Multisensory Facilitation of Working Memory Training

Anja Pahor[1,2] · Cindy Collins[1] · Rachel N. Smith-Peirce[2] · Austin Moon[1,2] · Trevor Stavropoulos[1] · Ilse Silva[1] · Elaine Peng[1] · Susanne M. Jaeggi[3,2] · Aaron R. Seitz[1]

Introduction

The human brain has evolved to learn in information-rich environments in which integration of information from multiple sensory inputs is ubiquitous. While perception and cognition are often studied within one sensory modality at a time, there is growing evidence that much of the neocortex is inherently multisensory (Ghazanfar & Schroeder, 2006; Murray et al., 2016) and that multisensory processing is behaviorally advantageous, particularly when the sources are temporally and semantically related (Diaconescu et al., 2011). Multisensory facilitation can benefit subsequent unisensory perceptual processing (Kim et al., 2008a; Shams et al., 2011; von Kriegstein & Giraud, 2006) and memory retrieval of unisensory information (Heikkilä et al., 2017; Lehmann & Murray, 2005; Murray et al., 2004; Thelen et al., 2015). Models suggest that multisensory facilitation can arise through cross-sensory connections between unisensory representations and/or feedback to unisensory representations supporting stronger encoding in those regions as well as by modification or formation of multisensory representations, so that later presentation of unisensory stimuli activates an expanded, multisensory network of brain regions (Shams & Seitz, 2008).

Working memory (WM) is an example of a system that is thought to be multisensory in nature (Quak et al., 2015). Indeed, parts of the WM circuitry such as the intraparietal and dorsolateral prefrontal cortices have been linked to multi-modal maintenance of information (Cowan et al., 2011). WM training has shown to lead to improved performance on tasks that rely on short-term and WM components, particularly when the untrained tasks share similar processes as the trained task (Holmes et al., 2019). It has been suggested that transfer effects only occur if the training and transfer tasks engage specific overlapping brain regions and processes (Dahlin et al., 2008) and that paradigm-specific effects may reflect changes in strategies developed throughout training, rather than an improvement in the efficiency of WM (Forsberg et al., 2020). It is worth noting that mechanisms leading to transfer, and the extent of transfer beyond the domain of WM, are topics of substantial controversy with even metanalyses reaching inconsistent conclusions (Au et al., 2015; Melby-Lervåg et al., 2016; Schwaighofer et al., 2015; Soveri et al., 2017; Weicker et al., 2016); however, this goes beyond the scope of the current paper.

In the present paper, we address issues of how to promote learning and (near) transfer within the same cognitive domain, namely, WM. A WM training protocol that incorporates multisensory objects would not only provide redundancy of information but might further enable stimuli to be encoded into richer, multisensory representations (Kim et al., 2008b; Shams & Seitz, 2008), which may facilitate learning and transfer (Deveau et al., 2014). However, WM training protocols tend to be restricted to the visual domain and/or are not specifically designed to induce multisensory facilitation. For example, many extant procedures, such as the dual n-back, require each modality stream to be processed separately, rendering multisensory integration disadvantageous to task performance. While this type of dual-task performance is certainly demanding, it was later shown that dual n-back training does not lead to greater learning and transfer compared to the single n-back training (Jaeggi et al., 2010; Jaeggi et al., 2014). To date, there are no procedures that specifically implement multisensory facilitation, and thus, the assumption that multisensory training can benefit WM training outcomes is yet to be tested.

Here we test the hypothesis that multisensory WM training will facilitate transfer to untrained tasks within the WM

✉ Anja Pahor
apahor@uci.edu

1 Department of Psychology, University of California, Riverside, Riverside, CA, USA

2 School of Education, University of California, Irvine, Irvine, CA, USA

3 Department of Cognitive Sciences, University of California, Irvine, Irvine, CA, USA

domain compared to visual WM training alone or training that contains visual and auditory stimuli but is presented as separate WM tasks. To address this hypothesis, we randomly assigned participants to one of 3 n-back training conditions and compared their learning and transfer to untrained WM tasks. The conditions consisted of a vision-only training, training that alternated between auditory and visual stimuli, or multisensory training where auditory and visual stimuli were presented concurrently. In the latter condition, pairs of visual-auditory stimuli were constant and congruent, for example, a picture of a bell was always accompanied with the sound of a bell. The performance of the three interventions was further compared with that of a passive control group. These conditions were chosen because they represent commonly used WM training protocols, with visual-only being the most common one, and very few studies include auditory-only WM training (Pergher et al., 2019).

Methods

Participants

Undergraduates from UC Riverside (UCR) and UC Irvine (UCI) were recruited to participate in the study via flyers, e-mails, and word of mouth between Fall 2018 and Fall 2019. A total of 306 students completed a consent form to participate in the study, 240 of which were subsequently enrolled based on their availability (5 times a week for 1 h). Participants were randomly assigned to one of 4 groups: 3 active training groups and a passive control group (see Table 1 for demographics). Forty-two participants were dropped due to attrition or technical errors (see Fig. 1): 6 from the Multisensory group, 10 from Visual Only, 15 from Alternating, and 11 from Passive control group. All procedures were approved by UCR and UCI Institutional Review Boards. Participants provided informed consent and received monetary compensation.

