Article:

Yang, T.-X. orcid.org/0000-0002-5359-0820, Allen, R.J. orcid.org/0000-0002-1887-3016, Waterman, A.H. orcid.org/0000-0001-9882-7206 et al. (3 more authors) (2023) Exploring techniques for encoding spoken instructions in working memory: a comparison of verbal rehearsal, motor imagery, self-enactment and action observation. Memory. ISSN 0965-8211

https://doi.org/10.1080/09658211.2023.2273763

Exploring techniques for encoding spoken instructions in working memory: A comparison of verbal rehearsal, motor imagery, self-enactment, and action observation

Tian-xiao Yang[1,2]*, Richard J. Allen[3], Amanda H. Waterman[3], Agnieszka J. Graham[4], Xiao-min Su[1,2], Yan Gao[1,2]

1. Neuropsychology and Applied Cognitive Neuroscience Laboratory, CAS Key Laboratory of Mental Health, Institute of Psychology, Chinese Academy of Sciences, Beijing, China

2. Department of Psychology, the University of Chinese Academy of Sciences, Beijing, China

3. School of Psychology, University of Leeds, Leeds, UK

4. School of Psychology, Queen's University Belfast, Belfast, UK

Author note:

Corresponding author: Tian-xiao Yang, yangtx@psych.ac.cn

ORCID: https://orcid.org/0000-0002-5359-0820

Neuropsychology and Applied Cognitive Neuroscience Laboratory, CAS Key Laboratory of Mental Health, Institute of Psychology, 16 Lincui Road, Chaoyang District, Beijing 100101, China.

Word count of the manuscript: 6799

Exploring techniques for encoding spoken instructions in working memory: A comparison of verbal rehearsal, motor imagery, self-enactment, and action observation

Abstract

Encoding and recalling spoken instructions is subject to working memory capacity limits. Previous research suggests action-based encoding facilitates instruction recall, but has not directly compared benefits across different types of action-based techniques. The current study addressed this in two experiments with young adults. In Experiment 1, participants listened to instructional sequences containing four action-object pairs, and encoded these instructions using either a motor imagery or verbal rehearsal technique, followed by recall via oral repetition or enactment. Memory for instructions was better when participants used a motor imagery technique during encoding, and when recalling the instructions by enactment. The advantage of using a motor imagery technique was present in both verbal and enacted recall. In Experiment 2, participants encoded spoken instructions whilst implementing one of four techniques (verbal rehearsal, motor imagery, observation of others' actions, or self-enactment), and then recalled the instructions by oral repetition or enactment. For both verbal and enacted recall, memory for instructions was least accurate in the rehearsal condition, while the other encoding conditions did not differ from each other. These novel findings indicate similar benefits of imagining, observation, and execution of actions in encoding spoken instructions, and enrich current understanding of action-based benefits in working memory.

Key words: working memory, rehearsal, observation, imagery, enactment

Introduction

Spoken instructions are frequently used as a guide to complete tasks that are organized as a series of successive action steps. For example, instructions on how to send an email might include: logging into one's email account, opening a new message box, specifying the email address of the recipient, including an informative subject line, and finally typing the message in the main field of the email before clicking 'send'. Such instructions can be encountered across many real-world contexts (Allen et al., 2023, for a review). The ability to successfully perform actions to command is particularly valuable in educational and work-related settings where failure to remember and act upon instructions can impede progress in individual learning activities, and over time delay the development of desired skills and proficiencies (Gathercole et al., 2008). Given the cognitive complexity of translating spoken commands into concrete action steps (Allen et al., 2023; Li et al., 2022), a range of factors determine whether instructions will be successfully implemented (e.g., sensory perception, language comprehension, working memory capacity). The primary focus of the two experiments reported here is on *encoding* techniques that can be incorporated to boost immediate memory for spoken multi-step instructions. Specifically, we sought to compare the efficacy of verbal encoding, self-enactment, action observation, and motor imagery in the following instruction task. This comparison allows us to provide further insights into the constraints of successful instruction-following, helping us tailor instruction-based information to facilitate its effectiveness.

How well an event is encoded predicts how well it is remembered, and verbal encoding is an important part of everyday memory. As the focus of this research is on spoken instructions, the nature of the task necessitates verbal encoding and, subsequently, verbal rehearsal for which adult participants repeat the information out loud or sub-vocally (e.g., Dunlosky & Kane, 2007; Friedman & Miyake, 2004; Unsworth, 2016). For both adults (e.g., Morrison et al., 2016) and children (Tam et al., 2010), verbal rehearsal is a frequently reported strategy when maintaining verbal information in working memory and accordingly studies of instruction-guided behavior have recognized a vital role for verbal working memory in maintaining the content of spoken instructions in children (e.g., Engle et al., 1991; Gathercole et al., 2008; Jaroslawska et al., 2016) and adults (e.g., Jaroslawska et al., 2018; Li et al., 2022; Yang et al., 2016). A number of studies conducted to date have used dual-task methodology to explore the potential role of different parts of the multiple-component working model (Baddeley, 1986; Baddeley & Hitch, 1974; Baddeley et al., 2021) in this context using adult participants. Across these studies, memory for instructions was disrupted by concurrent activities that taxed phonological short-term memory, spatial short-term memory, and executive control, indicating that the encoding and storage of verbal instructions depends on multiple sub-components of the system (e.g., Jaroslawska et al., 2018; Li et al., 2022; Yang et al., 2016; Yang et al., 2014).

A major source of insight regarding working memory for instructions has been drawn from use of self-enactment during encoding and/or recall, an experimental approach that has a long and productive history in the context of long-term memory for actions (e.g., Engelkamp & Zimmer, 1989; Kormi-Nouri et al., 1994; Mohr et al., 1989; Norris & West, 1993; Zimmer et al., 2000). Within the working memory literature, an encoding-based enactment advantage occurs when the physical performance of instructed action steps at the time of presentation improves subsequent recall in both adults (Allen & Waterman, 2015; Charlesworth et al., 2014; Lui et al., 2018) and children (Jaroslawska et al., 2016; Waterman et al., 2017; Wojcik et al., 2011). A related mnemonic effect of physical movement is the enacted recall advantage which is observed when sequences of instructions are physically performed at recall, rather than verbally repeated at recall; and this advantage shows in children (Gathercole et al., 2008; Jaroslawska et al., 2016; Yang et al., 2017) as well as in adults (e.g., Allen & Waterman, 2015; Koriat et al., 1990; Yang et al., 2014). When the two manipulations are combined in a single study protocol, the benefit of self-enactment during encoding typically is reduced or removed when performance is also tested through enactment. This pattern of findings has been observed in primary school-aged children (Jaroslawska et al., 2016; Waterman et al., 2017) and young adults (Allen & Waterman, 2015) but not in older adults who appear not to benefit from enactment at encoding (Coats et al., 2021; Jaroslawska et al., 2021). The interactive effect of enactment at encoding and recall can be explained in terms of the generation of motor representation in working memory: When performing physical actions during encoding, or planning for action recall, participants may actively construct action plans that incorporate spatial-motoric information and mental representations of intended movements (Koriat et al., 1990). Consequently, planning to execute action sequences does not improve recall to the same extent when the instructions had already been enacted at encoding, because the motoric memory trace generated through action execution has already been taken advantage of. Recent findings from adult participants suggest that the enacted recall advantage may rise from active action planning for enacted recall (Jaroslawska et al., 2018; Li et al., 2022), which facilitates the binding of action and related objects in the environment (Makri & Jarrold, 2021; Yang et al., 2016). In line with this, there is evidence that concurrent motor activity can disrupt encoding-based enactment effects in adults (Plancher et al., 2019).

Many procedural tasks are taught and learned through face-to-face or video demonstration, which is another technique of interest. Demonstration typically consists of a set of verbal instructions through which the different steps required to accomplish a particular task are described, accompanied by manual demonstration of how each of these steps are to be carried out. This form of instruction is particularly popular in the context of manual skills such as teaching surgical procedures (e.g., Mentis et al., 2014), driving (e.g., Deppermann, 2015), or bike repair (e.g., Arnold, 2012). Different forms of demonstration have also recently been explored in tasks requiring short-term storage and recall of multi-step instruction sequences. For example, Yang et al. (2015) showed that, when compared to purely auditory presentation of verbal instructions, silent video demonstration led to substantially better recall in adults (see also, Yang et al., 2019, 2022) as well as in children (Yang et al., 2017). Subsequently, Waterman et al. (2017) studied how planning for action, self-enactment, and researcher-demonstration influenced children's performance on the task. They found that the benefits of self-enactment at encoding were inconsistent and contingent on sequence complexity, whereas demonstration enhanced recall even when the sequences were complex. More recently, Allen et al. (2020) showed that video demonstration boosted adults' memory performance, compared to spoken presentation, while self-enactment only had relatively modest effects (see also, Coats et al., 2021). Together these findings illustrate that visual demonstration can substantially benefit performance in working memory.

It is worth noting that it is presently unclear whether the benefit of demonstration varies as a function of recall modality. On the one hand, some studies reported larger demonstration benefit when verbal rather than enacted recall was required in adults (Lui et al., 2018; Yang et al., 2022). This implies that, as was the case in the context of enactment at encoding, demonstration may provide additional visuo-spatial and motor information that boosts verbal recall but not enacted recall. On the other hand, other evidence suggests that demonstration has a comparable facilitating effect for both verbal and enacted recall (Waterman et al., 2017, adults; Yang et al., 2015, children), indicating that demonstration could provide additional benefits above action planning (for enacted recall), such as provision of an integrated representation without consuming limited working memory resources for action planning.

As should be clear from the above, studies of encoding-based manipulations have so far focused almost exclusively on actual or observable activity, either in the form of self-enactment or demonstration. However, a more recent investigation attempted to utilize the possibility that the enacted recall effect is at least partly driven by the generation of an *imagined* action plan during encoding. Yang et al. (2021) examined the potential benefit of imagination on following instructions in working memory by explicitly asking child participants to imagine performing the instructions during presentation. It was found that recall performance was superior when children imagined the actions compared to verbal encoding, providing evidence for benefits of forming an action-based representation via motor imagery in a developmental population. Furthermore, the motor imagery benefit was similar in magnitude irrespective of the type of recall required, contrary to the interactive pattern found when self-enactment/demonstration at encoding is followed by enactment at recall. Thus, directing participants to imagine (rather than verbally rehearse) during encoding may provide benefits that go beyond those derived from planning to carry out the instructions at recall.

The present study

Based on the findings outlined above, we sought to address two gaps in the extant literature. First, the benefit of motor imagery versus verbal rehearsal in children requires replication and more importantly, should be investigated in adults. For instance, the same type of action-related benefits may vary in children and adults, with self-enacting during encoding improving performance of following spoken instructions in young adults (Allen & Waterman, 2015; Coats et al., 2021; Jaroslawska et al., 2021; Lui et al., 2018), but the effect being weaker and less consistent in children (Jaroslawska et al., 2016; Waterman et al., 2017) and older adults (Coats et al., 2021; Jaroslawska et al., 2021). Indeed, whether adults show similar benefits from motor imagery strategy and rehearsal is less investigated and remains unclear. Second, no study has directly compared the benefits of motor imagery, action observation, and self-enactment, relative to rehearsal strategy in following spoken instructions. The present study is therefore novel in contrasting imagery with other forms of encoding condition, and in examining the impacts of each of these encoding conditions on either verbal or enacted recall. It thus provides a more detailed picture concerning how different forms of action-based processing interact in working memory.

