The personalisation method applied to a working memory task: Evidence of long-term working memory effects

Alessandro Guida
Laboratoire de Psychologie et Neurosciences Cognitives, Université Paris Descartes, Paris, France

Hubert Tardieu
Laboratoire de Psychologie et Neurosciences Cognitives, Université Paris Descartes, and Ecole Pratique de Hautes Études, Paris, France

Serge Nicolas
Laboratoire de Psychologie et Neurosciences Cognitives, Université Paris Descartes, Paris, France

Ericsson and Kintsch (1995) proposed that, in situations of expertise, individuals can overcome working memory limitations by using long-term working memory. It allows a greater capacity than working memory thanks to long-term memory encoding and retrieving. To test this characteristic, an adaptation of Daneman and Carpenter’s (1980) reading span was used. To operationalise expertise, the personalisation method (Guida & Tardieu, 2005) was employed. In Experiment 1, a personalised group, which read reading span sentences that mentioned familiar locations, was compared to a nonpersonalised group, which read sentences with unfamiliar locations. In Experiment 2, a personalised group, which read reading span sentences with neutral locations, was encouraged to mentally personalise these locations by thinking about known locations. This group was compared to a nonpersonalised group, which was encouraged to think about unknown locations. The personalised groups were expected to store and retrieve information in long-term memory via long-term working memory more easily than the nonpersonalised groups, which had to count massively on working memory. The results showed that personalisation enhanced reading span and confirmed one implication of the
Keywords: Long-term working memory; Working memory capacity; Personalisation; Reading-span; Domain knowledge.

Ericsson and Kintsch (1995) presented a new theory of memory, with the goal of resolving the paradox concerning the severe limitations on short-term memory and the ease with which people face the demands of complex tasks in everyday life. It is difficult to understand how very complex cognitive activities can be processed in a restricted storage space such as working memory. As Kintsch (1998, p. 215) says, "How can people live with such a terrible memory?" The purpose of the long-term working memory theory is to solve this problem by proposing that, under certain conditions, individuals can use retrieval structures to overcome the limits of working memory and perform high-level cognitive tasks.

After presenting the theory of long-term working memory and the personalisation method we used to operationalise this construct, we will describe our experimental study.

LONG-TERM WORKING MEMORY THEORY

Long-term working memory (LT-WM) theory proposes that part of long-term memory (LTM) can be used as working memory (WM), thus expanding an individual's memory storage and processing capacities. LT-WM is defined as the part of LTM that can be accessed rapidly and reliably by means of retrieval cues in WM. In other words, the available items in WM are the cues that activate a portion of LTM.

This proposal is very close to other theoretical proposals that include immediate access to LTM elements that are available for processing but are not actively maintained by attention processes (e.g., Anderson, Reder, & Lebiere, 1996; Cowan, 1995; Just & Carpenter, 1992; Oberauer, 2002). It differs from these models in terms of the description and influence of LTM structures. Based on Chase and Ericsson’s (1982) Skilled Memory Theory, Ericsson and Kintsch (1995) tried to describe how LTM structures allow and influence immediate access to LTM elements. They highlighted the role of retrieval structures that link people’s knowledge with the information they are processing. In the view of these authors, access to LTM elements becomes truly immediate only when retrieval structures have been built,
which occurs after extensive practice (for the importance of retrieval structures, see also Gobet, 2000a, 2000b; Richman, Staszewski, & Simon, 1995). These mechanisms allow LT-WM to explain data that are difficult to interpret with models based on the classical definition of WM as “temporary storage of information that is being processed in any of a range of cognitive tasks” (Baddeley, 1986, p. 34) especially if “temporary storage” means “active maintenance” or “rehearsal”. It is hard to understand how skilled activities can be interrupted and resumed later without detrimental effects on performance (for an example in text comprehension, see Glanzer, Dorfman, & Kaplan, 1981; Glanzer, Fisher, & Dorfman, 1984; in chess, see Charness, 1976; Cooke, Atlas, Lane, & Berger, 1993; Frey & Adesman, 1976). However, if one considers that elements in WM are rapidly linked to LTM elements, transferred to LTM, and later retrieved, the absence of a disruption effect becomes coherent.

LT-WM can also clarify certain problems related to short-term storage capacity, which is about 7 ± 2 units (for the most optimistic, see Miller, 1956; but see Cowan, 2001; Garavan, 1998; Gobet & Clarkson, 2004), considering only storage, or 3 to 4 items, considering storage plus processing (as can be measured, for example, by reading span, Daneman & Carpenter, 1980). These limitations are too small to explain performance on complex tasks or skilled activities. For example, they do not clearly explain the span size of mnemonists (for a complete review, see Ericsson & Kintsch, 1995). And current production systems such as ACT (Adaptive Control of Thought: Anderson, 1983; Anderson & Lebiere, 1998) and Soar (State, Operator, and Result; Newell, 1990) require more working space than seven elements. Again, if one assumes that elements in WM are rapidly linked to LTM elements and later retrieved from LTM, which means that a part of LTM can be used as WM, then these problems disappear.

Compared with WM, LT-WM allows greater and longer storage capacity, since LT-WM implies that it is possible to encode and retrieve from LTM. The first aim of this paper is to test this contrast by using the personalisation method (Guida & Tardieu, 2005).

THE PERSONALISATION METHOD

This method is based on Kintsch, Patel, and Ericsson’s (1999) definition of expertise in text comprehension. According to these authors, three conditions must be fulfilled if individuals are to use a part of LTM as WM while reading. First, the individuals have to be confirmed readers; second, the material has to be well written; and third, it has to have familiar content. In the personalisation method, the first two requirements are fulfilled; the last
characteristic is used as a variable with two modalities: a familiar content for a first group of participants, in order to make it possible for this group to use LT-WM, versus a less-familiar content for a second group, less convenient to use LT-WM.

The familiar and less-familiar content are obtained by personalising—or not personalising—the linguistic material. For the familiar content, participants read personalised material that includes very well-known locations (their university, their apartment, etc.). In the less-familiar version, participants read exactly the same material with exactly the same types of locations but located in a city they have never been to (e.g., an Armenian university, an Armenian apartment, etc.). In both groups, one or two objects are associated with the locations; the participants’ task is to recall the objects.

The idea of personalising data has been used before, especially to study its impact on maths problems in education (e.g., Cordova & Lepper, 1996; Davis-Dorsey, Ross, & Morrison, 1991; Ku & Sullivan, 2000; Lopez & Sullivan, 1992; Wright & Wright, 1986), but never to simulate expertise and therefore to operationalise LT-WM. It is important to emphasise that the situation of expertise is generated conjointly by the personalisation and the structure of the linguistic material. The material is constructed in such a way that reading it in the familiar condition is like using the Loci method (Yates, 1966). The ancient Greeks used the Loci method to remember important words during speeches. Before a speech, Greek orators would visualise familiar locations and use them to mentally store the important words. During the speech, they could retrieve those words simply by thinking about the familiar locations. The method of personalisation places participants reading a personalised version in a similar situation, where they can use schemas of known locations to store words and later retrieve them.

In a preliminary study, Guida and Tardieu (2005) attempted to confirm that LT-WM allowed a greater and longer storage capacity than WM. Using the personalisation method, they contrasted a personalised group, which read texts featuring familiar locations, with a nonpersonalised group, which read texts with unfamiliar locations. One or two objects were positioned in each location. At the end of each text, participants had to recall the objects, using the locations as cues. The rationale was that participants in the personalised group were able to store information in LTM via LT-WM, whereas participants in the nonpersonalised group would be less able to use LT-WM and therefore had to count massively on WM to recall the objects (a LT-WM vs. WM contrast). Thus, subjects in the former group were expected to recall more objects. The results confirmed these predictions, suggesting that the personalisation method was valuable.

