Components of Working Memory Updating

Ullrich K. H. Ecker (ullrich.ecker@uwa.edu.au)
School of Psychology, University of Western Australia
35 Stirling Hwy, Crawley, WA 6009, Australia

Stephan Lewandowsky (lewan@psy.uwa.edu.au)
School of Psychology, University of Western Australia
35 Stirling Hwy, Crawley, WA 6009, Australia

Klaus Oberauer (K.Oberauer@bristol.ac.uk)
Department of Experimental Psychology, University of Bristol
12a Priory Rd, Bristol, BS8 1TU, UK

Abstract

Working memory updating (WMU)—the ability to maintain accurate representations of information changing over time—has been successfully used in individual differences research to predict higher cognitive abilities. For instance, WMU has been found to predict fluid intelligence and reading comprehension. However, little is known about the underlying component processes of WMU or the relationship between WMU and working memory capacity (WMC). A decomposition of WMU into three distinct components—retrieval, transformation, and substitution—was implemented into a standard WMU paradigm. Experimental conditions featured every possible combination of these components. The decomposition was used to analyze the relationship between WMU subcomponents and WMC. We utilized structural equation modeling in a novel way, in that both interindividual variability and experimental effects on mean performance measures (RT and accuracy) were accounted for concurrently. Results suggest that the proposed components make distinct and additive contributions to WMU. We found that WMC reliably predicts WMU in general, but also that some components of WMU are independent of WMC. Hence, WMU and WMC may make independent contributions in predicting higher mental abilities.

Keywords: working memory updating; working memory capacity; individual differences; structural equation modeling

WM Updating and Individual Differences

The pattern of correlations between tasks can address issues relating to the number and nature of processes or components involved in WM. For instance, Oberauer and colleagues (e.g., Oberauer, Süß, Schulze, Wilhelm, & Wittmann, 2000) have used such an individual differences approach to decompose the WMC construct into functional components, which were then related to higher cognitive functions.

We already noted that WMU has also been found to predict such higher functions; however, most studies have measured WMU by the accuracy on WM tasks that require updating, such as a running memory task. This does not supply a pure measure of WMU, and therefore we do not know whether the correlation between the “updating” factor and intelligence reflects a relationship between intelligence and WMU, or intelligence and WMC. Indeed, some researchers have claimed that there is only a weak link between WMU and WMC (e.g., Radvansky & Copeland, 2001), so on the one hand, WMU and WMC may be dissociable dimensions.

On the other hand, prominent theories of WM that assume a tight link between WMC and executive functions (e.g., Kane et al., 2004) should predict that WMU and WMC are closely related because WMU is regarded as one executive function.

The Present Study

We began our examination by conducting a task analysis of previously-used WMU paradigms. We identified 3 putative sub-processes of WMU: retrieval (R), transformation (T), and substitution (S). Our experiment orthogonally manipulated these three components, thus permitting their empirical identification and assessment of their interrelationship. Further, our study included 4 independent working memory tasks, thus permitting a reliable estimate of each participant’s WMC. These WMC estimates were then statistically related to the individual components of WMU isolated by our experimental manipulations.

Method

Participants

Ninety-seven psychology students from the University of Western Australia participated in the two experimental ses-
sions for partial course credit (13 males, age range 19-41, mean age 21.2 years).

Working Memory Capacity Measures

The first session involved measurement of each subject’s WMC using a battery of 4 standard WMC tasks, taken with slight modifications from Oberauer (2005). The tasks were: Memory Updating (MU), Operation Span (OS), Sentence Span (SS), and Spatial Short-Term Memory (SSTM).

Working Memory Updating Experiment

The WMU experiment was carried out in the second session. The task was to encode a set of 3 letters, each presented in a separate frame, and to subsequently update these letters. Each trial involved 6 updating steps. All updating operations were cued by displaying the appropriate prompt (see below) in the to-be-updated frame. Subjects keyed in the result of the update at every updating step. To avoid an influence of frame switching (Oberauer, 2002), we held switching constant by moving to a new (randomly chosen) frame on each step.