The present study used an adaptation of the paradigm implemented with children by Yang et al. (2021), itself a development of the following instructions task introduced by Gathercole et al. (2008, see Allen et al., 2023 for a review). This task requires immediate memory for sequences of instructions containing features (action, color, and object type) that are arbitrarily paired and repeatedly re-presented across trials. This minimizes contributions of episodic and semantic long-term memory and emphasizes the need to create new working memory representations on each trial (see Baddeley et al., 2009, for a similar approach in the exploration of prose memory). Research using both dual-task and individual difference approaches has shown that this task relies on working memory capacity (Jaroslawska et al., 2018; Yang et al., 2016; Yang et al., 2014). We used this paradigm to address two unresolved issues. Experiment 1 built on the developmental study by Yang et al. (2021) by contrasting verbal rehearsal and motor imagery strategies. This was replicated and extended in Experiment 2, with four encoding techniques (verbal rehearsal, motor imagery, action observation, and self-enactment) compared. In both experiments, young adult participants recalled the instructions by repetition or enactment to examine the potential interaction between encoding technique and recall modality (i.e., verbal vs enacted recall). Hypotheses are presented in each experiment. As in the previous studies with adults (Allen & Waterman, 2015; Yang et al., 2019), serial position curves were examined to assess how different encoding techniques support memory at different points of the instruction sequence.

Experiment 1

In this experiment, we aimed to compare two encoding techniques: motor imagery versus verbal rehearsal on following spoken instructions. Based on the previous work in children (Yang et al., 2021), we predicted an advantage of motor imagery compared with a verbal rehearsal technique, an enacted recall advantage (i.e., better performance under verbal than enacted recall condition), and no interaction between the encoding technique and recall modality.

Methods

Participants

We used the G*Power 3.1.9.7 (Faul et al., 2009) to estimate the required sample size to detect a motor imagery effect. Yang et al. (2021) observed an effect size of η_p^2 = 0.29 for the motor imagery benefit (compared with verbal rehearsal) in children. This effect size was transformed to *Cohen's f* = 0.64, and finding this effect size in a F test with repeated-measure and as a within-subject factor, alpha = .05 and 80% power required an estimated total sample size of 23. A total of 24 native Mandarin Chinese speakers (50% female, 50% male) with a mean age of 23.08 (*SD* = 2.78) years participated in the experiment. Participants had an average of 16.25 years (*SD* = 2.07) of education. The present study was approved by the ethics committee of the first author in this study (the name of the university/institute was omitted for blind review), and all the participants provided informed consent.

Design and Procedure

The study used a 2 (encoding technique: motor imagery vs. verbal rehearsal) × 2 (recall modality: verbal vs. enacted) within-participants design[1]. The two encoding technique conditions were performed in separate blocks, and all conditions were counterbalanced across participants.

---

[1] This experiment was not pre-registered.

The main dependent variable was proportion of action phrases recalled correctly (i.e., the proportion of action-object pairs recalled in their correct serial position in the sequence). The proportion of correct action-object pairs for each serial position was also calculated (range from 0 to 1).

Instruction task

The instruction task was adapted from Yang et al. (2021). There were 12 objects, distributed in the same configuration for each participant (the lower row from left to right: a white eraser, a yellow ruler, a blue ruler, a green eraser, a red pencil, a black pencil; the upper row from left to right: a white basket, a yellow basket, a blue folder, a green folder, a red bag, a black bag) and 6 types of actions (pull, push, tap, lift, flip, spin). The task included four lists (see Appendix A), with each list containing 12 trials (2 practice trials and 10 formal trials) and each trial comprising 4 action-object pairs (e.g., "*tap the black bag, flip the blue folder, spin the red pencil, pull the yellow basket*")[2]. Sequences were recorded by a native Chinese female speaker at a moderate speed.

Participants were tested individually by an experimenter in a quiet room. In each condition, participants sat at a desk facing an array of objects, and the locations of objects remained the same throughout the test. The experimenter first familiarized participants with the names of the objects and actions, and demonstrated the actions until all the names were recalled correctly and all the actions were copied correctly by the participants. The experimenter then demonstrated a practice trial and allowed participants to have a practice on the same trial, and any incorrect behavior was

---

[2] We used a fixed length to ensure enough trials (power) to examine the serial position curves. Most previous studies involving fixed length of instruction adopted the four action-object pairs (Jaroslawska et al. 2018; Li et al., 2022; Yang et al., 2019 Experiment 3), and therefore we used this length in this study.

corrected during the time. In the verbal rehearsal condition, the experimenter demonstrated how to correctly repeat the action-object during encoding. In the motor imagery condition, the experimenter demonstrated the process of imagining actions by verbally describing the enacting process in detail. The exact wordings of all task instructions can be found in Appendix B.

The procedure for each trial is displayed in Figure 1(b). At the beginning of each trial, the experimenter said ‘prepare’, followed by the first action-object pair (approx. 2.5 s), and a silent 2.5 s interstimulus interval (ISI) for the required encoding activity (verbal rehearsal or motor imagery, depending on the condition). The second action-object pair and another ISI then followed. During the ISI, the participants imagined themselves performing the actions on the objects in the motor imagery condition, or they quietly repeated the instructions in the verbal rehearsal condition. The participants were not allowed to touch the objects during the entire presentation phase. At the end of the presentation phase, participants heard a beep sound and recalled the instructions either by physical enactment or verbal recall. The participants were required to recall the pairs in the same order in which they were presented and used “blank” to replace any forgotten item for both verbal and enacted recall.

All scoring was based on the formal trials (i.e., excluding practice trials). An action-object pair was considered as correct (score = 1) when both features (i.e., action and object) and their serial positions were correct. The total number of correct action-object pairs was divided by 40 (i.e., the total action-object pairs in a list), which led to the proportion correct of action-object pairs of the list/condition (range: 0 to 1). For each serial position, the number of correct action-object pairs across trials were summed up and divided by 10 (number of formal trials), yielding the proportional correct of action-object pairs in that serial position (range: 0 to 1). The data and analytic files are available through this link (https://www.scidb.cn/anonymous/UWZhNkIz).

Results

All analysis was carried out in JASP (https://jasp-stats.org/). Data were analyzed with repeated measures ANOVA, firstly on the overall 2 × 2 design, and then with the additional factor of serial position (i.e., 2 × 2 × 4). In addition to frequentist analysis, a Bayesian ANOVA was also implemented for the 2 × 2 and 2 × 2 × 4 models, with Bayes factors (BF) reported as an estimation of the strength of evidence for the data under the null and alternative hypotheses. These correspond to BF_{incl}, i.e., the strength of evidence for the inclusion of each factor and interaction in the model. $BF < 1$ indicates support for the null hypothesis, and $BF > 1$ supports for the alternative hypothesis. Bayes factors provide a continuous estimate of strength of evidence, but to aid interpretation we use the classification scheme in which BF 1–3 equates to anecdotal evidence, *BF* 3–10 as moderate evidence, and $BF > 10$ as strong evidence (Jeffreys, 1961; Lee & Wagenmakers, 2013).

Proportion correct of action-object pairs in the list/condition

Proportion correct of action-object pairs as a function of encoding technique and recall modality, is presented in Table 1 and Figure 2. A 2 (encoding technique: motor imagery, verbal rehearsal) × 2 (recall modality: verbal vs. enacted) repeated measures ANOVA was conducted. There was a significant main effect of encoding technique, showing as superior performance in the motor imagery than the verbal rehearsal condition, $F(1, 23) = 17.26$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = 0.43$, $BF_{10} > 10000$. There was also a significant main effect of recall modality, showing as superior enacted recall than verbal recall, F (1, 23) = 100.16, MSE = 0.01, p < .001, η_p^2 = 0.81, BF_{10} > 10,000. The interaction between encoding technique and recall modality was not significant, p = .160, BF_{10} = 1.97. The effect sizes for each of these effects are presented in Table 1.

Serial position curve

The proportion correct of action-object pairs of the four serial positions was calculated and shown as serial position curves. The serial position curves as a function of presentation modality and recall modality are presented in Figure 3. A 2 (encoding technique) × 2 (recall modality) × 4 (serial position) repeated measures ANOVA was conducted. Here, we only report effects relating to serial position. There was a significant main effect of position, $F(3, 69)$ = 34.60, MSE = 0.02, p < .001, η_p^2 = 0.60, BF_{10} > 10,000. Serial position interacted significantly with recall modality, $F(3, 69)$ = 36.43, MSE = 0.02, p < .001, η_p^2 = 0.61, BF_{10} > 10,000. Post-hoc comparisons with Bonferroni corrections revealed continuous decrement of performance between adjacent positions in the verbal recall conditions (all p values < .05), in contrast to a slower decrement from the first to third position (p =.003, no significant decrement between the first to second position, and the second to third position) and a recency effect (increase between the third and last position, p =.009) in the enacted recall conditions. Serial position did not interact with encoding technique, p =.334, BF_{10} = 0.19 and there was no significant three-way interaction, p = .779, BF_{10} = 0.04.

Discussion

Consistent with the hypotheses, we found a significant benefit of motor imagery at encoding compared with verbal rehearsal, an enacted recall advantage, and no interaction between encoding technique and recall modality. These findings are consistent with a recent study in children (Yang et al., 2021), suggesting that for both children and young adults, motor imagery improves verbal and enacted recall to a similar extent.

We also examined serial position curves in this experiment. These patterns varied with recall modality, replicating previous findings in adults (Allen & Waterman, 2015; Yang et al., 2019). In the verbal recall conditions, there was a large primacy effect, followed by continuous decrement of recall from the first to the last serial positions, indicating rapid loss of information by decay or interference during verbal output. In the enacted recall conditions, the decrement across serial positions was slow and there was a large recency effect. These differences also reflected an increased enacted recall advantage from the first to later positions, suggesting this advantage is particularly useful for maintaining later actions in the sequence. This unique advantage for later actions may be associated with a more robust action-based representation and a more efficient retrieval process of enacted recall, as evidenced by reduced preparation time and recall duration in enacted than verbal recall (Yang et al., 2019). In contrast, motor imagery facilitated all items in the sequence to a similar extent but did not alter the pattern of rapid memory decay in verbal recall. As both motor imagery and enacted recall involve action planning that strengthens the memory representation, the additional boost to later items by enacted recall may thus reflect the advantage of efficient retrieval (Yang et al., 2019), though Allen and Waterman (2015) observed a similar pattern of later position improvement for self-enactment during encoding. These findings, therefore, indicate both overlap and differences between different forms of action-based advantage.

Experiment 2

There were two aims of this second experiment. First, we aimed to replicate the findings of Experiment 1 (i.e., benefits of motor imagery at encoding and an enacted recall advantage, with no interaction between encoding technique and recall modality). Second, we added two further encoding conditions (action observation and self-enactment), and compared the resulting four encoding techniques. We predicted superior performance in the three conditions involving action-related processing (motor imagery, action observation and self-enactment) compared with the verbal rehearsal condition. Given the inconsistent findings relating to observing and performing actions during encoding, and absence of studies comparing the two techniques with the motor imagery technique, we made no specific hypotheses regarding this.

