Improving Older Adults' Working Memory: the Influence of Age and Crystallized Intelligence on Training Outcomes

Hering, Alexandra; Meuleman, Ben; Burki, Céline; Borella, Erika; Kliegel, Matthias

How to cite

HERING, Alexandra et al. Improving Older Adults" Working Memory: the Influence of Age and Crystallized Intelligence on Training Outcomes. In: Journal of Cognitive Enhancement, 2017, vol. 1, n° 4, p. 358–373. doi: 10.1007/s41465-017-0041-4

This publication URL: https://archive-ouverte.unige.ch/unige:112044

Publication DOI: 10.1007/s41465-017-0041-4

J Cogn Enhanc (2017) 1:358–373
DOI 10.1007/s41465-017-0041-4

ORIGINAL ARTICLE

Improving Older Adults' Working Memory: the Influence of Age and Crystallized Intelligence on Training Outcomes

A. Hering[1] · **B. Meuleman**[2] · **C. Bürki**[3] · **E. Borella**[4] · **M. Kliegel**[1,5,6]

Received: 31 March 2017 / Accepted: 28 September 2017 / Published online: 10 October 2017

Abstract To counter age-related decline in cognitive abilities, interventions such as working memory trainings have shown some promising results in old age. Yet, findings are mixed and there is enormous interindividual variability in training and transfer effects. Thus, it is still an open question which person-specific factors may moderate training and transfer effects in working memory interventions in older adults. The present study investigated this issue in the context of an established verbal working memory training. Eighty-eight participants (age range 60–82 years) performed either four sessions of the Borella et al. (*Psychology and Aging 25*(4):767–778, 2010) working memory training or of a visual search training as active control condition or belonged to a passive control group. Before and after the training, participants performed a test battery to assess different cognitive abilities and everyday competence. Furthermore, we included questionnaires on personality factors and intrinsic motivation as possible covariates. The training group showed a substantial training gain in the working memory criterion task that was not found in the active control group. Furthermore, only participants of the working memory training showed also near transfer to another working memory task. No far transfer effects including everyday competence emerged. In terms of possible moderators, age and crystallized intelligence influenced the training and transfer gain in the training group. In conclusion, our results showed that working memory can be improved in older adults and improvements transfer to a non-trained working memory task. However, person-specific factors have to be considered to understand who benefits from the training and why.

Keywords Cognitive training · Older adults · Working memory · Transfer · Prospective memory

✉ A. Hering
alexandra.hering@unige.ch

1 Faculté de Psychologie et Sciences de l'Education, Université de Genève, Geneva, Switzerland

2 Swiss Center for Affective Sciences, Université de Genève, Geneva, Switzerland

3 University Center for Medicine of Aging, Felix Platter-Hospital, Basel, Switzerland

4 Department of General Psychology, University of Padova, Padova, Italy

5 Center for the Interdisciplinary Study of Gerontology and Vulnerability, Université de Genève, Geneva, Switzerland

6 Swiss National Center of Competences in Research LIVES–Overcoming Vulnerability: Life Course Perspectives, Lausanne and Geneva, Switzerland

Introduction

Calculating the tip at a restaurant, following the message of a long paragraph in the newspaper, or adhering to a medication plan describe all typical everyday tasks that involve working memory. Working memory, that is the ability to maintain and manipulate information in a short term storage (e.g., Baddeley 2003), shows substantial decrements with increasing age across the adult lifespan (e.g., Park et al. 2002) and—at the same time—is highly relevant for everyday functioning in older adults (Tomaszewski Farias et al. 2009). Cognitive trainings aim to address this discrepancy by improving basic cognitive abilities such as working memory (e.g., Morrison and Chein 2011). Therefore, the main aim of the present study was to investigate the efficiency of working memory training in older adults.

The general assumption guiding cognitive training research is that training of underlying cognitive core functions or processes also stimulates improvements in other related cognitive functions that rely on the trained one (Lövdén et al. 2010). For example, training working memory with the help of an n-back training should not only improve performance in the n-back task but also in other working memory tasks as well as in tasks related to working memory such as reasoning. The training-related improvements in tasks that are different from the trained one but target the same cognitive function are called *near transfer* (e.g., after a verbal working memory training participants show improvements also in spatial working memory tasks), whereas improvements after the training in tasks that target different but related cognitive functions are called *far transfer* (e.g., after a verbal working memory training, participants show improvements in a matrices reasoning task; see Morrison and Chein 2011 for an overview).

There exist several meta-analyses that investigated the efficacy of cognitive trainings in older adults (e.g., Karbach and Verhaeghen 2014; Kueider et al. 2012; Lampit et al. 2014). The results of these systematic reviews indicate that older adults generally benefit from cognitive trainings and improve in the trained tasks. Furthermore, Karbach and Verhaeghen (2014) also demonstrated large near transfer effects but smaller far transfer effects. This general positive view on cognitive trainings in older adults is, however, challenged by meta-analyses that have not found far transfer effects such as improvements in fluid intelligence after working memory training (Melby-Lervag and Hulme 2015; Soveri et al. 2017). These contradictory results stimulated a search on possible moderators on the efficacy of cognitive trainings.

Moderators of Working Memory Training Outcomes

So far, available meta-analyses mainly focused on methodological aspects and investigated factors such as the influence of different types of control groups or the length of the training. In terms of control groups, active control groups usually use an alternative or placebo training (e.g., Redick et al. 2013) or practice the training task without increases in difficulty (e.g., Brehmer et al. 2012). Passive control groups (i.e. no-contact control groups) only receive the pre- and post-test but no treatment in between. Interestingly, while active control groups represent the methodological standard for cognitive training studies (Green et al. 2014)—they control possible retest effects but also motivational influences (e.g., expectancy to benefit from the procedure) or pure treatment effects (e.g., coming regularly to the lab for the training sessions)—recent meta-analyses did not find differences in training or transfer effects depending on the type of control group to which the training group is compared (Karbach and Verhaeghen 2014; Soveri et al. 2017; Weicker et al. 2016).

Furthermore, regarding the influence of the training duration, the number of training hours or sessions did not influence transfer effects (Au et al. 2016; Karbach and Verhaeghen 2014; Soveri et al. 2017). However, Weicker et al. (2016) reported a positive relation between training duration and training outcome specifying that the number of sessions but not the number of hours influences the training benefit.

Taken together, meta-analytic evidence on the efficacy of cognitive trainings, in particular for working memory have identified important methodological moderators. However, the targeted factors focused mainly on design aspects and have so far neglected to some extent individual differences in the recipients of the trainings.

Person-Specific Influences on Working Memory Training Outcomes

In their framework, von Bastian and Oberauer (2013) proposed that the interplay of the training and individual differences influences both training progress and the mechanisms of transfer including factors like age, personality, and motivation or initial cognitive abilities. The question on *who* benefits *why* from a working memory training is still an open question and only a few studies investigated person-specific factors on training outcomes (Borella et al. 2017; Brose et al. 2010; Minear et al. 2016; for an overview see von Bastian and Oberauer 2013). Among the mediators studied so far, participants' age has been identified as an important factor that influences the training outcome. Single training studies suggested that the benefit of working memory trainings decreases with increasing age (e.g., Borella et al. 2017; Brehmer et al. 2012; Heinzel et al. 2014; Zinke et al. 2014). For example, Brehmer et al. (2012) showed larger training improvements and near transfer effects in younger adults compared to older adults after 5 weeks of working memory training. However, in one of the training and transfer tasks, the authors found no age differences. Bürki et al. (2014) also found that younger adults improved more over the working memory training sessions than older adults. The authors further showed that individual differences in fluid intelligence predicted the initial training performance independent of age. However, no age differences were found in training gains and near transfer effects. Similarly, meta-analyses (Karbach and Verhaeghen 2014; Soveri et al. 2017) did not find differences regarding training improvement between younger and older adults. However, comparing young-old with old-old adults, Zinke et al. (2014) revealed that the moderating effect of age did emerge within the group of older adults.

In terms of other moderators, surprisingly little research has been done on possible motivational factors. Brose et al. (2010) showed a positive association between motivation and working memory performance over the course of a working memory training; however, the link was reduced in older adults compared to younger adults. They used the Intrinsic Motivation Inventory (Ryan and Deci 2000) and measured subjective enjoyment and effort on every training day. Furthermore, the authors found less variability in motivation ratings in older adults. Brose et al. (2010) focused on the training performance only and did not assess the influence on transfer effects. It is still an open question how motivational factors influence transfer effects in older adults.

To investigate the influence of personality on training benefits, some studies examined the influence of need for cognition on training benefits. Need for cognition describes a personality variable that characterizes the enjoyment when doing cognitively challenging activities or tasks (Cacioppo and Petty 1982). One would expect that people with a higher need for cognition should be more engaged in a cognitive training and benefit more from it. In fact, studies have not yet found a relation between need for cognition and training performance or transfer effects in younger adults (Jaeggi et al. 2014; Minear et al. 2016). However, there is some initial evidence in older adults. Carretti et al. (2011) could show that maintenance of the training benefit following a strategy memory training was predicted by need for cognition in older adults. Focusing on general personality traits such as the Big Five (e.g., Costa and McCrae 1992) showed only small relations between training gains and different traits. Studer-Luethi et al. (2012) reported that participants with high scores on neuroticism showed reduced transfer effects after a working memory training, whereas participants with higher scores on conscientiousness showed greater training benefits. However, a more recent study could not follow up on these results and found no associations between the Big Five and working memory training gains (Minear et al. 2016). In summary, empirical evidence on the influence of person-specific variables such as motivation or personality is still scarce and mainly limited to studies on younger adults.