There were 8 experimental conditions created by fully crossing the factors R, T, and S in a within-subjects design. Conditions involving retrieval required subjects to retrieve the most recent letter of the cued frame from memory to perform the current operation. In contrast, this letter was provided with the cue in no-retrieval conditions. Transformation conditions involved a transformation of the selected letter by alphabet arithmetic. Only positive operations of +1 and +2 were used. Substitution conditions resulted in the replacement of memory content with new information, whereas the outcome of no-substitution steps was identical to the information already held in memory. Conditions are summarized in Table 1, which presents the stimuli shown on a given updating step assuming the letter “C” is the currently remembered letter of the targeted frame; as in the table, conditions are referred to by numbers (1) to (8) in the following.

Condition (8) involved neither of the three processes, and the currently remembered letter was presented again (baseline). Condition (4) was identical but a different letter was presented (pure S without R or T). In condition (7), a “?” prompted subjects to retrieve the currently held letter and report it (pure R). Condition (5) was designed as a transformation that does not substitute memory content ( + 0). Conditions (1), (2), and (6) involved standard alpha-arithmetic operations. For instance, in (1), subjects had to add a number to whatever they currently remembered for that particular frame. The result of the transformation in (6) was identical to the remembered letter, hence no substitution. In (3), an arrow from one frame to another indicated that subjects should retrieve and then copy the letter from one frame to the other, thus requiring retrieval and substitution but no transformation. Note that conditions (3) and (5) were designed to permit the orthogonal combination of all three experimental variables, but that this required the use of peculiar operations (e.g., “+0” to avoid a substitution). We deal with the implications of these design decisions during data analysis.

Table 1: Conditions (Condition Numbers in Parentheses) and Sample Prompts Used in the WMU Experiment

<table>
<thead>
<tr>
<th>R yes</th>
<th>R no</th>
<th>T yes</th>
<th>T no</th>
</tr>
</thead>
<tbody>
<tr>
<td>S yes</td>
<td>(1) ?+1</td>
<td>(2) A+1/B+2</td>
<td>(3) →</td>
</tr>
<tr>
<td>S no</td>
<td>(5) ?+0</td>
<td>(6) A+2/B+1</td>
<td>(7) ?</td>
</tr>
</tbody>
</table>

Legend: R, Retrieval; T, Transformation; S, Substitution.

Note. Examples of prompts assume that the letter “C” is currently memorized. See text for details.

Table 2: Accuracy of Updating (in Upper Rows) and Reaction Times (in ms, in Lower Rows) in WMU Experiment

<table>
<thead>
<tr>
<th>R yes</th>
<th>R no</th>
<th>T yes</th>
<th>T no</th>
</tr>
</thead>
<tbody>
<tr>
<td>S yes</td>
<td>.76 (.016)</td>
<td>.87 (.011)</td>
<td>.82 (.014)</td>
</tr>
<tr>
<td>S no</td>
<td>3115 (63)</td>
<td>3009 (69)</td>
<td>2141 (50)</td>
</tr>
<tr>
<td>S yes</td>
<td>.85 (.013)</td>
<td>.89 (.011)</td>
<td>.87 (.012)</td>
</tr>
<tr>
<td>S no</td>
<td>1599 (35)</td>
<td>2803 (69)</td>
<td>1297 (24)</td>
</tr>
</tbody>
</table>

Legend: R, Retrieval; T, Transformation; S, Substitution.

Note. Standard Errors in Parentheses.

Results

Descriptive Data from the WMU experiment—mean updating performance and RTs—are shown in Table 2.

Structural Equation Modeling

To investigate the relation between WMC and WMU, we used structural equation modeling (SEM). SEM is typically used to capture individual differences and correlational dependencies between latent variables, without regard to experimental manipulations or differences between means. In the present case, we extended this standard approach by also modeling mean RT and accuracy for each experimental condition. We thus constructed an SEM model that concurrently captured both inter-individual variation and experimental effects.

Preliminary multilevel regression analyses suggested that the effects of transformation, retrieval and substitution were additive and showed no sign of interaction for either accuracy or RT (i.e., model fits were not improved by adding interactions); hence we initially focused on models that preserved the additive structure among experimental variables.

The SEM models for Accuracy and RT are depicted in Figure 1. The models consisted of two measurement models—one for the WMC part and one for the WMU part. The WMC part had 4 observed variables—one for each task of the WMC test battery—which were linked to a single WMC latent factor. The WMU part had 8 observed variables—one corresponding to each of the 8 experimental conditions (explained in Table 1)—that were connected to latent factors corresponding to the experimental variables R, T, and S, and a

\footnote{Given the similarity of the WMC task MU and the WMU experiment, analyses were also carried out excluding MU from the indicators of WMC. These generally supported the results from the full models.}

348
general factor accounting for the general level of performance (in regression terms: the intercept). Accuracy measures were probit-transformed to compensate for the lack of normality.