Procedure

Training and assessments were administered on tablet computers via software developed in-house ("Recollect the Study"; available on Google Play and Apple App Store). 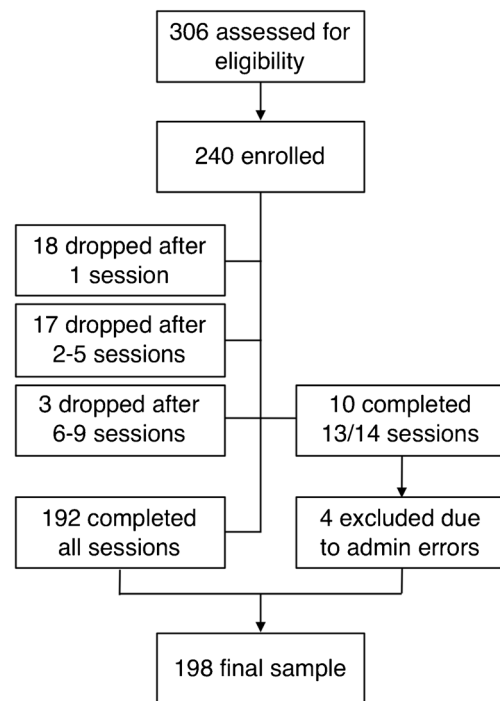In order to prevent fatigue associated with long assessment sessions, participants completed 3 sessions of pretest and 3 sessions of posttest separated by at least 8 working days. The active groups conducted training tutorials and practice (day 1), short 20-min training sessions (days 2, 3, 12, and 13) to ease into training, as well as full 40-min training sessions on days 4–11, for a total of 400 min of training (excluding day 1 practice), whereas the passive control group only took part in the assessments (see Fig. 2). An added benefit of administering testing sessions over 6 days is management of expectations of the passive control group (Green et al., 2019). To further probe expectations, at the end of the study (after the final post-test session), participants responded to the following question using a 5-point Likert scale: "Do you think that the sessions you completed during the study helped you perform better on any of the tasks you completed in the last 3 days?", wherein 1 = not at all, 2 = not really, 3 = cannot say, 4 = quite a bit, and 5 = very much. If the answer selected was 3 or higher, they were also asked to select which task(s), if any, they thought they had performed better on because of earlier sessions.

Fig. 1 Enrollment

Table 1 Demographics of the four groups

	Multisensory	Visual Only	Alternating	Control
Age (years (SD))	20.09 (2.57)	20.35 (5.23)	20.98 (3.98)	20.46 (4.86)
Female (%)	58	62	54	62
Male (%)	38	34	44	32
Other or gender unknown (%)	4	4	2	6
N	48	50	50	50
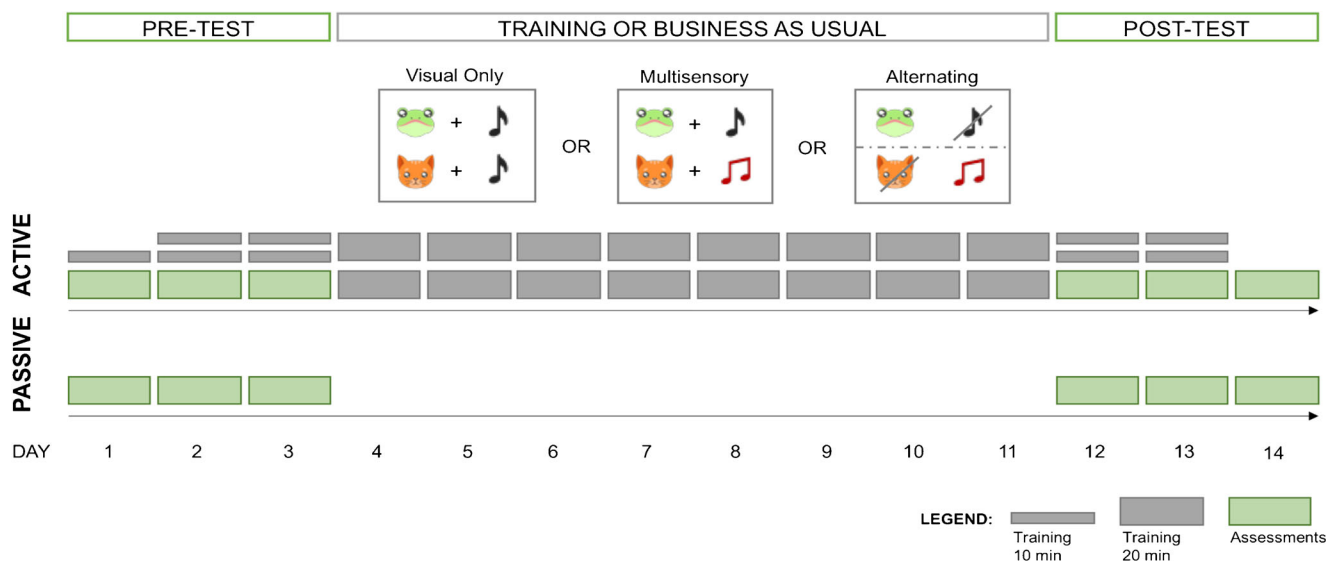
Fig. 2 Timeline of training and testing for passive and active groups. A 5- to 7-min break was given between testing and training blocks and between two 20-min training blocks

Training

The adaptive n-back training game was designed based on principles thought to promote learning (Deveau et al., 2014; Mohammed et al., 2017). In the "Visual Only" condition, all visual stimuli were paired with an identical auditory cue, representing a single n-back training protocol commonly used in the literature (Heinzel et al., 2014; Jaeggi et al., 2010; Küper & Karbach, 2016; Miró-Padilla et al., 2019). An "Alternating" visual/auditory condition consisted of 2 unisensory n-back blocks per day: visual n-back with a placeholder sound and auditory n-back with a visual cue, the order of which was counterbalanced across days (Fig. 2). In the "Multisensory" training condition, each n-back visual stimulus type was paired with a different matching sound.

Transfer Tasks

To assess near transfer, untrained n-back tasks were administered at pre- and post-test, as well as three other WM tasks. Table 2 shows a correlation matrix of baseline performance on transfer tasks. While this dataset is part of a larger study in which a number of other cognitive tasks were administered,[1] here we test the specific hypothesis that multisensory facilitation improves WM training outcomes. For all assessments, the app provided instructions, examples, as well as performance-gating practice trials with feedback.