Influences of Working Memory Trainings on Everyday Life

Besides understanding the mechanisms behind training gains and transfer effects an important long-term aim of cognitive training research is to improve everyday life functioning. So far, the influence of training outcomes is sparsely investigated and less than a handful studies report transfer effects to everyday life or everyday-like tasks in older adults. The studies found far transfer effects to tasks that mimic everyday problems (e.g., household management; Cantarella et al. 2017) or to everyday skills as reading comprehension (Carretti et al. 2013). Furthermore, participants who received a working memory training reported less memory problems in everyday life after the training compared to the control group (Brehmer et al. 2012).

Although these results seem promising that cognitive trainings indeed do transfer to everyday skills, none of the studies directly investigated possible improvements in the older adults' everyday life. However, in a training study on prospective memory (i.e., the ability to remember delayed intentions) in older adults, Rose et al. (2015) showed not only far transfer to a laboratory-based task on instrumental activities of daily living but also to a naturalistic prospective memory task assessed at the home of the participants. Participants had to remember to call the experimenter after a certain amount of time while they were engaged in the everyday routine at home.

The Present Study

The present study aimed to understand which factors and variables influence the efficacy of a working memory training in older adults. We chose a working memory training that has been applied successfully in older adults and that had shown not only training improvements in working memory but also near and far transfer effects to other working memory tasks, fluid intelligence, and executive functions (Borella et al. 2010; Cantarella et al. 2017; Carretti et al. 2013).

To address our main research aim, we identified three subgoals that aimed (1) to replicate an established training procedure in older adults, (2) to identify individual differences that influence the training outcome, and (3) to explore potential improvements in everyday functioning.

In a first step, we applied an existing working memory training in healthy older adults and compared it to an active control group, where participants trained a visual search task as well as a passive control group, where participants only came for the pre- and post-assessment. We expected to replicate the earlier findings by Borella et al. (2010) and to show greater training gains and transfer effects as both control groups.

For the second subgoal, we included a set of questionnaires to assess person-specific factors that might be implied in training changes in older adults. The aim was to understand potential individual difference variables that might influence the training and transfer effects and thus to answer the question who benefits from the working memory training and why. We assessed factors that were mainly investigated in young adults, such as intrinsic motivation and need for cognition, to examine their impact in older adults. Additionally, we included specific questionnaires on lifestyle factors of older adults like their physical and social activities.

Finally, as a third step, we investigated potential transfer to everyday-like tasks in the laboratory and outside the lab. One criticism of previous training studies is the lack of transfer to everyday life functioning. Of course, the benefit of cognitive trainings is also measured in respect to the actual improvements outside the laboratory or training environment in real-life of the participants. Therefore, we included the previously used everyday problems test and a naturalistic prospective memory task, where participants had to call back the experimenter at specific times during the day. As both tasks showed promising results in previous studies (Cantarella et al. 2017; Rose et al. 2015), we expected a benefit for the training group in both tasks compared to the control groups.

In the working memory training, participants were listening to lists of five words and their task was to remember a target word of each list. However, to increase working memory demands and reduce rehearsal, participants had to do a secondary task. Whenever participants were listening to an animal word, they had to press the space bar. The training difficulty was manipulated in terms of the position of the target word within the lists and the frequency of the animal words presented. Importantly, the secondary task resembles a prospective memory task, thus we investigated also transfer effects to prospective memory tasks. We assessed prospective memory with a laboratory task as far transfer measure because prospective memory involves also other cognitive abilities besides working memory (e.g., episodic memory, task switching; see Kliegel et al. 2011; Kliegel et al. 2000; Schnitzspahn et al. 2013). Moreover, prospective memory is also a cognitive ability that predicts everyday functioning and quality of life in older adults (Woods et al. 2015; Woods et al. 2012). Therefore, we included a naturalistic measure to investigate possible transfer of our training to everyday-like prospective memory.

Methods

Participants

Eighty-eight healthy older adults participated in the present study and were randomly assigned to three different experimental groups. Twenty-nine older adults aged 60 to 82 years (4 males, 25 females) received the working memory training based on an adaptation of the categorization working memory span task (*CWMS group*), 29 older adults aged 60 to 74 years (7 males, 22 females) received an alternative training regime based on a visual search task (VS *group*) and finally, 30 older adults aged 60 to 78 years (11 males, 19 females) participated only for the pre- and the post-test (*control group*) with on average of 16 days between the two sessions. The three groups did not differ regarding age, years of education, general cognitive abilities (assessed with the French version of the telephone interview for cognitive status, F-TICS-m; Vercambre et al. 2010), and crystallized intelligence (assessed with the Mill Hill vocabulary test; Raven et al. 1996). Table 1 gives an overview of the demographic characteristics for the three experimental groups. All participants reported normal or corrected to normal vision and hearing and were native-French speakers or have spoken French in their everyday life for 10 or more years. Furthermore, we included only participants who were not taking psychotropic medication (e.g., anti-depressant), had no neurologic or psychiatric treatment within the last 5 years, and received a higher score than 26 in the F-TICS-m. All interested volunteers, who fulfilled our inclusion criteria, signed an informed consent prior to study participation. The study was approved by the local ethics committee of the University of Geneva.

Material

Participants performed a battery of tasks before (pre-test) and after the training (post-test). We included three different measures of working memory representing near transfer tasks. For far transfer, we included tasks on cognitive inhibition, speed, cognitive flexibility, reasoning, prospective memory (laboratory task and call-back task), and everyday competence. Furthermore, participants answered questionnaires on intrinsic motivation, need for cognition, lifestyle factors, and their subjective impression of change after the training.

Criterion Task

Categorization Working Memory Span Task The Categorization Working Memory Span task (CWMS; Borella et al. 2008; Borella et al. 2010) was was used as criterion task for the training progress. We used the adapted version by Borella et al. (2010) and translated it into French. The task consisted 20 word lists with each list containing five words of high to medium word frequency (e.g., house, mother, dog, word, night). The lists were organized in sets of different list lengths ranging from two to six lists. Furthermore, the lists contained zero, one, or two animal nouns, presented in any position. In total, the task consisted of 100 words presented in 20 lists, and 28% of the words were animal words.

The task was programmed and presented on a computer using Eprime 2.0 software (Psychology Software Tools, Pittsburgh, PA). The words were audio-recorded and presented with a rate of 2 s per word. The task was to memorize the last word of each list. An acoustic tone signaled the end of each list and a prompt on the screen indicated when the participant had to recall aloud all memorized words. The experimenter recorded the repeated words by hand. The working memory span increased with each set to a maximum of six words to memorize (maintenance demand). Additionally, whenever participants heard an animal word, they had to press the space bar as secondary task (processing demand). Participants started the task with a set of two lists, followed by a set of three lists and so on until the last set of six lists. Before the experiment started, participants practiced the task

Table 1 Demographic characteristics for the three experimental groups

	CWMS group		VS group		Control group		$F(2, 85)$	p
	M	SD	M	SD	M	SD		
Age[1]	67.28	6.08	67.21	3.63	67.50	4.57	0.029	.971
Education[1]	15.21	4.52	16.31	5.28	18.77	7.53	2.788	.067
General cognitive ability[2]	34.24	3.10	35.79	3.89	35.53	3.39	1.665	.195
Vocabulary[3]	27.14	3.24	27.86	2.99	28.03	4.07	0.549	.580

CWMS categorization working memory span, *VS* visual search

[1] In years

[2] Assessed with the F-TICS-m (maximum = 43)

[3] Assessed with the Mill Hill (maximum = 33)

with two practice trials containing two word lists. Following the procedure by Borella et al. (2010), we used the total number of correctly recalled words (maximum 20). Additionally, we computed the total number of correctly achieved sets (maximum 5), which corresponded more to a working memory span measure similar to the transfer measures (Robert et al. 2009).

Near Transfer Measures

Reading Span The French version of the reading span task was used to assess verbal working memory (Daneman and Carpenter 1980; Delaloye et al. 2008). The task consisted of two parts. During the first part, the semantic judgment, participants had to decide by button press if sentences presented on the screen were semantically correct or not. The first part consisted of four practice trials and 16 experimental trials (one sentence represented one trial).

During the second part, the reading span, participants had to make again semantic judgments for sentences, but they also had to memorize the last word of each sentence. The sentences were presented in sets of varying number ranging from two sentences (i.e., memorizing two words) up to five sentences (i.e., memorizing five words). A white triangle on the computer screen indicated the end of each set and served as prompt for the word recall. In total, the second part consisted of 16 trials with four trials per set. The dependent variable was a score based on the number of correct recalled words divided by the number of trials; the score could vary between 0 and 3.5.

Digit Span Backwards Numerical working memory was assessed using the subtest digit span backwards of the French version of the Wechsler Adult Intelligence Scale (WAIS; Wechsler 2011). The experimenter read aloud sequences of digits of increasing number. Participants had to repeat the sequence in reversed order. The sequences contained two to seven digits, with two trials per level. The number of correct repeated sequences was used as dependent variable (maximum 14).

