

# Training working memory updating in young adults

Rocío Linares<sup>1</sup> · Erika Borella<sup>2</sup> · M. Teresa Lechuga<sup>1</sup> · Barbara Carretti<sup>2</sup> · Santiago Pelegrina<sup>1</sup> 

Received: 29 July 2016 / Accepted: 24 January 2017 / Published online: 9 March 2017  
© Springer-Verlag Berlin Heidelberg 2017

**Abstract** Working memory updating (WMU) is a core mechanism in the human mental architecture and a good predictor of a wide range of cognitive processes. This study analyzed the benefits of two different WMU training procedures, near transfer effects on a working memory measure, and far transfer effects on nonverbal reasoning. Maintenance of any benefits a month later was also assessed. Participants were randomly assigned to: an adaptive training group that performed two numerical WMU tasks during four sessions; a non-adaptive training group that performed the same tasks but on a constant and less demanding level of difficulty; or an active control group that performed other tasks unrelated with working memory. After the training, all three groups showed improvements in most of the tasks, and these benefits were maintained a month later. The gain in one of the two WMU measures was larger for the adaptive and non-adaptive groups than for the control group. This specific gain in a task similar to the one trained would indicate the use of a better strategy for performing the task. Besides this nearest transfer effect, no other transfer effects were found. The adaptability of the training procedure did not produce greater improvements. These results are discussed in terms of the training procedure and the feasibility of training WMU.

## Introduction

Working memory (WM) is a system of limited capacity that enables content in our memory to be maintained and manipulated as necessary for complex cognition (Morrison and Chein 2011; Salminen, Strobach and Schubert 2012; von Bastian and Oberauer 2014). WM can hold only a certain amount of information at a given moment in time, so it has to constantly replace old representations with new ones. This updating process is essential to the human mental architecture (Friedman et al. 2008; Schmiedek, Hildebrandt, Lövdén, Wilhelm and Lindenberger 2009), and it is the best predictor of several higher level cognitive functions, such as fluid intelligence (Belacchi, Carretti and Cornoldi 2010; Chen and Li 2007; Friedman et al. 2006). WM updating (WMU) is related to cognitive performance in several areas. For example, it mediates the relationship between WM and reading comprehension (Gernsbacher, Varner and Faust 1990), accounting for individual differences in reading skills (Borella, Carretti and Pelegrina 2010; Carretti, Cornoldi, De Beni and Romanò 2005; Palladino, Cornoldi, De Beni and Pazzaglia 2001), and also in mathematics (Passolunghi and Pazzaglia 2004; Pelegrina, Capodici, Carretti and Cornoldi 2014). Increasing evidence of the feasibility of training WM has recently emerged, prompting an increasing number of studies on this issue.

Training studies are based on practice with complex WM tasks with a view to improving performance in tasks similar to those trained, but also to identifying transfer effects to other tasks that may involve related processes. According to a conceptually-based continuum of nearest to far transfer tasks (i.e., Willis, Blieszner and Baltes 1981; Noack, Lövdén, Schmiedek and Lindenberger 2009), nearest transfer involves the performance benefits produced by training on tasks that tap the same process as the trained

✉ Santiago Pelegrina  
spelegri@ujaen.es

<sup>1</sup> Department of Psychology, University of Jaén, Paraje las Lagunillas s/n, 23071 Jaén, Spain

<sup>2</sup> Department of General Psychology, University of Padova, Via Venezia, 8, 35131 Padova, Italy

task, but with different types of stimuli. Near transfer instead occurs when training enhances performance on tasks that measure the same broad ability but involve different requirements (Chein and Morrison 2010; Jausovec and Jausovec 2012; Sprenger et al. 2013; von Bastian, Langer, Jäncke and Oberauer 2013; Westerberg and Klingberg 2007). Far transfer concerns the benefits that training may produce on tasks intended to measure other cognitive abilities that are correlated or share processes with WM. Some studies have found benefits of training in different populations on measures of cognitive control (Borella, Carretti, Riboldi and De Beni 2010; Chein and Morrison 2010; Westerberg and Klingberg 2007), fluid intelligence (Jausovec and Jausovec 2012; Westerberg and Klingberg 2007), reading comprehension (Carretti, Borella, Zavagnin and De Beni 2012; Chein and Morrison 2010), or reasoning (Borella et al. 2010b; Jausovec and Jausovec 2012; Sprenger et al. 2013). It has been suggested that stronger relations between a trained process and an untrained ability are indicative of greater transfer effects (Waris, Soveri and Laine 2015). However, not all studies have found support for transfer effects, especially with respect to intelligence or reasoning abilities (Shipstead, Redick and Engle 2010, 2012). Some meta-analyses seem to support the idea that WM training can be effective to enhance other cognitive skills in adulthood (Au et al. 2015) and in old age (Karbach and Verhaeghen 2014). Nonetheless, when found, transfer effects tend to be rather modest. In addition, there are some controversies regarding these conclusions, as different factors, the roles of which remain unclear, may be involved (see Dougherty, Hamovitz and Tidwell 2016; Melby-Lervåg and Hulme 2016).

The present study aims to train the WMU process using different tasks to ascertain to what extent the training produces transfer effects on other cognitive tasks.

### Training working memory updating

Several training studies on young adults included some WMU tasks among the training activities (Jausovec and Jausovec 2012; Sandberg, Rönnlund, Nyberg and Neely 2015; Sprenger et al. 2013; Westerberg and Klingberg 2007), and found some benefits in terms of WMU, WM, inhibition, attention, and reasoning tasks. It is difficult to assess the specific contribution of the WMU training to the transfer effects, however, because these studies usually included other tasks that do not involve updating as well.

Some other studies in young adults used mostly WMU tasks to examine the specific influence of the updating process on other cognitive functions. Some of them found improvements in WMU tasks (Dahlin et al. 2008a, b; Li et al. 2008; Lilienthal, Tamez, Shelton, Myerson and Hale

2013; Salminen et al. 2012; Waris et al. 2015; Xiu, Zhou and Jiang 2015), and also in some short-term memory measures that essentially involved the information maintenance (Jaeggi, Buschkuhl, Jonides and Perrig 2008; Waris et al. 2015). Some authors showed that it is difficult to identify gains in measures different to those trained (Dahlin et al. 2008a, b; Küper and Karbach 2015; Li et al. 2008; Redick et al. 2013; Thompson et al. 2013), although other studies succeeded in inducing improvements in tasks that tap task switching, attention (Salminen et al. 2012), emotional regulation (Xiu et al. 2015), or fluid intelligence (Jaeggi et al. 2008).

There may be several explanations for the variability in the reported results. It could be argued that transfer effects are due to an improvement in the updating process, or to having learned suitable strategies for a given task paradigm (von Bastian and Oberauer 2014). An improvement in the process would produce gains in performance irrespective of the tasks involved (Dahlin et al. 2008a, b; Lilienthal et al. 2013; Waris et al. 2015)—near and far transfer effects—, whereas an improvement due to having learned a strategy would only produce benefits in very similar tasks—specific training gain and nearest transfer effect—, in which the same strategy could be usefully applied, and they would not be maintained over time (Küper and Karbach 2015; Li et al. 2008; Salminen et al. 2012).

A first goal of the present study was to ascertain the extent to which a training program focusing on WMU would induce transfer effects on other updating tasks that share more or less the same features as the tasks used in the training.

A second possible determinant of the efficacy of training relates to the cognitive demands of the training tasks, and more specifically to the level of difficulty of the tasks administered during the training sessions. In some studies, the training sessions involved tasks with the same level of difficulty (Dahlin et al. 2008a, b; Li et al. 2008). These studies showed a subsequent improvement in tasks assessing WMU, but transfer effects were not always found. An alternative is to use an adaptive training procedure, in which the difficulty of the task is adjusted to individual performance (Jaeggi et al. 2008; Küper and Karbach 2015; Redick et al. 2013; Salminen et al. 2012; Waris et al. 2015; Xiu et al. 2015). Adaptive training has been shown to induce training gains and transfer effects on different tasks (Jaeggi et al. 2008; Salminen et al. 2012; Xiu et al. 2015; but see also; Redick et al. 2013; Waris et al. 2015).