---

[1] Three additional tasks were administered (math test, countermanding task and matrix reasoning) that do not assess WM, and furthermore, some of the tasks were not consistent across conditions. Therefore, they were not included in the paper. Those tasks were included for pilot purposes in preparation for a larger scale study in which we are targeting 30,000 participants and a wide range of training conditions and outcome measurers (https://www.scientificamerican.com/article/does-brain-training-actually-work/).

Untrained n-back

A visual n-back task with two versions, featuring pictures of animals or pictures of vehicles (counterbalanced across participants), was administered to test generalization to untrained stimuli in a non-game setting. On a given trial, a picture appeared on a black screen for 2500 ms with a 500-ms ISI, and participants were asked to tap the picture if it matched the n-back rule. Each block consisted of 30 + N trials with 9 targets (30% target and 30% lure rates). All participants completed N-levels 1–3 but could progress to 4-back (and beyond) if no more than 2 errors were made on the previous level. The main dependent measure was the highest N-level reached at pretest and posttest.

WM Measures

Three WM tasks were used to assess transfer of n-back training. A simple span task, *Corsi Blocks*, was used to assess WM storage, and two complex span tasks, *Sequencing* and *Symmetry Span*, were used to measure WM storage-and-processing (Cowan, 2008). In *Corsi Blocks*, participants viewed a sequence of 12 Gy squares turning blue and were asked to reproduce that sequence by tapping on the displayed squares in the same order in which they appeared. *Sequencing* is a tablet version of the WMS-III Letter-Number Sequencing test (Wechsler 1997), in which participants see a mixed sequence of letters and numbers appearing one by one (e.g., "5K8G2"), and are then prompted to enter the numbers in numerical order ("258") followed by letters in alphabetical order ("GK"). A given sequence did not include any of the characters in the previous trial or consecutive numbers and letters. Symmetry Span is a tablet-based version of

Table 2 Spearman's rank correlation coefficients of performance on working memory measures at pre-test ($N = 172$)

	Sequencing	Symmetry span	Corsi blocks	Untrained n-back
Sequencing	1.00			
Symmetry span	.25**	1.00		
Corsi blocks	0.10	.27**	1.00	
Untrained n-back	0.13	0.03	0.06	1.00

$**p < 0.01$

Automated Symmetry Span (Unsworth et al., 2009), in which participants are required to recall sequences of red squares in a 4 × 4 matrix while performing a symmetry judgment task.

In all three span tasks, the ISI was 500 ms while stimulus duration was 1000 ms for Corsi Blocks and Sequencing, and 650 ms in Symmetry Span. All three tasks had 2 trials per set size and started at the lowest set size of 2 items. The next set size (e.g., 3 items) was displayed if at least 1 trial was correct, and the task ended when both trials in a set size were incorrect. The highest possible set size was 10 for Corsi Blocks and 15 for the Symmetry Span and Sequencing tasks, ensuring that high performing individuals and/or training-related improvements were not masked by ceiling effects. An individual's span was defined as the highest set size at which at least 1 of the 2 trials was correct.

Data Analysis

IBM SPSS Version 24 (IBM corp.) and JASP 0.9.2 (JASP Team, 2019) were used for statistical analyses. For untrained n-back, data was available for all 198 participants and no outliers were removed based on $|z| > 3$. For the three span tasks, task data were removed for participants whose individual span was at a minimum (span = 2) either on pretest or posttest, indicating poor understanding of task demands or inattention. For Corsi Blocks, 1 participant had missing data, but no data points were removed. For Sequencing and Symmetry span, data were available for 196 and 191 participants, respectively, 15 of which were removed from each group (see Table 3 for sample sizes per group). The highest span achieved at pretest was 10 for Corsi Blocks, 9 for Sequencing and 8 for Symmetry Span.

Gain scores were obtained by subtracting pretest from posttest scores. Since the gain scores were not normally distributed, nonparametric tests were conducted on all transfer tasks. As a first step, we investigated whether the gain itself (pre- vs. posttest change) was significant in any of the groups. If at least one group showed significant gain, Kruskal-Wallis H tests were used to investigate how the four groups, on average, differed in gains, followed by Mann-Whitney U tests.

Results

Training

Participants in the Visual Only and Multisensory groups showed similar training progress, outperforming the Alternating group towards the end of training. In Fig. 3, visual and auditory n-back blocks of the Alternating group are shown separately. The highest N-level gain, calculated as average N-level in the last session minus the average N-level in the first session, was observed for the Multisensory group ($M = 1.95$, SD = 0.21), followed by the Visual Only ($M = 1.75$, SD = 0.18) and Alternating group ($M = 0.99$, SD = 0.12). An independent-samples Kruskal-Wallis test showed that the 3 groups significantly differed in terms of N-level gain ($H(2) = 15.82$, $p < 0.001$). Pairwise comparisons with adjusted p values showed that there was no difference in training gain between Multisensory and Visual Only groups ($p = 1.0$, $r^2 = 0.03$), in contrast, the Multisensory group achieved significantly higher N-level gain than the Alternating group ($p = 0.001$, $r = 0.36$) and the same pattern was observed for the Visual Only group ($p = 0.003$, $r = 0.33$).

Transfer

There is no evidence that the 4 groups differed significantly at baseline on any of the measures as indicated by Kruskal-Wallis one-way tests, which was expected given the random assignment to groups (see Table 4).

