed. In the current classification, this manipulation amounts to imposing a passive load. If this load is indeed truly passive, it will not interfere with concurrent processing in the focus of attention, and therefore miss its aim.

EXPERIMENT 1

As outlined in the introduction, one outstanding question in the literature concerns the retrieval dynamics of the outer store: Is the focus-switch cost in the N -count task stable over set-size, as found for our original N -back task, or does it covary with set-size, as found in Oberauer's arithmetic updating task? If the former is the case, a tradeoff between storage and processing demands is a likely explanation—more complex processing tasks necessitate controlled search, which will lead to a slope. If the latter is the case, the predictable/unpredictable nature of probing is the likely culprit.

Method

Participants—The sample consisted of 24 students (mean age = 18.75 years, $SD = 0.99$, ranging from 18 to 21 years; mean years of education = 12.79, $SD = 0.88$; 20 females and four males) from Syracuse University, who received course credits in return of their effort.

Procedure—We used the N -count task as described previously, with four levels of set-size ($N = 1$ to 4). For $N = 1$, all sequences are by definition nonswitch; for $N > 1$, half of the sequences were nonswitch trials, and half were switch trials. We used starting counts for each of the shapes; these were randomly assigned for each participant; they ranged from 1 to 4. Shapes were displayed along an imaginary horizontal line at the centre of the screen, in N imaginary columns, one for each shape. To keep the length of the longest possible saccades equal across set-size, the horizontal separation between two the two extreme column positions was kept constant at 12 cm for any sequence where $N > 1$. Distance from the monitor was not fixed; participants were encouraged to choose a distance that felt comfortable to them.

The four shapes used were a red square, a green triangle, a blue circle, and an orange star. For $N = 1$, the choice of shape was randomised for each sequence; for $N = 2$, we used red squares and blue circles, left to right; for $N = 3$, we used red squares, blue circles, and green triangles, left to right; and for $N = 4$, we used red squares, green triangles, blue circles, and orange stars, left to right. Consistent colours and column positions were assigned to each shape to facilitate perceptual processing of each new stimulus. The size of the shapes was approximately $2 \text{ cm} \times 2 \text{ cm}$.

The sequence of events within a sequence was as follows (see Figure 1 for an illustration for $N = 2$). First, a fixation cross was shown at the centre of the screen for 1000 ms. Immediately after the fixation cross disappeared, N starting counts were presented, left to right. This was done by showing each of the shapes in sequence, with the starting count presented right above it, projected in Courier New font, bold style and 1.5 cm tall. Participants paced presentation by pressing the spacebar in between each starting count. Once the N starting counts were presented, the sequence proper started, and stimuli were presented one by one

in their appropriate columns. The participant advanced to the next stimulus by pressing the spacebar; this is our measure of RT. The task was to add one to the current associated count for each of the shapes shown. The shapes were jittered by one pixel either left, right, up or down from the stimulus previously presented in that location; this was done to make the appearance of a new stimulus notable in nonswitch trials. Each sequence consisted of 12-16 to-be-responded-to items. The order of presentation of shapes in each sequence for each N was randomised beforehand, and then kept constant across participants. At the end of the sequence, the participant was prompted to type in the final count of the N shapes, left to right. In nonswitch trials, the shape could be repeated once, twice or thrice, each with equal probability.

A total of 22 sequences (mean length of a sequence was 14) for each N was presented in the following pattern: first a block of 11 sequences for $N=1$, then a block of 11 sequences for $N=2$, then a block of 11 sequences for $N=3$, then a block of 11 sequences for $N=4$, followed by a 15-minute break where participant completed a few paper-and-pencil tests (demographic information, Symbol Digit Substitution Test, Shipley's Vocabulary test, a brief vision test), followed by a block of 11 sequences each for $N=4$, $N=3$, $N=2$, and $N=1$. The first sequence for each block was considered practice and discarded from the analyses. Response times of the first item in each sequence, that is, the first item following the items presented by the experimenter, were also discarded from the analyses, because RT for these items might be contaminated with task switching, that is, switching from encoding to retrieval and updating. At the end of the computerised testing, the participants were asked to fill out a brief strategy questionnaire, which probed for vocal repetition and rehearsal in each of the sequences. The experiment was conducted in a single session, lasting between 90 and 120 minutes. Participants were encouraged to take breaks when needed between the sequences.

Results

Response times were collected for each individual item. Accuracy was recorded for the final counts at the end of the sequence. A sequence was considered correct when all the counts were reported correctly. Response time analyses were conducted on correct sequences only. For each condition, within each individual, the median response time of each sequence was used. These median response times were averaged across the sequences for each N for each individual. Alpha level for all statistical testing was set at .05. Throughout the paper, p -values were Green-house-Geisser corrected for sphericity.