Methods

Participants

We used the G*Power 3.1.9.7 (Faul et al., 2009) to estimate the required sample size to detect a motor imagery effect (η_p^2 = 0.43) based on Experiment 1. This effect size was transformed to *Cohen's f* = 0.87, and finding this effect size in an F test with repeated-measure and as a within-subject factor, alpha = .05 and 80% power required an estimated total sample size of 14. Given that recall modality was a between-participants variable, we increased sample size to 40 (i.e., N = 20 in each recall condition). A total of 40 native Mandarin Chinese speakers (67.5% female, 32.5% male) with a mean age of 19.90 (*SD* = 1.75) years, and an average of 14.95 years (*SD* = 1.88) of education, participated in the experiment and provided informed consent.

Design

The study used a 4 × 2 mixed design with encoding technique manipulated within-participants (i.e., verbal rehearsal, motor imagery, action observation and self-enactment) and recall modality (i.e., verbal vs. enacted recall) manipulated between-participants[3]. The dependent variable was the proportion of action-object pairs correctly recalled in their correct serial position in the sequence. The order of encoding conditions was counterbalanced across participants.

Instruction task

The paradigm was the same as in Experiment 1. The instructions lists were similar to Experiment 1 except that the action “push” was replaced with “shake” to avoid confusion of the actions “pull” and “push” during self-enactment condition (see Appendix A for the lists). For demonstration, short video clips were taken for all possible combinations of action-object pairs. Each video clip was taken from the perspective of the person performing the action, showing as a hand manipulating an object (see Figure 1a), which lasted about 2.5 sec. These short clips of actions were combined with the audio files to form a coherent video file for each instruction trial. To show the video clips, a computer screen was placed on the desk behind an array of the objects throughout the experiment.

The procedures of the motor imagery and verbal rehearsal conditions were the same as those

---

[3] This experiment was not pre-registered.

used in Experiment 1. In the self-enactment condition, the procedure was similar except participants performed the action after the audio play of each action-object pair. In the action observation condition, participants watched the video clip of the action after hearing the action-object pair. As in Experiment 1, recall was either verbal or via enactment. The exact wordings of all task instructions can be found in Appendix B.

Results

Action-object pairs

Descriptive results of action-object pairs are presented in Table 1 and Figure 4. A 4 (encoding technique: verbal rehearsal, motor imagery, action observation, self-enactment) $\times$ 2 (recall modality: verbal vs. enacted) mixed ANOVA was conducted. There was a significant main effect of encoding technique, $F(3, 114) = 26.03$, $MSE = 0.01$, $p < .001$, $\eta_p^2 = 0.41$, $BF_{10} > 10{,}000$. Post-hoc analysis with Bonferroni corrections indicated a similar level of performance among the motor imagery, action observation, and self-enactment conditions (all p values $> .05$, $BF_{10} < .5$), which were all superior to performance in the verbal rehearsal condition (all p values $< .001$, $BF_{10} > 10{,}000$). There was also a significant main effect of recall type, showing as superior enacted recall than verbal recall, $F(1, 38) = 13.55$, $MSE = 0.01$, $p = .001$, $\eta_p^2 = 0.26$, $BF_{10} = 30.61$. The interaction between recall modality and encoding technique was not significant, $p = .923$, $BF_{10} = 0.31$. The effect sizes of these effects are presented in Table 1.

Serial positions curve

A 4 (encoding technique) × 2 (recall modality) × 4 (serial position) mixed ANOVA was conducted. The descriptive results are shown in Figure 5. Here, we only report effects relating to serial position. There was a significant main effect of position, $F(2.31, 87.80) = 43.12$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = 0.53$, $BF_{10} > 10{,}000$. Position interacted significantly with recall modality, $F(2.31, 87.80) = 12.21$, $MSE = 0.02$, $p < .001$, $\eta_p^2 = 0.24$, $BF_{10} = 22142$. Post-hoc comparisons with Bonferroni corrections revealed a rapid decrement of performance in the first three positions (all p values < .05) but not in the last position ($p = 1.000$) in the verbal recall conditions, in contrast to a slower decrement from the first to third position ($p = .002$, but no significant decrement between adjacent positions including the first to second position, and the second to third position) and no further decrement from the third to last position ($p = .351$) in the enacted recall conditions. Although there was a significant interaction between serial position and encoding technique, $F(6.00, 228.04) = 2.19$, $MSE = 0.02$, $p = .045$, $\eta_p^2 = 0.06$, $BF_{10} = 1.52$, Bayesian analysis indicated anecdotal evidence for this interaction. Further exploration indicated this encoding technique x serial position interaction to reflect improved recall for later sequence positions in the self-enactment condition. An analysis focusing only on the conditions from Experiment 1 (i.e., verbal rehearsal and motor imagery) replicated the null interaction between the encoding technique and serial positions from that experiment ($p = .437$, $\eta_p^2 = 0.02$, $BF_{10} = 0.24$). There was no significant three-way interaction, $p = .850$, $BF_{10} = 0.007$.

Discussion

All findings from Experiment 1 were replicated, including the enacted recall advantage, the benefit of motor imagery over verbal rehearsal, and the absence of an interaction between encoding techniques and recall modality. Enacting the actions during encoding led to better memory performance relative to using the verbal rehearsal technique, which is consistent with previous findings based on comparing self-enactment during encoding with no requirement for encoding in adults (i.e., baseline conditions; Allen et al., 2020; Allen & Waterman, 2015; Coats et al., 2021; Jaroslawska et al., 2021; Lui et al., 2018) as well as in children (Jaroslawska et al., 2016). In the present study, the self-enactment benefit at encoding was similar in the verbal and enacted recall conditions. Given enacted recall already facilitates the action planning process, the presence of the self-enactment advantage in enacted recall conditions thus suggests an additional benefit provided by self-enactment during encoding. It is possible that the experience of enactment upon objects during encoding may provide motor and tactile feedback, which were unavailable in action planning; this sensory feedback may then serve as additional memory cues during recall. However, this finding is in contrast to previous studies where the benefit of self-enactment relative to baseline conditions was reduced or eliminated in the enacted recall conditions (Allen & Waterman, 2015; Jaroslawska et al., 2021; Jaroslawska et al., 2016; Lui et al., 2018). It is worth noting that in both working memory and long-term memory domains, self-enactment benefits are robust when applied during encoding. In contrast, enactment effects at retrieval appear to be more likely in working memory than long-term memory tasks. The enacted-recall advantage in working memory has been attributed to active action planning during encoding and a more efficient retrieval process (Allen & Waterman, 2015(Waterman et al., 2017; Yang et al., 2019). These cognitive processes may be particularly beneficial when participants need to encode and recall action phrases in sequence. In contrast, long-term memory tasks examining this question have typically required participants to recall nouns based on verb cues (Kormi-Nouri et al., 1994; Saltz & Dixon, 1982), and physically performing the verb action was less helpful to recall nouns unless the enactment could be integrated with the nouns such as body parts (Kormi-Nouri et al., 1994). In a long-term memory study that did require participants to recall all action phrases (mostly involving body parts), enactment did benefit when implemented at recall (Norris & West, 1993). These findings suggest enactment at the retrieval stage may be particularly beneficial for recalling action-objects or verb-nouns combinations. Future research may compare these enactment effects in working memory and long-term memory to clarify the similarity and differences underlying the mechanisms in the two domains.

Observing actions during encoding also improved memory performance relative to rehearsal, replicating findings in adults (Coats et al., 2021; Lui et al., 2018; Yang et al., 2015; Yang et al., 2022). As previously stated, this benefit may be driven by a richer memory representation that contains additional visual, spatial and motor information from action observation (Allen et al., 2020). In the present study, the observation advantage was similar for verbal and enacted recall. This finding supports some previous studies of following instructions in adults (Waterman et al., 2017; Yang et al., 2015, but see Lui et al., 2018; Yang et al., 2022), indicating an additional benefit of action observation above action planning, such as providing a more integrated and errorless representation.

A key novel finding in this experiment emerged from the comparison of three forms of action-based processing during encoding, with instruction recall being equivalent in the imagery, observation and self-enactment conditions. The serial position curves also indicated similarity between the three forms of action-based encoding (see Figure 5). These findings fit with the empirical evidence of similar brain networks among motor imagery, observation and execution of actions (Hardwick et al., 2018). The comparisons among the three forms of action advantages are discussed in the next section.

General Discussion

In two experiments, we compared the effects of different forms of action-based encoding on following spoken instructions in adults. The first experiment showed a similar benefit of motor imagery in both verbal and enacted recall. The second experiment replicated these findings and further indicated a similar benefit of motor imagery, action observation, and self-enactment relative to verbal rehearsal, in both verbal and enacted recall.

In both experiments, we observed an advantage of using motor imagery to encode spoken instructions compared with the commonly used verbal rehearsal technique. Moreover, this advantage was present in both verbal and enacted recall, which extends the findings from children to young adults and indicates different age groups can benefit from motor imagery during encoding. The current findings also supplement the largely overlooked area of motor imagery in working memory, and link to the beneficial effect of motor imagery in long-term memory (Ecker &

Engelkamp, 1995; Ma et al., 2021) and skill learning (Kim et al., 2020) in adults; together, they emphasized the value of motor imagery for improving memory within both short-term and long-term timeframes.

The present study indicated similar effects of motor imagery and action observation in facilitating memory of instructions, compared with verbal rehearsal. This novel finding supplements limited research on comparing motor imagery and action observation in memorizing action sequences (Gatti et al., 2013, adults). On the one hand, adopting and implementing a motor imagery strategy might conceivably be more demanding, compared to the benefits gained from observing demonstration of actions being carried out on objects. However, for adults, action planning abilities may be relatively mature, and so a self-generated memory representation can be as effective in enhancing performance as a representation provided by demonstration. Motor imagery may also provide additional benefits such as a sense of effort, agency and simulated kinesthetic input, which may compensate for any deficiency of self-generated representation (Vogt et al., 2013). In contrast, the ability to plan actions is still developing in children (Spruijt et al., 2015), and so self-generated memory representations of action sequence may be less effective relative to representations derived from demonstration. Children might then benefit more from action observation than from motor imagery, a hypothesis that awaits examination.

In the present study, motor imagery also yielded a similar pattern of performance as self-enactment. Unlike action observation, both motor imagery and self-enactment require action planning. Compared with motor imagery, self-enactment has additional physical motor movement components, which may provide an additional boost to memory. In fact, a recent meta-analysis involving 18-60 years old adults indicates a moderate benefit of self-enactment relative to imagery in long-term memory (Roberts et al., 2022), with the authors suggesting that the additional advantage of self-enactment is associated with genuine action planning and real physical actions. The findings from Experiment 2 suggest this might not apply to a working memory context. However, without a direct comparison between short- and long-term memory tasks using otherwise equivalent methodology, confident conclusions cannot be drawn about the relative pattern of effects and possible underlying mechanisms. This points to a new avenue of research comparing action-based advantages in different memory contexts.