Untrained n-back

Wilcoxon Signed Ranks Tests demonstrated that as expected, all groups except the control group reached significantly higher N-levels on the untrained tasks at posttest relative to pretest (Visual Only: $z = -4.89$, $p = < 0.001$; Alternating: $z = -4.36$, $p = < 0.001$; Multisensory: $z = -4.12$, $p = < 0.001$; Control: $z = -1.00$, $p = 0.32$). The groups differed significantly in N-level gain as shown by a Kruskal-Wallis test ($H(3) = 31.07$, $p < 0.001$; see Fig. 4a): Visual Only vs. Control ($U = 568.50$, $z = -5.42$, $p = < 0.001$, $r = -0.54$), Alternating vs.

---

[2] r is an effect size estimate calculated as $z / \sqrt{N}$

Table 3 Descriptive statistics and within-group Pearson's r, Cohen's d and Bayes Factor for performance on pretest and posttest on transfer tests

	Pretest			Posttest			Paired comparisons (pre vs. post)		
	N	Mean	SD	N	Mean	SD	r	d	BF_{10}
Best N-level									
Visual only	50	3.06	0.24	50	4.02	1.02	−0.09	0.90	>1000*
Alternating V/A	50	3.04	0.20	50	3.86	1.07	0.03	0.76	>1000*
Multisensory	48	3.04	0.20	48	3.75	0.96	0.39	0.79	>1000*
Passive	50	3.06	0.24	50	3.12	0.44	0.32	0.14	0.25
Corsi blocks									
Visual Only	50	6.74	1.08	50	6.72	1.13	0.34	0.02	0.16
Alternating V/A	50	6.58	1.13	50	6.86	1.41	0.59	0.24	0.57
Multisensory	48	6.88	1.02	48	7.06	1.14	.43	0.16	0.28
Passive	50	6.66	0.96	50	6.74	1.61	0.36	0.05	0.16
Sequencing span									
Visual only	46	5.61	1.24	46	5.89	1.32	0.49	0.22	0.44
Alternating V/A	46	5.61	1.47	46	6.17	1.60	0.55	0.39	3.43*
Multisensory	42	5.26	1.23	42	5.98	1.47	0.47	0.51	16.30*
Passive	47	5.36	1.05	47	5.70	1.25	0.50	0.29	0.98
Symmetry span									
Visual only	44	4.80	1.25	44	4.84	1.12	0.19	0.03	0.17
Alternating V/A	44	4.82	1.21	44	4.71	1.59	0.38	0.07	0.18
Multisensory	40	4.45	1.01	40	5.13	1.34	0.56	0.59	49.05*
Passive	48	4.44	1.22	48	4.69	1.08	0.33	0.19	0.35

*= p < 0.05, BF_{10} < 1 = Evidence for H_0

Significant evidence in favor of the alternative hypothesis is marked with an asterisk (BF_{10} >3); strong to extreme evidence is marked bold (BF_{10} > 10)

Accounts for the correlation between pre- and post-test measures: $(Mean_{Post}− Mean_{Pre})/\sqrt{(SD^2_{Pre} + SD^2_{Post} − 2r_{PrePost} * SD_{Pre} * SD_{Post})}$

BF_{10} grades the intensity of the evidence that the data provide for H_1 versus H_0. BF_{10} between 1 and 3 is considered to be only anecdotal evidence for H_1; 3–10: moderate evidence; 10–30: strong evidence; 30–100: very strong evidence; > 100 extreme evidence (Wagenmakers et al., 2018).

Control ($U = 711.00$, $z = −4.51$, $p = <0.001$, $r = −0.45$), and Multisensory vs. Control ($U = 730.00$, $z = −4.26$, $p = <0.001$, $r = −0.43$). In the active sample ($N = 148$), N-level gain throughout training showed a significant positive correlation with pre-/post-N-level gain on untrained N-back tasks ($rho = 0.45$, $p < 0.001$), which was also observed separately for each group (Visual Only: $rho = 0.41$, $p = 0.003$; Alternating: $rho = 0.30$, $p = 0.03$; Multisensory: $rho = 0.59$, $p < 0.001$).

Corsi Blocks

We next examined transfer to Corsi Blocks Forward, a measure of WM capacity. The change from pre- to posttest on span was not significant in any of the groups as indicated by Wilcoxon Signed Rank Tests (Visual Only: $z = −0.30$, $p = 0.77$; Alternating: $z = −1.76$ $p = 0.08$; Multisensory: $z = −1.17$, $p = 0.24$; Control: $z = −1.15$, $p = 0.25$); hence, no further analyses were conducted (Fig. 4b).

Table 4 Kruskal-Wallis H tests to assess potential group differences at pretest

Pretest measure	Kruskal-Wallis H test
Best N-level	$H(3) = 0.38$, $p = 0.94$
Corsi blocks	$H(3) = 3.68$, $p = 0.30$
Sequencing span	$H(3) = 2.69$, $p = 0.44$
Symmetry span	$H(3) = 5.05$, $p = 0.17$
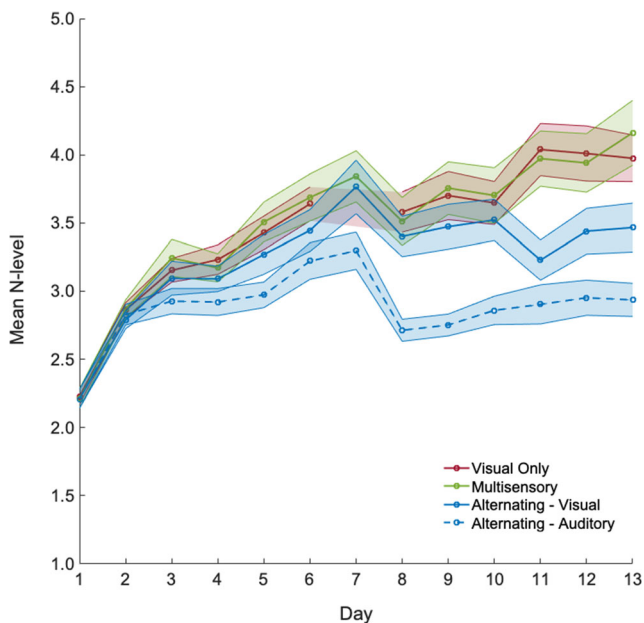
Figure 3 Mean N-level achieved on a given day, split by visual and auditory blocks for the Alternating group. A dip on day 8 reflects transition from less abstract stimuli to more abstract stimuli (equivalent across groups).