1 The description of Guida and Tardieu’s (2005) experiment is simplified in order to present only the crucial points.
However, the task used by Guida and Tardieu (2005)—a cued recall task following a text—was not merely a WM task; in all likelihood, LTM storage occurred. This led to a problem that affected the comparison between LTM-WM and WM. Even though participants in the nonpersonalised group were less able to use LTM-WM and had to count primarily on WM to recall the objects, they also undoubtedly used LTM to encode and retrieve information, since the task was not a WM task. The comparison between LTM-WM and WM was therefore invalid. To really pinpoint the LTM-WM vs. WM contrast and show the impact of LTM storage during WM activities—which is the definition of LTM-WM\(^2\)—the personalisation method would have to be applied to a strictly WM task. In that way, by comparing personalised subjects and nonpersonalised subjects, one can make the best possible comparison of LTM-WM storage versus WM storage. This is what the first part of this paper addresses. The personalisation method has therefore been applied to Daneman and Carpenter’s (1980) reading span task. In this way, any advantage the personalised group achieves can be interpreted as an extension of WM, since this task is supposed to measure WM.

**APPLYING PERSONALISATION TO A READING SPAN**

In our reading span task, participants had to read sentences aloud while remembering the last word of each sentence. The sentences were presented in groups that started with the smallest (two sentences) and increased to the largest group (six sentences). Participants were given three trials at each group size. The task stopped when participants failed to recall correctly the last words of the sentences in two trials out of three. Our reading span was based on a French adaptation of Daneman and Carpenter’s (1980) reading span task (Desmette, Hupet, Schelstrate, & van der Linden, 1995).

In Experiment 1, we developed a personalisable reading span task by inserting locations in each sentence. These locations could be personalised, meaning that the reader would be familiar with them, for example a well-known lecture theatre (“She enters the Lagache lecture theatre . . .”), or these locations could be nonpersonalised, for example with the name of a lecture theatre in an Armenian university (“She enters the Dovassar lecture theatre . . .”). This procedure created a personalised reading span and a nonpersonalised reading span. In Experiment 2, we also inserted locations into each sentence of a reading span, but this time they were neutral locations (“She enters the lecture theatre . . .”). To personalise these

\(^2\) Ericsson and Kintsch (1995, p. 211) wrote: “We propose that a general account of working memory has to include another mechanism based on skilled use of storage in long-term memory (LTM) in addition to the temporary storage of information.”
locations, we used a different kind of personalisation by inducing subjects to mentally personalise the locations. Before taking the reading span test, we asked our participants to read a personalisable text. Half of the participants read a personalised text that included known locations; the other half read a nonpersonalised text that included Armenian locations. The personalisable text was used in such a way that it was intended to encourage the participants to mentally personalise the neutral locations in the subsequent reading span task. Therefore, we encouraged one group of participants (the personalised group) to mentally personalise the locations of the reading span sentences by thinking about known locations while they read about the neutral locations. We encouraged the other group (the nonpersonalised group) to think about unknown locations in Armenia while reading about the neutral locations. This procedure also created a personalised reading span group and a nonpersonalised reading span group.

In both experiments, we expected that participants working with the personalised version would have a bigger reading span than participants using the nonpersonalised version. If this result was obtained, it could be interpreted in the framework of Ericsson and Kintsch’s theory as indicating that LT-WM was used to extend WM.

But this was not the only hypothesis our two experiments allowed us to formulate. The structure of the experiments and the use of a WM task allowed us to address a proposal—derived from LT-WM theory—that Ericsson and Kintsch (1995) had made. They suggested that the reading span task (Daneman & Carpenter, 1980) does not measure WM in the form of a temporary storage, but more as a capacity to retrieve information from LTM. This idea goes against the dominant point of view that considers cognitive resources as a pool that must be shared by short-term storage and processing, the “resource-sharing model”, as Hitch, Tows, and Hutton (2001) have coined it. This resource-sharing hypothesis was put forward by Daneman and Carpenter (1980) and has been dominant in the WM literature. Although other points of view have emerged (Engle, Kane, & Tuholski, 1999; Hasher & Zacks, 1988; MacDonald & Christiansen, 2002; Maehara & Saito, 2007; Towse, Hitch, & Hutton, 1998, 2000; Waters & Caplan, 1996; for a review, see Miyake, 2001) that do not invoke the idea of resource sharing, they “are still based on a core assumption, which is also at the root of the resource sharing hypothesis, that memory items are ‘actively’ maintained during performance on working memory span tests” (Saito, 2006, p. 54).

Ericsson and Kintsch’s (1995) point of view is different. They agree that WM is the activated part of LTM and that there is a limit in terms of resources, but they propose that this limit can be overcome thanks to elements in LTM, such as schemas and retrieval structures, in the case of experts. The consequence of this tenet is a redefinition of the differences
between high-span and low-span individuals. High-span individuals perform well because they are able to use LTM, whereas low-span individuals are less able to use LTM; from Ericsson and Kintsch’s point of view, the key to the utilisation of LTM is knowledge. In the case of the reading span, it is linguistic expertise that matters; thus, high-reading-span people perform better because they are more expert linguistically. For example, Daneman and Green (1986) showed that high-reading-span subjects have a larger lexicon than low-reading-span subjects, and Pearlmutter and MacDonald (1995) showed that only high-reading-span individuals used their knowledge of language to detect linguistic cues that could guide their reading of ambiguous sentences.

Ericsson and Kintsch’s position (1995, p. 229) can be summarised as follows: “our proposal for LT-WM emphasizes good readers’ use of more sophisticated, more complex comprehension strategies—procedures for the construction of mental representations—that result in the generation of more extensive retrieval structures and hence a larger effective working memory”.

The second part of our research addresses this idea. In order to do this, we added a classic reading span measure that gave the reading span size of every participant. This measure was used as a continuous variable to predict the result of the personalisable reading span scores. Adding this variable to the personalisation variable resulted in two independent variables: the personalisation used as an ordinal variable and the reading span size used as a continuous variable. These two independent variables were used in a multiple regression analysis to predict the personalisable reading span scores.

We made two hypotheses. The first one was presented earlier and concerned only the contrast between “personalised” and “nonpersonalised”; we expected that participants in the personalised group would have a larger reading span than participants in the nonpersonalised group, since personalisation allows participants to use their LT-WM whereas the nonpersonalised group would need to count heavily on WM.

The second hypothesis is based on the premise that high-reading-span participants have a larger reading span because of the use of LT-WM, allowing storage in LTM, whereas participants with a lower reading span can only count on WM. If this premise is true, we hypothesised that the personalisation of the reading span would have less effect on the performance of participants with a higher reading span, because they already benefit from LT-WM, compared to the effect on participants with a lower reading span, who use mainly their WM. It seemed to us that, in the case of participants with a lower reading span, the change of status from extensive reliance on WM to the use of both WM and LTM (LT-WM) should boost their performance more than individuals with a higher span, whose status would not change, since they are already supposed to use both
WM and LTM (LT-WM) extensively. Therefore, the personalisation of the reading span should allow subjects with a lower reading span to increase their span size more than individuals with a higher reading span. An interaction between personalisation and reading span size is therefore expected.

Two experiments were conducted in order to test these two hypotheses. In Experiment 1, the personalisation method was applied to a reading span task by inserting a personalisable location into each sentence. In Experiment 2, the personalisation was slightly different, as neutral locations were inserted into reading span sentences. The personalisation was not directly applied to the reading span task, but we induced half of the participants to mentally personalise the neutral locations by thinking about known locations; the other half were encouraged to think about unknown locations in Armenia.