The above-mentioned results could be due to the fact that keeping the task challenging during the training sessions produces more benefits than repeating the same task with the same level of difficulty, because it preserves some disparity between the demands of the task and the individual's capacity (Lövdén, Bäckman, Lindenberger, Schaefer

and Schmiedek 2010). Lilienthal et al. (2013) study compared the benefits of adaptive versus non-adaptive training with a dual  $n$ -back task. The results of this study showed that both groups improved in an  $n$ -back task, but the benefits of the training were greater for the adaptive than for the non-adaptive group. On the other hand, a recent study (von Bastian and Eschen 2015) found no difference between an adaptive training and an alternative—although not non-adaptive—training in which the level of difficulty was varied randomly throughout the task. In the same vein, Karbach and Verhaeghen (2014) conducted a meta-analysis focusing on older people and found no greater effects of adaptive training vis-à-vis non-adaptive procedures. There are no reports in the literature concerning younger adults, however, and this is an issue that remains to be clarified. To contribute to the debate on this issue, the second goal of the present study was to examine the specific transfer effects of an adaptive WMU training as compared with a non-adaptive WMU training.

### The present study

The purpose of this study was thus to examine transfer effects (from nearest to farthest) of a WMU training comprising two numerical updating tasks. Classic studies on training hold that introducing a degree of practice variability, as opposed to repetitive practice, during training leads to a higher transfer (for a review, see Schmidt and Bjork 1992). Recently, this approach has been used in some WM training researches (e.g., Waris et al. 2015). In the present study, the two training tasks were designed to be structurally similar, in the sense that they comprised the same WMU components, which had to be applied in analogous sequences. They differed mainly in terms of the stimuli and specific numerical operations used. In both tasks, participants had to complete different operations on previously-memorized information to obtain a result that could subsequently be updated. Two training regimens were included: an adaptive training regimen, in which the levels of load and suppression was varied, and a non-adaptive regimen that had consistent load and suppression levels. The possible transfer effects of the training were analyzed in terms of performance in WMU (nearest transfer), WM (near transfer), and fluid intelligence (far transfer).

To investigate nearest transfer effects, two WMU tasks were administered: an Odd–Even Number Updating task developed specifically for the present study and a Number Comparison Updating task (Carretti, Cornoldi and Pelegri 2007). These two tasks differ from the training tasks in terms of either the criterion for updating the information or the structure of the task (e.g., whether or not specific cues were given to indicate which element to retrieve

or substitute). Including two different updating tasks would enable us to see whether any gains were indicative of respondents using strategies they had learned or due to an improvement in the efficiency of their WMU ability. Any near transfer effect on WM was assessed by means of the Operation Span task (Turner and Engle 1989). Two fluid intelligence tasks were also administered to assess far transfer: the Standard Progressive Matrices (Raven, Court and Raven 1977); and the Cattell test (Cattell and Cattell 1963). A follow-up assessment 1 month after finishing the training was also included to establish whether any transfer effects were maintained.

For the training sessions, the same updating tasks were used, but in two different regimens, one adaptive and the other non-adaptive, to analyze the effects of different cognitive demands. In the adaptive regimen, the level of difficulty of the tasks was adjusted to individual performance, whereas in the non-adaptive regimen, the tasks were always administered with a fixed and relatively low level of difficulty. In addition to the two trained groups, an active control group was also included to distinguish the effects of the training from other effects relating to the experience of participating in the experiment. This active control group practiced with computer games unrelated to WM.

An improvement was expected for both training groups, in the tasks that assessed WMU (nearest transfer) and WM (near transfer), in line with the previous studies (Dahlin et al. 2008a; Jaeggi et al. 2008; Salminen et al. 2012; Waris et al. 2015; Xiu et al. 2015). More specifically, if the transfer effects were due to an improvement in WMU ability, then benefits would be seen in both updating tasks. If they were due to a strategy learning process, they would be detected mainly for the tasks most similar to those used in the training. Improvement in the fluid intelligence tasks (far transfer) could also be expected, but to a less extent for both trained groups if the training gains were more strategy-related rather than being due to changes in information processing per se. Previous studies have produced inconclusive results regarding this type of transfer (Dahlin et al. 2008b; Jaeggi et al. 2008; Küper and Karbach 2015; Redick et al. 2013; Salminen et al. 2012; Sandberg et al. 2014; Waris et al. 2015).

In general, greater improvements were expected for the adaptive training group than for the non-adaptive group (or the control group). Lilienthal et al. (2013) showed that the adaptive training was superior to the non-adaptive training in terms of training gains. Even though von Bastian and Eschen (2015) did not find any differences between an adaptive group and another in which the difficulty levels were administered in a random way, as we included an adaptive and a non-adaptive group our hypothesis should be compatible with the Lilienthal et al. (2013) study. If this was the case, a maintenance of the nearest and near transfer

effects after 1 month can be predicted for both training groups (Dahlin et al. 2008b; Li et al. 2008), with, however, greater gains for the adaptive training group.

## Method

### Participants

The study involved 81 university students of the University of Padova aged 18–35 ( $M=20.22$ ,  $SD = 1.99$ ) (71 females and 10 males): 27 (18–22 years,  $M=20.11$ ,  $SD = 1.05$ , 23 females and 4 males) were randomly assigned to the adaptive group; 27 to the non-adaptive group (18–35 years,  $M=21.19$ ,  $SD = 3$ , 25 females and 2 males); and 27 to the active control group (19–21 years,  $M=19.37$ ,  $SD = 0.63$ , 23 females and 4 males). All participants were native Italian speakers, volunteered for the study and gave their informed consent before taking part in the study. Participants attended two introductory psychology courses, and they were given course credits for their participation.

## Materials

### Training

#### *WMU training (adaptive and non-adaptive groups)*

**Arithmetical updating task** The task is similar to the one described by Oberauer, Wendland and Kliegl (2003), and Salthouse, Babcock and Shaw (1991). Ten lists of numbers and arithmetical operations were presented in different boxes. Participants had to memorize some numbers, apply different operations to these numbers, and remember the last number obtained. Each list started with a number of initial items corresponding to the load, as specified in the following, of each list. These initial items were displayed in boxes, from left to right, for 2000 ms. Then, arithmetical operations consisting of an addition or subtraction sign and a number (0, 1, 2 or 3) were displayed randomly in different boxes. The number of arithmetical operations presented varied from 6 to 9. There were two types of item: those requiring or not requiring updating. The items involving no updating included arithmetical operations with  $-0$  and  $+0$ , while those requiring an updating step included arithmetical operations, such as  $-1$ ,  $+1$ ,  $-2$ ,  $+2$ ,  $-3$ , and  $+3$ . Only one operation was presented at a time and remained on screen for 2000 ms. Participants had to mentally apply the arithmetical operation presented in a given box to the number previously displayed in the same box, and then remember the result to use it in the next arithmetical operation appearing in the same box. They, therefore, always had to apply the

new operation to the latest result obtained in a given box. At the end, silver boxes cued participants to type in the last result obtained in each box. Feedback was provided after each list.

The lists of numbers differed in terms of the memory load and degree of suppression required to fulfil the task aims. There were four memory load levels, which varied according on the number of elements (from 2 to 5) that had to be memorized. The degree of suppression was manipulated across two levels (low versus high) by varying the proportion of non-updating versus updating items. In the low-suppression condition, two-thirds of the items did not involve updating (with operations of  $+0$  or  $-0$ ), while the remaining third required updating (items with other operations). In the high-suppression condition, one-third of the items were non-updating and two-third were updating. Therefore, in the latter condition, participants were required to update the information more frequently.

The adaptive group performed this task on different levels of difficulty: after a participant completed two lists correctly, the load was increased for the next list. When participants failed in two consecutive lists, the load was reduced in an element. The first list for a given load involved a low suppression, then successful participants were administered a high-suppression list.

The non-adaptive group always performed the task on the same level of difficulty, using a memory load of two and in the high-suppression condition. These levels and conditions were chosen to make the task more enjoyable for this group.

The dependent variable was the percentage of correct lists recalled.