In terms of the two encoding techniques that are often compared together in previous studies (i.e., action observation and self-enactment), the present study reported similar levels of memory performance whereas the previous findings are rather inconsistent. Three studies of following instructions reported a larger or a more stable action-observation benefit than self-enactment effect in both adults (Allen et al., 2020; Coats et al., 2021) and children (Waterman et al., 2017), whereas another study showed larger benefit of self-enactment than action observation in adults (Lui et al., 2018). Observation superiority has mainly been found when using simple objects (i.e., 2D geometric shapes). Engelkamp's multimodal theory suggests self-enactment facilitates item specific processing while action observation improves item relational processing in long-term memory (Engelkamp, 2001). For instance, research on goal-related action sequence found that action observation of goal-directed actions improved memory for order recall and subjective organization of action sequence, whereas self-enactment enhanced recognition of actions during delayed recall (Schult et al., 2014). In the long term memory research using pure lists and free recall, self-enactment improves item specific information at the cost of relational processing whereas action observation facilitates relational processing at the expense of item-specific information, and these benefits may cancel each other out (Steffens et al., 2015). Recent evidence suggests that self-enactment facilitates binding of actions onto to-be-performed objects in tasks involving following spoken instructions in young adults (Makri & Jarrold, 2021). Therefore, self-enactment may particularly facilitate binding of features within an action phrase in the present following instruction paradigm. In terms of action observation, while its mechanism in working memory remain unclear, we suspect the manipulation may also improve relational processing in working memory as in long-term memory. For instance, action observation may help associate actions in the sequence to form an integrated action representation. Therefore, the unique advantages of action observation and self-enactment have led to equivalent memory performance in the present study. As the advantage of self-enactment over action observation tend to vary with task paradigms in long-term memory such as the test type (recognition vs. free recall) and the list type (mixed vs. pure) (Steffens et al., 2015), working memory research may also consider this issue. For instance, when the study employs a recognition test that emphasizes item-specific information, self-enactment might then lead to better memory than action observation.

In short, future work is especially warranted to clarify the similarities and differences between the mechanisms underlying action observation and self-enactment in working memory. In addition, previous research indicated a developmental changes of enactment effects in long-term memory, showing as younger children (< 10 years old) tending to benefit more from action observation while older children (> 12 years old) benefit more from self-enactment (Badinlou et al., 2018; Badinlou et al., 2017; Yang & Wang, 2021), similar to adults (Roberts et al., 2022). These long-term findings concord with the aforementioned working memory study in which school-age children showed larger and more stable benefits from action-observation than self-enactment when encoding spoken instructions (Waterman et al., 2017). Together, these findings point out a potentially fruitful route for investigating different enactment effects in both working memory and long-term memory from a lifespan perspective.

The action-based encoding conditions implemented in this study in each case were compared against each other (Experiment 2), and against rehearsal (both experiments). The latter comparison was intended to serve as a condition in which verbal encoding was emphasized, and action-based processing minimized. Verbal rehearsal is commonly implemented as a spontaneous strategy by participants (e.g., Morrison et al., 2016, adults; Tam et al., 2010, children), although there is little evidence that it serves to enhance performance (Oberauer, 2019, a review). Typically, when participants are not instructed to use any directed encoding strategy in a working memory task, they will likely implement a myriad of techniques, with this varying between and even within individuals, tasks, and trials (e.g., Gonthier, 2021, a review). This is of a piece with the view that working memory ‘capacity’ reflects contributions from various contributory components drawn from a ‘cognitive toolbox’ (Logie et al., 2021). Thus, a ‘no-strategy condition’ would not in fact be strategy-free. Our use of an articulatory rehearsal condition in the present study therefore serves to provide a verbal contrast to imagery, observation, and self-enactment. Nevertheless, future work should explore the extent to which different forms of directed encoding strategy contrast both with each other and with conditions in which participants are free to spontaneously adopt strategic approaches to the task.

It remains to be seen how motor imagery might be incorporated along with action planning, enactment, and observation effects into theoretical approaches to working memory. Perceptual and motor storage and processing may be fundamentally important aspects of working memory when considering a ‘standard model’ of the mind (Laird et al., 2017), and several broad theoretical frameworks within working memory have at least begun to consider how such information may be handled (e.g., Baddeley et al., 2021; Barrouillet & Camos, 2021; Logie et al., 2021 see Li et al., 2022 for a brief discussion). Within the multicomponent model as described by Baddeley et al. (2021), for example, visuospatial and motoric information might be incorporated into a more detailed specification of the visuospatial sketchpad and combined with other sources of information within a modality-general episodic buffer (Allen & Waterman, 2015; Li et al., 2022), while the addition of a specialized motor store has also been suggested (Jaroslawska et al., 2018). Alternatively, the impacts of action-based processing on working memory may be considered by a theoretical approach that rejects the need for specialized buffers and instead views working memory as reflecting the operation of an object-oriented action system that co-opts perceptual-motor processing (e.g., Macken et al., 2015), though the extent to which such an approach primarily differs in terminology and levels of explanation remains controversial. At a broad level, for reasons of parsimony and simplicity we presently prefer the Baddeley et al. (2021) framework as related to memory for action and instruction by Li et al. (2022). However, the current study was not designed to differentiate between the broad leading frameworks that are available. Indeed, they are likely not sufficiently detailed to support meaningful speculation in this regard, though most do incorporate or acknowledge the likely availability of modality specific storage and processing capabilities along with a means of pulling together such information into conscious awareness. Effectively capturing how rich, multi-dimensional processing that includes different forms of action planning, input, and output might be achieved in working memory will require considerable theoretical development and supporting empirical evidence.

Research has started to explore to what extent multiple types of action-based intervention might improve long-term memory. For instance, researchers observed that combining action observation and motor imagery lead to superior superior cognitive and skill performance, relative to independent action-observation or motor imagery training (Kim et al., 2020). Similary, adding motor imagery during physical practice can further boost mental representations of motor skills in adults (Frank et al., 2014). The combined benefit of action observation and motor imagery has also been observed in rehabilitation in patients with Parkinson’s disease (Caligiore et al., 2017). As these findings focused on long-term memory, it remains unclear whether mutiple forms of action-based encoding can also boost following instructions in working memory, and whether it adds additional benefit above a single form of action-based encoding (e.g., self-enactment, motor imagery, action observation). Establishing this will have useful theoretical and practical implications.

In addition to encoding-based action benefits, we also found an enacted recall advantage, which occurred in all encoding conditions. The observation of an enacted recall advantage in action-encoding conditions might suggest additional benefits occurring at retrieval. Indeed, a previous study (Yang et al., 2019) has found that enactment retrieval led to more rapid access of memory representation and more efficient output production than oral repetition. To examine this possibility further, we re-analyzed the reaction time data of that study and found the enacted-recall advantage for preparation duration (*d* = 1.22) was larger than for recall duration (*d* = 0.73) in the spoken instruction conditions. This would imply a relatively larger contribution of representation accessibility than retrieval efficiency to the enacted-recall advantage. However, as we did not measure reaction time in the present study, we do not know whether the same results hold true for action-based encoding conditions (i.e., motor imagery, self enactment, action observation). Future research should include reaction time data to understand the potential contributions of retrieval process in enacted-recall advantages in working memory.

In conclusion, the present study provided empirical evidence that motor imagery, action observation and self-enactment all facilitate adults' ability to follow spoken instructions relative to rehearsal. These action-based benefits are present for both verbal and enacted recall and appear similar in size. It will be important to establish whether these different action-based encoding techniques have a similar effect on recall for children, or whether children obtain different levels of benefit from motor imagery, enactment and observation, relative to non-action based conditions. Understanding the presence and size of relative benefits to memory of different strategies in different populations will contribute to our understanding of how working memory develops, as well as informing applied contexts such as the education sector where the requirement to follow instructions is a frequent and critical part of the task environment (Gathercole et al., 2008).

Data availability statement

The data and analytic files are available via this link (https://www.scidb.cn/anonymous/UWZhNklz).

Acknowledgements

This study was supported by the National Natural Science Foundation of China (31400873) and the CAS Key Laboratory of Mental Health, Institute of Psychology. The funding agents had no further role in the study design; in the collection, analysis, and interpretation of the data; in the writing of the manuscript; or in the decision to submit the manuscript for publication.

Declaration of interest statement

The authors have no known conflict of interest to disclose.

References

Allen, R. J., Hill, L. J. B., Eddy, L., & Waterman, A. (2020). Exploring the effects of demonstration and enactment in facilitating recall of instructions in working memory. *Memory & Cognition*, *48*(3), 400-410. https://doi.org/10.3758/s13421-019-00978-6

Allen, R. J., & Waterman, A. H. (2015). How does enactment affect the ability to follow instructions in working memory? *Memory & Cognition*, *43*(3), 555-561. https://doi.org/10.3758/s13421-014-0481-3

Allen, R. J., Waterman, A. H., Yang, T., & Jaroslawska, A. J. (2023). Working memory in action: Remembering and following instructions. In R. H. Logie, Z. Wen, S. E. Gathercole, N. Cowan, & R. W. Engle (Eds.), *Memory in science for society: There is nothing as practical as a good theory*. Oxford University Press.

Arnold, L. (2012). Reproducing actions, reproducing power: Local ideologies and everyday practices of participation at a California community bike shop. *Journal of linguistic anthropology*, *22*(3), 137-158. https://doi.org/10.1111/j.1548-1395.2012.01153.x

Baddeley, A. D. (1986). *Working Memory*. Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G.H.Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47-89). Academic Press.

Baddeley, A. D., Hitch, G. J., & Allen, R. (2021). A multicomponent model of working memory. In R. H. Logie, V. Camos, & N. Cowan (Eds.), *Working memory: State-of-the science* (pp. 10-43). Oxford University Press.

Baddeley, A. D., Hitch, G. J., & Allen, R. J. (2009). Working memory and binding in sentence recall. *Journal of Memory and Language*, *61*, 438-456.

Badinlou, F., Kormi-Nouri, R., & Knopf, M. (2018). A study of retrieval processes in action memory for school-aged children: the impact of recall period and difficulty on action memory [Article]. *Journal of Cognitive Psychology*, *30*(8), 792-802. https://doi.org/10.1080/20445911.2018.1535495

Badinlou, F., Kormi-Nouri, R., Mousavi Nasab, S. M., & Knopf, M. (2017). Developmental differences in episodic memory across school ages: evidence from enacted events performed by self and others. *Memory*, *25*(1), 84-94. https://doi.org/10.1080/09658211.2015.1126607

Barrouillet, P., & Camos, V. (2021). The time-based resource-sharing model of working memory. In *Working memory: State of the science*. Oxford University Press, Incorporated.

Caligiore, D., Mustile, M., Spalletta, G., & Baldassarre, G. (2017). Action observation and motor imagery for rehabilitation in Parkinson's disease: A systematic review and an integrative hypothesis. *Neuroscience and Biobehavioral Reviews*, *72*, 210-222. https://doi.org/https://doi.org/10.1016/j.neubiorev.2016.11.005

Charlesworth, L. A., Allen, R. J., Morson, S., Burn, W. K., & Souchay, C. (2014). Working memory and the enactment effect in early Alzheimer's disease. *ISRN Neurology*, *2014*, 694761. https://doi.org/10.1155/2014/694761

Coats, R. O., Waterman, A. H., Ryder, F., Atkinson, A. L., & Allen, R. J. (2021). Following instructions in working memory: Do older adults show the enactment advantage? *The Journals of Gerontology: Series B*, *76*(4), 703-710. https://doi.org/10.1093/geronb/gbaa214

Deppermann, A. (2015). When recipient design fails: Egocentric turn-design of instructions in driving school lessons leading to breakdowns of intersubjectivity. *Gesprächsforschung-Online-Zeitschrift zur verbalen Interaktion*, *16*, 63-101.