**EXPERIMENT 1**

In this experiment, a first classic reading span task was used to measure the participants’ reading span size. Then the personalisation method was applied to a second reading span task by inserting locations in each sentence. These locations could either be personalised (meaning that the reader was familiar with them) or not, creating a personalised reading span or a nonpersonalised reading span. The purpose was to show (1) that the personalisation globally enhanced the reading span performance and (2) that the reading span performance of participants with a lower reading span size was enhanced more than that of participants with a higher reading span size.

**Method**

**Participants**

Forty undergraduate students at Paris Descartes University participated in this experiment. All were native French speakers (mean age 21.9 years, SD = 3.11).

**Experimental design**

Two independent variables (an ordinal and a continuous variable) were used in a multiple linear regression analysis in order to predict a dependent variable: the personalisable reading span scores. Personalisation, the ordinal independent variable, was a two between-subjects factor: One half of the participants completed a personalisable reading span task in a personalised form and the other half completed the personalisable reading span task in a nonpersonalised form. Reading span size, the continuous independent
variable, was the reading span size of every participant measured with a classic reading span test.

**Material**

Two types of reading span test were employed: a classic reading span and a personalisable reading span. Each reading span test included five sequences corresponding to five levels of sentences: a two-sentences level, a three-sentences level, a four-sentences level, a five-sentences level, and a six-sentences level. The first one included two sentences with three trials (total of six sentences), the second one included three sentences with three trials (total of nine sentences), and so on, until the last one, which included six sentences with three trials (total of 18 sentences). In total, 60 sentences were employed for each reading span test. Each sequence included three trials or attempts; to move on to the following sequence, a subject had to make two successful attempts.

Classic reading span task. The classic reading span task was constructed using the French adaptation (Desmette et al., 1995) of the Daneman and Carpenter (1980) reading span test. Some sentences from the classic version were modified in order to make the classic version and the personalisable version as similar as possible. The two versions were similar concerning the mean frequency of the last word—21.28 (mean) and 0.42 (standard deviation) for the personalisable reading span, and 21.28 and 0.32 for the classic version—and the number of syllables of the last word—1.82 (mean) and 0.07 (standard deviation) for the personalisable version, and 1.73 and 0.11 for the classic version. The two versions were not completely similar for the number of syllables in the sentences (30.88 and 0.18 for the personalisable version, and 21.74 and 0.79 for the classic version).

Personalisable reading span task. The sentences for the personalisable reading span test were adapted from Guida and Tardieu’s (2005) material. The sentences employed in the personalised version and in the nonpersonalised version were the same except for the names of the locations and the characters. These are familiar in the case of the personalised version (known locations in Paris) and unfamiliar in the case of the nonpersonalised version (unknown locations in Armenia). For example, a sentence (translated from French) in the nonpersonalised version might be, “For the second time this month, there was an accident in Takor Street involving a car and a scooter”.
which would become “For the second time this month, there was an accident in Danjou Street\textsuperscript{4} involving a car and a scooter” in the personalised version (see also Appendix 1).

We made sure in the personalisable reading span that the sentences for each trial did not have any kind of thematic or semantic link. The last word of each sentence was always a manufactured object. Its frequency and number of syllables were controlled so that they would be equivalent for each sequence.

**Pre-experimental questionnaire.** In order to find out exactly what locations were familiar to the participants in the personalised group, a questionnaire was used. For each participant, familiar locations were collected. This information was then included in the personalised version. For example, in order to personalise the following sentence (translated from French) “In a post office in Rovan, a young man wearing a Halloween costume was proud to show off his beautiful cape”, we asked the participants in the personalised group where they lived and replaced Rovan with the name of their city.

In order to equalise the personalised condition and the nonpersonalised condition, the same type of information was requested in the nonpersonalised condition; for example, instead of asking participants where they lived, we asked them if they knew the city of Rovan.

**Procedure**

The experiment took place in three stages.

1. Participants took the classic reading span task.
2. Participants answered the pre-experimental questionnaire.
3. Half of the participants took the personalised version of the personalisable reading span task and the remaining half took the nonpersonalised version of the personalisable reading span task.

The procedure for both the personalisable reading span task and the classic reading span task was as follows. Every participant sat in front of a computer screen and the instructions for the reading span task were given to him or her on the computer screen. Participants were asked to read aloud the sentences displayed at their own pace and to memorise the last word of each sentence. Sentences appeared one by one, in sets of two, three, four, five, or six, depending on the level at which they were displayed. The experiment started with a set of two sentences. When a participant pronounced the last

\textsuperscript{4}Danjou Street is a street near Paris 5 University that is well known to Paris 5 students.
word of a sentence, the following sentence appeared on the screen. At the end of the last sentence of each set, rappel (“recall” in French) was displayed on the screen. Participants were then supposed to recall the last word of each sentence in the set.

Participants were warned to expect the number of sentences per set to increase during the course of the test. At each level, three trials could be displayed. To move on to a more difficult set (first level: two sentences; second level: three sentences; third level: four sentences; fourth level: five sentences; and fifth level: six sentences), participants had to correctly recall the last words of two trials (if a participant correctly recalled the first two trials, the third was not displayed). If the participant did not succeed, the experiment was stopped.

The last level at which a participant performed correctly on two out of three trials was taken as a measure of the reading span, which corresponded to the set preceding the last set. Half a point was added if the participant was correct on one out of three trials for the last set. For example, if a participant performed correctly on two out of three four-sentence sets and none of the three five-sentence sets, the reading span assigned was 4, while if a participant performed correctly on two out of three four-sentence sets and one out of three five-sentence sets, the reading span assigned was 4.5.

Results

Results of the personalisable reading span task. Participants taking the nonpersonalised version of the personalisable reading span task had 2.95 (SD = 1.08) as average score, whereas participants taking the personalised version had 3.55 (SD = 0.9) as average score. The difference was significant, Student’s $t$: $t(38) = 2.07$, $p < .04$.

Results of the classic reading span task. The mean performance on the classic reading span task was 3.3 (SD = 0.88). Concerning this variable, no difference was observed between the personalised group and the nonpersonalised group, Student’s $t$: $t(38) = 1.18$, $p > .1$. This result seems to suggest that the personalised group and the nonpersonalised group were comparable in terms of working memory capacity as it can be measured by a reading span task.

Hierarchical regression analyses. A hierarchical regression analysis was conducted using personalisation, reading span size, and their interaction as predictors of the personalisable reading span score. Personalisation was an ordinal variable, we coded it using dummy variable coding (see for the choice of coding systems, Aiken & West, 1991). This regression analysis was used to test the effect of each predictor on the Personalisable reading span score.
when the others predictors were controlled (Cohen, Cohen, West, & Aiken, 2003). The hierarchical analysis was conducted in two steps: personalisation and reading span size were entered in the first step to evaluate the main effect of each predictor variable on the personalisable reading span score, and the cross product was entered in the second step (for a similar analysis procedure, see Hambrick & Engle, 2002; Hambrick & Oswald, 2005). We reported for each step the increment in variance accounted for (inc. $R^2$) by the variables entered in that step. We also presented each squared semipartial correlation ($sr^2$) for each variable. The results of the hierarchical regression are presented in Table 1.

Personalisation and reading span size taken together accounted for 27%, $F(2, 37) = 6.85, p < .005$, of the variance on the personalisable reading span score. Concerning the proportion of variance in the personalisable reading span score uniquely accounted for by personalisation ($sr^2 = .244$) or uniquely accounted for by reading span size ($sr^2 = .405$), only reading span size was significant ($p < .05$), the effect of personalisation on the personalisable reading span score was marginal. These results indicated that (1) a larger reading span size was associated with higher level of personalisable reading span scores and that (2) personalisation (as opposed to nonpersonalisation) was associated with higher personalisable reading span scores, but only marginally in terms of inferential statistics.