**Number size updating task (based on Carretti et al. 2007 and Lendínez, Pelegrina and Lechuga 2011)** As in the previous task, ten lists of numbers were presented in different boxes. Participants had to memorize the last number in each box according to a given criterion. This task was similar to the Arithmetical Updating task, the only difference being the stimuli presented, and the requirement of the task. The same load levels of and suppression sublevels were also used. In the low-suppression lists, one-third of the items were updating items, whereas in the high-suppression lists, two-thirds of the items were updating items. Thus, depending on the level of suppression, the lists differed in terms of the amount of no-longer relevant information that had to be discarded.

Each list was preceded by a message specifying whether participants had to remember the smaller or larger numbers that were presented (each number was presented in a discrete box). The first numbers (between 10 and 99) were presented consecutively, each in a discrete box, from left to right for 2000 ms. Then sequences of two-digit numbers were successively displayed in the various boxes at random,

again for 2000 ms. There were two types of items (updating and non-updating), depending on whether the number presented had to be updated according to the larger or smaller number rule. Participants therefore had to either update or remember the (larger or smaller) number in each box in a given item in order to meet the criterion established at the beginning of a given list. At the end of a list, silver boxes appeared for participants to type in the last numbers recalled that met the criterion. After each list, feedback was provided.

As with the previous task, the adaptive group was administered the task on different levels of difficulty, adjusting the memory load and level of suppression, while the non-adaptive group always performed the task with a load of two items and in the high-suppression condition.

The dependent variable was the percentage of correct answers.

#### Active control group

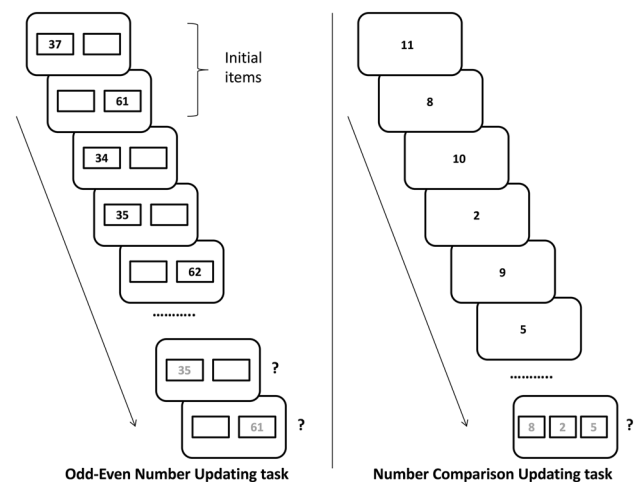
**Tetris** In this computer game, participants have to rotate and move blocks falling from the top of the screen to create horizontal rows without any gaps at the bottom of the screen. When a row is completed, it disappears, and when a certain number of rows have been cleared, the game enters a new difficulty level. If no rows are completed, the blocks pile up until they reach the top of the screen when the game ends. Thus, the level of difficulty varied according to the participant's ability to play the game.

**Crossword** In this game, participants are given clues to solve and have to place the words in horizontal and vertical boxes in a square grid. After completing a given crossword, a new and more difficult crossword, in terms of the number of words that the participant was required to find, was presented. In both games, participants could see their scores on the game markers.

#### Cognitive assessment at baseline, post-training, and follow-up

##### Nearest transfer tasks: working memory updating

**Odd–even number updating task** In this task, developed for the present study, different lists of numbers between 10 and 99 were presented in a variable number of boxes, depending on the level of memory load (from 2 to 5 boxes). The lists comprised a number of initial items corresponding to the load of each list (that is, three items corresponded to a load level of 3) and nine additional study items. Each list started with a message indicating whether participants had to remember the last odd or the last even number appearing in each box. At the outset, all of the initial items ful-



**Fig. 1** Working memory updating (WMU) assessment tasks. In the *left panel*, the Odd–Even Number Updating Task is illustrated for a list of load two. In this task, an initial number were presented in each box. Then, a sequence of numbers appeared randomly in the different boxes. Participants had to recall the last odd/even number for each box following a previous criterion. In the example, the criterion was the odd numbers. In the *right panel*, a schematic representation of a list in the Number Comparison Updating task is displayed. A sequence of ten numbers was presented in the center of the screen. Participants had to remember, depending on the criterion, the three smallest or largest numbers, in the same order at that in which they were presented. In the example, the criterion was the smallest numbers

filled the criterion of being updated, such that they had to be memorized. For example, they were odd numbers when the instructions were to recall the final odd numbers. The initial items were presented in separate boxes from left to the right. The rest of the items were presented in boxes selected at random. All items remained on the screen for 3000 ms. Participants had to update the number presented in each box according to the odd–even rule. At the end, when participants saw silver boxes, they had to type in the last even or odd numbers, as instructed. An example list is shown on the left side of Fig. 1.

As in the trained tasks, there were four memory load levels (levels 2–5). All lists were low suppression; that is, two-thirds of the items involved no updating, and the remaining third required updating. In the items that required updating, the number presented met the specified criterion (e.g., it was an even number when the last even number had to be retained), while in the non-updating items, it did not.

This task had the same load and structure as both the training tasks, with information displayed in different boxes. In all these tasks, the various WMU processes, such as the retrieval, transformation, and substitution of the information, were cued in a comparable way. The task started with four practice lists, two with load one and two with load two, and ended when the participant failed in two

consecutive lists to avoid frustration, as often done in the literature with complex WM test (e.g., Borella et al. 2010b; Borella and Ribaupierre 2014). Thus, the dependent variable was the number of correct answers for the whole task. The maximum score was 28.

*Number comparison updating task (Carretti et al. 2007)* Ten lists of ten numbers were presented in the center of the screen. Each number was displayed for 2 s followed by a mask (##) for 1 s. Participants had to remember the three smallest numbers on each list. At the end of the list, participants had to type the three smallest numbers in their order of presentation. Unlike the training tasks, in this case, the information was presented at the center of the screen, so the element to be substituted was not cued and participants had to decide which element of the list to retrieve and when to substitute it. An example is shown on the right side of Fig. 1. The task started with four practice lists. The dependent variable was the average of the numbers recalled in the correct order. The maximum score was 30.

#### *Near transfer task: working memory*

*Operation span task (Turner and Engle 1989)* This task consisted of sets of pairs of arithmetical expressions (e.g.,  $4 + 3 - 1 = 2$ ) and single-digit numbers (e.g., 8). The arithmetical expressions were presented in sets of 2–7 items, depending on the level of memory load. Each list started with the presentation of an arithmetical expression for 5000 ms. Participants had to check it and indicate whether it was correct or not. Then, a blue number appeared in the center of the screen for 3000 ms. At the end of each set, participants had to indicate the blue numbers that had appeared after each operation in their order of presentation. Two sets were displayed for each level of memory load. The task ended when a participant failed in two sets with the same load to avoid frustration. There were four practice lists.

As in the Odd–Even Number Updating task, the number of correct answers in all presented lists was considered. The maximum possible score was 54.

#### *Far transfer tasks: fluid intelligence*

*Standard progressive matrices (Raven et al. 1977)* Twenty visual patterns with a part missing were presented in this paper-and-pencil task. Different pieces were presented and participants were asked to identify the one needed to complete the pattern. Participants were not allowed to move on to the next pattern before answering for the one currently displayed, nor were they allowed to use a pencil and paper to solve the problem. The dependent variable was the number of correctly-solved problems. The maximum score was

20. Two parallel versions (as in Shipstead et al. 2012) were used, which were counterbalanced across testing sessions.

*Culture fair test, scale 3 (Cattell and Cattell 1963)* Scale 3 of the Cattell test consists of two parallel forms (A and B), each containing four subtests to be completed in 2.5–4 min, depending on the subtest. In the first subtest, series, participants saw incomplete series of abstract shapes and figures, and had to choose one of six options that best completed the series. In the second subtest, classifications, participants saw 14 problems consisting of abstract shapes and figures, and had to choose which two of the five differed from the other three. In the third subtest, Matrices, participants were presented with 13 incomplete matrices containing four-to-nine boxes of abstract figures and shapes plus an empty box and six options: they had to select the answer that correctly completed each matrix. In the final subtest, Conditions, participants were presented with ten sets of abstract figures, lines and a single dot, along with five options, and they had to assess the relationship between the dot, figures, and lines, then choose the option in which a dot could be positioned in the same relationship. The dependent variable was the number of correctly-solved items across the four subsets (maximum score 50).