Response times for nonswitch and switch items as a function of N —RT results are presented in Figure 2A. In a first analysis, we conducted a one-way ANOVA on nonswitch trials only. The main effect of set-size was significant, $F(3, 69) = 164.16$, $MSE = 55,507.54$, $p < .001$; RT increased as set-size increased. This increase had a significant quadratic component, $F(1, 23) = 4.87$, $MSE = 34,916.97$, $p = .038$, as well as a significant linear component, $F(1, 23) = 238.13$, $MSE = 112,272.17$, $p < .001$. A significant increase in RT with set-size was also noted when the analysis was restricted to set-sizes 2 to 4, $F(2, 46) = 75.26$, $MSE = 52,246.72$, $p < .001$. Both the linear, $F(1, 23) = 102.10$, $MSE = 75,202.63$, $p < .001$.

001, and the quadratic component, $F(1, 23) = 6.35$, $MSE = 29,290.80$, $p = .019$, were significant, indicating upwards curvature.

The next set of analyses examined whether focus switching interacted with set-size. The RTs were submitted to a 3 (working memory set-size: 2, 3, and 4) $\times$ 2 (switch vs. nonswitch) ANOVA. The main effects of switching and set-size, as well as their interaction, were significant, $F(1, 23) = 123.64$, $MSE = 206,304.49$, $p < .001$; $F(2, 46) = 110.73$, $MSE = 105,414.84$, $p < .001$, and $F(2, 46) = 17.88$, $MSE = 21,612.37$, $p < .001$, respectively. Switch trials were slower than nonswitch trials, and larger set-sizes took longer to respond to. The interaction implies that the increase in RT over set-size was larger in switch trials than in nonswitch trials, signifying an increase in focus-switch costs with set-size. RT in switch trials increased linearly over and above RT in non-switch trials, $F(1, 23) = 27.53$, $MSE = 56,070.42$, $p < .001$; the quadratic component was not significant, $F < 1$.

Accuracy as a function of working memory set-size—Accuracy decreased steadily with set-size (see Figure 2B). This decline was monotonic and precipitous throughout the entire range of set-sizes (from 95% at set-size 1 to 40% at set-size 4), $F(3, 69) = 40.99$, $MSE = 0.020$, $p < .001$.

Focus-switch cost as a function of set-size: Alternative hypotheses—So far, the analyses yield two consistent results: (1) There is a significant difference between RTs of nonswitch and switch trials even for the smallest relevant set-size (i.e., 2-count), and (2) the difference between the two types of RT increases linearly as set-size increases.

This section evaluates two primary hypotheses to explain these findings. The first primary hypothesis is that the difference between the switch and nonswitch RTs is due to a focus-switch cost, resulting from the fact that the focus of attention can hold only a single item. The alternate hypothesis is that the responses to the nonswitch trials are simply faster because of *repetition priming*: When the same shape is seen repeatedly, responses become faster. If repetition priming were a viable explanation of the shorter RTs on nonswitch trials (and if repetition priming accumulates over trials), one would expect a monotonic decline in RT, with the largest drop in RT for items repeated twice versus items that were not repeated only once. These RTs were submitted to a univariate ANOVA with repetition (varying from 1 to 3) as the factor. The main effect of repetition was not significant, $F(2, 46) = 2.66$, $MSE = 15,225.76$, $p = .079$.

The second primary hypothesis is that a search process occurs in the outer store to determine which count in working memory has to be accessed; this in turn increases the focus-switch cost when the set-size becomes larger because more units (counts) have to be searched. There are (at least) three alternative hypotheses to this second primary hypothesis, all pointing at artifacts in the data.

First, our design *confounds shapes and columns*: Each value of $N > 1$ was always associated with the same set of shapes. If, by accident, larger counts contain shapes associated with longer switch RTs, but not longer nonswitch RTs, an increase of focus-switch costs over set-size would result. This differential effect of shapes may not be found in the RTs of

nonswitch trials because no switching between shapes occurs. To test for this artifact, a series of 2 (switch vs. nonswitch) by N (column position) ANOVAs were conducted for each set-size, except for set-size 1. The main effect of column position was not significant for set-sizes 2 and 3, $F(1, 23) = 1.72$, $MSE = 21,770.32$, $p = .303$, and $F(1, 23) = 0.50$, $MSE = 37,679.87$, $p = .610$, respectively, but it was significant for set-size 4, $F(1, 23) = 3.35$, $MSE = 118,900.36$, $p = .043$; marginal means suggest a decrease in RT for both switch and nonswitch trials of the fourth column compared to the other three. Importantly, the interaction between switching and column position was not significant for any of the three set-sizes, all F -values $(1, 23) < 1$. That is, the switch RTs are parallel to the nonswitch RTs for all shapes in all set-sizes. Therefore, the increase in focus-switch cost over set-size is not a product of differential discriminability of the shapes used.