Corsi Block Span To assess spatial working memory, we used the forward version of the Corsi block span task (Milner 1971). The task was administered by using a wooden board with nine cubes arranged irregularly on the board (see Berch et al. 1998 for a review on the task and graphical depictions of the board). The experimenter pointed at a sequence of cubes that the participant had to repeat. The first sequence started with three cubes. The number of cubes per sequence increased from three to seven, and three trials were performed per level of sequence. The number of correctly pointed sequences was used as dependent variable (maximum 21).

Far Transfer Measures

Digit Symbol Test Processing speed was administered using the digit symbol subtest of the French version of the WAIS (Wechsler 2011). Participants had to copy symbols to the nine digits as fast as possible during 120 s. The combination of symbols and numbers was given on top of the sheet. Number of correct copied symbols was used as dependent variable.

Stroop Test The Stroop test was used to assess cognitive inhibition (Stroop 1935). The test consisted of three parts, each containing 36 items presented on three separate sheets. First, participants had to read aloud as fast as possible the color words red, blue, green, and yellow. Second, participants had to name the four colors from a sheet with colored squares. The third sheet contained the color words red, blue, green, and yellow but printed in different colors (e.g., the word red printed in blue). Participants had to name as fast as possible the color of the ink instead of reading aloud the word. For all three parts, the experimenter recorded the time used. The dependent variable was the difference score for the time in seconds of part 3 minus part 2.

Trail Making Test To assess cognitive flexibility, we used the Trail Making test (Reitan 1979). Participants had to connect first, the numbers 1 to 25 in increasing order (test A) and second, the numbers 1 to 13 and letters A to L in an alternating order (test B). The quotient of reaction times to complete both tests was used as indicator for cognitive flexibility.

Raven Standard Progressive Matrices The Raven standard progressive matrices (SPM) was used to assess fluid intelligence and reasoning (Raven et al. 1996). Participants saw incomplete patterns of abstract shapes and had to find the missing piece among six to eight possible responses. The test consisted of 29 items with increasing difficulty taken from the original version (Raven et al. 1996). Participants had 10 min to work on the task. A similar procedure was reported by Zinke et al. (2014). The dependent measure was the number of correct answers (maximum 29).

Laboratory Prospective Memory Task The laboratory prospective memory task was taken from Hering et al. (2016). The task consisted of two subtasks, the prospective memory task and an ongoing task, where the prospective memory task was embedded. For the ongoing task, participants had to make semantic judgments. They saw two words in the middle of the screen and had to decide if they belonged to the same or different categories. The word pairs appeared in different colors and were presented for 3 s. During the ongoing task, encoding trials appeared that contained two lines of either the letter "C" or "V" in either the color magenta or green (e.g., two lines of C in green). For these encoding trials, participants had to memorize the letter-color combination. The encoding trials appeared for 4 s. For the prospective memory task, participants had to press a button with either the letter "C" or "V" on it, whenever they saw a word pair in the previously learned color (e.g., when a word pair appeared in green, participants had to press "C"). After executing the prospective memory task, participants continued working on the ongoing task until they received a new encoding trial. Participants practiced first the ongoing task on an ongoing task only block of 70 trials followed by the instruction of the prospective memory task and a practice block of 32 trials including two prospective memory encoding trials. The experimental block consisted of 190 trials among them 12 prospective memory encoding trials and 12 execution trials. Between encoding and execution there were six to ten ongoing task trials. As dependent variables we used the accuracy rate for the ongoing task trials (ongoing task score) and the prospective memory execution trials (prospective memory task score).

Transfer Measures to Everyday Performance

Everyday Problems Test The everyday problems test was used to assess everyday functioning in a laboratory-based task. The test was based on the everyday problems test by Willis (1994) and adapted to a Swiss context (for a similar procedure see Cantarella et al. 2017). The test assessed several different domains of everyday life such as finances, household management, shopping, and medication adherence. For example, participants had to read the laundry symbols or calculate the ticket price from a table of public transport. The test consisted of seven everyday problems. The number of correct responses was used as dependent variable (maximum 19).

Call-Back Task The call-back task was used to assess prospective memory in a naturalistic setting in everyday life (for a similar procedure see Rose et al. 2015). Participants had to call back the experimenter at five different times over the next 2 days. The specific times were arranged with participants at the pre-test and post-test. The participants had to leave a message with their name and the time, when they were calling. The average difference between the called time and the intended time was used as dependent variable.

Questionnaire on Subjective Changes in Everyday Life The questionnaire on subjective changes in everyday life was a questionnaire designed to assess the subject feeling of improvement after the training. Participants were asked about 14 everyday situations and how they might have improved after the testings (e.g., Do you have the impression that you can solve puzzle like Sudoku more easily? Can you focus better on reading a text while the TV or radio is on?). The questions were answered on a 5-point Likert scale (1 = no, no difference, 5 = yes, big difference).

Possible Moderators

Mill Hill The Mill Hill was used to assess vocabulary or crystallized intelligence (Raven et al. 1996). Participants had to find a synonym out of six options for a given word. Number of correct answers was used as dependent measure (maximum = 33). The Mill Hill was only assessed once at the post-test.

Questionnaires We also examined several personality and motivational variables using questionnaires. These questionnaires were assessed one time either at the pre- or the post-test.

Need for Cognition Need for cognition, a personality trait that describes how much people enjoy working on cognitively challenging tasks (Cacioppo and Petty 1982), was examined with the French need for cognition scale (Ginet and Py 2000). Participants had to evaluate 19 statements on a 0 to 10 Likert scale (0 = that is absolutely wrong; 10 = that is absolutely right).

Motivation We assessed different types of motivation with two different questionnaires. At the pre-test, participants received a questionnaire on five different types of motivation ranging from amotivation to extrinsic and intrinsic motivation: amotivation, external regulation, introjected regulation, identified regulation, and intrinsic motivation (Ryan and Deci 2000; Vallerand et al. 1989). The questionnaire consisted of 17 items that had to be evaluated on a 5-point Likert scale.

At the post-test, participants received a second questionnaire on motivation. This time, the questionnaire was on intrinsic motivation only assessing the subscales interest/enjoyment, value/usefulness, and perceived choice based on the Intrinsic Motivation Inventory (Deci et al. 1994; McAuley et al. 1989). The questionnaire consisted of 14 items that had to be evaluated on a 5-point Likert scale.

Lifestyle Factors The lifestyle questionnaire assessed the usual frequency of social, physical, and intellectual everyday activities (e.g., cycling, gardening, playing chess, reading, going to the cinema, traveling). The activities were evaluated on a 5-point Likert scale (every day, several times per week, one time per week, one time per month, never). These lifestyle factors are linked to the concept of cognitive reserve (Stern 2002).

Training

CWMS Group The WM training was adapted from Borella et al. (2010) and translated into French with some changes related to duration and presentation format. The present training consisted of four sessions in which participants were trained on modified versions of the CWMS task (note that the original version of the training by Borella et al. contains only three sessions). As for the CWMS task, the basic instructions were to recall the target words and press spacebar whenever an animal noun arose. An acoustic tone signaled, when the participant had to memorize the target word and a prompt on the screen indicated when the participant had to recall aloud all memorized words. The experimenter wrote down the repeated words manually. However, some manipulations to the task were made during the four sessions to favor generalized transfer and limit the development of task-specific strategies (see Borella et al. 2010). The maintenance demand of the CWMS task was manipulated by increasing the number of words to be recalled in the case of success, and presenting the lowest memory load in the case of failure (first training session). Moreover, the task instructions were varied, requiring recall of (i) the last or first word of each series (first training and third training session) and (ii) words that were followed by a signal tone at any place within the list (second training session and fourth training session). The processing demand, pressing spacebar for animal words, was also manipulated by varying the frequency of these animal words in the lists (second training session and fourth training sessions). There were no instructions to use specific strategies and no feedback was provided.

In the first training session, the word lists were grouped into sets of different lengths of lists from two to five word lists. For the lengths of two to four lists, there were three different sets (i.e., trials) and for the length of five word lists there were two different sets (i.e., trials). The first session comprised three parts: (1) participants had to recall the last word of the word lists, (2) they had to recall the first word of the word lists, and (3) they had to recall again the last word. Participants started the first part with the length of two word lists and had thus to recall two words. The length increased to the next level (i.e., length of three word lists), if the participant performed at least two sets correctly. The length was increased until a length of five word lists. If the participant did not succeed, the part ended and the participant continued with the second part. For the second and third part, the same procedure was applied until all three parts were finished. Each part consisted of 185 words with 20% animal words.

In the second training session, the CWMS task consisted of four sets of word lists per level of length ranging from two to five. Two sets of each length contained only few animal words, whereas the other two sets contained many animal words. Furthermore, participants had to remember the word that was followed by a signal tone and not at specific position. The task difficulty was not adjusted to participants' performance; all participants had to work on all sets. Of the total number of words (280) included in the task, either 13% (few animals) or 32% (many animal words) were animal words.

In the third training session, the CWMS task consisted of four sets of word lists per level of length ranging from two to five. This time, participants had to recall alternately the words in either the last or the first position of each list. The task difficulty was not adjusted to participants' performance; all participants had to work on all sets. Of the total number of words (280) included in the task, 20% were animal words.

In the fourth training session, the procedure of session two was repeated with a new set of word lists. The whole training procedure was presented using the experimental programming software Eprime 2.0 (Psychology Software Tools, Pittsburgh, PA).