There were two parallel versions (A and B) of each task. The versions were counterbalanced across assessment sessions following an ABA design that has been used frequently in other training studies (e.g., Borella et al. 2010b).

## Procedure

Participants in the three groups attended seven sessions. The first and sixth were the pre- and posttest sessions, and the last one was the follow-up session (1 month later); these sessions were administered individually. During the other four sessions (from the second to the fifth), the training or control activities were administered to pairs of participants that had to do either the WMU training tasks or games. They were accommodated on opposite sides of a desk, so that they could not see what the other participant was doing. Although participants knew that they were enrolled in a training study, they did not receive information as to what group they belonged to.

The pretest session aimed to assess the baseline level of each participant in each task. The training started on the day after the pretest session. The posttest session took place at least 1 day after completing the training. The comparison between the pre–posttest sessions enabled us to ascertain any changes induced by the training. The follow-up session a month later was to establish whether any changes identified were maintained over time. Each session lasted an hour

**Table 1** Descriptive pretest, posttest, and follow-up data for the adaptive and non-adaptive training, and the active control groups

|                       |    | Pretest    |               | Posttest   |               | Follow-up  |               |            |
|-----------------------|----|------------|---------------|------------|---------------|------------|---------------|------------|
|                       |    | Max. score | <i>M</i> (SD) | Percentage | <i>M</i> (SD) | Percentage | <i>M</i> (SD) | Percentage |
| WMU odd–even numbers  | 27 |            |               |            |               |            |               |            |
| Adaptive              |    |            | 13.12 (5.31)  | 48.59      | 19.54 (5.79)  | 72.37      | 20.69 (4.71)  | 76.63      |
| Non-adaptive          |    |            | 15.26 (5.13)  | 56.52      | 17.48 (5.72)  | 64.74      | 17.52 (6.24)  | 64.89      |
| Control               |    |            | 13.52 (4.46)  | 50.07      | 12.07 (5.70)  | 44.70      | 14.52 (5.71)  | 53.78      |
| WMU number comparison | 30 |            |               |            |               |            |               |            |
| Adaptive              |    |            | 20.19 (4.46)  | 67.30      | 21.35 (3.39)  | 71.17      | 21.96 (4.06)  | 73.20      |
| Non-adaptive          |    |            | 20.74 (3.25)  | 69.13      | 21.30 (4.11)  | 71.00      | 22.33 (5.28)  | 74.43      |
| Control               |    |            | 18.52 (4.73)  | 61.73      | 20.33 (5.37)  | 67.77      | 21.11 (5.04)  | 70.37      |
| WM operation span     | 54 |            |               |            |               |            |               |            |
| Adaptive              |    |            | 19.77 (15.55) | 36.61      | 34.77 (17.98) | 64.39      | 37.42 (16.04) | 69.30      |
| Non-adaptive          |    |            | 25.70 (18.43) | 47.59      | 33.30 (16.02) | 61.67      | 38.56 (13.89) | 71.41      |
| Control               |    |            | 26.00 (16.91) | 48.15      | 31.52 (16.62) | 58.37      | 36.67 (14.47) | 67.91      |
| Raven                 | 20 |            |               |            |               |            |               |            |
| Adaptive              |    |            | 17.00 (1.57)  | 85.00      | 16.92 (1.94)  | 84.60      | 17.65 (1.60)  | 88.25      |
| Non-adaptive          |    |            | 16.15 (2.86)  | 80.75      | 16.89 (2.03)  | 84.45      | 16.81 (2.34)  | 84.05      |
| Control               |    |            | 16.59 (1.93)  | 82.95      | 16.56 (1.83)  | 82.80      | 16.33 (1.52)  | 81.65      |
| Cattell               | 41 |            |               |            |               |            |               |            |
| Adaptive              |    |            | 27.73 (4.07)  | 67.63      | 28.81 (3.53)  | 70.27      | 30.73 (4.77)  | 74.95      |
| Non-adaptive          |    |            | 27.48 (4.34)  | 67.02      | 29.19 (5.36)  | 71.20      | 30.89 (5.25)  | 75.34      |
| Control               |    |            | 25.85 (3.96)  | 63.05      | 26.52 (3.26)  | 64.68      | 28.22 (4.27)  | 68.83      |

Percentages were calculated by considering the maximum score obtained on each task

and the tasks were administered in the order in which they are described in “Materials”.

The adaptive and non-adaptive training groups completed the training over four 30-min sessions within a 2-week time frame, with a fixed 2-day break between sessions. The WMU training consisted of two numerical training tasks that participants performed twice at each session. First, they performed the Arithmetical Updating task and then the Number Size Updating task. The order of tasks was fixed across sessions and participants. Both groups (adaptive and non-adaptive) performed the same tasks but on different levels of difficulty, i.e., the adaptive training group moved on across different levels of difficulty, the non-adaptive group always performed tasks with a same memory load of two and in the high-suppression condition.

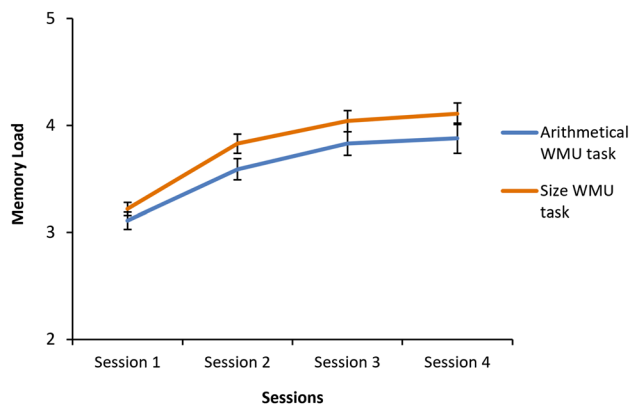
The active control group played Tetris during the first and third sessions and completed crosswords during the second and fourth sessions. The tasks had different levels of difficulty and participants could advance within a session. Thus, when a Tetris game was completed, a new, somewhat more difficult version was presented. When a crossword was solved, a new crossword with more words

was displayed. Participants could see their scores on the screen. Both games were played for the same amount of time as the other two groups’ training sessions.

## Results

### Baseline measures

To ensure there were no differences between the groups at the pretest stage, separate analyses of variance (ANOVAs) were run on the pretest performance in all tasks, with group (adaptive, non-adaptive and control) as the between-subjects factor. The results indicated that there were no baseline differences between the groups on each task: Odd–Even Number Updating task,  $F(2, 78)=1.44$ ,  $p=.244$ ,  $\eta^2=0.04$ ; Number Comparison Updating task,  $F(2, 78)=2.03$ ,  $p=.138$ ,  $\eta^2=0.05$ ; Operation span task,  $F(2, 78)=0.76$ ,  $p=.471$ ,  $\eta^2=0.02$ ; Raven,  $F(2, 78)=1.20$ ,



**Fig. 2** Mean memory load in the two training tasks across the four sessions of adaptive training. The *bars* represent one standard error of the mean

$p = .306$ ,  $\eta^2 = 0.03$ ; and Cattell,  $F(2, 78) = 1.84$ ,  $p = .165$ ,  $\eta^2 = 0.05$  (see Table 1)<sup>1</sup>.

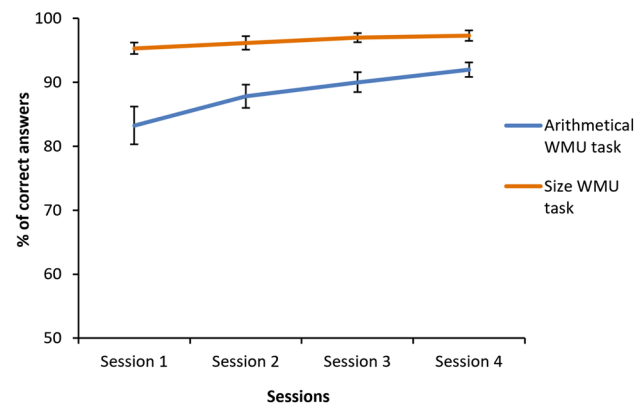
Raw and percentage means, for each group and in each assessment session, are shown in Table 1. The percentage values were calculated by considering the highest score obtained in the task for any given participant.