The second alternative hypothesis is that the increase in focus-switch cost with increased set-size might be due to the *increased number of intervening columns* associated with switching from one shape to another; switch trials might be slower when there are more intervening columns between two subsequent switches. The number of intervening columns does not affect the nonswitch trials because the two subsequent stimuli appear in the same column. To test whether the number of intervening columns (and/or the direction of switches) affects the switching time, nominal values were assigned to each switch trial on the basis of the number of intervening columns as well as directionality (left or right) from the preceding stimulus. A switch to the left was assigned a negative sign, and a switch to the right was assigned a positive sign. Note that these are post hoc analyses; the number of switches was not balanced across the different nominal values for each set-size when planning the experiment. RTs for each individual for the switch trials were submitted to a series of three univariate ANOVAs, one for each value of N , with intervening columns (2-count: $-1, +1$; 3-count: $-2, -1, +1, +2$; 4-count: $-3, -2, -1, +1, +2, +3$) as the factor (as an illustration, Figure 3 depicts RT as a function of directional intervening columns for the 4-count condition; the switch value of 0 indicates the nonswitch trials). The main effect of intervening columns was not significant for any set-size, largest $F(3, 69) = 2.22$, $MSE = 117,538.20$, $p = .116$. Therefore, RTs for switching did not increase with the number of intervening columns, irrespective of the switch direction.

The third alternative hypothesis is that the increase in the focus-switch cost over set-sizes 2 to 4 is a product of *lag effects* on the switch trials; the longer an item remains out of focus, the more resources might need to be expended to switch the item back into the focus of attention, or the larger the probability it has been subjected to time-based decay. Larger set-sizes are associated with longer average lags; therefore, switch RTs may be longer for larger set-sizes. There is no lag effect on nonswitch trials because the item already stays in the focus of attention. To test for lag effects, we computed the median response time for each individual for each lag value from each sequence. These RTs, for each lag value, were averaged across all sequences within a set-size for all individuals. They were further averaged across individuals for each set-size. The data are presented in Figure 4. We submitted the average RTs of each individual to a 3 (set-size 2 to 4) by 4 (lag 1 to 4) ANOVA. Lag 0 was not considered because it exists only for the nonswitch trials; lag values larger than 4 were also not considered because four is the longest lag value all set-sizes had

in common. The main effect of lag was not significant, $F(3, 69) = 0.67$, $MSE = 160,642.20$, $p = .572$. The interaction between lag and set-size, however, was significant, $F(6, 138) = 1.65$, $MSE = 132,802.70$, $p = .023$. The interaction appears to be due to the slight decrease in RTs as a function of lag in the 4-count condition; this result runs counter to the hypothesis that the increase in switch costs for higher counts might be due to lag effects. It seems then that once the item has disappeared from the focus of attention ($\text{lag} > 0$), the time needed to retrieve it is statistically independent of the number of intervening items.

Discussion

The main results obtained from Experiment 1 are: (1) The focus-switch cost is significant even for the smallest working memory load (set-size = 2); (2) there is a jump in RTs for nonswitch trials from set-size 1 to set-size 2; (3c) nonswitch RTs increase monotonically over set-sizes 2 to 4; and (4) the focus-switch cost increases linearly with set-size, at a rate of about 240 ms/item.

First, we found a significant difference in RT for switch and nonswitch trials at set-size 2—a cost of about 600 ms. This signifies that the focus of attention can hold no more than a single item. As expected, the focus-switch cost we obtained is comparable to the cost reported in the two previously published 2-count experiments (about 500 ms in Garavan, 1998; about 550 ms in Voigt & Hagendorf, 2002). Our data also suggest that there is little contribution of repetition priming to the nonswitch responses: There was no monotonic decline of RT over the number of repetitions in nonswitch trials. Note, however, that paradigms that have disambiguated repeated stimuli from repeated counts by assigning multiple shapes to a single count have noted small but reliable contributions of repetition priming to focus-switch costs (Gehring et al., 2003; Li et al., 2006).

Second, we found that the nonswitch RT is significantly larger at $N = 2$ than at $N = 1$. This increase is larger than the corresponding increase per set-size unit for set-sizes 2 to 4. This result replicates Oberauer's results from memory updating paradigm (2003). The difference seems attributable to the need to keep items active in the outer store in the 2-count condition, a requirement absent from the 1-count condition. This suggests that there is a trade-off of resources: The focus of attention works more efficiently when the outer store is not engaged. Alternatively, the finding could point at an attentional selection mechanism: When only one item is present, there is no competition; when two or more items are present, items outside the focus start competing for attentional resources. The focus then serves a filtering or selection function.