Dunlosky, J., & Kane, M. J. (2007). The contributions of strategy use to working memory span: A comparison of strategy assessment methods. *Quarterly Journal of Experimental Psychology*, *60*(9), 1227-1245. https://doi.org/10.1080/17470210600926075

Ecker, W., & Engelkamp, J. (1995). Memory for actions in obsessive-compulsive disorder. *Behavioural and Cognitive Psychotherapy*, *23*(4), 349-371. https://doi.org/10.1017/S1352465800016477

Engelkamp, J. (2001). Action memory: A system-oriented approach. In H. D. Zimmer, R.Cohen, M. Guynn, J. Engelkamp, R. Kormi-Nouri, & M. Foley (Eds.), *Memory for action: A distinct form of episodic memory*. Oxford University Press.

Engelkamp, J., & Zimmer, H. D. (1989). Memory for action events: A new field of research. *Psychological Research*, *51*(4), 153-157. https://doi.org/10.1007/BF00309142

Engle, R. W., Carullo, J. J., & Collins, K. W. (1991). Individual differences in working memory for comprehension and following directions. *Journal of Educational Research*, *84*(5), 253-262.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149-1160. https://doi.org/10.3758/brm.41.4.1149

Frank, C., Land, W. M., Popp, C., & Schack, T. (2014). Mental representation and mental practice: Experimental investigation on the functional links between motor memory and motor imagery. *PloS One*, *9*(4), e95175. https://doi.org/10.1371/journal.pone.0095175

Friedman, N. P., & Miyake, A. (2004). The reading span test and its predictive power for reading comprehension ability. *Journal of Memory and Language*, *51*(1), 136-158. https://doi.org/10.1016/j.jml.2004.03.008

Gathercole, S. E., Durling, E., Evans, M., Jeffcock, S., & Stone, S. (2008). Working memory abilities and children's performance in laboratory analogues of classroom activities. *Applied Cognitive Psychology*, *22*, 1019–1037. https://doi.org/10.1002/acp.1407

Gatti, R., Tettamanti, A., Gough, P. M., Riboldi, E., Marinoni, L., & Buccino, G. (2013). Action observation versus motor imagery in learning a complex motor task: A short review of literature and a kinematics study. *Neuroscience Letters*, *540*, 37-42. https://doi.org/10.1016/j.neulet.2012.11.039

Gonthier, C. (2021). Charting the diversity of strategic processes in visuospatial short-term memory. *Perspectives on Psychological Science*, *16*(2), 294-318. https://doi.org/10.1177/1745691620950697

Hardwick, R. M., Caspers, S., Eickhoff, S. B., & Swinnen, S. P. (2018). Neural correlates of action: Comparing meta-analyses of imagery, observation, and execution. *Neuroscience and Biobehavioral Reviews*, *94*, 31-44. https://doi.org/10.1016/j.neubiorev.2018.08.003

Jaroslawska, A. J., Bartup, G., Forsberg, A., & Holmes, J. (2021). Age-related differences in adults' ability to follow spoken instructions. *Memory*, *29*(1), 1-12. https://doi.org/10.1080/09658211.2020.1860228

Jaroslawska, A. J., Gathercole, S. E., Allen, R. J., & Holmes, J. (2016). Following instructions from working memory: Why does action at encoding and recall help? *Memory & Cognition*, *44*(8), 1183-1191. https://doi.org/10.3758/s13421-016-0636-5

Jaroslawska, A. J., Gathercole, S. E., & Holmes, J. (2018). Following instructions in a dual-task paradigm: Evidence for a temporary motor store in working memory. *Quarterly Journal of Experimental Psychology*, *71*(11), 2439-2449. https://doi.org/10.1177/1747021817743492

Jeffreys, H. (1961). *Theory of probability* (3rd ed.). Oxford University Press.

Kim, T., Frank, C., & Schack, T. (2020). The effect of alternate training of action observation and motor imagery on cognitive and skill performance. *International Journal of Sport Psychology*, *51*(2), 101-121. https://doi.org/10.7352/IJSP.2020.51.101

Koriat, A., Ben-Zur, H., & Nussbaum, A. (1990). Encoding information for future action: Memory for to-be-performed tasks versus memory for to-be-recalled tasks. *Memory & Cognition*, *18*(6), 568-578. https://doi.org/10.3758/BF03197099

Kormi-Nouri, R., Nyberg, L., & Nilsson, L.-G. (1994). The effect of retrieval enactment on recall of subject-performed tasks and verbal tasks. *Memory & Cognition*, *22*(6), 723-728. https://doi.org/10.3758/BF03209257

Laird, J. E., Lebiere, C., & Rosenbloom, P. S. (2017). A standard model of the mind: Toward a common computational framework across artificial intelligence, cognitive science, neuroscience, and robotics. *Ai Magazine*, *38*(4), 13-26. https://doi.org/10.1609/aimag.v38i4.2744

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian cognitive modeling: A practical course*. https://doi.org/10.1017/CBO9781139087759

Li, G., Allen, R. J., Hitch, G. J., & Baddeley, A. D. (2022). Translating words into actions in working memory: the role of spatial-motoric coding. *Quarterly Journal of Experimental Psychology*. https://doi.org/10.1177/17470218221079848

Logie, R. H., Belletier, C., & Doherty, J. M. (2021). Integrating theories of working memory. In R. H. Logie, V. Camos, & N. Cowan (Eds.), *Working memory: State of the science.* (pp. 389-429). Oxford University Press.

Lui, S., Yang, T., Ng, C., Wong, T., Wong, J., Ettinger, U., . . . Chan, R. C. K. (2018). Following instructions in patients with schizophrenia: The benefits of actions at encoding and recall. *Schizophrenia Bulletin*, *44*(1), 137-146. https://doi.org/10.1093/schbul/sbx026

Ma, J., Wang, L., Chen, L., & Zhang, Y. (2021). Imagery processing in action memory-mental imagery is necessary to the subject-performed task effect. *Journal of Cognitive Psychology (Hove, England)*, *33*(1), 12-23. https://doi.org/10.1080/20445911.2020.1862129

Macken, B., Taylor, J., & Jones, D. (2015). Limitless capacity: a dynamic object-oriented approach to short-term memory. *Frontiers in Psychology*, *6*, 293. https://doi.org/10.3389/fpsyg.2015.00293

Makri, A., & Jarrold, C. (2021). Investigating the underlying mechanisms of the enactment effect: The role of action–object bindings in aiding immediate memory performance. *Quarterly Journal of Experimental Psychology*, *74*(12), 137-146. https://doi.org/10.1177/17470218211019026

Mentis, H. M., Chellali, A., & Schwaitzberg, S. (2014). Learning to see the body: Supporting instructional practices in laparoscopic surgical procedures. SIGCHI conference on human factors in computing systems,

Mohr, G., Engelkamp, J., & Zimmer, H. D. (1989). Recall and recognition of self-performed acts. *Psychological Research*, *51*(4), 181-187.

Morrison, A. B., Rosenbaum, G. M., Fair, D., & Chein, J. M. (2016). Variation in strategy use across measures of verbal working memory. *Memory & Cognition*, *44*(6), 922-936. https://doi.org/10.3758/s13421-016-0608-9

Norris, M. P., & West, R. L. (1993). Activity memory and aging: The role of motor retrieval and strategic processing. *Psychology and Aging*, *8*(1), 81-86. https://doi.org/10.1037/0882-7974.8.1.81

Oberauer, K. (2019). Is rehearsal an effective maintenance strategy for working memory? *Trends In Cognitive Sciences*, *23*(9), 798-809. https://doi.org/10.1016/j.tics.2019.06.002

Plancher, G., Mazeres, F., & Vallet, G. T. (2019). When motion improves working memory. *Memory (Hove)*, *27*(3), 410-416. https://doi.org/10.1080/09658211.2018.1510012

Roberts, B. R. T., MacLeod, C. M., & Fernandes, M. A. (2022). The enactment effect: A systematic review and meta-analysis of behavioral, neuroimaging, and patient studies. *Psychological Bulletin*, *148*(5-6), 397-434. https://doi.org/10.1037/bul0000360

Saltz, E., & Dixon, D. (1982). Let's pretend: The role of motoric imagery in memory for sentences and words* 1. *Journal of Experimental Child Psychology*, *34*(1), 77-92.

Schult, J., Von Stülpnagel, R., & Steffens, M. C. (2014). Enactment versus observation: Item-specific and relational processing in goal-directed action sequences (and lists of single actions). *PloS One*, *9*(6), e99985-e99985. https://doi.org/10.1371/journal.pone.0099985

Spruijt, S., van der Kamp, J., & Steenbergen, B. (2015). Current insights in the development of children's motor imagery ability. *Frontiers in Psychology*, *6*(787). https://doi.org/10.3389/fpsyg.2015.00787

Steffens, M. C., Von Stülpnagel, R., & Schult, J. C. (2015). Memory recall after "learning by doing" and "learning by viewing": Boundary conditions of an enactment benefit. *Frontiers in Psychology*, *6*, 1907-1907. https://doi.org/10.3389/fpsyg.2015.01907

Tam, H., Jarrold, C., Baddeley, A. D., & Sabatos-DeVito, M. (2010). The development of memory maintenance: Children's use of phonological rehearsal and attentional refreshment in working memory tasks. *Journal of Experimental Child Psychology*, *107*(3), 306-324.

Unsworth, N. (2016). Working memory capacity and recall from long-term memory: Examining the influences of encoding strategies, study time allocation, search efficiency, and monitoring abilities. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(1), 50-61. https://doi.org/10.1037/xlm0000148

Vogt, S., Rienzo, F. D., Collet, C., Collins, A., & Guillot, A. (2013). Multiple roles of motor imagery during action observation. *Frontiers in Human Neuroscience*, *7*, 807. https://doi.org/10.3389/fnhum.2013.00807

Waterman, A. H., Atkinson, A. L., Aslam, S. S., Holmes, J., Jaroslawska, A., & Allen, R. J. (2017). Do actions speak louder than words? Examining children's ability to follow instructions. *Memory & Cognition*, *45*(6), 877-890. https://doi.org/10.3758/s13421-017-0702-7

Wojcik, D. Z., Allen, R. J., Brown, C., & Souchay, C. (2011). Memory for actions in autism spectrum disorder. *Memory*, *19* (6), 549-558. https://doi.org/10.1080/09658211.2011.590506

Yang, J., & Wang, L. (2021). Could actions improve retrieval performance in 4-to 6-year-old children? *European Journal of Developmental Psychology*, *18*(3), 396-411. https://doi.org/10.1080/17405629.2020.1794808

Yang, T., Allen, R. J., & Gathercole, S. E. (2016). Examining the role of working memory resources in following spoken instructions. *Journal of Cognitive Psychology*, *28*(2), 186-198.

https://doi.org/10.1080/20445911.2015.1101118

Yang, T., Allen, R. J., Holmes, J., & Chan, R. C. K. (2017). Impaired memory for instructions in children with attention-deficit hyperactivity disorder is improved by action at presentation and recall. *Frontiers in Psychology*, *8*(39). https://doi.org/10.3389/fpsyg.2017.00039

Yang, T., Allen, R. J., Waterman, A. H., Zhang, S., Su, X., & Chan, R. C. K. (2021). Comparing motor imagery and verbal rehearsal strategies in children's ability to follow spoken instructions. *Journal of Experimental Child Psychology*, *203*, 105033. https://doi.org/10.1016/j.jecp.2020.105033

Yang, T., Allen, R. J., Yu, Q., & Chan, R. C. K. (2015). The influence of input and output modality on following instructions in working memory. *Scientific Reports*, *5*(1), Article 17657. https://doi.org/10.1038/srep17657