The second step of the analysis indicated that the Personalisation × Reading span size interaction accounted for an additional 5.4% but only marginally, $F(1, 36) = 2.87, p < .1$.

As shown in Figure 1, the pattern suggested that the effect of reading span size on the personalisable reading span performance was more important for the nonpersonalised group than for the personalised group. Or taken in the other way, the pattern suggested that the effect of personalisation on the personalisable reading span performance was more important for the participants with a lower reading span size than

<p>| TABLE 1 | Hierarchial regression analysis predicting personalisable reading span score |
|-----------------|-----------------|------------|------------|------------|</p>
<table>
<thead>
<tr>
<th>Predictor variable</th>
<th>Inc. $R^2$</th>
<th>$F$ change</th>
<th>$B$</th>
<th>$t$ value</th>
<th>$sr^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>.270</td>
<td>6.847**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PERS</td>
<td>.515</td>
<td>1.738 ≈*</td>
<td>-</td>
<td></td>
<td>.244</td>
</tr>
<tr>
<td>RSS</td>
<td>.494</td>
<td>2.881*</td>
<td>-</td>
<td></td>
<td>.405</td>
</tr>
<tr>
<td>Step 2</td>
<td>.054</td>
<td>2.868 ≈*</td>
<td>-</td>
<td></td>
<td>- .232</td>
</tr>
<tr>
<td>PERS × RSS</td>
<td>- .573</td>
<td>-1.694 ≈*</td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Inc. $R^2$ = increment in variance accounted for; $B$ = unstandardised regression coefficient; $sr^2$ = squared semipartial correlation; PERS = personalisation (nonpersonalised or personalised); RSS = reading span size measured by a classic reading span task. ≈*p < .10, *p < .05, **p < .005.
participants with a higher reading span size. The pattern was the one expected but the effect was only marginal.

Even if the interaction was only marginal, we analysed the slope of the two regression lines in order to know if the slopes were significantly different from horizontal. Concerning the regression line of the personalised group, the line is not significantly different from zero, Student’s $t$: $t(18) = 0.521$, $p > .1$, whereas the regression line of the nonpersonalised group is significantly different from zero, Student’s $t$: $t(18) = 4.79$, $p < .0001$.

**Discussion**

Our results can be summarised under three points: (1) An effect of personalisation as it can be measured by a $t$-test, (2) a marginal interaction between personalisation and reading span size, and (3) an effect of the reading span size for participants in the nonpersonalised group versus a lack of effect of the reading span size for participants in the personalised group.

Our first result confirms what had been shown by Guida and Tardieu (2005), who found with an analysis of variance that participants reading a personalised version of a text recalled more objects than subjects who received the nonpersonalised version. In this experiment, the personalised group also achieved superior performance but with a WM task. Therefore,
the difference between the two groups (personalised vs. nonpersonalised) can now be interpreted in terms of a LT-WM versus WM contrast.

This result was predictable within LT-WM theory. In fact, when individuals are in a situation that allows them to use their expertise, they should be able to use their LT-WM and therefore to benefit from an extension of WM. This is the case for the participants in the personalised group because we have fulfilled Kintsch et al.’s (1999) definition of expertise. Therefore, we can analyse the difference between the two groups as resulting from LT-WM’s allowing an apparent extension of WM in the personalised group. We think that the retrieval structures involved are “schemas of locations”. In fact, each sentence in the reading span task included a location and an object. In the personalised version, the locations were very familiar to participants. Thus, when the participants read a sentence with the aim of memorising the last word, they were able to use their episodic and autobiographic knowledge of the location. This facilitated the association between the location and the word. During the recall, participants were able to use the location as a cue to find the final word of each sentence; the richer the cue, the easier the task is for participants.

However, before advancing further in the interpretation of the results, we believe that this first experiment raises two issues. The first is about the interpretation of the superiority of the personalised group and the second one is related to the statistical marginality of our results.

Concerning the first point, and as suggested by a reviewer, an alternative interpretation of the superiority of the personalised group can be proposed. It is possible that the familiar locations in the case of the personalised group were read more easily and hence processed faster, leaving more resources available for storage of the final words. This could explain why the personalised group stored more objects and therefore recalled more objects. Experiment 2 was designed to evaluate this interpretation.

The second issue concerns the marginal effect we observed analysing the interaction between personalisation and reading span size and the apparent inconsistency concerning the effect of personalisation. As reported above the effect of personalisation is significant with a t-test but only marginally significant in the regression analysis. This is due to the fact that in the regression analysis we carried out, the result was obtained controlling the effect of the reading span size. That means that part of the effect in the t-test is due to the reading span size, and when this is partialled out the effect is only marginal.

We think that the two marginal effects could be due to the small sample size ($N = 40$). This can be problematic in two ways (Howell, 2002): the reliability of the measure and its statistical power. Concerning the first point, Harris (1985) suggested that the number of participants should exceed the number of predictors by 50 and Howell reported 40 as a more liberal
suggestion. In our experiment, the number of participants only exceeded the number of predictors by 37. Concerning the second point, the statistical power of tests in multiple regression analysis increases as the sample increases (e.g., Cohen et al., 2003).

Therefore, in accordance with these two last points, Experiment 2, which was first designed to evaluate an alternative explanation of the superiority of the personalised group, was conducted with a 50% increase in the number of participants.

**EXPERIMENT 2**

The major purpose of this second experiment was (1) to apply the personalisation method while giving the nonpersonalised group and the personalised group the exact same reading span material and (2) to increase the number of participants in the second experiment.

The rationale concerning the first point was that, if the two groups had exactly the same personalisable reading span sentences with exactly the same names of locations, then the possibility that the familiar locations in the personalised version were read more easily and faster, leaving more resources for storage, would no longer be justified. Therefore, instead of directly personalising the material, we provided neutral material for both groups; this task was therefore called the “reading span with neutral locations”. That means that, for example, the sentence “In the library of Paris 5 University, he found a really old manuscript ...”, taken from the material of the personalised group in Experiment 1, and the equivalent sentence taken from the nonpersonalised group (“In the library of Erevan University, he found a really old manuscript ...”) were both replaced by the following neutral sentence: “In the library of the University, he found a really old manuscript ...” (all the sentences are translated from French).

Since the reading span sentences were neutral, we had to personalise the task in another way. We therefore encouraged half of the participants (the personalised group) to think about familiar locations when reading about the neutral locations and we encouraged the other half (the nonpersonalised group) to think about Armenian locations when reading about the neutral locations. To do this, we had them first read a personalisable text. For the nonpersonalised group, the text story concerned locations in Armenia; for the personalised group, the text story contained known locations. The stories gave our participants a mental frame that was intended to continue into “the reading span with neutral locations” task—an instance of “inertia”, as defined in physics.

Therefore, our expectations of this second experiment were as follows. If the superiority of the personalised group in the first experiment was really
due to the fact that the familiar locations were processed faster, we should not find an effect of personalisation in this second experiment with the reading span with neutral locations because all the sentences from the span test were the same and the personalisation occurred before “the reading span with neutral locations”. But if the effects found in Experiment 1 were due to a long-term working memory effect then we should replicate all the results of Experiment 1 in a more controlled fashion.

Method

Participants

Sixty undergraduate students at Paris 5 René Descartes University participated in this experiment. All were native French speakers (mean age 22.7 years, $SD = 3.19$).