### Training gains

Different dependent variables (because of the different training procedures, adaptive versus non-adaptive) were analyzed to assess the training gains in each group. The mean level reached at each session was considered to analyze the gains in the adaptive training group.

The group given adaptive training gradually improved in the performance of both the training tasks in subsequent sessions (see Fig. 2). There was an effect of session,  $F(3, 78) = 36.98$ ,  $p < .001$ ,  $\eta^2 = 0.59$ , for performance in the Arithmetical Updating task. Post-hoc Bonferroni comparisons showed that there were differences between all sessions except between the third and fourth, i.e., performance improved in the second session with respect to the first ( $M_{diff} = 0.48$ ,  $p < .001$ ), and improved further in the third ( $M_{diff} = 0.72$ ,  $p < .001$ ), and then in the fourth ( $M_{diff} = 0.77$ ,  $p < .001$ ). Performance was also higher in the third ( $M_{diff} = 0.24$ ,  $p = .045$ ) and fourth sessions ( $M_{diff} = 0.29$ ,  $p = .023$ ) than in the second.

Similar results were obtained for the Number Size Updating task. The main effect of session was also significant,  $F(3, 78) = 92.99$ ,  $p < .001$ ,  $\eta^2 = 0.78$ . Post-hoc



**Fig. 3** Percentage of correct answers in the two training tasks across the four sessions of non-adaptive training. The *bars* represent one standard error of the mean

Bonferroni comparisons showed that there were differences between all sessions, except between the third and fourth, i.e., performance was better in the second ( $M_{diff} = 0.61$ ,  $p < .001$ ), third ( $M_{diff} = 0.81$ ,  $p < .001$ ), and fourth ( $M_{diff} = 0.88$ ,  $p < .001$ ) sessions than in the first. Performance also improved in the third ( $M_{diff} = 0.21$ ,  $p = .01$ ) and fourth ( $M_{diff} = 0.28$ ,  $p < .001$ ) sessions by comparison with the second.

The non-adaptive group was given lists on the same level of difficulty across all sessions. Instead of considering the performance reached at each session, the percentage of correct answers obtained on each session was analyzed (see Fig. 3). There was an effect of session for performance in the Arithmetical Updating task,  $F(3, 69) = 6.29$ ,  $p = .01$ ,  $\eta^2 = 0.22$ ; and post-hoc Bonferroni comparisons showed that performance in the fourth session was higher than in the first ( $M_{diff} = 8.75$ ,  $p = .01$ ). No differences were found between the first and fourth sessions for the Number Size Updating task in this group.

### Transfer effects

To assess the effect of training, a 3 (group: adaptive, non-adaptive, control)  $\times$  3 (session: pretest, posttest, follow-up) mixed-design ANOVA was run for each dependent variable, with group as a between-subjects factor, and sessions as repeated measures. One participant in the adaptive training group did not complete the follow-up assessment<sup>2</sup>. Significant main effects and interactions were analyzed using post-hoc pairwise comparisons with Bonferroni's correction at  $p < .05$ , adjusted to multiple comparisons.

<sup>1</sup> There was a participant in the non-adaptive group whose age differed from the mean age of the other participants. All the analyses described were also run with this participant excluded and the results were largely identical.

<sup>2</sup> The participant was excluded listwise from the analyses. Additional analyses were run with this participant excluded and the pattern of results was the same to that described above.



**Table 2** Results of mixed-design 3×3 ANOVA for the measures of interest, with group (adaptive, non-adaptive, and control) as the between-subjects factor, and session (pretest, posttest, and follow-up) as repeated measures

|                       | <i>F</i> | <i>df</i> | <i>p</i> | $\eta^2_p$ |
|-----------------------|----------|-----------|----------|------------|
| Transfer effect       |          |           |          |            |
| WMU odd–even numbers  |          |           |          |            |
| Group (G)             | 7.88     | 2,77      | 0.001    | 0.17       |
| Session (S)           | 15.49    | 2,15      | <0.001   | 0.17       |
| G × S                 | 7.36     | 4,15      | <0.001   | 0.16       |
| WMU number comparison |          |           |          |            |
| Group (G)             | 1.49     | 2,77      | 0.233    | 0.04       |
| Session (S)           | 5.85     | 2,15      | 0.004    | 0.07       |
| G × S                 | 0.24     | 4,15      | 0.916    | 0.01       |
| WM operation span     |          |           |          |            |
| Group (G)             | 0.13     | 2,77      | 0.874    | 0.00       |
| Session (S)           | 29.01    | 2,15      | <0.001   | 0.27       |
| G × S                 | 1.29     | 4,15      | 0.277    | 0.03       |
| Raven                 |          |           |          |            |
| Group (G)             | 1.24     | 2,77      | 0.296    | 0.03       |
| Session (S)           | 1.61     | 2,15      | 0.204    | 0.02       |
| G × S                 | 2.14     | 4,15      | 0.079    | 0.05       |
| Cattell               |          |           |          |            |
| Group (G)             | 3.28     | 2,77      | 0.043    | 0.08       |
| Session (S)           | 23.33    | 2,15      | <0.001   | 0.23       |
| G × S                 | 0.34     | 4,15      | 0.851    | 0.01       |

*G* group, *S* session

Descriptive statistics are given in Table 1, and the results of the ANOVAs are summarized in Table 2.

#### Nearest transfer effect

For the Odd–Even Number Updating task, the main effect of session was significant,  $F(2, 154)=15.49$ ,  $p<.001$ ,  $\eta^2=0.17$ , indicating a better performance at posttest (Mdiff=2.40,  $p=.001$ ) and follow-up (Mdiff=3.61,  $p<.001$ ) than at pretest. The posttest performance was maintained at follow-up. The main effect of group was also significant,  $F(2, 77)=7.88$ ,  $p=.001$ ,  $\eta^2=0.17$ , showing that the adaptive group (Mdiff=4.41,  $p=.001$ ) and the non-adaptive group (Mdiff=3.38,  $p=.014$ ) performed better overall than the control group. The significant interaction between session and group,  $F(4, 154)=7.36$ ,  $p<.001$ ,  $\eta^2=0.16$ , revealed that the improvement from pretest to posttest was larger in the adaptive (Mdiff=7.78,  $p<.001$ ) and non-adaptive (Mdiff=5.41,  $p=.003$ ) groups than in the control group. A month later, at the follow-up session, only the superiority of the adaptive group relative to the active control group was maintained (Mdiff=6.17,  $p<.001$ ).

When the effect of session was analyzed independently for each group, performance was

better at posttest (Mdiff=6.42,  $p<.001$ ) and follow-up (Mdiff=7.58,  $p<.001$ ) than at pretest for the adaptive group, whereas only an improvement from posttest to follow-up (Mdiff=2.44,  $p=.038$ ) emerged for the non-adaptive group.

In the Number Comparison Updating task, the main effect of session was significant,  $F(2, 154)=5.85$ ,  $p=.004$ ,  $\eta^2=0.07$ , showing that performance was better at follow-up (Mdiff=1.99,  $p=.005$ ) than at pretest, with no differences between pretest and posttest (Mdiff=1.12,  $p=.110$ ) and posttest and follow-up ( $p=.509$ ). Neither the effect of group ( $p=.233$ ), nor the interaction ( $p=.916$ ) reached statistical significance, however.

#### Near transfer effect

In the Operation Span task, the main effect of session was significant,  $F(2, 154)=29.01$ ,  $p<.001$ ,  $\eta^2=0.27$ , indicating that performance was better at posttest (Mdiff=9.37,  $p<.001$ ) and follow-up (Mdiff=13.72,  $p<.001$ ) than at pretest. Performance also continued to improve from posttest to follow-up (Mdiff=4.35,  $p=.020$ ). Neither the effect of group ( $p=.874$ ), nor the interaction ( $p=.270$ ) were significant.