Yang, T., Gathercole, S. E., & Allen, R. J. (2014). Benefit of enactment over oral repetition of verbal instruction does not require additional working memory during encoding. *Psychonomic Bulletin & Review*, *21*(1), 186-192. https://doi.org/10.3758/s13423-013-0471-7

Yang, T., Jia, L., Zheng, Q., Allen, R. J., & Ye, Z. (2019). Forward and backward recall of serial actions: Exploring the temporal dynamics of working memory for instruction. *Memory & Cognition*, *47*(2), 279-291. https://doi.org/10.3758/s13421-018-0865-x

Yang, T., Su, X.-m., Allen, R. J., Ye, Z., & Jia, L.-x. (2022). Improving older adults' ability to follow instructions: benefits of actions at encoding and retrieval in working memory. *Memory*, *30*(5), 610-620. https://doi.org/10.1080/09658211.2022.2035768

Zimmer, H. D., Helstrup, T., & Engelkamp, J. (2000). Pop-out into memory: A retrieval mechanism that is enhanced with the recall of subject-performed tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(3), 658-670. https://doi.org/10.1037/0278-7393.26.3.658

Tables

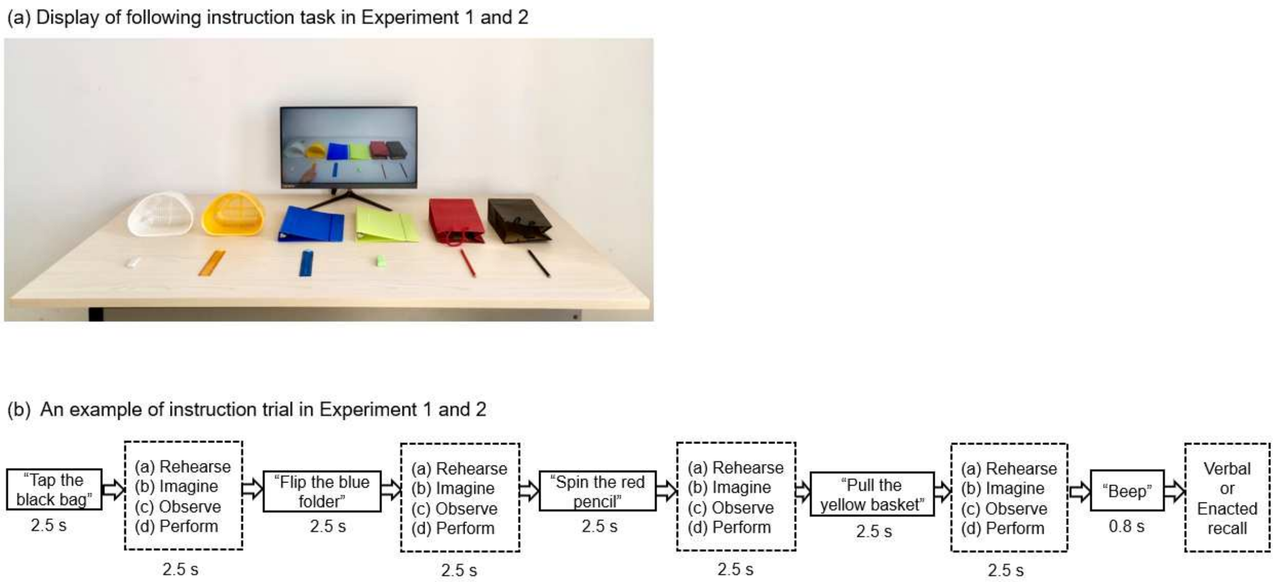
Table 1. Descriptive results of proportion correct of action-object pairs and action advantages in Experiment 1 and 2

	Verbal Recall		Enacted recall		Enacted vs Verbal recall
	M (SD)	*Cohen's d*	*M (SD)*	*Cohen's d*	*Cohen's d*
Experiment 1					
Verbal rehearsal condition	0.54 (0.14)		0.71 (0.16)		1.15
Motor imagery condition	0.64 (0.16)		0.85 (0.12)		1.57
Motor imagery vs verbal rehearsal		0.67		1.06	
Experiment 2					
Verbal rehearsal condition	0.57 (0.16)		0.69 (0.15)		0.81
Motor imagery condition	0.72 (0.19)		0.85 (0.10)		0.88
Action observation condition	0.74 (0.15)		0.89 (0.11)		1.15
Self-enactment condition	0.73 (0.17)		0.85 (0.10)		0.90
Motor imagery vs verbal rehearsal		0.61		0.91	
Action observation vs verbal rehearsal		0.78		1.12	
Self-enactment vs verbal rehearsal		0.69		0.90	

Figures

Figure 1

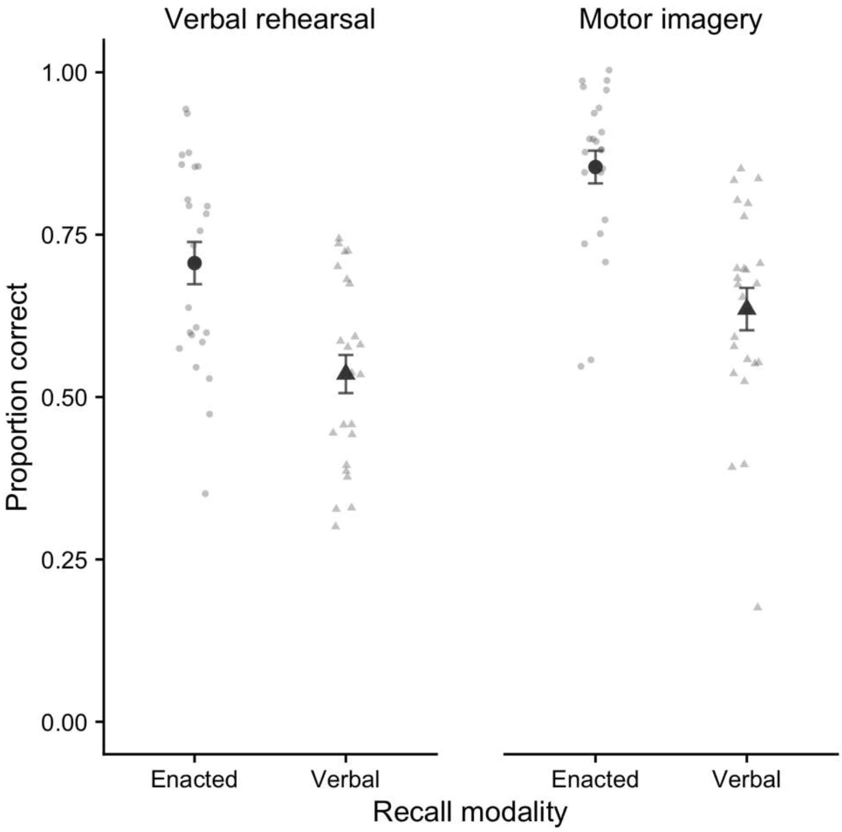
The following instruction task



Notes. Experiment 1 contained the (a) Rehearse and (b) Imagine conditions. Experiment 2 contained all four conditions. The computer screen was turned off in all conditions in Experiment 1 and 2 except for the action observation condition in Experiment 2.

Figure 2

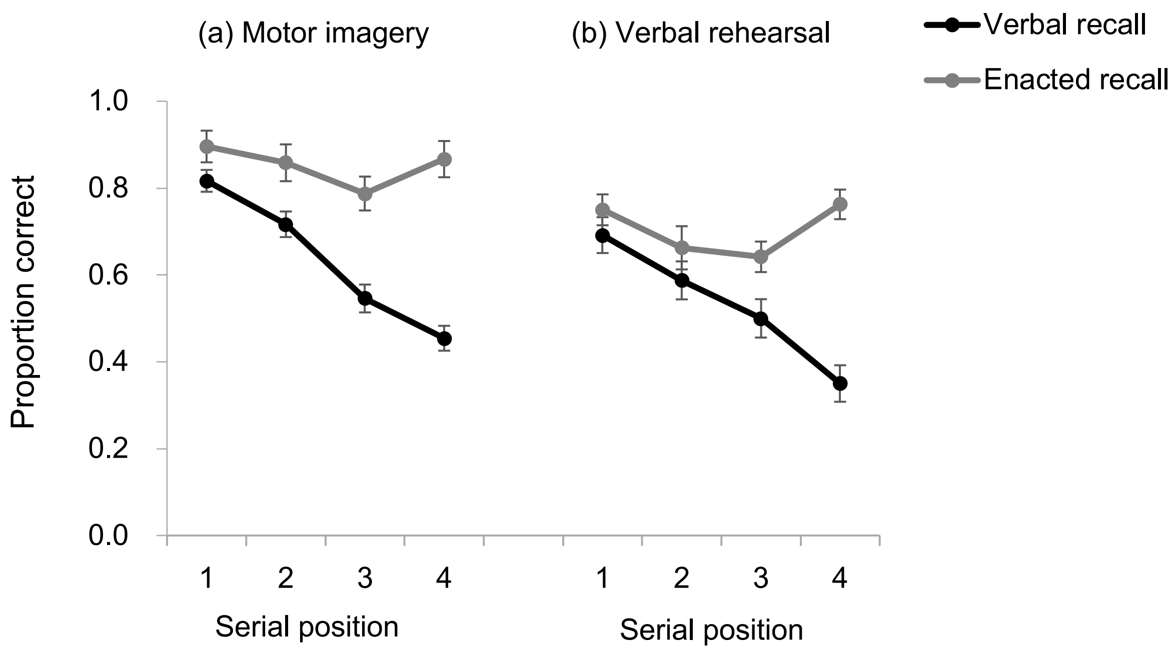
Proportion correct of action-object pairs as a function of encoding technique and recall modality in Experiment 1



Notes. Error bars represent standard errors, and grey points show individual participants.

Figure 3

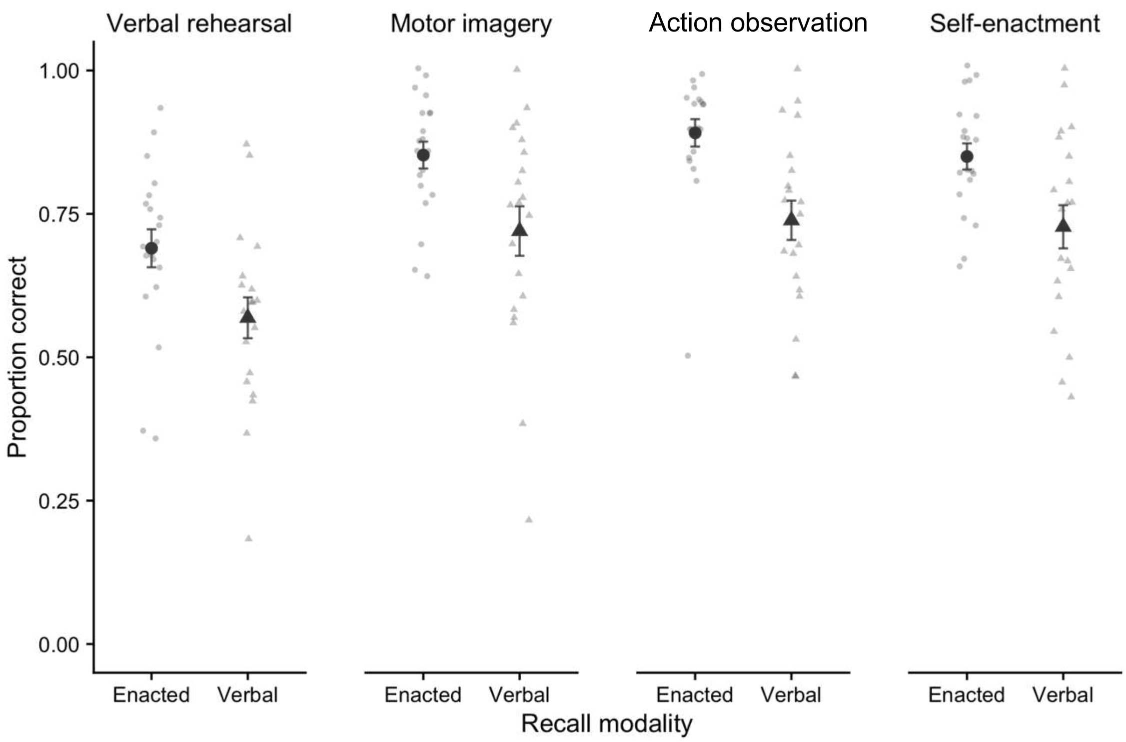
Proportion correct of action-object pairs as a function of serial position, encoding technique, and recall modality in Experiment 1



Notes. Error bars represent standard errors.