Third, we found that nonswitch RTs increase monotonically and near-linearly with set-size (see also Oberauer, 2003). This appears to be a side-effect of rehearsal: Twenty-three out of twenty-four participants used the rehearsal strategy at least sometime during the experiment; nineteen of these twenty-three participants started using it somewhere in either the 2-count or 3-count sequences in the first half of the experiment; eleven of these nineteen participants used rehearsal as a strategy from the second block onwards. Rehearsal would naturally give rise to a linear function if each unit is rehearsed at the same rate.

Fourth, we found that the focus-switch cost increases with set-size, at a rate of about 240 ms/ item. This result goes contrary to our previous findings with the identity-judgement N -back task (Verhaeghen & Basak, 2005; Verhaeghen et al., 2004), where the focus-switch cost remained stable over the $N=2$ to $N=5$ range. Before we could conclude that this increase in focus-switch cost over working memory load is real and attempt to pin down the source of the discrepancy, we needed to rule out the possibility that the increased cost is due to artifacts induced by our experimental paradigm. We investigated three such possible artifactual interpretations.

The first alternative hypothesis is that the focus-switch cost increases over set-sizes 2 to 4 because particular shapes affect the switch and nonswitch RTs differently. That is, it is possible that participants can access the count of some shapes faster than others (e.g., it might be easier to switch to more common shapes, like circles and squares). This would leave the nonswitch trials unaffected, but increase switch RT if by sheer accident the larger values of N contain less accessible shapes. This hypothesis was clearly found to have no merit: Switch and nonswitch RTs were parallel for all shapes for all set-sizes.

The second alternative hypothesis is that the switch RTs are sensitive to the number of intervening columns (either towards the left or the right) between two successive different shapes. This would yield larger averaged switch RTs for larger sets, where there are on average more intervening columns. This hypothesis was likewise falsified: Switch RTs were clearly independent of the number of intervening columns between two consecutive counts.

The third alternative hypothesis is that the switch response time may increase with set-size because increasing set-size increases the average lag between updates of any single count. It is possible that when the time delay grows longer or the number of intervening items increases, the activation of the counts declines; this can lead to an increase in access time (for a similar argument see Oberauer, 2003). This hypothesis was again not supported by the data: RTs did not vary meaningfully across lag.

We are left with the conclusion that focus-switch cost increases as a direct and linear function of the number of items held in the outer store. As explained in the introduction, the most likely explanation for the discrepancy between this finding and our previous results is the predictable and/or forward vs. random nature of probing: In the identity-judgement N -back task, the participant knows which element to access, namely the item in the N th position back, which is also the item following the item probed last; in the N -count task, the participant is completely uncertain which element will be probed next. Predictability, then, might help participants bypass the search process altogether; or the forward nature of probing might help subjects rely on an efficient chaining mechanism. (Note that implementing a predictable/forward version of the N -count task is not feasible—the task could then be done strategically by memorising the starting counts, keeping track of only one count, and then adding that count to all starting counts.)

EXPERIMENT 2

In Experiment 1, we investigated the dynamics of focus switching in the outer store when only active items are present. In Experiment 2, we turn our attention to the (lack of)

interplay between active and passive items in the outer store. We again use the N -count task. As in the Oberauer (2002) study, a subset of counts was designated as active (i.e., to be updated) and another subset as passive (i.e., not to be updated). The shapes associated with passive counts were shown as coloured outlines rather than as coloured shapes; subjects were instructed to increment the count for the coloured shapes only, and simply remember the starting count for the outlined shapes. In Oberauer's task, the passive items were a truly passive load—stored away and never accessed until final retrieval. We decided to increase competition between active and passive items by showing to-be-ignored shapes during the course of each trial. If the passive items are truly held in a state that is impermeable to the focus of attention and the outer store, the number of passive items should not affect either the switch or nonswitch RTs for active items, or their accuracy; likewise, intervening passive items should not influence either the switch or the nonswitch RTs for two active items.

Method

Participants—The sample consisted of 25 students (mean age = 18.40 years, $SD = 0.76$; mean number of years of education = 12.40, $SD = 0.71$; 17 females), who received course credit in return for their 1.5-2 hours of single-session testing.