Experimental design

The experimental design was the same as in Experiment 1. Two independent variables—personalisation as ordinal and reading span size as continuous variable—were used in a multiple linear regression analysis in order to predict a dependent variable: the personalisable reading span scores. The only difference was the adjunction of a secondary dependent variable: the percentage of correct responses at a text comprehension task that followed the personalisable text.

Material

Like in Experiment 1 a classic reading span was used and unlike Experiment 1 a personalisable text and a “reading span with neutral locations” were displayed. Both reading span tests included five sequences corresponding to five levels of sentences (cf. Experiment 1).

Classic reading span task. See Experiment 1.

Personalisable text. The text was composed of 21 sentences. Fifteen locations were presented in the text. They were chosen in such a way that one of these 15 locations was included in each trial of each sequence of the reading span with neutral locations. In the nonpersonalised version, the locations were in Armenia; in the personalised version the locations were known to the subjects (see Appendix 2). In the latter case, we personalised the text using the information from the pre-experimental questionnaire. The character’s name changed as a function of the participant’s group and sex. If the
participant was in the personalised group, the character had a French name, whereas if he or she was in the nonpersonalised group, the name was Armenian. The sex of the character in the text was the same as that of the participant reading the text.

*Text comprehension task.* To find out whether the participants were paying attention to the text, they were asked four factual questions that did not concern the locations mentioned in the text.

*Reading span with neutral locations.* The sentences of the reading span with neutral locations were adapted from the personalisable reading span of Experiment 1. We neutralised each sentence of the personalisable reading span by erasing the name of each location. For example, the sentence (translated from French) “For the second time this month, there was an accident in *Takor Street* involving a car and a scooter” in the nonpersonalised version and the sentence “For the second time this month, there was an accident in *Danjou Street* involving a car and a scooter” in the personalised version became “For the second time this month, there was an accident in a street near the university involving a scooter” in the reading span with neutral locations. The need to introduce a neutral location led us to add some details like “a street near the university” instead of *Danjou Street* or *Takor Street*, but to avoid having overly long sentences, we eliminated some details from the sentences from Experiment 1 (in this example, we cut “a car and”). The average number of syllables in the sentences was 25.65 ($SD = 0.21$), whereas in Experiment 1 the average number of syllables was 30.88 ($SD = 0.18$).

*Pre-experimental questionnaire for the personalisable text.* To find out exactly which locations were familiar to the participants in the personalised group and in order to insert those locations into the personalisable text, a questionnaire was used. Familiar locations were collected for each participant. This information was then included in the personalised version of the personalisable text. For example, in order to personalise the following sentence (translated from French) “To go to the University, he had to leave from [name of the station of departure] Station”, we asked the participants in the personalised group which subway station they left from in the morning to go the university and introduced it into the sentence. In order to equalise the personalised and nonpersonalised conditions, the same type of information was requested in the nonpersonalised condition; for example, instead of asking participants where they lived, we asked them whether they had heard of Narëk subway station.
**Procedure**

The experiment took place in four stages.

1. Participants took the classic reading span task.
2. Participants answered the pre-experimental questionnaire.
3. Half of the participants read the personalised version of the personalisable text and the remaining participants read the nonpersonalised version of the personalisable text. The personalisable text was displayed with a RSVP (Rapid Serial Visual Presentation, see Forster, 1970) presentation. Participants saw the text segment by segment on the screen at the rate of 40 ms per character. Each segment corresponded to a unit of meaning. This presentation speed allowed for fast but comfortable reading, with no possibility of rereading the text. This type of presentation was used to display texts in Guida and Tardieu’s (2005) study.
4. All the participants took the reading span task with neutral locations.

The procedure for the classic reading span task and the reading span with neutral locations was the same as in Experiment 1.

**Results**

*Results of the personalisable text.* Statistical analysis showed that the personalised group and the nonpersonalised group had similar results: 98.3% ($SD = 6.34$) of correct answers for the personalised group, and 96.7% ($SD = 10.9$) for the nonpersonalised group. The difference is not significant, Student’s $t$: $t(58) = 0.73$, $p > .1$. These percentages indicate that the participants really read the text with the intent to understand it.

*Results of the reading span with neutral locations.* Participants who read the nonpersonalised version of the personalisable text scored 3.27 ($SD = 0.99$) at the reading span with neutral locations, while participants who read the personalised text scored 3.93 ($SD = 0.92$) at the reading span with neutral locations. This difference was significant, Student’s $t$: $t(58) = 2.84$, $p < .004$.

*Results of the classic reading span task.* In the classic reading span task, the average score was 3.4 ($SD = 0.87$). No difference was observed between the personalised group and the nonpersonalised group in the reading span with neutral locations, Student’s $t$: $t(58) = 1.41$, $p > .1$. This result seems to suggest that the personalised group and the nonpersonalised group were comparable in terms of working memory capacity as it can be measured by a reading span task.
Hierarchical regression analyses. As in Experiment 1, the same hierarchical regression analysis was conducted using Personalisation, Reading span size, and their interaction as predictors of the reading span score with neutral locations (see Experiment 1). The results of the hierarchical regression are presented in Table 2.

Personalisation and reading span size accounted together for 25.5%, $F(2, 57) = 9.76, p < .0005$, of the variance on the reading span score with neutral locations. The proportion of variance in the reading span score with neutral locations that was uniquely accounted for by personalisation and reading span size were $sr^2 = .296$, and $.348$, respectively. Both effects of personalisation ($p < .01$) and reading span size ($p < .005$) on the reading span score with neutral locations were significant. These results indicated (1) that a larger reading span size was associated with higher level of reading span scores with neutral locations and (2) that personalisation (as opposed to nonpersonalisation) was also associated with higher reading span scores with neutral locations.

The second step of the analysis showed that the Personalisation × Reading span size interaction accounted for an additional 5.3%, $F(1, 56) = 4.29, p < .05$.

As shown in Figure 2, the pattern suggested that the effect of reading span size on the reading span with neutral locations performance was greater for the nonpersonalised group than for the personalised group. Or similarly, the personalisation on the reading span performance with neutral locations was more important for the participants with a lower reading span size than participants with a higher reading span size.

As in Experiment 1, we analysed the slope of the regression lines in order to know if the slopes were significantly different from zero (the horizontal). Concerning the regression line of the personalised group, the line is not significantly different from zero, Student’s $t$: $t(28) = 0.59, p > .1$, and

<table>
<thead>
<tr>
<th>Predictor variable</th>
<th>Inc. $R^2$</th>
<th>$F$ change</th>
<th>B</th>
<th>$t$ value</th>
<th>$sr^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>.270</td>
<td>9.757****</td>
<td>.618</td>
<td>2.591**</td>
<td>.296</td>
</tr>
<tr>
<td>PERS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RSS</td>
<td>.418</td>
<td>3.045***</td>
<td></td>
<td></td>
<td>.348</td>
</tr>
<tr>
<td>Step 2</td>
<td>.053</td>
<td>4.291*</td>
<td>-.555</td>
<td>-2.072*</td>
<td>-.230</td>
</tr>
<tr>
<td>PERS × RSS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Inc. $R^2$ = increment in variance accounted for; $B$ = unstandardised regression coefficient; $sr^2$ = squared semipartial correlation; PERS = personalisation (nonpersonalised or personalised); RSS = reading span size measured by a classic reading span task. *$p < .05$, **$p < .01$, ***$p < .005$, ****$p < .0005$. 

TABLE 2 Hierarchical regression analysis predicting personalisable reading span score
concerning the nonpersonalised group, the regression line of is significantly different from zero, Student’s \( t \): \( t(28) = 3.96, p < .0005 \).

Discussion

The results of Experiment 2 replicated the results of the first experiment in a more convincing way, they can be summarised under three points: (1) an effect of personalisation even partialling out the effect of the reading span size, (2) an interaction between personalisation and reading span size, and (3) an effect of the reading span size for participants in the nonpersonalised group versus a lack of effect of the reading span size for participants in the personalised group.