#### Far transfer effect

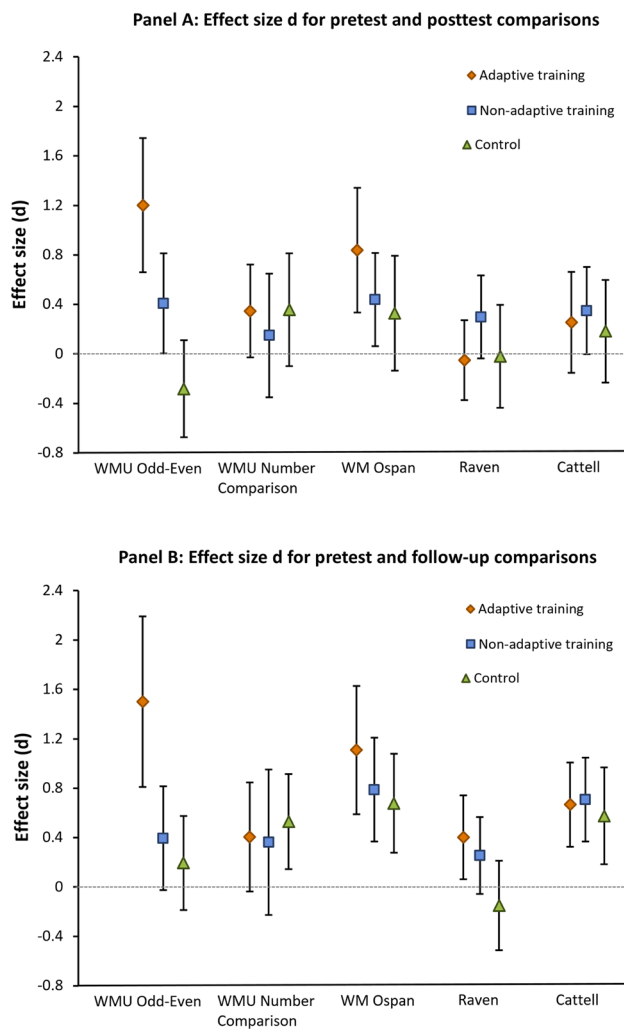
For the Raven task, the main effects of session and group and the interaction between them were not significant ( $p=.204$ ,  $p=.296$ ,  $p=.079$ , respectively).

As for the Cattell task, the main effect of session was significant,  $F(2, 154)=23.33$ ,  $p<.001$ ,  $\eta^2=0.23$ , indicating that performance was better at posttest (Mdiff=1.15,  $p=.047$ ) and follow-up (Mdiff=2.93,  $p<.001$ ) than at pretest. Performance also continued to improve from posttest to follow-up (Mdiff=1.78,  $p<.001$ ). The main effect of group was also significant,  $F(2, 77)=3.28$ ,  $p=.043$ ,  $\eta^2=0.08$ . None of the comparisons reached statistical significance when post-hoc comparisons were run, however ( $p>.078$ ). The interaction between session and group was not significant ( $p=.573$ ).

For the purpose of analyzing any transfer effects, Cohen's *d* (1988) (expressing the effect size of the comparisons) was calculated to establish the effect size by comparing pretest and posttest, and pretest and follow-up measures (see Fig. 4). Values higher than 0.80 were considered as large effects.

## Discussion

The present study considered the effects on young adults of a WMU training program under two regimens, adaptive,



**Fig. 4** Effect size *d* for pretest and posttest comparisons (a) and for pretest and follow-up comparisons (b) showing transfer effects by group (adaptive, non-adaptive and control). Error bars represent 95% confidence intervals

and non-adaptive, comparing both with an active control group that practiced with tasks unrelated to WM. Overall results showed that the training produced a specific improvement in a WMU task similar to the one used in the training, prompting a better performance in both the trained groups (adaptive and non-adaptive) with respect to the control. This means that the gains were due to the training and not to the effects of test–retest practice. There was also some evidence of maintenance effects a month after completing the training.

### Transfer effects

The training produced a specific improvement in one of the updating tasks (Odd–Even Number Updating), which measured nearest transfer effects, and this benefit persisted

a month later. This task has the same structure as the tasks used in the training (and the Number Comparison Updating task in particular), although it differs in terms of the updating criterion, so it loads the trained WMU components. For each item, participants had to retrieve an element from their WM, decide whether or not to substitute the information based on a specific criterion and, where appropriate, update the element. As in the training tasks, participants were cued about the element they needed to retrieve, focus their attention on, and possibly substitute, by means of boxes on the screen that identified each of the elements maintained in WM.

Unlike the previous updating task, the Number Comparison Updating task—assumed to measure nearest transfer effects—revealed no training effects. This task is structurally dissimilar to the training tasks because participants had to identify the element to retrieve, and then substitute it if necessary, but the element was not indicated as in the Odd–Even Number Updating task.

An additional difference between the two WMU assessment tasks was that only the Odd–Even Number Updating varied in memory load, similar to the trained tasks. Such a pattern of results seems to suggest that training with different loads leads to specific improvement in the maintenance component of updating. In this case, an effect would be easier to detect with the Odd–Even Number versus Number Comparison Span task, which has a fixed load. However, as discussed in the following, the lack of a transfer effect to the other WM measure—the Operation Span task (which varies also in memory load)—indicates that the gains observed cannot be attributed to improvement in the maintenance component per se.

These results may thus mean that the training had more effect on the performance of the task itself than on the WMU process. The enhanced efficiency identified could be mediated by strategy use. In other words, participants may have acquired a strategy suited to the tasks used during the training and then applied it to the Odd–Even Number Updating task (which shared the same characteristics as the training tasks). Improvements due to strategy use are generally task-specific, however, and lead to a narrow effect (von Bastian and Oberauer 2014). However, we have no information concerning strategy use, including the type of strategy, by our participants. Future studies should thus purpose to investigate this issue more directly, by assessing the strategy used, for example.

It is also worth mentioning that the training included two different WMU tasks, and therefore, it may be difficult to disentangle the specific contribution of each one to the benefits observed in training gain. Given that the training tasks were structurally similar, they may have conferred equivalent benefits on the structurally similar WMU task. Common strategies used during training could be applied to this

transfer task. At the same time, the training tasks differed in the specific numerical operations used to determine the updating criterion, as well as in the numerical stimuli used. This training variability may have produced the nearest transfer effect by fostering more flexible approach to tasks similar in terms of processes involved. Further research should improve our understanding of transfer moderators in WM training studies.

The absence of specific gains in performance in the Number Comparison Updating task suggests that the WMU training tasks did not influence such underlying processes, such as access to information, substitution of information, or inhibition of irrelevant elements. These processes are shared by all the updating tasks used here, so improvements should have been seen in both WMU tasks if the related processes had really been affected by the training, whereas this was not the case.

There may be other explanations for the lack of generalization to Number Comparison Updating task, however. It could be that participants opted for a passive strategy in performing this task, i.e., they might wait until the end of the sequence of numbers and then retrieve all the elements presented from their long-term memory and select the three smallest ones. Different studies have shown that this passive strategy may sometimes be adopted in updating tasks, such as the running task (Botto, Basso, Ferrari and Palladino 2014; Bunting, Cowan and Saults 2006; Palladino and Jarrold 2008; Ruiz, Elosúa and Lechuga 2005).

Other measures used to assess transfer effects showed no specific improvements after the training. For instance, no specific gains were observed for the Operation Span. A number of WMU training studies also failed to identify any transfer effect on this type of complex measure of WM, which involves both retaining and processing information (Dahlin et al. 2008b; Jaeggi et al. 2008; Lilienthal et al. 2013; Redick et al. 2013). Similarly, the absence of specific far transfer effects to the fluid intelligence measures is consistent with numerous other studies (Dahlin et al. 2008b; Küper and Karbach 2015; Redick et al. 2013; Salminen et al. 2012; Waris et al. 2015; although see; Jaeggi et al. 2008).