Procedure—We used the N -count task as described previously, with four levels of set-size ($N = 1-4$). The number of active items (A) ranged from 1 to 4 and the number of passive items (P) ranged from 0 to 3, such that the total memory load ($N = A + P$) was always less than or equal to 4, yielding a total of 10 conditions, presented in increasing order of task difficulty: ($A = 1, P = 0$), ($A = 1, P = 1$), ($A = 1, P = 2$), ($A = 1, P = 3$), ($A = 2, P = 0$), ($A = 2, P = 1$), ($A = 2, P = 2$), ($A = 3, P = 0$), ($A = 3, P = 1$), and ($A = 4, P = 0$). For $N = 1$, all sequences are by definition non-switch; for $N > 1$, half of the sequences were nonswitch trials, half were switch trials. The starting counts for the shapes, ranging from 1 to 4, were randomly assigned. Shapes were displayed along an imaginary horizontal line at the centre of the screen, in N imaginary columns, one for each shape. To keep the length of the longest possible saccades equal across set-size, the horizontal separation between the two extreme column positions was kept constant at 12 cm for any sequence where $N > 1$. All active items were colour-filled shapes projected in the relevant leftmost columns; all passive items were unfilled shapes with coloured boundaries projected in the relevant rightmost columns. The shapes used (2 cm $\times$ 2 cm in size) were a red square, a green triangle, a blue circle, and an orange star. For $N = 1$, the choice of shape was randomised for each sequence; for $N = 2$, we used red squares and blue circles, left to right; for $N = 3$, we used red squares, blue circles, and green triangles, left to right; and for $N = 4$, we used red squares, blue circles, green triangles, and orange stars, left to right. Consistent colours and column positions were assigned to each shape to facilitate perceptual processing of each new stimulus.

Each condition consisted of 10 sequences and each sequence consisted of 12-16 (average = 14) to-be-responded-to items, yielding a total of 140 RTs for each condition. The sequence of events within a sequence was as follows (Figure 5 provides an illustration). After the disappearance of the central fixation cross, the shapes appeared in sequence, with the starting count presented right above it, in boldface, 1.5 cm tall. Participants paced

presentation by pressing the spacebar. Once the N starting counts were presented, the sequence proper started, and stimuli were presented one by one in their appropriate columns. The participant advanced to the next stimulus by pressing the spacebar; this is our measure of RT. The task was to update the count of the current shape by 1. To make the appearance of a new stimulus notable in nonswitch trials, the shapes were jittered by one pixel either left, right, up, or down from the stimulus previously presented in that location.

At the end of each sequence, the participant typed in the final counts of the active shapes and the initial counts of the passive shapes, left to right.

At the end of the computerised testing, the participants filled out a strategy questionnaire. Only RTs for correct sequences were included in the RT analyses; a sequence was considered accurate if the final counts of each item was reported correctly or deviated by +1 or -1 (see Garavan, 1998, and Voigt & Hagendorf, 2002, for similar treatment).

Results

RT and accuracy for sequences with active items only—Our first set of analyses concerned trials with only active items ($P = 0$, Figure 6). The RTs for $A > 1$ were submitted to a 3 (A : 2, 3, and 4) $\times$ 2 (switch and nonswitch) ANOVA. Switch trials were slower than nonswitch trials, $F(1, 24) = 141.64$, $p < .001$, and larger set-sizes took longer to respond to, $F(2, 48) = 7.34$, $p = .002$. Increase in RT over set-size was larger in switch trials than in nonswitch trials, $F(2, 48) = 6.23$, $p = .004$. This increase in focus-switch costs with set-size was linear, $F(1, 24) = 15.03$, $p < .001$, at 169 ms/set-size; the quadratic component was not significant, $F < 1$. Proportion of accurate sequences declined as a function of set-size, $F(3, 72) = 19.74$, $p < .001$, from 9.8 ($A = 1$) to 6.1 ($A = 4$).

In a follow-up analysis on nonswitch items only, now including the $A = 1$ condition as well, we found that nonswitch RTs increased as set-size increased, $F(3, 72) = 20.42$, $p < .001$. Nonswitch RT still increased with set-size when the analysis was restricted to set-sizes 2-4, $F(2, 48) = 3.47$, $p = .039$. To determine whether the increase in RT from set-size 1 to 2 was significantly longer than the increase per set-size for larger set-sizes, difference scores for successive pairs of set-sizes were computed for each individual and were submitted to a repeated measures ANOVA, which was significant, $F(2, 48) = 23.46$, $p < .001$; difference contrasts showed that the difference scores of set-sizes 2 and 1 were significantly different from the difference scores of set-sizes 3 and 2, $F(1, 24) = 36.81$, $p < .001$, but that the difference scores of set-sizes 3 and 2 were not significantly different from the difference scores of set-sizes 4 and 3, $F(1, 24) = 1.44$, $p = .242$.