The aim of this second experiment was to test an alternative interpretation of the superiority of the personalised group in the first experiment. Our first interpretation was that the subjects from the personalised group were better able to use their LT-WM than the subjects from nonpersonalised group; the alternative one proposed that the familiar locations used for the personalised group were easier to read and therefore were processed faster, which left more resources available for storage. In order to rule out that interpretation, we presented exactly the same reading span to the persona-
lised and nonpersonalised groups without presenting any kind of familiar location. The rationale was that since the locations were exactly the same in the two versions, then any advantage detected could no longer be imputed to a difference in processing speed for familiar words. The fact that an effect of personalisation was still observed therefore seems to favour the first explanation.

As in the first experiment, we think that the difference between the personalised and nonpersonalised groups can be explained in terms of “retrieval structures” and “schemas of locations”. For both groups, we explicitly stated in the instructions for the reading span that the neutral locations were related to the locations mentioned in the personalisable text. Thus, the personalised group knew that the neutral locations in the reading span test related to familiar locations. Consequently, when they read a sentence in the reading span with the aim of memorising the last word, they were able to mentally interpret the neutral location in terms of a known location and therefore use their episodic and autobiographic knowledge related to that known location. This must have facilitated the association between the location (the schema of the location) and the last word to be recalled. We think that, during the recall, each schema of location could have been used as a cue to find the relevant final word. And the fact that the cues in the personalised version were richer made it easier to recall the final words, resulting in a higher percentage of recalled words. In the case of the nonpersonalised participants, the neutral locations in the reading span test could not be interpreted in a way that allowed participants to efficiently use their episodic and autobiographic knowledge.

In our interpretation, the locations are integrative schemas, and the mechanism involved is a retrieval structure. In a way, retrieval structures can be compared to integrative schemas, scripts or frames (Bartlett, 1932; Minsky, 1975; Rumelhart, 1975; Schank & Abelson, 1977), since they allow incoming information to be reorganised according to familiar patterns. However, LT-WM needs to be distinguished from these latter constructs, among other reasons because its purpose is to describe how, in the case of expertise, individuals can use overpractised schemas to encode and rapidly retrieve information in LTM. This may occur during WM tasks, giving the impression that WM has been expanded.

Even if we believe that the above interpretation is genuine and accounts for the difference between the personalised and nonpersonalised group, we would like to add a caution. It is not implausible that a part of this difference may be due to the fact that participants in the nonpersonalised group had to imagine unfamiliar locations during the reading span with neutral locations, which could have placed significant demands on processing. And these significant demands could have had a deleterious effect on
performance explaining part of the differences. In the General Discussion, we will propose a future experiment to rule out this possibility.

Even if the results of Experiment 2 replicate those of Experiment 1, it is also worth noting that globally the reading span with neutral locations led to higher scores than the personalisable reading span in the first experiment (3.6 vs. 3.25). We think that two factors may explain this difference. The first and most important is the length of the sentences. The reading span sentences in Experiment 1 were longer than those in Experiment 2; the average number of syllables per sentence in Experiment 2 was 25.65 versus 30.88 in Experiment 1. But we also think that the text that was read before the reading span with neutral locations may have helped participants with the subsequent reading span. The personalisable text contained locations that were subsequently included in the reading span; this could have benefited participants by activating those locations and perhaps also by implicitly preactivating some locations that were not mentioned in the text but were linked—for example, topographically—to the locations of the text.

Since the pattern of results from the first experiment is comparable to that in Experiment 2 for the remaining data, we will discuss it in the General Discussion. We will simply add a final comment concerning the fact that in the second experiment the personalisation was done indirectly by personalising a previous text. This makes it comparable to experiments that have shown that differences in WM span could be observed using strategies, instructions or training. McNamara and Scott (2001), for instance, have shown that reading span performance can be increased if participants are trained to use a story-formation strategy by chaining the words to be recalled. Turley-Ames and Whitfield (2003) have shown that a rehearsal strategy with an operation span measure could increase the span measure. And, more recently, Ridgeway (2006) improved participants’ performance on a spatial span memory test by encouraging them to use a chunking method. But although our experiments can be linked to these results and considered as confirming them, we would like to stress that we did not ask our participants to use a strategy, but simply informed them that the neutral locations were related “to a known locations like those of the text” (see Appendix 3) for the personalised group and “to an Armenian locations like those of the text” (see Appendix 3) in the nonpersonalised group. The aim was to establish a mental frame and continuity between the personalised text and the reading span with neutral locations.

**GENERAL DISCUSSION**

The purpose of these experiments was to test the predictions of the LT-WM theory using the personalisation method (Guida & Tardieu, 2005).
Experiment 1 examined the possibility that WM can be expanded with expertise by using part of LTM. To do this, we used the personalisation method, inserting familiar locations into each sentence of a reading span task for a personalised group and unknown locations into each sentence for a nonpersonalised group. Participants in the personalised group were expected to use the known locations as schemas to associate them to the final word (an object) of each reading span sentence. During the recall phase, participants in the personalised group could retrieve the final words directly from WM (meaning “active maintenance”) or they could use their LTM via the schemas of known locations to retrieve the words that had previously been associated with those schemas. The first experiment led to results in favour of the LT-WM theory, showing an expansion of reading span for the personalised group, but they also raised the possibility of an alternative explanation. It was possible that subjects in the personalised group read the familiar locations inserted in each sentence more easily and therefore processed them faster than the unfamiliar locations used in the nonpersonalised version. If this was the case, it would have left more resources available for storage.

Experiment 2 was designed to test this possibility and increase the sample size. Instead of inserting names of locations in a reading span, we inserted neutral locations, giving the two groups of participants identical material. We used a previously presented personalised text to encourage one group (the personalised group) to mentally personalise the neutral location by thinking about familiar locations. Similarly, we used a nonpersonalised text to encourage the other group (the nonpersonalised group) to mentally associate the neutral locations with unknown locations in Armenia. The results replicated those of Experiment 1 in a more convincing way, showing a significant difference between the personalised group and the nonpersonalised group. Since the difference could no longer be imputed to the speed of processing the familiar locations in the reading span, we attributed it to a LT-WM effect. However as mentioned in the previous discussion, an alternative interpretation could account in part for the difference between the personalised and nonpersonalised group. It is not impossible that since participants in the nonpersonalised group had to imagine unfamiliar locations during the reading span with neutral locations, it could have for consequence to put high demands on processing compared to the personalised group who had solely (this is an oversimplification in order to understand the problem) to activate their knowledge of known locations. These demands could account at least in part for the personalisation effect. One way to rule out this interpretation could be to conduct an experiment with three groups, a personalised group and a nonpersonalised group like in the second experiment plus a third group that (1) would not read a personalisable text before the reading span with neutral locations and (2)
would not be encouraged to imagine unfamiliar locations. We will certainly include this experiment in a following study.

Confirmation of our two hypotheses

The major results of these two experiments confirmed our two hypotheses: (1) an effect of personalisation, suggesting that personalisation and therefore LT-WM (as it has been operationalised by us and with the caution we have just added) enhanced reading span scores; (2) an interaction between personalisation and reading span size, with a greater effect of personalisation for participants with a lower reading span than for participants with a larger reading span.