Concerning the corollary aim of our studies, that comparison of two different training regimens (adaptive and non-adaptive), performance in the Odd–Even Number Updating task improved independent of the training regimen in comparison with the active control group. There was no evidence to suggest that the adaptive training was superior, and this is consistent with the findings of aging studies (Karbach and Verhaeghen 2014). It is worth bearing in mind, nevertheless, that the effect size was numerically larger for the adaptive group than for the non-adaptive group, not only in this task but also in the Operation Span task; this difference was maintained even after 1 month.

Although the confidence intervals substantially overlap, this result seems to indicate that adaptive training promotes improved performance in the Operation Span task, which was not seen with the non-adaptive treatment. Such a result suggests that this benefit may be attributable to the adaptive regimen, but further confirmation is required.

All in all, these findings seem to indicate that the training might be moving in the right direction, possibly producing this specific transfer effect after more training sessions. In addition, a novel finding of the present study lies in that, whatever the training regimen used, WMU can be specifically trained and improved in some degree at least in tasks with the same structure.

Apart from the previously-described specific effect, no other transfer effects came to light. The performance of the three groups involved in this study improved in several tasks, however, and these gains were maintained after a month. To be specific, improvements were found in both the WMU tasks (Odd–Even Number Updating and Number Comparison Updating), in the working memory task (Operation Span) and in a fluid intelligence measure (Cattell). These nonspecific effects of the training may be due to practice and also to participants' familiarization with the experimental setting during the training sessions, which could make them better disposed to dealing with the tasks administered at the posttest session. These findings highlight the importance of always including in the same training setting an active control group involved in different tasks from those used in the training. As for the specific tasks used in the present study, Tetris can be considered mainly as a visuospatial task, while crosswords principally involve searches of long-term memory. It could be argued that the games used for the active control group, as with most of the tasks, required a degree of working memory; however, neither game appears to load specifically on WM resources. In addition, it might also be a good idea to include a passive control group to ascertain whether any gains seen in the active control group are due to the tasks performed or to other factors, such as the assessment setting.

### Strengths and weakness of the present study

This training study has the novelty of having analyzed the effect of different cognitive demands imposed by a WMU training in young adults by considering two training regimens, adaptive and non-adaptive. The study also included WMU tasks not previously used in training studies. The results show some evidence of transfer—albeit limited to a specific nearest transfer effects—and some maintenance effects. To complete these tasks correctly, participants had to access information retained in their memory, apply the

necessary comparison or other operation, and substitute the previously-stored number with the resulting new number.

It is worth emphasizing that performance in a specific updating task could be improved with a limited number of sessions and a total of 2 h of training. In the present study, a steady increase in performance was observed only until the third session. It is possible that the amount of improvement between the third and the fourth sessions was not large enough to induce significant differences in the measures used.

The results obtained in the present WMU training study are in line with other reports on short-lived WM training studies that were also successful. For example, Van der Molen, van Luit, van der Molen, Klugkist and Jongman (2010) provided training on complex WM tasks for less than 2 h and found improvements afterwards in a simple arithmetic task, a story recall task and a visual WM task. More recently, Küper and Karbach (2015) also found some gains after five 30-min sessions in which participants practiced with two versions of the *n*-back task. The present training regimen, although short in terms of the number of sessions, may have promoted a method for completing the WM updating task, in the training tasks, in a more flexible way.

The present study underscores the importance of including different tasks to assess near transfer effects. This enables specific information to be obtained on the potential mechanism underlying any such effects. It would be interesting to analyze whether other WMU training methods can improve the process itself, or whether the benefits are limited to the generalization of the strategy used, since we did not examine what types of strategy were used by our participants. As argued by Chein and Morrison (2010), moreover, it would be useful to include at least two tasks loading the processes whose transfer we wish to assess to ensure that any effects identified are due to the process itself and not to the characteristics of the task.

Finally, no far transfer effects were found in this study. Using a different WMU training procedure (e.g., Jaeggi et al. 2008), other authors identified benefits in fluid intelligence, as assessed with Raven's matrices. The current results point toward the importance of understanding how to improve the efficacy of the training regimen to be presented to produce more profound changes, i.e., in terms of transfer gains. A clear result of the present study is that the adaptivity of the training is not a key aspect; however, an aspect to consider more closely is the interval between sessions, and their intensity: Jaeggi et al. (2008), for example, reported transfer to fluid intelligence after daily training over 12 sessions.

A possible limitation of the present study was that the training tasks were numerical, whereas the games used for the active control group included verbal and visuospatial

components. The difference in material types may have led to underestimation of the effect of training in the fluid intelligence tasks comprising visuospatial stimuli. Future studies should assess whether WMU training with other types of stimuli (e.g., spatial) produces equivalent results, for instance, in terms of learning a strategy.

## General conclusion

To conclude, this study examined the transfer effects of a WMU training program under two regimens, one adaptive, and the other with a fixed level of difficulty. Although some other studies used updating tasks, few of them focused on training this process alone. Our training program produced a specific transfer effect on a numerical updating task that was similar to the one used in the training, in terms of the requirements of the task, and this effect was maintained a month later. Even though neither near nor far transfer effects were found, this study shows that WMU training can improve performance in other structurally similar WMU tasks, irrespective of whether an adaptive or fixed-difficulty training regimen is adopted, probably via strategy learning.

**Acknowledgements** This study was supported in part by a grant from the Spanish Ministry of Economy and Competitiveness (PSI2012-37764).

## Compliance with ethical standards

All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

**Conflict of interest** The authors declare that they have no conflict of interest.

**Informed consent** Informed consent was obtained from all individual participants included in the study.

## References

- Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuhl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: A meta-analysis. *Psychonomic Bulletin and Review*, 22(2), 366–377.
- Belacchi, C., Carretti, B., & Cornoldi, C. (2010). The role of working memory and updating in coloured raven matrices performance in typically developing children. *European Journal of Cognitive Psychology*, 22(7), 1010–1020.
- Borella, E., Carretti, B., & Pelegrina, S. (2010a). The specific role of inhibition in reading comprehension in good and poor comprehenders. *Journal of Learning Disabilities*, 43(6), 541–552.
- Borella, E., Carretti, B., Riboldi, F., & De Beni, R. (2010b). Working memory training in older adults: Evidence of transfer and maintenance effects. *Psychology and Aging*, 25(4), 767–778.