Visual Search Group The active control group performed an alternative training regime that was independent of working memory. We used the adaptive visual search (VS) task by Redick et al. (2013), that has proven to be an appropriate alternative control task for a working memory training. During the task, participants had to detect the letter "F" and indicate the direction it is facing by pressing the corresponding left or the right arrow key. The target letter was presented with an array of two distractor letters (E, T). The VS array was presented for 1500 msec, followed by a mask screen of 2000 msec and a fixation cross of 500 msec until the next visual search array appeared. The letter F and the distractors were arranged as a grid. Increasing the number of distractor letters increased the difficulty level of the visual search task. The participants started with a 2 × 2 grid. The next grid size was 4 × 4—adding two columns and rows and so forth. Furthermore, the array of distractors varied: in uniform trials, there were only one type of distractor or the distractors were facing all in the same direction, whereas in mixed trials, the distractors and their orientation were mixed. The first trial level consisted of uniform trials and started with three distractors and the target arranged as a 2 × 2 grid. The second trial level consisted of mixed trials in a 2 × 2 grid followed by a uniform 4 × 4 grid (third level) followed by a mixed 4 × 4 grid (fourth level) and so forth. If the participants succeed the level with 87.5% or more accuracy rate, the difficulty increased, if the participants performed between 75 and 87.5%, the same level was repeated and if the participants scored below 75%, the difficulty level decreased. Each level consisted of 24 trials.

The VS group performed four sessions like the CWMS group. During the first session, the task was explained to the participants and the first three levels were practiced. Afterwards, the participants performed the task again starting with the first level until they finished all 15 blocks. During the training sessions two to four, the participants of the VS group performed the visual search task directly without further practice.

Procedure

Participants of all three groups started the testing with signing an informed consent. The first session was the pre-test that lasted around 2 h with breaks. During the pre-test, all participants performed the following tests and questionnaires: pre-test motivation questionnaire, Corsi block span, Trail Making test, CWMS task, laboratory prospective memory task, everyday problems test, Stroop test, digit span backwards, digit symbol test, reading span, Raven SPM, need for cognition questionnaire, and lifestyle questionnaire. At the end of the pre-test, participants received instructions for the call-back task that was administered on the following 2 days.

In the following 2 weeks, the four training sessions for the CWMS group and the VS group took place at the university laboratory (mean number of days between pre- and post-test M = 14.93 days; SD = 3.40). In both groups, the training sessions lasted around 45 min. In both training regimes, the sessions were around 2 to 3 days apart from each other (M = 2.61 days; SD = 1.47). The control group was neither invited to university nor contacted between pre- and post-test (mean number of days between pre- and post-test M = 16.23 days; SD = 5.33).

During the post-test, all participants were tested again on the following tests and questionnaires: Mill Hill test, Corsi block span, Trail Making test, CWMS task, laboratory prospective memory task, everyday problems test, Stroop test, digit span backwards, digit symbol test, reading span, Raven SPM, post-test questionnaire on motivation, and questionnaire on subjective change. At the end of the post-test, participants received again instructions for the call-back task. For the CWMS task, the two prospective memory tasks, the everyday problems test, the reading span, and the Raven SPM parallel versions were used at pre- and post-test to minimize task-specific practice effects. However, for the digit span and the Corsi block span we used the same task. Half of the participants started with one version at pre-test, the other half with the second version.

Data Analysis

The data analysis consisted of three steps: in a first step, the training benefit and transfer effects were analyzed; in the second step, a principal component analysis was run to identify meaningful groups of tasks with respect to training benefits; in a third step, the moderating effect of covariates on training benefit was examined by a stepwise interaction selection. Step 1 constituted the primary inferential analyses for these data, whereas steps 2 and 3 constituted a more exploratory follow-up.

For the first step, mixed ANOVAs were conducted with group as between factor (3: CWMS group, VS group, control group) and time as within factor (2: pre-test, post-test) for the criterion task and the different near and far transfer tasks. The main focus of the analyses was on the interaction effects of group by time. Significant interactions were explored using post-hoc t tests for contrasts between the groups. Post hoc tests were corrected for multiple testing using the Bonferroni adjustment: for the subsequent t tests, all p values were multiplied by the number of comparisons made (indicated as p_{adj}).

In the second step, we used principal component analysis (PCA) to identify meaningful groups of tasks in terms of training benefit. To do this, we first computed a matrix of difference scores for each task (i.e., post-test minus pre-test) to summarize the "training progress" effect. Next, we conducted Horn's parallel analysis (Horn 1965) to determine the number of meaningful principal components (PCs) in the data. This criterion compares the observed eigenvalues of the PCs against simulated eigenvalues obtained by repeated data permutations. Components are only considered meaningful when their eigenvalue exceeds the average eigenvalue of the same component in random data. Once the number of PCs was identified, we ran the final PCA, using an oblique promax rotation to preserve any correlation between different task groups. The obtained scores on these PCs were then submitted to a group (3: CWMS group, VS group, control group) MANOVA, testing whether overall progress in groups of tasks was significantly different in the three training groups. When significant, univariate ANOVAs of group on PC score were conducted by each PC score separately. PCs that were identified as showing significant group differences were then retained for the next analysis step.

Finally, in the third step, we investigated the moderating influence of demographical variables and questionnaire covariates on the group by training progress effect (see Tables 1 and 2 for an overview of covariates). We included the following covariates: age, F-TICS, education, Mill Hill, need for cognition, subjective changes in everyday life, motivation, and lifestyle factors. For each progress PC that had shown a significant group effect in the previous analysis, we conducted a forward stepwise selection of interaction effects. At each step of the algorithm, the model selected the effect that most reduced the *generalized R-squared* (GR^2) criterion. This criterion selects effects according to their predictive strength and is approximately equivalent to selecting effects by confirmatory cross-validation (Friedman 1991; Hastie et al. 2009). We chose this strategy because of its reliance on effect size rather than significance and, as such, avoided multiple testing issues associated with stepwise regression using *p* values as the decision criterion. For the final selected interactions, we report inferential statistics, however.

One task measure, call-back task performance, was removed entirely from analysis steps 2 and 3, due to having too many missing values (42 out of 88 subjects). Of the remaining data (without the call-back task), approximately 1.3% were missing. Although this fraction was small, the default of many analyses is to handle missingness by list-wise deletion, which means that every subject with *at least one* missing value (18 out of 88 subjects in this data set) is removed entirely from the analysis. In order to make maximal use of observed data, we therefore replaced all missing values by zeroes. This can be considered as the most conservative imputation strategy, and allowed all 88 subjects to be used for all analyses.

All analyses were run using the R statistical software (version 3.3.3), using the packages "car" (Fox and Weisberg 2011) for general linear modeling, "paran" (Dinno 2012) for Horn's parallel analysis, "psych" (Revelle 2016) for PCA, and "earth" (Milborrow 2017) for automatic interaction selection.

Results

Analyses of Training Outcomes and Transfer Effects

Baseline Performance First, we compared the three groups regarding their baseline performance at the pre-test for the different outcome measures. We conducted a MANOVA with group as between factor and the set of pre-test scores as multivariate dependents. The MANOVA showed no significant group differences at the pre-test, Wilks' Lambda = .597, $F(26, 98) = 1.110$, $p = .346$. Descriptive information for all tasks is given in Table 2.

Training Effects We analyzed the change from pre-test to post-test in the CWMS task as indicator for the training effect. When analyzing the number of correctly recalled words, we did not find a significant interaction of group by time, $F(2, 85) = 2.004$, $p = .141$, $\eta_p^2 = .045$. However, the repeated measures ANOVA for the total number of achieved lists revealed a significant interaction of group by time, $F(2, 85) = 5.079$, $p = .008$, $\eta_p^2 = .107$. Post hoc *t* tests showed, that both, the CWMS group and the control group improved over time and achieved more lists at the post-test, CWMS group: $t(28) = 5.975$; $p_{adj} < .001$; $d = 1.108$; control group: $t(29) = 3.319$; $p_{adj} = .006$; $d = 0.565$. The VS group did not show a significant improvement, $t(28) = -1.907$, $p_{adj} = .201$, $d = 0.372$. Furthermore, comparing the effect sizes of the CWMS group and the control group showed a great difference with larger improvements in the CWMS group indicating a substantial training gain in this group following the working memory training (see Fig. 1a).

Near Transfer Effects To investigate near transfer effects, we compared the change from pre- to post-test for the different working memory measures. We assessed verbal working memory with the reading span task, numerical working memory with the digit span backwards task and spatial working memory with the Corsi block span task. The repeated measures ANOVAs revealed only a significant interaction of group by time for the reading span task, $F(2, 84) = 6.003$, $p = .004$, $\eta_p^2 = .125$. Follow-up *t* tests showed that there was only a significant improvement from pre-test to post-test for the CWMS group, $t(28) = 3.241$; $p_{adj} = .009$; $d = 0.506$, but not for the two other groups, VS group: $t(27) = -.354$; $p_{adj} > .999$; $d = 0.045$; control group: $t(29) = -1.417$; $p_{adj} = .501$; $d = -0.159$; see Fig. 1b. For the other two working memory measures, we did not find specific improvements in the training group over time. For both tasks the interactions of group by time did not turn significant (all $p > .460$).