The effect of passive items (P) on active items (A)—Judging from Figure 6, the number of passive items does not affect either nonswitch or switch RT for active items. To test this assertion, we conducted a series of ANOVAs, one for each level of A . For $A = 1$, RTs were submitted to a one-way ANOVA with the number of passive items as the within-subject factor. The main effect of the number of passive items was not significant, $F(3, 72) = 2.43$, $p = .073$. RTs were submitted to a 2 (switch vs. nonswitch) $\times$ 3 (number of passive items) ANOVA for both $A = 2$ and $A = 3$. Switch responses were slower than nonswitch responses: $A = 2$, $F(1, 24) = 127.32$, $p < .001$; $A = 3$, $F(1, 24) = 197.22$, $p < .001$. There was

neither a significant main effect of the number of passive items: $A=2$, $F(2, 32) = 2.33$, $p = .11$; $A=3$, $F(2, 32) = 0.19$, $p = .67$, nor a significant interaction between number of passive items and switching, for both $A=2$ and $A=3$, $F < 1$. Thus, the number of passive items had no discernible effect on RT for active items.

Likewise, accuracy did not decline as a function of the number of passive items, either inside or outside the focus of attention: In a formal test using a set of univariate ANOVAs, one for each value of A , all F -values for the effect of P were smaller than 1.

The effects of intervening passive items—We restricted our analyses of the effects of intervening items to the ($A=2$, $P=1$) condition, because this is the first condition where the participant was exposed to switch trials and a mix of passive and active items. We assume that the participant is most likely to experience interference from the passive items when relatively inexperienced— with more practice the participant might get better at ignoring passive items. The results are presented in Figure 7.

In a first analysis, we investigated three-item sequences in which two *identical* active counts bookended a passive item, constituting, in effect, a nonswitch, except for the intervening passive item. We compared the final RT of this sequence with the final RT for active-active two-item *nonswitch* sequences, and found no significant difference between the RTs, $t(23) = 0.37$, $p = .713$. In a second analysis, we investigated three-item sequences with a *different* active count book-ending a single intervening passive item. We compared the final RT of this sequence with the final RT for active-active two-item *switch* sequences, and found no significant difference, $t(24) = 1.72$, $p = .099$. Thus, RT in neither switch nor nonswitch trials was influenced by the presence of an intervening passive item.

Discussion

Before we examine the results regarding passive items, we would like to note that Experiment 2 closely replicates the main results of Experiment 1. First, switch RTs were larger than nonswitch RTs. Second, nonswitch RTs increased over set-sizes 2-4. Third, the focus-switch cost in RT increased over set-sizes 2-4. Fourth, there was a monotonic decline of accuracy over set-size.

Experiment 2 was designed to investigate the fate of a specific type of memory representation, namely representations of items that are to be recalled at the end of a trial, but not to be updated during it. Our results support Oberauer's (2002) conclusion that passive items indeed receive special status within the working memory system.

First, whereas the number of active items influences both nonswitch RT and focus-switch costs, the number of passive items does not. The finding that focus-switch costs are not influenced by the number of passive items replicates the results Oberauer obtained with his memory updating paradigm (2002). The finding that the number of passive items does not influence the RT by set-size slope for nonswitch items is quite new, and telling (results from Oberauer, 2002, show the same pattern, but no formal analysis was undertaken in his paper). One could argue (see Experiment 1) that the slope in active items is associated with rehearsal of the active counts. The present result then suggests that the passive items are

exempt from such rehearsal; if they were not, the slope would covary with the total number of items and not with the number of active items only. Thus, passive items are truly functionally distinct from active items—these counts are not even accessed during the course of a sequence. Note that this result stands in stark contrast with the results obtained for active items in both Experiments 1 and 2: Even the presence of a single active item in the outer store slows down nonswitch times considerably.

Second, we found that intervening passive items influenced neither RT for the item stored in the focus of attention, nor RT for the active items stored in the outer store. Again, this result provides evidence for the position that the passive count is not accessed at all, that is, it does not enter either the focus of attention or the active region of the outer store. If it did, it would lead to a switch cost when the passive item intervenes between two identical counts; if it intervened between two different active counts, an additional cost for retrieving it from the passive store into the active store would likely be observed.

Third, we found that accuracy declines solely as a function of the number of active items, not as a function of the number of passive items or the total item load. The decline in accuracy over set-size for active items has been observed before in this and related tasks (e.g., Experiment 1; Oberauer, 2002; Verhaeghen & Basak, 2005; Verhaeghen et al., 2004), and is typically attributed to item decay, interitem interference, or both. The finding that the presence of passive items does not contribute further to the decline suggests, again, that these passive items are functionally impermeable to the active items held in both the focus of attention and the outer store; passive items do not compete for activation with active items at all.