Sequencing

Wilcoxon Signed Rank Tests demonstrated significant gain in sequencing span for Alternating ($z = -2.21$, $p = 0.03$) and Multisensory ($z = -2.85$, $p = 0.004$) groups whereas Visual Only ($z = -1.37$, $p = 0.17$) and Control groups ($z = -1.84$, $p = 0.07$) did not show significant gain (Fig. 4C). Overall, the groups did not differ significantly in span gain ($H(3) = 4.27$, $p = 0.23$); however, there was a trend for the Multisensory group to show larger span gain compared to the Visual Only group ($U = 759.50$, $z = -1.78$, $p = 0.075$, $r = -0.19$) and to the Control group ($U = 778.50$, $z = -1.77$, $p = 0.077$, $r = -0.19$). Training gain did not predict span gain ($rho = 0.01$, $p = 0.96$) in the active sample ($N = 134$).

Symmetry Span

The only group that showed significant improvement in symmetry span was the Multisensory group ($z = -3.31$, $p = 0.001$), whereas the other groups did not improve (Visual Only: $z = -0.16$, $p = 0.88$; Alternating: $z = -0.58$, $p = 0.56$; Control: $z = -1.34$, $p = 0.18$). While a Kruskal-Wallis test did not show significant differences in span gain among the groups ($H(3) = 6.33$, $p = 0.10$), a comparison of gain in the three active groups was marginally significant ($H(2) = 5.96$, $p = 0.05$), prompting comparisons between pairs of active groups. As can be seen in Fig. 4d, the multisensory group showed higher span gain than the Visual Only group ($U = 667.50$, $z = -1.96$, $p = 0.05$, $r = -0.25$) and the Alternating group ($U = 634.00$, $z = -2.26$, $p = 0.03$, $r = -0.21$), albeit

the effect sizes are small. Moreover, training N-level gain showed a significant positive correlation with symmetry span gain in the active sample ($rho = 0.23$, $p = 0.01$, $N = 128$); however, this was not observed for the individual groups (Visual Only: $rho = 0.11$, $p = 0.47$; Alternating: $rho = 0.20$, $p = 0.20$; Multisensory: $rho = 0.25$, $p = 0.13$).

Self-Reported Transfer

Out of 182 participants that answered the question "Do you think that the sessions you completed during the study helped you perform better on any of the tasks you completed in the last 3 days?", the most common answer across all four groups was 4 ("Quite a bit"). The highest average score was reported by the Multisensory group (M = 3.74, SD = 0.93, $N = 46$), followed closely by Visual Only (M = 3.71, SD = 0.76, $N = 45$), Alternating (M = 3.52, SD = 0.85, $N = 48$), and the Passive control group (M = 3.47, SD = 0.80, $N = 43$). A Kruskal-Wallis H test indicated that the four groups did not show a statistically significant difference in their answer to this question ($H(3) = 5.33$, $p = .15$). These results provide tentative evidence that the four groups perceived a similar rate of change in post-test performance as a result of participating in the study (even the passive control group). Figure 5 shows the percentage of participants per group that selected a given task as one in which they thought they had improved on as a result of participating in the study. Note that this question was only presented to participants who answered 3 ("Cannot say") or higher on the previous question, i.e., to 163 participants (89.6%) and that they were not required to select any tasks.

Discussion

Based on the rich literature on benefits of multisensory exposure to learning and memory (Matusz et al., 2017; Quak et al., 2015; Shams & Seitz, 2008), it was hypothesized that WM training could be facilitated by introducing multisensory objects in an n-back training task. In order to test this, performance on multisensory n-back training was contrasted against Visual Only and Alternating audio-visual n-back training, as well as a passive control group. Results showed that participants in the Multisensory group had equal N-level gain compared to Visual Only training yet showed significant transfer to some untrained WM tasks, whereas the Visual Only condition did not. This is in line with studies suggesting that multisensory experience can facilitate memory (Thelen et al., 2014) and learning (Shams & Seitz, 2008). In fact, it has been demonstrated that even a single multisensory exposure can influence memory for visual and auditory objects (Matusz et al., 2017). Multisensory cues can also improve delayed retention of incidental learning (Broadbent et al., 2019) and aid perceptual learning (Seitz et al., 2006). The main question