In summary, we only found near transfer effects for a verbal working memory task in the CWMS group but not to spatial or numerical working memory tasks.

Far Transfer Effects To assess far transfer effects of our working memory training, we investigated pre-post-test changes in measures for processing speed (digit symbol test), cognitive inhibition (Stroop test), cognitive flexibility (Trail Making test), reasoning (Raven SPM test), everyday competence (everyday problems test), and prospective memory (laboratory task, call-back task). In none of the tests, the

Table 2 Descriptive statistics for the training and transfer measures as well as the questionnaires

	Pre-test						Post-test					
	CWMS group		VS group		Control group		CWMS group		VS group		Control group	
	M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Criterion task												
CWMS task—achieved lists	1.97	1.12	1.72	.84	2.00	.91	3.10	.90	2.07	1.00	2.57	1.07
CWMS task—recalled words	13.31	2.90	12.41	3.63	13.70	3.14	15.86	2.52	13.83	2.77	14.77	3.70
Near transfer												
Reading span	2.49	.49	2.57	.55	2.83	.47	2.71	.37	2.61	.50	2.76	.44
Digit span backwards	5.97	2.16	6.07	1.87	6.27	2.52	6.59	1.59	6.59	2.35	6.80	2.46
Corsi block span	8.38	2.09	7.90	1.88	8.40	1.92	8.83	2.36	7.90	1.68	9.03	1.96
Far transfer												
Digit symbol test	63.90	12.01	66.17	12.11	66.27	13.08	64.86	11.54	70.59	12.68	70.33	15.10
Stroop test	20.24	8.27	19.52	8.25	20.33	9.30	18.31	8.06	17.17	6.14	18.30	8.75
Trail-making test	2.25	.90	2.37	.77	2.22	.64	2.04	.60	2.43	.81	2.53	1.42
Raven standard progressive matrices	16.28	4.09	18.10	2.86	18.50	3.34	16.93	4.57	18.28	3.42	19.10	3.58
LPM—ongoing task	.79	.09	.82	.06	.81	.06	.81	.06	.84	.06	.84	.05
LPM—prospective memory task	.45	.28	.43	.29	.51	.31	.52	.33	.57	.32	.76	.25
Everyday problems test	11.81	2.25	11.62	2.28	12.33	1.51	11.88	2.12	12.14	2.07	12.28	2.22
Call-back task	3.14	4.21	18.56	20.23	12.41	21.29	11.72	15.71	13.16	15.31	15.87	20.05
Questionnaires												
Need for cognition	6.35	1.36	6.67	1.00	6.63	1.31						
Motivation												
Amotivation	2.06	.76	2.02	.89	1.99	.81						
External regulation	1.50	.84	1.29	.50	1.45	.61						
Introjected regulation	1.90	.91	1.79	.83	1.86	.70						
Identified regulation	2.99	.92	3.07	.86	2.91	1.04						
Intrinsic motivation	3.60	.84	3.64	.72	3.48	1.01						
Interest/enjoyment							4.06	.85	4.01	.71	4.25	.67
Value/usefulness							3.86	.84	3.57	.98	3.77	.90
Perceived choice							4.83	.48	4.84	.35	4.86	.36
Lifestyle factors												
Social activities							2.98	.64	3.05	.49	3.15	.64
Physical activities							2.95	.72	2.77	.65	3.17	.78
Intellectual activities							2.27	.70	2.26	.61	2.28	.50
Subjective change in everyday life							1.55	.72	1.68	.86	1.22	.48

CWMS categorization working memory span, *VS* visual search, *LPM* laboratory prospective memory task

interaction of group by time reached significance indicating that there was no far transfer in our study (all $p > .083$; see Table 3 for the inferential statistics).

Principal Component Analysis

First, we computed a matrix of difference scores for each individual task (i.e., post-test minus pre-test) to summarize their "training progress" effect. This matrix was submitted to a principal component analysis which explored whether progress between different tasks was correlated. Horn's parallel analysis for component selection indicated that five principal components (PCs) were meaningful. A PCA extracted these five components and rotated the loadings according to an oblique promax rotation (see Table 4). PC1 loaded strongly on the Raven progressive matrices task, the ongoing activity of the laboratory prospective memory task, and the everyday problems test. It can be considered an index of reasoning or problem solving. PC2 loaded strongly on the two criterion task measures (CWMS recalled words and achieved lists), and the reading span task. It can be considered a primary index of the training and of general verbal working memory. The remaining three PCs each loaded strongly on just a single task: the laboratory prospective memory task (prospective memory score) for PC3, the digit symbol test for PC4, and the Trail Making test for PC5.

A MANOVA of group on PC scores indicated that there were significant group differences in at least one of the progress PCs, *Pillai's trace* = .263, $F(10, 164) = 2.478$, $p = .009$. Follow-up univariate ANOVAs with Bonferroni correction indicated that group differences were only significant for PC2, $F(2, 85) = 6.610$; $p_{adj} = .011$, which encoded primarily progress in the training task and some verbal working memory in general, and hence replicated the progress tests in the first analysis step (see Fig. 2). Pairwise comparisons with Bonferroni correction confirmed that the CWMS group differed significantly in PC2 progress from the VS group, $t(85) = -3.197$, $p_{adj} = .006$, and from the control group, $t(85) = -3.107$, $p_{adj} = .008$.

As a follow-up, we also checked for transfer between progress PCs, by running an ANCOVA of group by PC2 on the other PCs. That is, we tested whether group differences in training progress were modified by the magnitude of progress in PC2. However, all ANCOVA interaction tests were found to be insignificant, with all $F < 1$.

Covariate Analyses

In the final step of the analysis, we investigated whether group differences in training progress were modified by demographical characteristics or questionnaire covariates. Since the previous analysis identified only PC2 as having meaningful progress by group, we considered only this PC as the dependent in this analysis. A forward stepwise selection of interaction effects by minimizing of GR^2 revealed three meaningful interaction effects involving group, and one meaningful interaction effect not involving group. The group interactions involved, age, Mill Hill score, and subjective change in everyday life, respectively (see Fig. 2).

The moderation effect of age suggested that, in the CWMS group, older subjects progressed more due to the training, $t(76) = 1.947$, $p = .055$. The moderation effect of Mill Hill score suggested that, in the CWMS group, higher crystallized intelligence predicted less training progress, $t(76) = -3.128$, $p = .002$. Finally, the moderation effect of subjective change in everyday life effect suggested that, in the CWMS group, greater impression of improvement in everyday life was paradoxically associated with less training progress, $t(76) = -2.867$, $p = .005$.

Discussion

The major goal of the present study was to investigate the efficiency of verbal working memory training in older adults. To do so, we formulated three subgoals: (1) investigating the training benefit and transfer effects to cognitive abilities such as reasoning, other types of working memory and different executive functions; (2) investigating which person-specific factors might moderate training or transfer gains; and (3) examining the possible influence of the training to everyday competences and daily life functioning. We chose the training of Borella et al. (2010) because it had shown to be an age appropriate training regime for older adults. We compared performance of our working memory training group (CWMS group) to an active control group working on a visual search training (VS group; Redick et al. 2013) and a passive no-contact control group (control group).

In a first step, we assessed the training gain of the working memory training and possible transfer effects for various cognitive abilities. We included near transfer measures for verbal, numerical and spatial working memory and far transfer measures for reasoning, executive functions (cognitive inhibition and cognitive flexibility) and prospective memory. The results revealed a clear training benefit for the working memory training group. The CWMS group showed a large improvement in

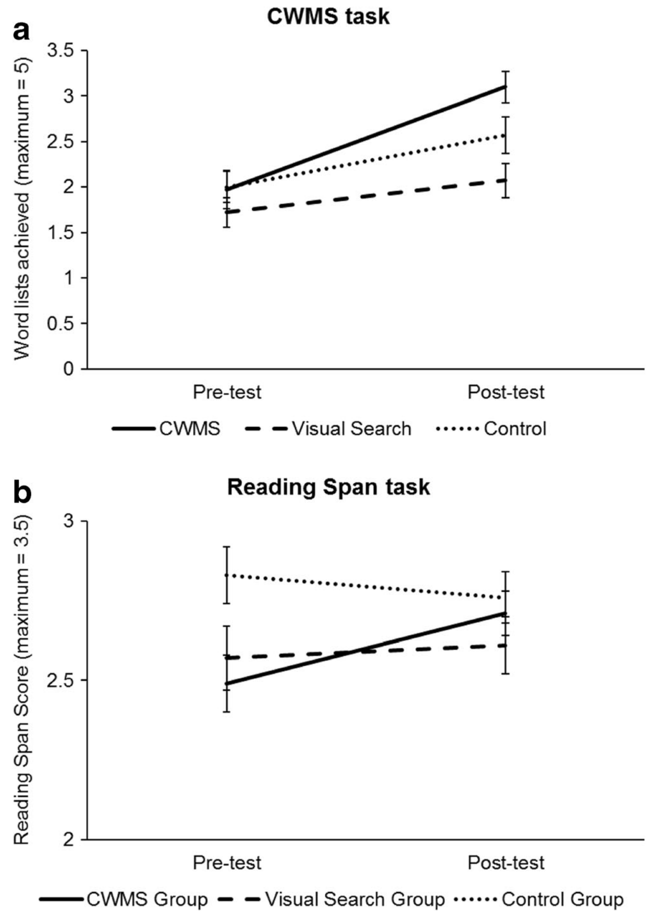

Fig. 1 Pre-post changes for the three experimental groups. **a** Training effect for the CWMS task (criterion task). **b** Near transfer effect for verbal working memory in the reading span task. *Note.* Error bars represent the standard error of the mean. CWMS categorization working memory span