Figure 4

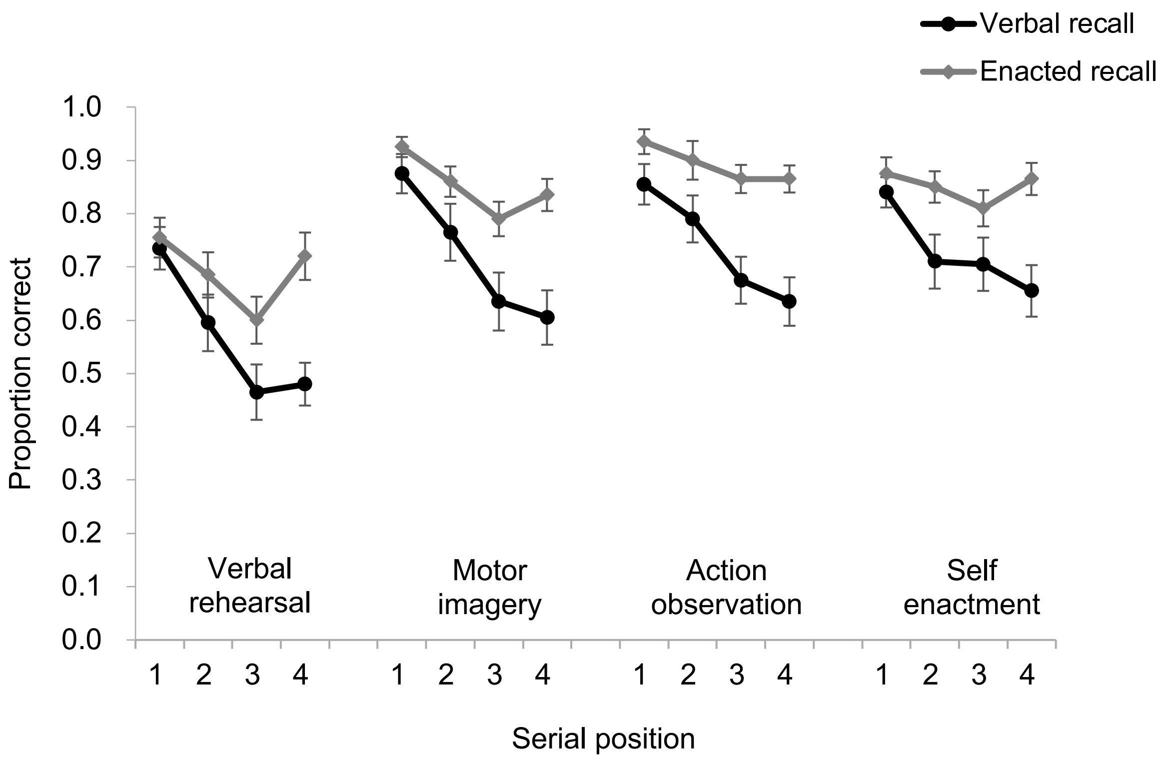
Proportion correct of action-object pairs as a function of encoding technique and recall modality in Experiment 2



Notes. Error bars represent standard errors, and grey points show individual participants.

Figure 5

Proportion correct of action-object pairs as a function of serial position, encoding technique, and recall modality in Experiment 2



Notes. Error bars represent standard errors.

Figure captions as a list

Figure 1. The following instruction task

Figure 2. Proportion correct of action-object pairs as a function of encoding technique and recall modality in Experiment 1

Figure 3. Proportion correct of action-object pairs as a function of serial position, encoding technique, and recall modality in Experiment 1

Figure 4. Proportion correct of action-object pairs as a function of encoding technique and recall modality in Experiment 2

Figure 5. Proportion correct of action-object pairs as a function of serial position, encoding technique, and recall modality in Experiment 2

Appendix A

Lists of instructions for Experiment 1

Practice trials

1. Tap the white basket, pull the yellow ruler, lift the red bag, push the black pencil
2. Push the black bag, flip the blue folder, spin the red pencil, pull the yellow basket

Formal trials

1. Pull the green folder, tap the blue ruler, push the white basket, flip the black bag
2. Spin the red bag, pull the white eraser, tap the yellow basket, push the green folder
3. Flip the yellow ruler, spin the black pencil, lift the blue folder, tap the white eraser
4. Lift the black pencil, pull the red pencil, flip the white eraser, spin the blue ruler
5. Tap the green eraser, flip the yellow basket, push the blue ruler, pull the white basket
6. Flip the blue ruler, spin the green folder, pull the black bag, lift the yellow ruler
7. Pull the blue folder, push the green eraser, lift the black pencil, tap the red bag
8. Flip the red pencil, lift the white basket, spin the green eraser, push the blue folder
9. Lift the white eraser, push the red bag, tap the green folder, spin the yellow ruler
10. Spin the yellow basket, tap the black bag, lift the red pencil, flip the green eraser

List 2

Practice trials

1. Lift the blue folder, push the yellow ruler, tap the white eraser, pull the red pencil
2. Pull the yellow basket, spin the blue ruler, push the green folder, flip the white basket

Formal trials

1. Push the red bag, flip the green eraser, pull the black pencil, tap the black bag
2. Tap the yellow ruler, push the red pencil, spin the white basket, Pull the blue ruler
3. Lift the black bag, tap the blue folder, flip the green eraser, spin the red bag
4. Flip the white eraser, spin the yellow basket, lift the blue ruler, pull the green eraser

5. Push the blue ruler, pull the green folder, tap the black bag, flip the yellow ruler
6. Pull the green eraser, lift the red bag, flip the blue folder, spin the yellow basket
7. Lift the green folder, tap the white basket, pull the red pencil, push the white eraser
8. Spin the black pencil, push the white eraser, flip the red bag, lift the blue folder
9. Tap the yellow basket, spin the red pencil, lift the black bag, push the black pencil
10. Lift the white basket, flip the black pencil, spin the yellow ruler, tap the green folder

List 3

Practice trials

1. Pull the red bag, Tap the black pencil, push the blue folder, lift the white eraser
2. Push the green folder, flip the yellow basket, pull the white eraser, spin the blue ruler

Formal trials

1. Tap the red pencil, pull the blue folder, flip the green folder, push the black bag
2. Pull the green eraser, push the blue ruler, spin the red pencil, tap the black pencil
3. Spin the white basket, flip the green eraser, tap the yellow ruler, lift the blue folder
4. Pull the black pencil, lift the yellow ruler, spin the yellow basket, flip the red bag
5. Flip the yellow ruler, tap the green folder, pull the black bag, push the white basket
6. Spin the blue folder, flip the red pencil, lift the white basket, Pull the green eraser
7. Push the white basket, pull the green folder, tap the black bag, lift the blue ruler
8. Lift the white eraser, flip the red bag, push the green eraser, spin the yellow ruler
9. Spin the blue ruler, lift the black bag, tap the yellow basket, push the red bag
10. Tap the white eraser, spin the yellow basket, flip the red pencil, lift the black pencil

List 4

Practice trials

1. Tap the green eraser, spin the yellow basket, push the black pencil, flip the white eraser
2. Push the green folder, flip the yellow ruler, spin the red bag, pull the blue folder

Formal trials

1. Pull the blue ruler, lift the white basket, push the green eraser, spin the red pencil
2. Lift the blue folder, flip the red bag, pull the black bag, tap the green folder
3. Pull the yellow basket, spin the green eraser, tap the white eraser, lift the yellow ruler
4. Tap the black bag, pull the red pencil, flip the blue folder, push the white basket
5. Spin the yellow ruler, flip the red bag, push the yellow basket, pull the black bag
6. Flip the white basket, lift the blue folder, spin the red pencil, tap the white eraser
7. Pull the white eraser, tap the black pencil, flip the blue ruler, lift the yellow basket
8. Flip the red pencil, push the blue ruler, lift the green folder, spin the black pencil
9. Spin the green eraser, lift the black bag, tap the white basket, push the red bag
10. Lift the black pencil, push the green folder, pull the yellow ruler, tap the blue ruler

Lists of instructions for Experiment 2

List 1

Practice trials

1. Tap the white basket, pull the yellow ruler, lift red bag, shake the black pencil
2. Shake the black bag, flip the blue folder, spin the red pencil, pull the yellow basket

Formal trials

1. Pull the green folder, tap the blue ruler, shake the white basket, flip the black bag
2. Spin the red bag, pull the white eraser, tap the yellow basket, shake the green folder
3. Flip the yellow ruler, spin the black pencil, lift blue folder, tap the white eraser
4. Lift black pencil, pull the red pencil, flip the white eraser, spin the blue ruler
5. Tap the green eraser, flip the yellow basket, shake the blue ruler, pull the white basket
6. Flip the blue ruler, spin the green folder, pull the black bag, lift yellow ruler
7. Pull the blue folder, shake the green eraser, lift black pencil, tap the red bag
8. Flip the red pencil, lift white basket, spin the green eraser, shake the blue folder
9. Lift white eraser, shake the red bag, tap the green folder, spin the yellow ruler
10. Spin the yellow basket, tap the black bag, lift red pencil, flip the green eraser

List 2

Practice trials

1. Lift blue folder, shake the yellow ruler, tap the white eraser, pull the red pencil
2. Pull the yellow basket, spin the blue ruler, shake the green folder, flip the white basket

Formal trials

1. Shake the red bag, flip the green eraser, pull the black pencil, tap the black bag
2. Tap the yellow ruler, shake the red pencil, spin the white basket, pull the blue ruler
3. Lift black bag, tap the blue folder, flip the green eraser, spin the red bag
4. Flip the white eraser, spin the yellow basket, lift blue ruler, pull the green eraser
5. Shake the blue ruler, pull the green folder, tap the black bag, flip the yellow ruler
6. Pull the green eraser, lift red bag, flip the blue folder, spin the yellow basket
7. Lift green folder, tap the white basket, pull the red pencil, shake the white eraser
8. Spin the black pencil, shake the white eraser, flip the red bag, lift blue folder
9. Tap the yellow basket, spin the red pencil, lift black bag, shake the black pencil
10. Lift white basket, flip the black pencil, spin the yellow ruler, tap the green folder

List 3

Practice trials

1. Pull the red bag, tap the black pencil, shake the blue folder, lift white eraser
2. Shake the green folder, flip the yellow basket, pull the white eraser, spin the blue ruler