Our first result confirms, to some extent, those of Fincher-Kiefer, Post, Greene, and Voss (1988). These authors adapted Daneman and Carpenter’s (1980) reading span. They varied the sentences content by introducing—or not introducing—elements that concerned baseball. In a first experiment with a standard reading span task, Fincher-Kiefer et al. did not find any difference between baseball novices and experts (in both sentences that concerned baseball and those that did not). In their second experiment, the authors asked their participants to recall not just the last word of each sentence (as in a standard reading span test) but the entire sentence. In this case, they found a difference between the two groups of participants. They interpreted the effect as indicating the need for participants to reorganise the information because of its complexity and scope, as compared to the necessity to recall only the last words. In this situation, the baseball experts’ knowledge helped them to reorganise the incoming information and then to use it as a cue for recall.

The difference between the absence of effect in the first experiment by Fincher-Kieffer et al. (1988) and the effect we obtained by personalising the reading span task could be due to the fact that our sentences have a structure that is supposed to help participants to use their knowledge, even though we did not ask them to recall the sentences in full. We think that the structure of each sentence (one location + one object) encouraged the participants in the personalised group to use the locations as cues and to associate them with the objects. The consequence is that participants reorganise the sentences using their knowledge to restructure the incoming information.

Our second major result showed that participants with a higher reading span benefited less from personalisation than participants with a lower reading span. Our general interpretation of this finding can be considered a consequence of Ericsson and Kintsch’s (1995) idea, which we presented in the introduction. In their view, what differentiates participants with high reading spans from low-span participants is not short-term storage capacity...
but the ability to encode better in LTM by using LT-WM. Based on this premise, we think that it is possible to explain why individuals with a higher reading span benefited less from personalisation, because these kinds of individuals already use their LT-WM in the nonpersonalised version. On the other hand, personalisation is crucial for participants with a lower span, since without it they are unable, or only partially able, to use their LT-WM to encode information. That is why personalisation was more effective for participants with a lower reading span than for participants with a higher reading span in our experiment.

This result is very important because it gives us some indication of the impact of WM and the use of domain knowledge on memory performance. In fact, the experimental situation caused by personalisation can be considered similar to a situation of domain knowledge. In our experiment, the personalised group can be considered a knowledgeable group (they are able to use their knowledge of locations in the texts); the other group is less knowledgeable since the locations are in Armenia.

**Types of interactions between working memory capacity and domain knowledge**

Several studies have examined the impact of working span size (or capacity) and knowledge on the retention/comprehension of texts or passages, with a view to determining the best predictor (Britton, Stimson, Stennett, & Gülgöz, 1998; Haenggi & Perfetti, 1992, 1994; McNamara, 2004), but very few have tested the interplay between domain knowledge and WM size and its effect on memory performance. To the best of our knowledge, only Hambrick and colleagues (Hambrick & Engle, 2002; Hambrick & Oswald, 2005) have presented results concerning this interaction. This is a shame because it is an important question for cognition in general, since investigating how WM capacity constrains the use of knowledge or how the use of knowledge changes the impact of WM size is primordial. But it is also a crucial issue for LT-WM theory, because how WM and LTM interact via LT-WM is the core of Ericsson and Kintsch’s theory. Hambrick and Engle (2002) presented three possible models of interaction. The first model, which can be called the “compensation model”, puts forward the idea that “high levels of domain knowledge can compensate for low levels of working memory capacity” (p. 342); thus, low-span individuals can benefit more from high domain knowledge, catching up with the performance of high-span individuals. The second model is additive: The utilisation of domain knowledge is assumed to be independent of WM capacity, with the two factors allowing subjects to improve their performance but with no interaction. The last model was termed the “rich-get-richer” model, meaning
that the bigger WM is, the more able one is to use domain knowledge. Before examining Hambrick and colleagues’ results, let us try to analyse this interaction in the LT-WM framework. What type of interaction should one expect?

Ericsson and Kintsch (1995) did not write specifically about this kind of interaction. However, we know that in their theoretical framework, items in WM are the cues that allow people to retrieve information from LTM. Based on this premise, one can assume that the larger WM is, the higher the number of items in WM will be, allowing one to activate a larger portion of LTM. In other words this should increase the possibility of accessing elements in LTM and therefore making use of domain knowledge, corresponding to the “rich-get-richer” model.

However, this rationale does not take into account the structure of LTM, which is a key part of LT-WM theory. Indeed, it seems to us that the decisive element in Ericsson and Kintsch’s (1995) theory is not the number of elements in WM but the number of elements in LTM linked to the elements in WM, which depends crucially on knowledge and expertise. For instance, in the case of an expert, a single item in WM could allow via a retrieval structure to retrieve a whole block of knowledge. The prediction in this case would favour the additive model, with no interaction between WM size and the use of domain knowledge.

Based on these considerations, the additive model or the “rich-get-richer” model should be predicted. And indeed, Hambrick and colleagues’ results tend to show that WM capacity and domain knowledge (knowledge of baseball) operate independently on memory performance, confirming the additive model. They performed regression analyses using baseball experts and novices, using WM span as a variable and varying the content of the task by measuring memory performance, which could either concern baseball or not. Hambrick and Engle (2002) used a counting span and an operation span to measure WM capacity and a LTM task to measure memory performance, asking the participants to answer questions about a baseball radio broadcast. Hambrick and Oswald (2005) used four tasks (a counting span, an operation span, a matrix span, and a rotation span) to measure WM capacity and asked participants to track and recall the movements of players on a baseball field. In neither case was an interaction between WM capacity and domain knowledge observed.

5 The relation between domain knowledge and WM size could be more complicated; for instance, it could vary. For a low level of knowledge, the relation between domain knowledge and WM size could be of the “rich-get-richer” type, whereas for a high level of expertise, the size of WM might cease to influence the capacity to use domain knowledge, because the benefit for performance via the size of WM would become negligible.
Our view about working memory and domain knowledge interactions: The involvement of knowledge in working memory tasks

The question that stems from the previous results is, why did we not find the same lack of interaction in our experiment and, more importantly, why did we hypothesise that participants with a high reading span would benefit less from personalisation than low-reading-span participants? This hypothesis runs contrary to the absence of interaction and corresponds instead to the compensation model. We made this hypothesis because one crucial element has not yet been taken into consideration, namely the fact that WM measures do not correspond solely to “temporary storage” (i.e., intended as the focus of attention) but also to LTM storage. Several researchers share this view (for a recent review, see Guida, Tardieu, & Nicolas, in press). Of course, Ericsson and Kintsch (1995) are among them, but they are not the only ones. In the field of chess expertise, for example, Gobet and Simon (1996) have put forward the role of structures called “templates” (for a comparison of retrieval structures and templates, see Ericsson & Kintsch, 2000; Gobet, 2000a,b), with the idea that expert individuals can use their knowledge to store information in LTM at a pace that is similar to the pace for WM. This idea is also comparable to the proposal made by Cowan (1995, 1999) concerning virtual short-term memory, a component of LTM that can be used as short-term memory. And Saito and Ishii (2004) proposed a retrieval-based account of WM span performance, which assumes that individual differences in WM span are partly driven by individual differences in efficiency when retrieving items from long-term memory. More recently, Unsworth and Engle (2006a, 2007) have developed a dual-component view of performance in complex spans based on a primary memory, comparable to the focus of attention, and a secondary memory, which corresponds to LTM. In this framework, WM tasks and complex activities are processed using secondary memory to a great extent (for another similar proposal, see Woltz & Was, 2006).

Therefore, to really judge the interaction between WM and domain knowledge, one needs to take into account the fact that, when measuring WM, one does not obtain the number of elements that are in WM (intended as the focus of attention or short-term working memory in Ericsson and Kintsch’s, 1995, terminology). Instead, one obtains a mixture of two constructs: focus of attention and LTM. And the involvement of LTM (in other words, the percentage of the WM measure that depends on LTM) undoubtedly varies according to the type of task. The more knowledge can intervene in a WM task, the more LTM may become involved. For example, Masunaga and Horn (2000) showed that, if one uses a short-term memory (STM) task adapted to the GO game with GO players, then the STM span
can be increased, exactly as in the recall of chess positions by chess experts (Chase & Simon, 1973). And Cowan (2001) has proposed four ways of limiting the impact of LTM in a STM task.