Table 3 Inferential statistics for the near and far transfer measures

	Interaction of group by time			
	F	*df*	*p*	η_p^2
Criterion task				
CWMS task—achieved lists	5.079	2, 85	.008	.107
CWMS task—recalled words	2.004	2, 85	.141	.045
Near transfer measures				
Reading span	6.003	2, 84	.004	.125
Digit span backwards	0.017	2, 85	.983	.000
Corsi block span	0.782	2, 85	.461	.018
Far transfer measures				
Digit symbol test	1.423	2, 85	.247	.032
Trail Making test	1.450	2, 85	.240	.033
Stroop test	.044	2, 85	.957	.001
Raven standard progressive matrices	0.197	2, 85	.821	.005
LPM—ongoing task	0.420	2, 83	.659	.010
LPM—prospective memory task	2.255	2, 85	.084	.057
Everyday problems test	0.322	2, 85	.726	.008
Call-back task	0.503	2, 43	.608	.023

CWMS categorization working memory span, *LPM* laboratory prospective memory task

one of the two measures in the criterion task that was not found in the two other groups. The criterion task was similar to the trained tasks and served as a measure for the training gain. However, the control group also showed improvements from pre- to post-test in this task but to a lesser extent than the CWMS group probably indicating a retest effect. The VS group did not improve at all.

The large training benefit in the working memory training group follows up on earlier findings of studies using the same training regime (Borella et al. 2010; Borella et al. 2013; Cantarella et al. 2017; Carretti et al. 2013). Although the training was relatively short with only four training sessions, trained participants still achieved a substantial increase in performance that was not attributable to pure retest effects as seen in the control group, but to actual gains based on the working memory training. The results on the training progress are also in line with meta-analytic evidence on cognitive trainings in older adults (Karbach and Verhaeghen 2014) and show the general potential of older adults to still be able to improve their working memory.

Regarding the transfer effects, we found a clear near transfer effect to another—not trained—verbal working memory task. Only participants of the CWMS group improved significantly from pre- to post-test in the reading span task, whereas the two other groups did not show significant performance increases. Furthermore, the principal component analyses confirmed this finding and showed a clear component (PC2) that can be interpreted as verbal working memory and was related to the CWMS group only. The near transfer effect shows that the training progress generalized beyond the trained task. However, the reading span task resembles the trained task and could also be classified as "nearest transfer" measure, whereas the working memory measures—that served as near transfer measures—tapped on other modalities such as spatial or numerical working memory. The transfer effect to the reading span task might also indicate transfer of strategic task performance. In general, our transfer measures could be sorted continuously from nearest to far transfer. We did not find other near transfer effects to the other working memory tasks. Furthermore, we did also not find far transfer effects to any of our measures.

The limited findings of transfer were surprising given previous results of studies using the same training regime. In their first study, Borella et al. (2010) showed not only near transfer effects to other working memory tasks but also far transfer to inhibition, reasoning and processing speed. Similarly, in a recent study the authors showed again far transfer to two reasoning tasks after their verbal working memory training in older adults (Cantarella et al. 2017). One reason might be the differences in the control groups used between the studies. In the study by Borella et al. (2010), control participants worked on questionnaires, whereas our active control participants in the VS group worked on an adaptive computerized visual search task similar in general procedure as in the CWMS group. Verbal reports from participants of the visual search training indicated, that they enjoyed the visual search task, which might have encouraged them to make an effort. Furthermore, in general, transfer effects seem to be harder to detect when the training group is compared to treated control groups that worked on alternative training regimes (Melby-Lervag and Hulme 2013).

An alternative explanation for the discrepancies in transfer effects between the different studies leads to our second aim in investigating the influence of person-specific factors. Participants in the studies by Borella and colleagues (Borella et al. 2010; Cantarella et al. 2017) reported around 9 years of education, whereas the participants in our study reported around 16 years. The differences in education could have also influenced transfer effects. Although the number of years of education itself did not turn out to be an influencing factor, our results on possible covariates support this assumption for other demographic variables.

Participants in the CWMS group improved more when they had a lower score on the Mill Hill vocabulary test. It suggests that baseline abilities such as crystallized intelligence play an important role on determining who benefits from a training. Furthermore, we also found an interaction of group by age for the progress in the verbal working memory principal component. Although we tested only healthy older adults, the finding indicated that the older participants in the CWMS group improved more than younger participants among our sample. It supports also other work of Borella and

Table 4 PCA loadings for task progress variables. Loadings > 0.5 are presented in italics. Note that rotation of components can change their order from the unrotated solution

Task progress (post–pre score)	Component loadings				
	PC1	PC2	PC4	PC5	PC3
Criterion task					
CWMS task—achieved lists	− 0.131	*0.855*	− 0.057	0.048	− 0.070
CWMS task—recalled words	− 0.114	*0.796*	0.095	− 0.006	0.172
Near transfer measures					
Reading span	0.247	*0.506*	0.107	− 0.401	− 0.251
Digit span backwards	0.184	0.178	− 0.248	0.383	− 0.222
Corsi block span	0.207	0.252	− 0.117	0.231	0.437
Far transfer measures					
Digit symbol test	0.172	0.026	*0.839*	0.147	0.274
Trail-making test	− 0.056	− 0.026	0.269	*0.883*	− 0.042
Stroop test	.013	− .018	− *.748*	− .235	.345
Raven standard progressive matrices	*0.857*	− 0.019	0.185	0.012	0.089
LPM—ongoing task	− *0.655*	0.101	0.091	− 0.295	0.129
LPM—prospective memory task	0.012	− 0.018	0.090	− 0.090	*0.861*
Everyday problems test	*0.705*	− 0.177	0.010	− 0.286	0.024
Proportion variance explained	0.155	0.147	0.124	0.116	0.108
Cumulative variance explained	0.155	0.302	0.425	0.542	0.650

CWMS categorization working memory span, *LPM* laboratory prospective memory task

collaborators that showed significant training and transfer effects not only in a sample of young-old adults (age range 65 to 75 years; Borella et al. 2010) but also in old-old adults (age range 75 to 87 years; Borella et al. 2013). However, these findings contrast an earlier study that reported smaller training and transfer gains with increasing age in older adults and no influence of crystallized abilities (Zinke et al. 2014); however, Zinke et al. (2014) trained much older adults than our study. The findings of both studies together indicate that age and basic cognitive abilities such as crystallized intelligence might play an important role for training outcomes in younger older adults (see also Borella et al. 2017 for a similar finding).

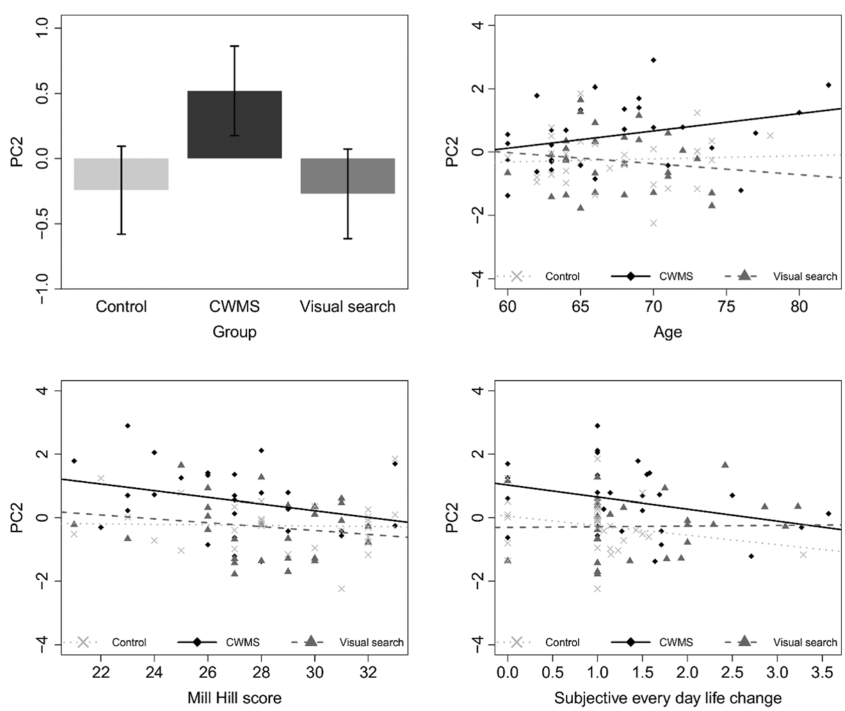

Fig. 2 Main effect of group on PC2 with 95% confidence intervals (top left panel), and moderation effects of group by age, Mill Hill score, and subjective everyday life change on PC2 (remaining panels)

Surprisingly, crystallized intelligence and age were the only covariates that were identified in our model. We also investigated the influence of intrinsic motivation and personality variables like need for cognition but these factors did not influence the training outcome or transfer effects. Brose et al. (2010) found motivational influences on their training even in older adults, although the association was smaller in older adults. In our study, the different aspects of motivation were only assessed once and all three experimental groups reported similar levels of intrinsic motivation, which might have hindered the detection of possible influences, but it is in line with findings that older adults report less variability in the motivation ratings (Brose et al. 2010).