We imposed strong constraints on our models that are typically absent in SEM applications. First, all loadings between manifest and latent variables in the WMU measurement models were fixed. Specifically, for both accuracy and RT models, all loadings on the general performance variables \((GenAcc)\) and \((GenRT)\) were set to 1. Loadings on the three factors \(R\), \(T\), and \(S\) were set to \(+1\) or \(−1\), respectively, as any additional need for retrieval, transformation, or substitution was assumed to increase RT but reduce accuracy. In addition, the intercepts of the manifest variables were fixed to zero, as were the means of the error terms associated with them.

The rationales for these constraints were as follows. First, the additivity suggested by the regression is captured by the fixed loadings because they imply that each factor has the same effect on all conditions that load on it—for instance, retrieval reduces accuracy by the same amount in all four design cells involving retrieval. Second, fixing the error means to zero implies that no individual condition had a mean higher or lower than that predicted from the additive model. Thus, the estimated mean of the General factor represents the baseline level of performance (i.e., the intercept), and the estimated means of the three factors representing components of updating reflect the mean effect of each experimental manipulation. Correspondingly, the estimated variances of the factors reflect the individual differences in baseline performance and in the magnitude of the experimental effects, respectively.

For both RTs and accuracies we had to relax these strong constraints of the purely additive model at a few points, either based on theoretical considerations or, on a few occasions, based on deviations of the data from the additive model, associated with conditions (3) and (5). For instance, the “arrow” condition (3) involved an additional attentional shift, which we estimated by fixing the associated error term \(e3\) of the RT model to an independent estimate (483 ms) provided by Garavan (1998). The accuracy model departed from the strict additive structure in that the link between \(T\) and \(Acc_{pb}.5\) was allowed to vary freely, reflecting the assumption that the “zero-transform” condition (5) entailed a smaller effect of \(T\) than the other three conditions involving a transformation. Moreover, the means of the error terms \(e4\) and \(e8\) were allowed to be different from zero (but were constrained to be equal) because they refer to “type-what-you-see” conditions (4) and (8), which differed from the remaining conditions in that they did not allow for much error other than erroneous copying of the screen display). This effectively means that these conditions, for obvious conceptual considerations, had intercepts larger than would be expected from the additive model.

In the overall structural models, the covariances between latent WMU variables and \(WMC\) reveal the extent to which individual variation in WMC relates to variation among subjects’ responses to our experimental WMU manipulations. Given the unusually strong constraints imposed on the models, we achieved acceptable fit \((χ^2 (26/24) = 37.92/63.28, \text{comparative fit index} = .95/94, \text{root-mean-square error of approximation} = .061/.086, \text{standardized root-mean-square residual} = .080/.068 \text{for the accuracy and RT model, respectively})\). The models are readily summarized: First, they confirmed the assumed additivity among experimental variables. Transformations had a strong impact on both accuracy and especially RT, whereas substitutions had small but reliable effects, and retrieval had a large effect on accuracy but no effect on RT. At the level of means, the models provided an excellent fit of the data from the updating experiment. Hence the estimated means shown in the SEM figures permit accurate reconstruction of all observed means (Table 2); for example, if an operation involves \(T\) and \(S\) but not \(R\), the RT predicted by the model is calculated as the \(GenRT\) mean plus \(T\) and \(S\) means, or 1.275 + 1.470 + .335 = 3.08—the observed level of performance was 3.01 seconds. Overall, with the exception of \(T\) estimates for conditions (3) and (5), SEM mean predictions were very close to the observed condition means, with average deviations of 30 ms and below .03 accuracy units, respectively.

Second, the models captured the variance among individuals’ WMC and related it to the variation in the magnitude of the experimental effects. That is, in the accuracy model, \(WMC\) correlated positively with overall accuracy on the updating task \((GenAcc)\) and negatively with \(T\) and \(S\) (the latter correlations are negative because the higher WMC, the lower the negative impact of \(R\) and \(T\) on updating accuracy). In the RT model, a relation between \(GenRT\) and \(T\) could be observed within the WMU part of the model, such that the higher the baseline RT reflected in \(GenRT\), the larger the impact of \(T\). The model additionally showed that \(WMC\) correlated negatively with both overall updating latency \((GenRT)\) and \(T\). In both models, \(S\) contributed significantly to WMU accuracy, but did not covary with \(WMC\).