Fig. 4 Difference scores for each group: V = Visual Only, A = Alternating Audio-visual, M = Multisensory, and C = Control (Passive); color coded as red, blue, green, and gray, respectively. (**a**) Difference between highest N-level achieved at post-test relative to pre-test on an untrained visual N-back task. (**b–d**) Change in span on Forward Corsi Blocks, Sequencing, and Symmetry Span, respectively. Error bars represent S.E.M. * $p < 0.05$, ** $p < 0.01$. Stars without lines indicate significant difference at post-test relative to pre-test in each group (Wilcoxon signed-rank test), whereas stars with lines indicate significant differences in change scores in one group relative to another (Kruskal-Wallis test)
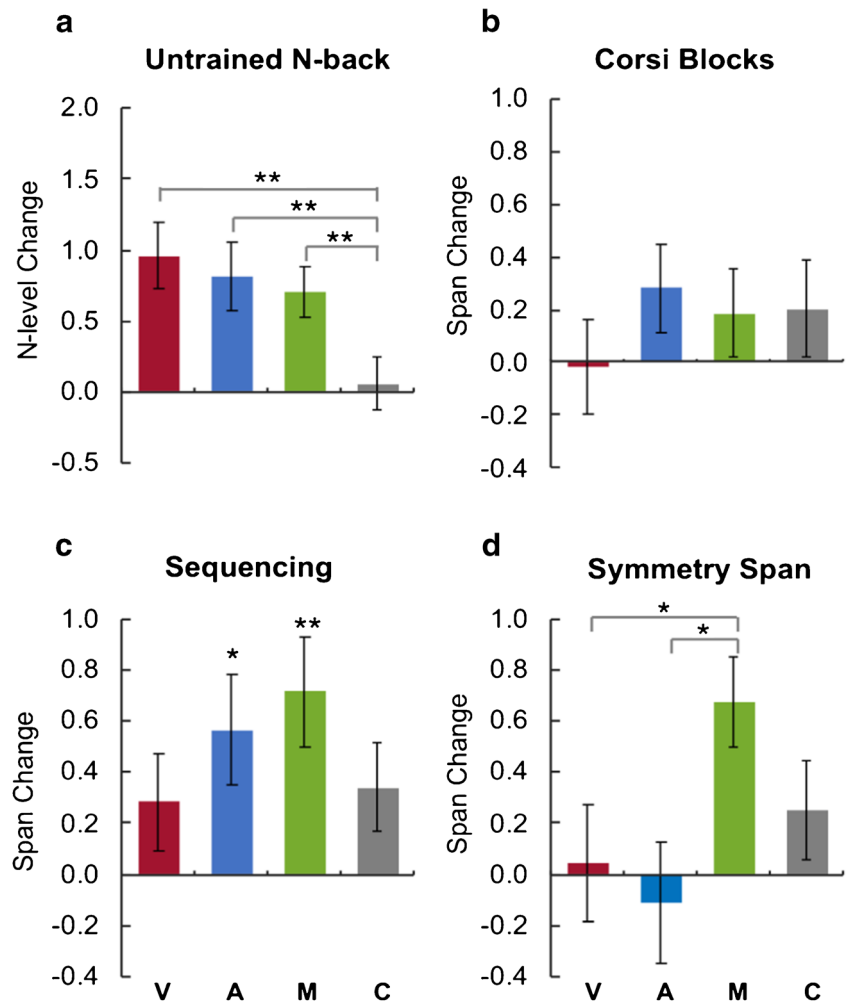
of interest here, however, is whether multisensory training can boost near transfer.

As expected, all three active training groups improved on an untrained N-back task compared to the control group. Training N-level gain predicted transfer to untrained N-back tasks in the active sample irrespective of condition. There was no evidence of training-related enhancement of WM capacity as measured by a simple span task for any of the groups, possibly due to the higher capacity used in this task compared to that experienced during training (e.g., typical N-back span was 4 whereas typical Corsi span was 7). However, the Multisensory group showed significant performance improvement on complex span tasks, particularly symmetry span, and to some extent, sequencing. Training on a task that requires rapid updating of temporary bindings in WM may have strengthened the processes involved in manipulation of information that are crucial for performance on complex span tasks. Indeed, research indicates that updating tasks, complex span tasks, and binding tasks share a large proportion of variance (Wilhelm et al., 2013).
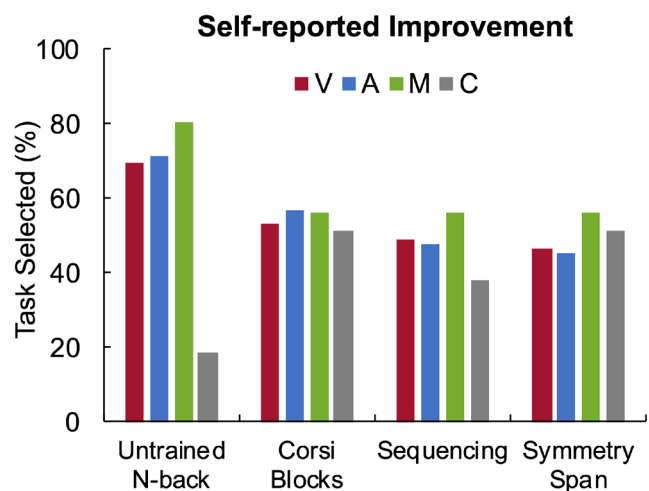
Fig. 5 Self-reported improvement on tasks at post-test as a result of earlier sessions, shown as percentage per group; V = Visual Only, A = Alternating Audio-visual, M = Multisensory, and C = Control (Passive)

We note that a key component of the Multisensory training employed here is that the auditory and visual stimuli at each time-point supported the same task response. This is a key element of multisensory facilitation (Shams & Seitz, 2008) and is in contrast to dual WM tasks in which participants have to perform a secondary task while maintaining information in WM (Fougnie & Marois, 2011), or dual N-back tasks that require updating of two separate streams of stimuli, one in the visual and the other in the auditory domain. We suggest that while these dual tasks may give rise to some benefits of multi-tasking (Anguera et al., 2013), they put the senses at competition with each other and do not give rise to facilitation (Deveau et al., 2014).