The results on crystallized intelligence and age also fit in the broader context of cognitive reserve. Cognitive reserve describes the potential to compensate for age-related changes (e.g., age-related decline) through (lifelong) intellectual activities and environmental factors (Barulli and Stern 2013; Stern 2002). Our findings suggest that working memory training can be especially beneficial for older adults with reduced preconditions such as lower levels of crystallized intelligence (see also Brom et al. 2014 for a similar finding on fluid intelligence).

The third aim of our study was to investigate the transfer of the working memory training on everyday competence and daily life. We included a laboratory task on everyday competence and a prospective memory task that was implemented in the daily life of the participants. However, both tasks did not show transfer effects. In contrast, Cantarella et al. (2017) could show transfer to the everyday problems test in their study. Again, general differences between the trained samples in the different studies might explain these diverging results. Future studies should investigate these general differences to identify under which conditions people improve following a working memory training. Also the published meta-analyses on working memory trainings focused so far mainly on methodological variables and less on person-specific factors such as crystallized intelligence or education.

Interestingly, we found an interaction between group and the subjective impression of change in everyday life after the training. Participants, who reported improvements in cognitive everyday tasks after the working memory training, showed less training progress in the verbal working memory component. It suggests that participants were not able to accurately estimate their training progress or that motivational expectancy effects biased their evaluation. This finding implies that subjective outcome measures should be treated carefully. More research is needed to objectively assess possible everyday improvements in the daily life of the participants.

There are also some limitations to consider when interpreting the present results. We only assessed one post-test, but did not investigate a follow-up test to examine maintenance effects or effects that may emerge with some delay. Previous studies using the same training regime found maintenance effects for their near and far transfer measures after 6 to 8 months (Borella et al. 2010; Carretti et al. 2013). Moreover, we modified the training procedure and task administration for the CWMS task in comparison to the study by Borella et al. (2010). We added a forth session to have three comparable training sessions. The procedure of the first training session differed from the following sessions and was adapted to the participants' performance whereas the other sessions had a fixed procedure that was the same for all participants. Furthermore, the training task and the CWMS task were administered using a presentation software that also allowed for a reaction time assessment of the secondary task (i.e., pressing spacebar when hearing an animal word). In the original version by Borella et al. (2010), audio-recorded sequences were played with a recorder by the experimenter and the participants had to tap with their hand on the table for the secondary task. It might be that these methodological differences had an influence on the results. More precisely, the changes in task administration of the training could have influenced the motivational engagement and metacognitive awareness of the participants in comparison to the participants in the other training studies (e.g., Borella et al. 2017; Borella et al. 2010), which might have influenced the obtained results. Finally, as mentioned earlier, we had a sample of relatively highly educated participants as compared to Borella et al. (2010), which might have limited the potential for improvement or transfer. However, in a more recent study by Borella et al. (2017), the authors found training and transfer effects in higher educated older adults. Future studies have to clarify the role of education on training outcomes.

In conclusion, our study showed that a verbal working memory training can improve working memory performance in older adults and even result in some near transfer indicating that working memory trainings are valuable tools to improve cognition in old age and thus, work against the general cognitive decline in this age group. However, there is still research needed to understand under which conditions these effects could also result in far transfer effects to other cognitive abilities and to more general effects in everyday life. Our analyses of covariates showed that interindividual differences in crystallized intelligence but also age itself might play an important role. Clearly, more research is necessary to fully investigate who is benefiting why under which circumstances from a cognitive training intervention.

Acknowledgements MK acknowledges the support from the Swiss National Science Foundation (SNSF). We thank Thomas Redick (Purdue University) for providing us the task for our active control training group.

Compliance with Ethical Standards All participants signed an informed consent prior to study participation. The study was approved by the local ethics committee of the University of Geneva.

References

Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuehl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: a meta-analysis. *Psychonomic Bulletin & Review*, 22, 366–377. https://doi.org/10.3758/s13423-014-0699-x.

Baddeley, A. (2003). Working memory: looking back and looking forward. *Nature Reviews Neuroscience, 4*(10), 829–839. https://doi.org/10.1038/Nrn1201.

Barulli, D., & Stern, Y. (2013). Efficiency, capacity, compensation, maintenance, plasticity: emerging concepts in cognitive reserve. *Trends in Cognitive Sciences, 17*(10), 502–509. https://doi.org/10.1016/j.tics.2013.08.012.

Berch, D. B., Krikorian, R., & Huha, E. M. (1998). The Corsi block-tapping task: methodological and theoretical considerations. *Brain and Cognition, 38*(3), 317–338. https://doi.org/10.1006/Brcg.1998.1039.

Borella, E., Carbone, E., Pastore, M., De Beni, R., & Carretti, B. (2017). Working memory training for healthy older adults: the role of individual characteristics in explaining short- and long-term gains. *Frontiers in Human Neuroscience, 11*(99). https://doi.org/10.3389/fnhum.2017.00099.

Borella, E., Carretti, B., & De Beni, R. (2008). Working memory and inhibition across the adult life-span. *Acta Psychologica, 128*(1), 33–44. https://doi.org/10.1016/j.actpsy.2007.09.008.

Borella, E., Carretti, B., Riboldi, F., & De Beni, R. (2010). Working memory training in older adults: evidence of transfer and maintenance effects. *Psychology and Aging, 25*(4), 767–778. https://doi.org/10.1037/A0020683.

Borella, E., Carretti, B., Zanoni, G., Zavagnin, M., & De Beni, R. (2013). Working memory training in old age: an examination of transfer and maintenance effects. *Archives of Clinical Neuropsychology, 28*(4), 331–347. https://doi.org/10.1093/Arclin/Act020.

Brehmer, Y., Westerberg, H., & Backman, L. (2012). Working-memory training in younger and older adults: training gains, transfer, and maintenance. *Frontiers in Human Neuroscience, 6*(63). https://doi.org/10.3389/Fnhum.2012.00063.

Brom, S. S., Schnitzspahn, K. M., Melzer, M., Hagner, F., Bernhard, A., & Kliegel, M. (2014). Fluid mechanics moderate the effect of implementation intentions on a health prospective memory task in older adults. *European Journal of Ageing, 11*(1), 89–98. https://doi.org/10.1007/s10433-013-0288-2.

Brose, A., Schmiedek, F., Lövdén, M., Molenaar, P. C. M., & Lindenberger, U. (2010). Adult age differences in covariation of motivation and working memory performance: contrasting between-person and within-person findings. *Research in Human Development, 7*(1), 61–78. https://doi.org/10.1080/15427600903578177.

Bürki, C. N., Ludwig, C., Chicherio, C., & de Ribaupierre, A. (2014). Individual differences in cognitive plasticity: an investigation of training curves in younger and older adults. *Psychological Research, 78*(6), 821–835. https://doi.org/10.1007/s00426-014-0559-3.

Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology, 42*(1), 116–131. https://doi.org/10.1037//0022-3514.42.1.116.

Cantarella, A., Borella, E., Carretti, B., Kliegel, M., & de Beni, R. (2017). Benefits in tasks related to everyday life competences after a working memory training in older adults. *International Journal of Geriatric Psychiatry, 32*(1), 86–93. https://doi.org/10.1002/gps.4448.

Carretti, B., Borella, E., Zavagnin, M., & De Beni, R. (2011). Impact of metacognition and motivation on the efficacy of strategic memory training in older adults: analysis of specific, transfer and maintenance effects. *Archives of Gerontology and Geriatrics, 52*(3), E192–E197. https://doi.org/10.1016/j.archger.2010.11.004.

Carretti, B., Borella, E., Zavagnin, M., & de Beni, R. (2013). Gains in language comprehension relating to working memory training in healthy older adults. *International Journal of Geriatric Psychiatry, 28*(5), 539–546. https://doi.org/10.1002/gps.3859.

Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13*(6), 653–665. https://doi.org/10.1016/0191-8869(92)90236-I.

Daneman, M., & Carpenter, P. A. (1980). Individual-differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6.

Deci, E. L., Egharri, H., Patrick, B. C., & Leone, D. R. (1994). Facilitating internalization—the self-determination theory perspective. *Journal of Personality, 62*(1), 119–142. https://doi.org/10.1111/j.1467-6494.1994.tb00797.x.

Delaloye, C., Ludwig, C., Borella, E., Chicherio, C., & de Ribaupierre, A. (2008). L'Empan de lecture comme épreuve mesurant la capacité de mémoire de travail : normes basées sur une population francophone de 775 adultes jeunes et âgés. *Revue Européenne de Psychologie Appliquée/European Review of Applied Psychology, 58*(2), 89–103. https://doi.org/10.1016/j.erap.2006.12.004.

Dinno, A. (2012). paran: Horn's test of principal components/factors (Version R package version 1.5.1).

Fox, J. G., & Weisberg, S. (2011). *An R companion to applied regression* (2nd ed.). Thousand Oaks: Sage.

Friedman, J. (1991). Multivariate adaptive regression splines (with discussion). *Annals of Statistics, 19*(1), 1–141.

Ginet, A., & Py, J. (2000). Need for cognition: a French scale for children and its consequences on a sociocognitive level. *L'Année Psychologique, 100*(4), 585–628.