General Discussion

We presented the first examination of the basic processes that govern the updating of WM. Our decomposition of WMU into three distinct components—retrieval (R), transformation (T), and substitution (S)—was tested by examining every possible combination of these components. The data indicate that these processes make distinct and additive contributions to WMU performance.

Further, we used the decomposition of WMU to analyze the relationship between the WMU components and WMC. As noted at the outset, both the view that WMC and WMU are strongly related and the view that they form dissociable dimensions of mental ability have been represented in the literature. Our study may help resolve this inconsistency because while we found that WMC did predict performance in the experimental WMU task to some degree, some processing components involved in that task were independent of WMC.
We now discuss the decomposition of WMU first, followed by some of its theoretical implications, and finally the relationship between WMU and WMC.

**Component Processes**

The analyses revealed specific costs associated with each component of WMU. Retrieval and transformation, and to a lesser degree substitution, affected the accuracy of updating. In contrast, transformation and—again to a lesser degree—substitution, but not retrieval, determined WMU response times. Hence, transformative processing seems to be the main determinant of WMU processing, retrieval processes seem to be error-prone, and substitution also has a significant, albeit smaller, impact on WMU.

The fit of our structural equation models (and preceding linear regression models) was good without including any interactions among components; hence we suggest that the contributions of R, T, and S to WMU are orthogonal and additive. This observation favors a strictly serial processing model because if processes were partially parallel (i.e., cascaded, cf. McClelland, 1979; Navarrete & Costa, 2005) interactions would be likely to occur.

One could argue, of course, that our decomposition of WMU processes is too coarse. Indeed, it is plausible to further deconstruct some of the components identified here. For example, the substitution process may in turn involve the removal of old information followed by the addition of new information (cf. Oberauer & Vockenberg, in press; Postle, 2003). We therefore expect that the proposed components will be subject to further examination and refinement; nonetheless, the fact that our data conformed to an additive model suggests that any further decomposition of processes would only take place within the proposed components.

**Instantaneous Retrieval**

One perhaps surprising aspect of the data is that retrieval did not affect updating times. What does this result tell us about the structure and representations of working memory? It implies that there is no separate process of retrieval that occurs only in the R but not in the no-R conditions. With each step involving a frame switch, each updating step brings a new object in WM into the focus of attention. Our data suggest that this involves obligatory retrieval, whether that object is needed for the particular processing step or not. We focus on two models in the literature that can accommodate such a finding.

Oberauer (e.g., 2002) proposed that there are three functionally distinct layers of representation in working memory: (1) the activated part of long-term memory, (2) the region of direct access, and (3) the focus of attention. In this model, the focus of attention can only select a single item from the direct-access region. Making a representation available for processing therefore involves two steps: retrieving a representation from long-term memory, thereby bringing it into the region of direct access, and accessing one representation among several candidates in the direct-access region. Within this model, we can explain our results as follows: All three letters are continuously held in the direct-access region. Ev-
Every updating step involves a frame switch, and the new frame serves as an automatic trigger for the focus of attention to focus on that frame. The critical assumption is that moving the focus of attention to a new frame means not merely to move visual attention to a new location in space, but rather involves focusing on a new object in WM. That object is a tightly bound composite of the spatial location of the frame and the letter currently bound to it; hence it is immaterial whether or not that letter has to be retrieved for the next updating step as it is already in the focus of attention.

An alternative model was proposed by McElree based mainly on examination of the time-course of retrieval (e.g., McElree, 2001). This model postulates that there are only two representational states: (1) passive representations and (2) a single item in the focus of attention. The latter item can be processed very quickly, whereas items outside of attention are retrieved in a time-consuming fashion with constant retrieval speed. Similar to Oberauer’s model, this view can also handle the present data by assuming that whenever people shift their focus of attention to a new frame (as they had to at all steps in our experiment), they automatically retrieve the corresponding letter (whether needed or not).