While the current study examined transfer to a variety of visual working memory tasks, we note that the question of transfer to auditory or multisensory working memory tasks has not been addressed. Further work will be required to test the hypothesis that the multisensory condition would also be advantageous to transfer to working memory tasks that involve auditory, and maybe even other stimulus modalities. We also acknowledge a limitation in that we did not directly assess placebo effects and/or enjoyment; however, posttest self-report results indicated that the four groups perceived a similar rate of transfer overall, meaning that even the passive control group felt some improvement—likely due to retest effects. When asked whether they observed improvement on specific tasks, the pattern of results mirrored the measured transfer effects, indicating good insight into individual outcomes of the study (Tsai et al., 2018). At present there is extremely limited and mixed data on whether expectation-based effects can drive transfer effects (Green et al., 2019); hence, this should be addressed by future research. A further limitation of the study is the use of a passive control group. While passive control groups only control for retest effects, active control groups can help manage participants' expectancies, equalize participant-experimenter contact time, and reduce demand characteristics (Boot et al., 2013). On the other hand, designing an appropriate control training protocol that does not rely on WM is no trivial task, and certainly not as easy as administering a placebo drug. Furthermore, a recent meta-analysis did not find any evidence that control group type meaningfully impacts effect sizes from cognitive measures (Au et al., 2020). Even though the use of active control groups is still strongly recommended, research involving passive control groups may still be informative and useful, especially in early, proof-of-concept studies such as this one (cf. Green et al., 2019).

In summary, the present results suggest that incorporating multisensory objects in a WM training protocol can benefit performance on the training task compared to training WM in each sense separately and that multisensory training can potentially facilitate transfer to complex WM span tasks. To gain a better understanding of multisensory facilitation of WM training, future research should include an auditory-only training group, auditory WM transfer tests, and explore whether multisensory facilitation affects far transfer tasks as well. Evidence of multisensory facilitation of WM training could inform development of training protocols, particularly those targeting older adults, where supporting learning via multiple senses may be advantageous to those who experience transient deficits in either sense.

Compliance with Ethical Standards

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., et al. (2013). Video game training enhances cognitive control in older adults. *Nature, 501*(7465), 97–101. https://doi.org/10.1038/nature12486.

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*(2), 366–377. https://doi.org/10.3758/s13423-014-0699-x.

Au, J., Gibson, B. C., Bunarjo, K., Buschkuehl, M., & Jaeggi, S. M. (2020). Quantifying the difference between active and passive control groups in cognitive interventions using two meta-analytical approaches. *Journal of Cognitive Enhancement, 4*(2), 192–210. https://doi.org/10.1007/s41465-020-00164-6.

Boot, W. R., Simons, D. J., Stothart, C., & Stutts, C. (2013). The pervasive problem with placebos in psychology: Why active control groups are not sufficient to rule out placebo effects. *Perspectives on Psychological Science : A Journal of the Association for Psychological Science, 8*(4), 445–454. https://doi.org/10.1177/1745691613491271.

Broadbent, H. J., Osborne, T., Mareschal, D., & Kirkham, N. Z. (2019). Withstanding the test of time: Multisensory cues improve the delayed retention of incidental learning. *Developmental Science, 22*(1), e12726. https://doi.org/10.1111/desc.12726.

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research, 169*, 323–338. https://doi.org/10.1016/S0079-6123(07)00020-9.

Cowan, N., Li, D., Moffitt, A., Becker, T. M., Martin, E. A., Saults, J. S., & Christ, S. E. (2011). A neural region of abstract working memory. *Journal of Cognitive Neuroscience, 23*(10), 2852–2863. https://doi.org/10.1162/jocn.2011.21625.

Dahlin, E., Neely, A. S., Larsson, A., Bäckman, L., & Nyberg, L. (2008). Transfer of learning after updating training mediated by the striatum. *Science, 320*(5882), 1510–1512. https://doi.org/10.1126/science.1155466.

Deveau, J., Jaeggi, S. M., Zordan, V., Phung, C., & Seitz, A. R. (2014). How to build better memory training games. *Frontiers in Systems Neuroscience, 8*, 243. https://doi.org/10.3389/fnsys.2014.00243.

Diaconescu, A. O., Alain, C., & McIntosh, A. R. (2011). The co-occurrence of multisensory facilitation and cross-modal conflict in

the human brain. *Journal of Neurophysiology, 106*(6), 2896–2909. https://doi.org/10.1152/jn.00303.2011.

Forsberg, A., Fellman, D., Laine, M., Johnson, W., & Logie, R. H. (2020). Strategy mediation in working memory training in younger and older adults. *Quarterly Journal of Experimental Psychology, 73*(8), 1206–1226. https://doi.org/10.1177/1747021820915107.

Fougnie, D., & Marois, R. (2011). What limits working memory capacity? Evidence for modality-specific sources to the simultaneous storage of visual and auditory arrays. *Journal of Experimental Psychology. Learning, Memory, and Cognition, 37*(6), 1329–1341. https://doi.org/10.1037/a0024834.

Ghazanfar, A. A., & Schroeder, C. E. (2006). Is neocortex essentially multisensory? *Trends in Cognitive Sciences, 10*(6), 278–285. https://doi.org/10.1016/j.tics.2006.04.008.

Green, S., Bavelier, D., Kramer, A. F., Vinogradov, S., Ansorge, U., Ball, K. K., et al. (2019). Improving methodological standards in behavioral interventions for cognitive enhancement. *Journal of Cognitive Enhancement, 3*(1), 1–28. https://doi.org/10.1007/s41465-018-0115-y.

Heikkilä, J., Alho, K., & Tiippana, K. (2017). Semantically congruent visual stimuli can improve auditory memory. *Multisensory Research, 30*(7–8), 639–651. https://doi.org/10.1163/22134808-00002584.

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. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition, 21*(2), 146–173. https://doi.org/10.1080/13825585.2013.790338.

Holmes, J., Woolgar, F., Hampshire, A., & Gathercole, S. E. (2019). Are working memory training effects paradigm-specific? *Frontiers in Psychology, 10*, 1103. https://doi.org/10.3389/fpsyg.2019.01103.

Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., Su, Y.-F., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning — Implications for training and transfer. *Intelligence, 38*(6), 625–635. https://doi.org/10.1016/j.intell.2010.09.001.

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.