Green, C. S., Strobach, T., & Schubert, T. (2014). On methodological standards in training and transfer experiments. *Psychological Research, 78*(6), 756–772. https://doi.org/10.1007/s00426-013-0535-3.

Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning* (2nd ed.). New York: Springer.

Heinzel, S., Schulte, S., Onken, J., Duong, Q. L., Riemer, T. G., Heinz, A., et al. (2014). Working memory training improvements and gains in non-trained cognitive tasks in young and older adults. *Aging Neuropsychology and Cognition, 21*(2), 146–173. https://doi.org/10.1080/13825585.2013.790338.

Hering, A., Wild-Wall, N., Gajewski, P. D., Falkenstein, M., Kliegel, M., & Zinke, K. (2016). The role of cue detection for prospective memory development across the lifespan. *Neuropsychologia, 93*(Part A), 289–300. https://doi.org/10.1016/j.neuropsychologia.2016.11.008.

Horn, J. L. (1965). A rationale and a test for the number of factors in factor analysis. *Psychometrika, 30*, 179–185.

Jaeggi, S. M., Buschkuehl, M., Shah, P., & Jonides, J. (2014). The role of individual differences in cognitive training and transfer. *Memory & Cognition, 42*(3), 464–480. https://doi.org/10.3758/s13421-013-0364-z.

Karbach, J., & Verhaeghen, P. (2014). Making working memory work: a meta-analysis of executive-control and working memory training in older adults. *Psychological Science*. https://doi.org/10.1177/0956797614548725.

Kliegel, M., Altgassen, M., Hering, A., & Rose, N. S. (2011). A process-model based approach to prospective memory impairment in Parkinson's disease. *Neuropsychologia, 49*(8), 2166–2177. https://doi.org/10.1016/J.Neuropsychologia.2011.01.024.

Kliegel, M., McDaniel, M. A., & Einstein, G. O. (2000). Plan formation, retention, and execution in prospective memory: a new approach and age-related effects. *Memory & Cognition, 28*(6), 1041–1049. https://doi.org/10.3758/Bf03209352.

Kueider, A. M., Parisi, J. M., Gross, A. L., & Rebok, G. W. (2012). Computerized cognitive training with older adults: a systematic review. *Plos One, 7*(7), e40588. https://doi.org/10.1371/journal.pone.0040588.

Lampit, A., Hallock, H., & Valenzuela, M. (2014). Computerized cognitive training in cognitively healthy older adults: a systematic review and meta-analysis of effect modifiers. *Plos Medicine, 11*(11), e1001756. https://doi.org/10.1371/journal.pmed.1001756.

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. https://doi.org/10.1037/A0020080.

McAuley, E., Duncan, T., & Tammen, V. V. (1989). Psychometric properties of the intrinsic motivation inventory in a competitive sport setting—a confirmatory factor-analysis. *Research Quarterly for Exercise and Sport, 60*(1), 48–58.

Melby-Lervag, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology, 49*(2), 270–291. https://doi.org/10.1037/A0028228.

Melby-Lervag, 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 & Review, 23*(1), 324-330. https://doi.org/10.3758/s13423-015-0862-z.

Milborrow, S. (2017). Earth: multivariate adaptive regression splines (Version R package version 4.4.9.).

Milner, B. (1971). Interhemispheric differences in localization of psychological processes in man. *British Medical Bulletin, 27*(3), 272–277.

Minear, M., Brasher, F., Guerrero, C. B., Brasher, M., Moore, A., & Sukeena, J. (2016). A simultaneous examination of two forms of working memory training: evidence for near transfer only. *Memory & Cognition, 44*(7), 1014–1037. https://doi.org/10.3758/s13421-016-0616-9.

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 & Review, 18*(1), 46–60. https://doi.org/10.3758/S13423-010-0034-0.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S., Smith, A. D., & Smith, P. K. (2002). Models of visuospatial and verbal memory across the adult life span. *Psychology and Aging, 17*(2), 299–320. https://doi.org/10.1037//0882-7974.17.2.299.

Raven, J. C., Court, J. H., & Raven, J. (1996). *Manual for Raven's progressive matrices and Mill Hill vocabulary scales. Standard progressive matrices*. Oxford: Oxford Psychologists Press.

Redick, T. S., Shipstead, Z., Harrison, T. L., Hicks, K. L., Fried, D. E., Hambrick, D. Z., … Engle, R. (2013). No evidence of intelligence improvement after working memory training: a randomized, placebo-controlled study. *Journal of Experimental Psychology: General, 142*(2), 359–379. https://doi.org/10.1037/a0029082.

Reitan, R. (1979). *Trail-making test*. Arizona: Reitan Neuropsychology Laboratory.

Revelle, W. (2016). Psych: procedures for personality and psychological research (Version R package version 1.6.12).

Robert, C., Borella, E., Fagot, D., Lecerf, T., & de Ribaupierre, A. (2009). Working memory and inhibitory control across the life span: intrusion errors in the reading span test. *Memory & Cognition, 37*(3), 336–345. https://doi.org/10.3758/Mc.37.3.336.

Rose, N. S., Rendell, P. G., Hering, A., Kliegel, M., Bidelman, G., & Craik, F. I. M. (2015). Cognitive and neural plasticity in older adults' prospective memory following training with the Virtual Week computer game. *Frontiers in Human Neuroscience, 9*(592). https://doi.org/10.3389/fnhum.2015.00592.

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55*(1), 68–78. https://doi.org/10.1037//0003-066x.55.1.68.

Schnitzspahn, K. M., Stahl, C., Zeintl, M., Kaller, C. P., & Kliegel, M. (2013). The role of shifting, updating, and inhibition in prospective memory performance in young and older adults. *Developmental Psychology, 49*(8), 1544–1553. https://doi.org/10.1037/a0030579.

Soveri, A., Antfolk, J., Karlsson, L., Salo, B., & Laine, M. (2017). Working memory training revisited: a multi-level meta-analysis of n-back training studies. *Psychonomic Bulletin & Review, 24*(4), 1077–1096. https://doi.org/10.3758/s13423-016-1217-0.

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society, 8*(3), 448–460. https://doi.org/10.1017/S1355617702813248.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*, 643–662. https://doi.org/10.1037/H0054651.

Studer-Luethi, B., Jaeggi, S. M., Buschkuehl, M., & Perrig, W. J. (2012). Influence of neuroticism and conscientiousness on working memory training outcome. *Personality and Individual Differences, 53*(1), 44–49. https://doi.org/10.1016/j.paid.2012.02.012.

Tomaszewski Farias, S., Cahn-Weiner, D. A., Harvey, D. J., Reed, B. R., Mungas, D., Kramer, J. H., & Chui, H. (2009). Longitudinal changes in memory and executive functioning are associated with longitudinal change in instrumental activities of daily living in older adults. *The Clinical Neuropsychologist, 23*(3), 446–461. https://doi.org/10.1080/13854040802360558.

Vallerand, R. J., Blais, M. R., Briere, N. M., & Pelletier, L. G. (1989). The construction and validation of a new measure of motivation toward education, Ecmelle-De-Motivation-En-Education (Eme). *Canadian Journal of Behavioural Science-Revue Canadienne Des Sciences Du Comportement, 21*(3), 323–349. https://doi.org/10.1037/h0079855.

Vercambre, M. N., Cuvelier, H., Gayon, Y. A., Hardy-Leger, I., Berr, C., Trivalle, C., et al. (2010). Validation study of a French version of the modified telephone interview for cognitive status (F-TICS-m) in elderly women. *International Journal of Geriatric Psychiatry, 25*(11), 1142–1149. https://doi.org/10.1002/Gps.2447.

von Bastian, C. C., & Oberauer, K. (2013). Effects and mechanisms of working memory training: a review. *Psychological Research, 78*(6), 803–820. https://doi.org/10.1007/s00426-013-0524-6.

Wechsler, D. (2011). *Échelle d'intelligence de Wechsler pour adultes (WAIS-IV) - 4ème édition*. Paris, France: ecpa.

Weicker, J., Villringer, A., & Thone-Otto, A. (2016). Can impaired working memory functioning be improved by training? A meta-analysis with a special focus on brain injured patients. *Neuropsychology, 30*(2), 190–212. https://doi.org/10.1037/neu0000227.

Willis, S. (1994). *Test administration manual for everyday problems for cognitively challenged elderly*. University Park Pennsylvania: The Pennsylvania State University.

Woods, S. P., Weinborn, M., Li, Y. Q. R., Hodgson, E., Ng, A. R. J., & Bucks, R. S. (2015). Does prospective memory influence quality of life in community-dwelling older adults? *Aging Neuropsychology and Cognition, 22*(6), 679–692. https://doi.org/10.1080/13825585.2015.1027651.

Woods, S. P., Weinborn, M., Velnoweth, A., Rooney, A., & Bucks, R. S. (2012). Memory for intentions is uniquely associated with instrumental activities of daily living in healthy older adults. *Journal of the International Neuropsychological Society, 18*(1), 134–138. https://doi.org/10.1017/S1355617711001263.

Zinke, K., Zeintl, M., Rose, N. S., Putzmann, J., Pydde, A., & Kliegel, M. (2014). Working memory training and transfer in older adults: effects of age, baseline performance, and training gains. *Developmental Psychology, 50*(1), 304–315. https://doi.org/10.1037/a0032982.