GENERAL DISCUSSION

Our two experiments add one new paradigm to the arsenal of paradigms used to study access times in working memory, the *N*-count task. This task was helpful in settling the question about the nature of access to items in the outer store and allowed for independent verification of the existence of a passive store, currently only investigated using the memory updating paradigm.

Before we go deeper into those two results, it is useful to point out that three paradigms have now confirmed that in tasks where stimuli are processed sequentially the focus of attention can hold exactly one item: the identity judgement *N*-back task (e.g., McElree, 2001; Verhaeghen & Basak, 2005), the memory updating task (e.g., Oberauer, 2002), and the running count/*N*-count task (Garavan, 1998; present study). This constitutes strong replication.

Regarding access to the outer store, the current results, as well as previous results by Basak (2006) and Oberauer (2006), clearly falsify our previous position that active items in the outer store are accessed through a content-addressable mechanism. It appears that not the complexity of the ongoing operations is the culprit, but the predictability of the item to be accessed. When items are tapped in a predictable sequence, identical to the input pattern, focus switch costs do not vary with set-size (e.g., Verhaeghen & Basak, 2005); when items are tapped unpredictably, focus-switch costs do increase with set size (as in the present

experiments). This increase is large: We found a cost of about 240 ms/item. This rate is much higher than the rate in classical memory scanning studies (e.g., Sternberg, 1966), which yield search times of about 40 ms/item. There are two possible explanations for this result. First, in classical scanning tasks, there are no ongoing operations performed on the items, so the scans happen at faster speed, indicated by a shallow slope; it is possible that the maintenance and update of multiple registers in memory in the *N*-back task and the *N*-count task increases the slope of search times, indicating slower and less efficient scanning. Another possibility is that the classical memory scanning task operates not on the outer store, but on an expanded focus of attention: Verhaeghen et al. (2004) found scanning rates close to Sternberg's in an expanded focus, holding four items after 10 hours of practice, and note that the literature points at an increase of scanning rates when the memory load is larger than four or five items, when the outer store is presumably tapped. Obviously, further research is necessary to settle this question. We also note here that our design does not allow us to probe into the nature of the search process occurring in the outer store— serial or parallel, exhaustive or self-terminating.

Our results confirm the existence of a third level of accessibility, outside the focus of attention and the active items held in the outer store. This third level consists of items that are tagged for later retrieval but are not subject to any kind of processing in the ongoing trial. This level is not only separate from the other two; our results suggest that a firewall exists between these passive items and both the focus of attention and the active items in the outer store: The number of passive items does not influence either RT or accuracy for active items, and intervening passive items do not create a switch cost. Such impermeability does not exist between the focus of attention and active items held in the outer store. First, the outer store acts as an overflow buffer for active items that cannot fit into the focus of attention: If the focus fills to capacity, items are by default relayed to the active outer store. It is not certain whether overflow from active items held in the outer store by necessity results in storage into the passive store. Second, items can be assigned passive status even if the outer store is not filled to capacity with active items: In Experiment 2, items could be easily relayed to passive status even in the presence of only one, two, or three active items. Third, and most importantly, even the presence of a single active item in the outer store slows down processing inside the focus of attention considerably.

Our results cannot determine what this third level is—whether it is truly part of the temporary buffer that working memory is defined as, or whether it is part of long-term memory. The strong compartmentalisation between this third level and the two inner tiers, combined with the crosstalk observed between the focus of attention and active items held in the outer store suggest that the third level may be part of long-term memory rather than an integral part of working memory.

One consequence of this finding, and Oberauer's (2002), is that the concept of cognitive load, often used in social or instructional psychology, might be in need of revision. Cognitive load experiments often provide participants with some information (usually a set of digits) to be retained while a concurrent task is being performed. In our classification, this manipulation amounts to imposing a passive load. If the goal of a cognitive load manipulation is to create interference within working memory and to investigate whether

processing is hampered by such interference, creating a passive load would be a suboptimal research strategy. Rather, investigators should impose loads that remain active, for instance by requiring frequent updating of the items held in working memory or by incorporating these items in the task at hand (e.g., Mayr, Kliegl, & Krampe, 1996).

Acknowledgments

This research was supported by a grant from the National Institute on Aging (AG 16201). We thank John Cerella for useful comments on an earlier draft. This work was conducted in partial fulfilment of the first author's doctorate degree requirements.