Because retrieval in McElree’s model is from long-term memory, it follows that in order to accommodate our results, retrieval of object-location bindings from long-term memory must be automatic and obligatory. We argue that there is ample evidence that associative retrieval from long-term memory is not automatic. Nobel and Shiffrin (2001) made a strong point that both associative recognition and cued recall differ from single-item recognition in that they involve a slow, non-automatic search process, and Gronlund and Ratcliff (1989) argued for a different time course of retrieval for single items and associative information. The neuropsychology literature has strongly corroborated these assumptions, and suggests that automatic retrieval from long-term memory is limited to the recognition of single items (or highly unitized chunks). By contrast, associative retrieval of arbitrary bindings, such as object–location, word–temporal position, or face–name pairings, rests on slower retrieval processes that have been shown to be strategically controlled (Cansino, Maquet, Dolan, & Rugg, 2002; Ecker, Zimmer, & Groh-Bordin, 2007).

To conclude, we suggest that in order to explain our finding of obligatory retrieval of location-bound information at each updating step, the McElree model needs to make assumptions about associative long-term memory retrieval that are difficult to reconcile with research in that area. By contrast, the Oberauer model avoids these issues by means of the intermediate stage of direct access.

Yet, why did our R conditions incur a cost in accuracy? If a focus shift entails obligatory retrieval, why is responding so much more accurate when the item is (re-)presented? The reason is simple: Memorized representations of the three letters are not always accurate—either because a previous updating step went wrong, or because the object, which was previously outside the focus of attention, has been degraded by the time it is accessed.

The Role of WMC for WMU
Another aim of this study was to examine the relationship between WMC and WMU component processes, using structural equation models. Not surprisingly, WMC was strongly related to the latent factors that represented baseline performance in the WMU task (i.e., GenRT and GenAcc); these factors simply represent performance on a WM task. Of greater theoretical interest is the relation between WMC and the specific factors representing experimental effects; we discuss these in turn.

Transformation had the strongest impact on WMU performance overall; conditions featuring a transformation were significantly slower and less accurate than conditions without transformation requirements. The size of this impact varied consistently with the size of the transformative operation, replicating the known effects of “problem size” with alpha-arithmetic operations in untrained subjects (e.g., Logan, 1988). Further, the SEM models demonstrated that the T latent variable showed consistent covariation with WMC. It could thus be concluded that WMC predicts transformation skills and that these primarily underpin the relation between WMC and any updating that involves transformation.

WMC strongly predicted the accuracy cost of retrieval, such that people with lower WMC were more likely to fail in retrieving the correct item. This finding echoes recent reports that WMC predicted the accuracy of retrieval in an updating task (Unsworth & Engle, 2008). These authors argued that low-span subjects have difficulties to correctly select items from outside the focus of attention during retrieval. Given that retrieval did not take any time (over and above the time necessary to encode new information in no-retrieval conditions), we cannot interpret the null covariation between R and WMC meaningfully. Any individual differences in retrieval time are either captured in our generic GenRT factor (among other processing aspects) or are absent altogether.

Finally, substitution had a small but reliable impact on WMU performance. Moreover, individual differences in the S factor were not correlated with WMC. In the accuracy model, the variance in S was too small for a covariance with WMC to become significant. For the RT model, however, it stands to reason that the null correlation is genuine, which implies that the encoding of a new letter into working memory and the removal of a previous letter from working memory are unrelated to WMC, and this is consistent with previous research (Vogel, McCollough, & Machizawa, 2005; Oberauer, 2005). We therefore conclude that unlike the other two components (retrieval and transformation), substitution uniquely contributes variance to WMU independent of WMC.

Our finding that substitution is unique to WMU may help clarify the interrelations between WMU, WMC, and higher cognitive functions. Specifically, if WMU tasks are utilized that draw heavily upon substitution (as opposed to retrieval and transformation) skills, then updating performance may
be more likely to be independent of WMC than if the WMU tasks draw on retrieval and transformation (e.g., compare Galletly, MacFarlane, & Clark, 2007 and Friedman et al., 2006).

Conclusions
In summary, we reported the first study that systematically decomposed working memory updating into constituent processes. Our results suggest that the processes of retrieval, transformation, and substitution make distinct and additive contributions to updating performance.

We further found a selective relationship between working memory capacity and working memory updating, whereby only the process of substitution seems to be unique to updating. This potentially reconciles inconsistencies in the literature, as some studies may have conflated updating measures with more capacity-related factors.

Acknowledgments
Preparation of this paper was facilitated by a Discovery Grant from the Australian Research Council to SL and KO, as well as an Australian Professorial Fellowship to SL. The lab’s web address: http://www.cogsciwa.com. We thank Abby Chee, Lee-Xieng Yang & Charles Hanich for research assistance.

References