Kim, R. S., Seitz, A. R., & Shams, L. (2008a). Benefits of stimulus congruency for multisensory facilitation of visual learning. *PLoS One, 3*(1), e1532. https://doi.org/10.1371/journal.pone.0001532.

Kim, R., Seitz, A. R., & Shams, L. (2008b). Neural mechanisms of multisensory perceptual learning. *Journal of Vision, 8*(6), 976a.

Küper, K., & Karbach, J. (2016). 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. https://doi.org/10.1080/20445911.2015.1118106.

Lehmann, S., & Murray, M. M. (2005). The role of multisensory memories in unisensory object discrimination. *Brain Research. Cognitive Brain Research, 24*(2), 326–334.

Matusz, P. J., Wallace, M. T., & Murray, M. M. (2017). A multisensory perspective on object memory. *Neuropsychologia, 105*, 243–252. https://doi.org/10.1016/j.neuropsychologia.2017.04.008.

Melby-Lervåg, M., Redick, T. S., & Hulme, C. (2016). Working memory training does not improve performance on measures of intelligence or other measures of "far transfer": Evidence from a meta-analytic review. *Perspectives on Psychological Science : A Journal of the Association for Psychological Science, 11*(4), 512–534. https://doi.org/10.1177/1745691616635612.

Miró-Padilla, A., Bueichekú, E., Ventura-Campos, N., Flores-Compañ, M.-J., Parcet, M. A., & Ávila, C. (2019). Long-term brain effects of N-back training: An fMRI study. *Brain Imaging and Behavior, 13*(4), 1115–1127. https://doi.org/10.1007/s11682-018-9925-x.

Mohammed, S., Flores, L., Deveau, J., Cohen Hoffing, R., Phung, C., Parlett, M., & C., … R. Seitz, A. (2017). The benefits and challenges of implementing motivational features to boost cognitive training outcome. *Journal of Cognitive Enhancement, 1*(4), 491–507. https://doi.org/10.1007/s41465-017-0047-y.

Murray, M. M., Michel, C. M., Grave de Peralta, R., Ortigue, S., Brunet, D., Gonzalez Andino, S., & Schnider, A. (2004). Rapid discrimination of visual and multisensory memories revealed by electrical neuroimaging. *Neuroimage, 21*(1), 125–135. https://doi.org/10.1016/j.neuroimage.2003.09.035.

Murray, M. M., Lewkowicz, D. J., Amedi, A., & Wallace, M. T. (2016). Multisensory processes: A balancing act across the lifespan. *Trends in Neurosciences, 39*(8), 567–579. https://doi.org/10.1016/j.tins.2016.05.003.

Pergher, V., Shalchy, M. A., Pahor, A., Van Hulle, M. M., Jaeggi, S. M., & Seitz, A. R. (2019). Divergent research methods limit understanding of working memory training. *Journal of Cognitive Enhancement*, 1–21. https://doi.org/10.1007/s41465-019-00134-7.

Quak, M., London, R. E., & Talsma, D. (2015). A multisensory perspective of working memory. *Frontiers in Human Neuroscience, 9*, 197. https://doi.org/10.3389/fnhum.2015.00197.

Schwaighofer, M., Fischer, F., & Bühner, M. (2015). Does working memory training transfer? A meta-analysis including training conditions as moderators. *Educational Psychologist, 50*(2), 138–166. https://doi.org/10.1080/00461520.2015.1036274.

Seitz, A. R., Kim, R., & Shams, L. (2006). Sound facilitates visual learning. *Current Biology, 16*(14), 1422–1427. https://doi.org/10.1016/j.cub.2006.05.048.

Shams, L., & Seitz, A. R. (2008). Benefits of multisensory learning. *Trends in Cognitive Sciences, 12*(11), 411–417. https://doi.org/10.1016/j.tics.2008.07.006.

Shams, L., Wozny, D. R., Kim, R., & Seitz, A. (2011). Influences of multisensory experience on subsequent unisensory processing. *Frontiers in Psychology, 2*, 264. https://doi.org/10.3389/fpsyg.2011.00264.

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.

Team, J. A. S. P. (2019). JASP 0.9.2 (0.10.2). *Computer software, computer software*.

Thelen, A., Matusz, P. J., & Murray, M. M. (2014). Multisensory context portends object memory. *Current Biology, 24*(16), R734–R735. https://doi.org/10.1016/j.cub.2014.06.040.

Thelen, A., Talsma, D., & Murray, M. M. (2015). Single-trial multisensory memories affect later auditory and visual object discrimination. *Cognition, 138*, 148–160. https://doi.org/10.1016/j.cognition.2015.02.003.

Tsai, N., Buschkuehl, M., Kamarsu, S., Shah, P., Jonides, J., & Jaeggi, S. M. (2018). (Un)great expectations: The role of placebo effects in cognitive training. *Journal of Applied Research in Memory and Cognition, 7*(4), 564–573. https://doi.org/10.1016/j.jarmac.2018.06.001.

Unsworth, N., Redick, T. S., Heitz, R. P., Broadway, J. M., & Engle, R. W. (2009). Complex working memory span tasks and higher-order cognition: A latent-variable analysis of the relationship between

processing and storage. *Memory, 17*(6), 635–654. https://doi.org/10.1080/09658210902998047.

von Kriegstein, K., & Giraud, A.-L. (2006). Implicit multisensory associations influence voice recognition. *PLoS Biology, 4*(10), e326. https://doi.org/10.1371/journal.pbio.0040326.

Wagenmakers, E.-J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., et al. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin & Review, 25*(1), 58–76. https://doi.org/10.3758/s13423-017-1323-7.

Weicker, J., Villringer, A., & Thöne-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.

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology, 4*, 433. https://doi.org/10.3389/fpsyg.2013.00433.