- Borella, E., & de Ribaupierre, A. (2014). The role of working memory, inhibition, and processing speed in text comprehension in children. *Learning and Individual Differences, 34*, 86–92.
- Botto, M., Basso, D., Ferrari, M., & Palladino, P. (2014). When working memory updating requires updating: Analysis of serial position in a running memory task. *Acta Psychologica, 148*, 123–129.
- Bunting, M., Cowan, N., & Sauls, J. S. (2006). How does running memory span work? *The Quarterly Journal of Experimental Psychology, 59*(10), 1691–1700.
- Carretti, B., Borella, E., Zavagnin, M., & De Beni, R. (2012). Gains in language comprehension relating to working memory training in healthy older adults. *International Journal of Geriatric Psychiatry, 28*(5), 539–546.
- Carretti, B., Cornoldi, C., De Beni, R., & Romanò, M. (2005). Updating in working memory: A comparison of good and poor comprehenders. *Journal of Experimental Child Psychology, 91*(1), 45–66.
- Carretti, B., Cornoldi, C., & Pelegrina, S. L. (2007). Which factors influence number updating in working memory? The effects of size, distance and suppression. *British Journal of Psychology, 98*(1), 45–60.
- Cattell, R. B., & Cattell, H. E. P. (1963). *Measuring intelligence with the culture fair tests*. Champaign: Institute for Personality and Ability Testing.
- Chein, J. M., & Morrison, A. B. (2010). Expanding the mind's workspace: training and transfer effects with a complex working memory span task. *Psychonomic Bulletin and Review, 17*(2), 193–199.
- Chen, T., & Li, D. (2007). The roles of working memory updating and processing speed in mediating age-related differences in fluid intelligence. *Aging, Neuropsychology, and Cognition, 14*(6), 631–646.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale: Lawrence Erlbaum.
- Dahlin, E., Neely, A. S., Larsson, A., Bäckman, L., & Nyberg, L. (2008a). Transfer of learning after updating training mediated by the striatum. *Science, 320*(5882), 1510–1512.
- Dahlin, E., Nyberg, L., Bäckman, L., & Neely, A. S. (2008b). Plasticity of executive functioning in young and older adults: immediate training gains, transfer, and long-term maintenance. *Psychology and Aging, 23*(4), 720–730.
- Dougherty, M. R., Hamovitz, T., & Tidwell, J. W. (2016). Reevaluating the effectiveness of n-back training on transfer through the Bayesian lens: Support for the null. *Psychonomic Bulletin and Review, 23*(1), 306–316.
- Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., DeFries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science, 17*(2), 172–179.
- Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General, 137*(2), 201–225.
- Gernsbacher, M. A., Varner, K. R., & Faust, M. E. (1990). Investigating differences in general comprehension skill. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*(3), 430–445.
- Jaeggi, S., Buschkuhl, M., Jonides, J., & Perrig, W. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, 105*(19), 6829–6833.
- Jausovec, N., & Jausovec, K. (2012). Working memory training: Improving intelligence—changing brain activity. *Brain and Cognition, 79*, 96–106.
- Karbach, J., & Verhaeghen, P. (2014). Making working memory work: A meta-analysis of executive control and working memory training in older adults. *Psychological Science, 25*(11), 2027–2037.
- Küper, K., & Karbach, J. (2015). Increased training complexity reduces the effectiveness of brief working memory training: Evidence from short-term single and dual n-back training interventions. *Journal of Cognitive Psychology, 28*(2), 199–208.
- Lendínez, C., Pelegrina, S., & Lechuga, T. (2011). The distance effect in numerical memory-updating tasks. *Memory and Cognition, 39*(4), 675–685.
- Li, S. C., Schmiedek, F., Huxhold, O., Röcke, C., Smith, J., & Lindenberger, U. (2008). Working memory plasticity in old age: Practice gain, transfer, and maintenance. *Psychology and Aging, 23*(4), 731–742.
- Lilienthal, L., Tamez, E., Shelton, J. T., Myerson, J., & Hale, S. (2013). Dual n-back training increases the capacity of the focus of attention. *Psychonomic Bulletin and Review, 20*(1), 135–141.
- Lövdén, M., Bäckman, L., Lindenberger, U., Schaefer, S., & Schmiedek, F. (2010). A theoretical framework for the study of adult cognitive plasticity. *Psychological Bulletin, 136*(4), 659–676.
- Melby-Lervåg, M., & Hulme, C. (2016). There is no convincing evidence that working memory training is effective: A reply to Au et al. (2014) and Karbach and Verhaeghen (2014). *Psychonomic Bulletin and Review, 23*(1), 324–330.
- Morrison, A. B., & Chein, J. M. (2011). Does working memory training work? The promise and challenges of enhancing cognition by training working memory. *Psychonomic Bulletin and Review, 18*(1), 46–60.
- Noack, H., Lövdén, M., Schmiedek, F., & Lindenberger, U. (2009). Cognitive plasticity in adulthood and old age: Gauging the generality of cognitive intervention effects. *Restorative Neurology and Neuroscience, 27*(5), 435–453.
- Oberauer, K., Wendland, M., & Kliegl, R. (2003). Age differences in working memory—the roles of storage and selective access. *Memory and Cognition, 31*(4), 563–569.
- Palladino, P., Cornoldi, C., De Beni, R., & Pazzaglia, F. (2001). Working memory and updating processes in reading comprehension. *Memory and Cognition, 29*(2), 344–354.
- Palladino, P., & Jarrold, C. (2008). Do updating tasks involve updating? Evidence from comparisons with immediate serial recall. *The Quarterly Journal of Experimental Psychology, 61*(3), 392–399.
- Passolunghi, M. C., & Pazzaglia, F. (2004). Individual differences in memory updating in relation to arithmetic problem solving. *Learning and Individual Differences, 14*(4), 219–230.
- Pelegrina, S., Capodieci, A., Carretti, B., & Cornoldi, C. (2014). Magnitude representation and working memory updating in children with arithmetic and reading comprehension disabilities. *Journal of Learning Disabilities, 48*(6), 658–668. doi:10.1177/0022219414527480.
- Raven, J., Court, J. H., & Raven, J. C. (1977). *Standard progressive matrices*. London: H.K.Lewis.
- Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., Kane, M.J. & Engle, R. W., et al. (2013). No evidence of intelligence improvement after working memory training: A randomized, placebo-controlled study. *Journal of Experimental Psychology: General, 142*(2), 359–379.
- Ruiz, M., Elosúa, M. R., & Lechuga, M. T. (2005). Old-fashioned responses in an updating memory task. *The Quarterly Journal of Experimental Psychology Section A, 58*(5), 887–908.
- Salminen, T., Strobach, T., & Schubert, T. (2012). On the impacts of working memory training on executive functioning. *Frontiers in Human Neuroscience, 6*, 166. doi:10.3389/fnhum.2012.00166.
- Salthouse, T. A., Babcock, R. L., & Shaw, R. J. (1991). Effects of adult age on structural and operational capacities in working memory. *Psychology and Aging, 6*(1), 118–127.

- Sandberg, P., Rönnlund, M., Nyberg, L., & Neely, A. S. (2014). Executive process training in young and old adults. *Aging, Neuropsychology, and Cognition*, *21*(5), 577–605.
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, *3*(4), 207–217.
- Schmiedek, F., Hildebrandt, A., Lövdén, M., Wilhelm, O., & Lindenberger, U. (2009). Complex span versus updating tasks of working memory: The gap is not that deep. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*(4), 1089–1096.
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2010). Does working memory training generalize. *Psychologica Belgica*, *50*(3–4), 245–276.
- Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin*, *138*(4), 628–654.
- Sprenger, A. M., Atkins, S. M., Bolger, D. J., Harbison, J. I., Novick, J. M., Chrabaszcz, J. S., Weems, S. A., Smith, V., Bobb, S., Bunting, M.F. & Dougherty, M. R. (2013). Training working memory: Limits of transfer. *Intelligence*, *41*(5), 638–663.
- Thompson, T. W., Waskom, M. L., Garel, K. L. A., Cardenas-Iniguez, C., Reynolds, G. O., Winter, R., Chang, P., Pollard, K., Lala, N., Alvarez, G. A. & Gabrieli, J. D. (2013). Failure of working memory training to enhance cognition or intelligence. *PLoS One*, *8*(5), e63614.
- Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task-dependent? *Journal of Memory and Language*, *28*(2), 127–154.
- Van der Molen, M., Van Luit, J. E. H., Van der Molen, M. W., Klugkist, I., & Jongmans, M. J. (2010). Effectiveness of computerised working memory training in adolescents with mild to borderline intellectual disabilities. *Journal of Intellectual Disability Research*, *54*(5), 433–447.
- von Bastian, C., & Eschen, A. (2015). Does working memory training have to be adaptive? *Psychological Research Psychologische Forschung*, *80*(2), 181–194.
- von Bastian, C., Langer, N., Jäncke, L., & Oberauer, K. (2013). Effects of working memory training in young and old adults. *Memory and Cognition*, *41*, 611–624.
- von Bastian, C., & Oberauer, K. (2014). Effects and mechanisms of working memory training: A review. *Psychological Research Psychologische Forschung*, *78*, 803–820.
- Waris, O., Soveri, A., & Laine, M. (2015). Transfer after working memory updating training. *PLoS One*, *10*(9), e0138734.
- Westerberg, H., & Klingberg, T. (2007). Changes in cortical activity after training of working memory—a single-subject analysis. *Physiology and Behavior*, *92*, 186–192.
- Willis, S. L., Blieszner, R., & Baltes, P. B. (1981). Intellectual training research in aging: Modification of performance on the fluid ability of figural relations. *Journal of Educational Psychology*, *73*(1), 41–50.
- Xiu, L., Zhou, R., & Jiang, Y. (2015). Working memory training improves emotion regulation ability: Evidence from HRV. *Physiology and Behavior*, *155*, 25–29.