REFERENCES

- Basak C. Capacity limits of the focus of attention and dynamics of the focus switch cost in the working memory. *Dissertation Abstracts International: Section B: The Sciences and Engineering*. 2006; 66(10 B):5715.
- Basak C, Verhaeghen P. Subitizing speed, subitizing range, counting speed, the Stroop effect, and aging: Capacity differences, speed equivalence. *Psychology and Aging*. 2003; 18:240–249. [PubMed: 12825774]
- Cowan N. The magical number 4 in short term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*. 2001; 24:87–185. [PubMed: 11515286]
- Garavan H. Serial attention within working memory. *Memory and Cognition*. 1998; 26:263–276. [PubMed: 9584434]
- Gehring WJ, Bryck RL, Jonides J, Albin RL, Badre D. The mind's eye, looking inward? In search of executive control in internal attention shifting. *Psychophysiology*. 2003; 40:572–585. [PubMed: 14570165]
- Li Z, Bao M, Chen X, Zhang DR, Han S, He S, et al. Attention shift in human verbal working memory: Priming contribution and dynamic brain activation. *Brain Research*. 2006; 1078:132–142. [PubMed: 17109826]
- Mayr U, Kliegl R, Krampe RT. Sequential and coordinative processing dynamics across the life span. *Cognition*. 1996; 59:61–90. [PubMed: 8857471]
- McElree B. Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2001; 27:817–835.
- Nee DE, Jonides J. Neural correlates of access to short term memory. *Proceedings of the National Academy of Sciences*. 2008; 105:14228–14233.
- Oberauer K. Removing irrelevant information from working memory: A cognitive aging study with the modified Sternberg task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2001; 27:948–957.
- Oberauer K. Access to information in working memory: Exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2002; 28:411–421.
- Oberauer K. Selective attention to elements in working memory. *Experimental Psychology*. 2003; 50:257–269. [PubMed: 14587173]
- Oberauer K. Is the focus of attention in working memory expanded through practice? *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2006; 32(2):197–214.
- Sternberg S. High speed scanning in human memory. *Science*. 1966; 153:652–654. [PubMed: 5939936]
- Sternberg S. Memory scanning: Mental processes revealed by reaction time experiments. *American Scientist*. 1969; 57:421–457.
- Unsworth N, Engle RW. The nature of individual differences in working memory capacity: Active maintenance in primary memory and controlled search from secondary memory. *Psychological Review*. 2007; 114(1):104–132.

- Vaughan L, Basak C, Hartman M, Verhaeghen P. Aging and working memory inside and outside the focus of attention: Dissociations of availability and accessibility. *Aging, Neuropsychology, and Cognition*. 2008; 15(3):1–22.
- Verhaeghen P, Basak C. Aging and switching of the focus of attention in working memory: Results from a modified *N* back task. *Quarterly Journal of Experimental Psychology*. 2005; 58A:134–154. [PubMed: 15881295]
- Verhaeghen P, Cerella J, Basak C. A working memory workout: How to expand the focus of serial attention from one to four items, in ten hours or less. *Journal of Experimental Psychology: Learning, Memory, and Cognition*. 2004; 30(6):1322–1337.
- Voigt S, Hagenhoff H. The role of task context for component processes in focus switching. *Psychologische Beiträge*. 2002; 44:248–274.
- Zhang Y, Verhaeghen P. Glimpses of a one speed mind: Focus switching and search for verbal and visual, and easy and difficult items in working memory. *Acta Psychologica*. 2009; 131:235–244. [PubMed: 19552896]

Figure 1.

Experiment 1: Illustration of a 2 count sequence. If the sequence were as long as depicted, the final count of the two shapes, square and circle, would be 4 and 6, respectively.

Figure 2.

Experiment 1: Average response time of switch and nonswitch trials (Figure 2A) as a function of working memory set size for the full dataset. Accuracy is depicted in Figure 2B. The error bars depict the standard error.

Figure 3.

Experiment 1: Average response time as a function of number of switches to the right or left of the previous shape, that is, directional intervening columns, for 4 count condition. The error bars depict the standard error.

Figure 4.

Experiment 1: Average response time for set sizes 2 to 4 as a function of lag. The error bars depict the standard error.

Figure 5.

Experiment 2: Illustration of a 3 count sequence with two active counts (square and circle) and one passive count (triangle). If the sequence were as long as depicted, the final count of the three shapes, square, circle, and triangle, would be 5, 6, and 7, respectively.

Figure 6.

Experiment 2: Average response time of switch (S) and nonswitch trials (NS) as a function of active set size. The error bars depict the standard error. The parameter P denotes the number of passive items.

Figure 7.

Experiment 2: Data showing that intervening passive items do not influence RT in active sequences. A1 → A1 denotes accessing the same active count twice in succession (i.e., a nonswitch trial); A1 → P1 → A1 denotes accessing the same active count twice in succession but with one passive item inserted in between access episodes; A1 → A2 denotes accessing two different active counts in succession (i.e., a switch trial); A1 → P1 → A2 denotes accessing two different active counts in succession but with one passive item inserted in between access episodes. The error bars depict the standard error.