But the involvement of LTM in WM also varies according to the type of participant. For example, one can assume that the more expertise participants have in a task, the more they can use their knowledge, for instance in terms of chunks (Chase & Simon, 1973; Cowan, 2005; Gobet & Simon, 1996). However, another dimension may also intervene: WM size. We have already presented Ericsson and Kintsch’s (1995) view that high-span individuals are able to use LTM storage more efficiently than low-span individuals; this could explain (at least partly) the difference between the two types of individuals. But new data from Cowan et al. (2003) can also be interpreted in a similar way. They examined children’s performance on sentence-based span tests (such as reading and listening span tests). A detailed response-time analysis was conducted for each word participants had to recall, in order to determine the length of the silence intervals between the words recalled. Their results were consistent “with the notion that the pauses were used at least partly for some sort of memory search process [in LTM], as Cowan et al. (1994, 1998) have suggested for STM tasks, and that this process was accomplished more efficiently in children with a higher span” (p. 120). This interpretation is also compatible with Unsworth and Engle’s (2006b) results. Their intention was to show that, to perform well on WM tasks, one needs to encode and retrieve information from secondary memory (LTM). Unsworth and Engle (2006b) found that individuals with a low reading span had a deficient searching process in LTM; they used temporal context cues badly, which increased proactive interference (for a link between the increase in proactive interference and the decrease in WM span performance, see also Lustig, May, & Hasher, 2001; May, Hasher, & Kane, 1999).

Therefore if one takes into account the fact that the involvement of LTM certainly varies according to (1) the type of task and (2) the type of participant, we think that another prediction can be made concerning the interaction between WM size and the use of knowledge. On the premise that the larger participants’ WM is, the more LTM in involved and therefore the more knowledge can intervene, we think that helping the participants to use more their knowledge via personalisation should have more impact on the performance of participants who do not use their knowledge or who naturally use their knowledge less: low-span individuals. That is why we proposed a hypothesis in favour of the compensation model.

Concerning the discrepancy between Hambrick and colleagues’ results and ours, we think that two reasons can be invoked. The first one concerns exclusively the way WM was measured. As mentioned earlier, Hambrick and Engle (2002) used a counting span and an operation span, and Hambrick
and Oswald (2005) used a counting span, an operation span, a matrix span, and a rotation span. In our experiments, we used exclusively a reading span. It is possible that the reading span involves more LTM knowledge than other WM tasks in all-comers participants. This could be because most individuals have more knowledge concerning the verbal content of reading span sentences than knowledge concerning the numerical content of an operation span or the visuospatial content of a rotation span.

An alternative (or additive) explanation of this discrepancy could concern the heterogeneity of the measures involved. Hambrick and Engle (2002) used two tasks (a counting span and an operation span) to measure WM capacity and used an LTM task to measure memory performance (answering questions about a baseball radio broadcast). Hambrick and Oswald (2005) used four tasks (a counting span, an operation span, a matrix span, a rotation span) to measure WM capacity and used another LTM measure as a dependent variable (tracking and recalling the movements of players on a baseball field). In our experiment, WM capacity was measured with a reading span task and memory performance with a reading span task; the domain knowledge was knowledge of the sentences in the reading span task. The conditions of our experiment were certainly more favourable for an interaction since all the measures were related to the reading span task. This could have important implications; in fact, if the interaction between domain knowledge and WM capacity varies depending on the measures, then new experiments should be proposed varying this dimension.

Original manuscript received April 2007
Revised manuscript received February 2008
First published online July 2008

REFERENCES


APPENDIX 1: Sample sentences from Experiment 1 (translated from French) used in the personalisable reading span test

Examples of sets of 2, 3, and 4 sentences. In the personalised form, two types of locations were used: (1) locations known to all the participants (e.g., their university), (2) locations that were personalised (e.g., X1, X2, X3).

X1: the name of the metro station near where the participant lives
X2: the name of a street in the participant’s neighbourhood
X3: the name of the city where the participant lives.

Two sentences

Personalised sentences:

- In the library of Paris 5 University, he found a really old manuscript; he liked it so much that he decided to buy the book.
- Coming out of X1 station, she asked an old woman if she could help her with her luggage.

Nonpersonalised sentences:

- In the library of Erevan University, he found a really old manuscript; he liked it so much that he decided to buy the book.
- Coming out of Narek station, she asked an old woman if she could help her with her luggage.

Three sentences

Personalised sentences:

- She observed that since her last visit, the Louvre museum had changed; the walls of the auditorium were full of mirrors.
- She enters the Lagache lecture theatre, but as she sits down, she knocks off her neighbour’s glasses.
- Because she was superstitious, she changed sidewalks on X2 to avoid walking under a ladder.
Nonpersonalised sentences:

- She observed that since her last visit, the Vanak museum had changed; the walls of the auditorium were full of mirrors.
- She enters the Dovassar lecture theatre, but as she sits down, she knocks off her neighbour’s glasses.
- Because she was superstitious, she changed sidewalks on David Yezni Street to avoid walking under a ladder.

Four sentences

Personalised sentences:

- She is sitting on a subway going towards Pont de Sèvres; she falls asleep, her hands resting on her coat.
- The cars were stuck in the parking lot of the Gaumont movie theatre and did not stop honking because the entry was blocked by a container.
- For the second time this month, there was an accident in Danjou Street involving a car and a scooter.
- Coming back from the city hall of X3, he suddenly realised that he had forgotten his brand new pen.

Nonpersonalised sentences:

- She is sitting on a subway going towards Mezadou; she falls asleep, her hands resting on her coat.
- The cars were stuck in the parking lot of the Torkom movie theatre and did not stop honking because the entry was blocked by a container.
- For the second time this month, there was an accident in Tukor Street involving a car and a scooter.
- Coming back from the city hall of Rovan, he suddenly realised that he had forgotten his brand new pen.

APPENDIX 2: Sample personalisable text in Experiment 2 (translated from French)

Nonpersonalised version (for a female participant)

The following story is about Ménia.
She is a student at the Central University of Armenia.
She likes the teaching that is done there pretty well, except maybe for one or two classes that she hates.
The place where she lives is Erevan.
Winter can be difficult in this city, especially in the morning when Ménia leaves.
To go to the university, she takes the metro at Narèk Station.
...
Personalised version (for a female participant)

The following story is about Françoise.
She is a student at the Paris Descartes University.
She likes the teaching that is done there pretty well,
except maybe for one or two classes that she hates.
The place where she lives is X1.
Winter can be difficult in this city,
especially in the morning when Françoise leaves.
To go to the university, she takes the metro at X2 Station.

X1: the name of the city where the participant lives.
X2: the name of the metro station near where the participant lives.

APPENDIX 3: Instructions for the reading span with neutral sentences

This is another memory test. As in the previous one, you will have to read some sentences aloud. You also have to remember the last word of each sentence, and when you get to the end of a group of sentences, you have to list the words you have been remembering. As in the previous test, try to recall them in the right order, but if you can't, just recall as many words as possible.

But there is one change. In this memory test, the sentences are linked to the text you have just read. More specifically, the sentences in the reading span concern the same character and the locations where the character goes to in his (her) everyday life. Even if the names of the locations are not explicitly stated, in every sentence, the location corresponds to a known/an Armenian location like those of the text.

In bold: “known” was inserted in the case of the personalised group and “Armenian” was in the case of nonpersonalised group.