Formal trials

1. Tap the red pencil, pull the blue folder, flip the green folder, shake the black bag
2. Pull the green eraser, shake the blue ruler, spin the red pencil, tap the black pencil
3. Spin the white basket, flip the green eraser, tap the yellow ruler, lift blue folder
4. Pull the black pencil, lift yellow ruler, spin the yellow basket, flip the red bag
5. Flip the yellow ruler, tap the green folder, pull the black bag, shake the white basket
6. Spin the blue folder, flip the red pencil, lift white basket, pull the green eraser

7. Shake the white basket, pull the green folder, tap the black bag, lift blue ruler
8. Lift white eraser, flip the red bag, shake the green eraser, spin the yellow ruler
9. Spin the blue ruler, lift black bag, tap the yellow basket, shake the red bag
10. Tap the white eraser, spin the yellow basket, flip the red pencil, lift black pencil

List 4

Practice trials

1. Tap the green eraser, spin the yellow basket, shake the black pencil, flip the white eraser
2. Shake the green folder, flip the yellow ruler, spin the red bag, pull the blue folder

Formal trials

1. Pull the blue ruler, lift white basket, shake the green eraser, spin the red pencil
2. Lift blue folder, flip the red bag, pull the black bag, tap the green folder
3. Pull the yellow basket, spin the green eraser, tap the white eraser, lift yellow ruler
4. Tap the black bag, pull the red pencil, flip the blue folder, shake the white basket
5. Spin the yellow ruler, flip the red bag, shake the yellow basket, pull the black bag
6. Flip the white basket, lift blue folder, spin the red pencil, tap the white eraser
7. Pull the white eraser, tap the black pencil, flip the blue ruler, lift yellow basket
8. Flip the red pencil, shake the blue ruler, lift green folder, spin the black pencil
9. Spin the green eraser, lift black bag, tap the white basket, shake the red bag
10. Lift black pencil, shake the green folder, pull the yellow ruler, tap the blue ruler

Appendix B

Procedure and instructions for Experiment 1 and 2

Most instructions were the same for Experiment 1 and 2, with differences indicated in the text or in the footnotes.

1. Practice of naming objects

The experimenter said, ‘This is a memory test. Before testing, let us get familiar with these items. The objects include: a white eraser, a yellow ruler, a blue ruler, a green eraser, a red pencil, a black pencil, a white basket, a yellow basket, a blue folder, a green folder, a red bag, a black bag. Now please tell me the name of these objects from left to right.’ If the participants provided an incorrect answer, the experimenter corrected it immediately and required the participants to name it again until all the names of objects were provided correctly.

2. Practice of the actions

The experimenter said, ‘Next, let us get familiar with the actions.’ The experimenter demonstrated each action while speaking the names (*push/shake*[4]*, pull, spin, tap, lift, flip*). While speaking the action *push*, *pull*, and *lift*, the experimenter moved the object approximately 5 cm away from the original position. For the action *tap*, the experimenter patted the object from above. For the action *flip*, the

[4] In Experiment 2, the action “push” was replaced with “shake” to avoid confusion of the actions “pull” and “push” during self-enactment condition. While speaking the action *shake*, the experimenter picked and jiggled the object.

experimenter picked up the object and turned it upside down. For the action *spin*, the experimenter rotated the object clockwise by 45 degrees (only for demonstration, the exact rotation angle was not required in the formal test). For the action *shake*, the experimenter picked up the object and jiggled it from side to side. After demonstration, the experimenter said, 'Now, I will do each action and you tell me the name of the action.' Six action-object pairs were provided (*push/shake the white eraser; pull the yellow ruler; spin the blue ruler; tap the green eraser; lift the red pencil; flip the black pencil*). The experimenter then performed each action one at a time and the participants gave the name of the action, any inaccurate answer was immediately corrected by the experimenter. Finally, the experimenter said, 'Next, I will say the names of the actions, and you do the actions.' Six action-object pairs were provided (*push/shake the white basket; pull the yellow basket; spin the blue folder; tap the green folder; lift the red bag; flip the black bag*). The experimenter needed to check whether the participants could understand and operate the actions correctly and smoothly, as well as checking that the six types of actions were distinguishable. The details of the movement such as the distance of movement and the angle of rotation were not strictly required to be exactly the same as the experimenter's demonstration. Any mistake when performing actions during practice was immediately corrected by the experimenter.

3. Overall introduction of the experiment

The experimenter said, 'There were four conditions of the test, and I will introduce the overall rules of the test. You will hear spoken instructions or see videos of demonstrated instructions played on the computer. Each instruction contains a series of action-object pairs such as *push/shake*[5] *the white eraser*, *tap the yellow ruler*. Then, the experimenter said, ‘There are silent pauses between action-object pairs, and you need to memorize the actions using the strategy required by the experimenter. At the end of an instruction, when you hear a reminding tone *du*, you need to recall the entire instruction in order.’

Next, the experimenter said, ‘There are some rules for the task. First, you need to put your hands on your lap before each instruction starts. Second, while the instruction is playing, you need to listen carefully and avoid making any hand movements. Third, you need to remember all the action-object pairs in the correct order. If you forget an element, you can say “*blank”* instead. A blank represents a forgotten action or object. For example, if the instruction is shake the white eraser and tap the yellow ruler, and 1) you forget the action shake, you need to say *blank* the white eraser and tap the yellow ruler; 2) you forget the object white eraser, you need to say shake the *blank* and tap the yellow ruler; 3) you forget the first action-object pair, you need to say *blank* and tap the yellow ruler. Now, let’s start the first game.’ In the formal test, if a participant forgot to say *blank*, the experimenter would remind the participant after the trial ended.

4. Task introduction and practice instructions for different conditions

There were four conditions in Experiment 1 (i.e., conditions 1-4), and eight conditions in Experiment 2 (i.e., conditions 1-8). The descriptions of the task introduction and practice are presented below.

[5] In Experiment 2, the action “push” was replaced with “shake” to avoid confusion of the actions “pull” and “push” during self-enactment condition. While speaking the action *shake*, the experimenter picked and jiggled the object.

The instructions for the formal test were the same and are presented in part 5.

Condition 1：Verbal rehearsal-verbal recall condition

The experimenter first introduced the task requirement by saying, ‘I will play the spoken instructions on my computer. Each instruction contains a series of action-object pairs and there is a silent pause after each one. During this time, you need to *repeat the action-object pairs quietly*. At the end of the instruction, you will hear the reminding tone *du*, and you need to *speak out* loud all the action-object pairs of the instruction in order.’

The experimenter then demonstrated the procedure, and said, ‘I will demonstrate, and you must observe how I perform it.’ The experimenter played the sample recording and repeated each action *quietly*, and *repeated* the entire instruction after hearing the *du* sound. After the demonstration, the experimenter said, ‘This time, I will play the instructions, and you do the task as I just demonstrated. After each action, you must repeat it quietly, and when you hear the *du* sound, say out loud all the action-object pairs in order. Let’s practice.’ The experimenter played two practice trials, and observed and corrected any rule-breaking behaviors until the participants could perform the task as required. The experimenter paid attention to the typical errors that might occur during the encoding stage (e.g., repeating the instructions aloud or silently; pointing to or touching objects; simulating enactment). These rule-breaking behaviors were prohibited by the experimenter.

After the experimenter had ensured the participants could understand and perform the task correctly, the experimenter started the formal test.

Condition 2：Verbal rehearsal-enacted recall condition

The experimenter first introduced the task requirement by saying, ‘I will play the spoken instructions on my computer. Each instruction contains a series of action-object pairs and there is a silent pause after each one. During this time, you need to *repeat the action-object pairs quietly*. At the end of the instruction, you will hear the reminding tone *du*, and you need to *perform* all the action-object pairs of the instruction in order. Please be aware that you can use one hand or both hands to complete the task, but do not use both hands at the same time on the same object.’

The experimenter then demonstrated the procedure, and said: “I will demonstrate, and you must observe how I perform it.’ The experimenter played the sample recording and *repeated each action quietly*, and *performed* the entire instruction sequentially after the *du* sound. After recall, the experimenter put all the objects back in their initial locations. After the demonstration, the experimenter said, ‘This time, I will play the instructions, and you do as I just demonstrated. After each action, you should repeat it quietly, and when you hear the *du* sound, perform all the action-object pairs in order. Let’s practice.’ The practice procedure was similar to that in Condition 1 except for the additional typical error (i.e., placing the object back immediately after enactment in the middle of recalling process). After practice, the formal test began.

Condition 3：Motor imagery-verbal recall condition

The experimenter first introduced the task requirement by saying, ‘I will play the spoken instructions on my computer. Each instruction contains a series of action-object pairs and there is a silent pause after each one. During this time, you need to *imagine completing the action in your mind*. At the end of the instruction, you will hear the reminding tone *du*, and you need to *speak out loud* all of the action-object pairs of the instruction in order.'

The experimenter then demonstrated the procedure, and said, 'I will demonstrate and describe how I imagine the actions; but in the formal test, you do not need to speak out loud what you are doing'. The experimenter then played the first sample recording *pull the white eraser*, and said 'In my mind, I imagine my hand moves away from lap, stretches towards the white eraser, takes hold of the white eraser and pulls it towards me, and the white eraser leaves the previous location and moves toward me'. The experimenter then played the second sample recording *tap the yellow ruler*, and said 'In my mind, I imagine my hand moves away from lap, stretches towards the yellow ruler, and taps it'.

After this demonstration, the experimenter said, 'This time, I will play the instructions, and you do as I just demonstrated. Please try to imagine the process of completing the action after hearing it. You don't need to speak out loud what you are doing as I just demonstrated, just imagining it in your head is fine. When you hear the *du* sound, *speak out* loud all the action-object pairs in order'. The experimenter then played the two practice trials (same as demonstrations). After the participants recalled the instructions, the experimenter asked, 'To what extent do you think you can imagine the actions in the instruction?' then the participants selected from the following options (1=cannot imagine the actions at all; 2=have difficulty imagining the actions; 3=can imagine the actions a little; 4=can imagine the actions quite a lot; 5=can imagine the actions completely). If the participants reported difficulty with motor imagery (option 1 or 2), the participants had further motor imagery practice with eyes closed, or speaking out loud as they engaged in the imagery process. In the formal test, all participants were required to imagine the actions silently with eyes open.

After that, the experimenter said, ‘Let’s have some more practice.’ The experimenter played two practice trials, and observed and corrected any rule-breaking behaviors until the participants could perform the task as required. The formal test then began.

Condition 4：Motor imagery-enacted recall condition

The experimenter first introduced the task by saying, ‘I will play the spoken instructions on my computer. Each instruction contains a series of action-object pairs and there is a silent pause after each one. During this time, you need to *imagine completing the action in your mind*. At the end of the instruction, you will hear the reminding tone *du*, and you need to *perform* all of the action-object pairs of the instruction in order. Please be aware that you can use one hand or both hands to complete the task, but do not use both hands at the same time on the same object.’

After that, the experimenter demonstrated motor imagery, with the process being the same as that described in Condition 3. Given the counterbalancing of conditions across participants, the experimenter only needed to demonstrate motor imagery technique in the first motor imagery condition and skipped this process in the second motor imagery condition.

Then, the experimenter repeated the requirements of the task, and asked participants to practice. The practice process was similar to Condition 3, including two trials and the question about the clarity of motor imagery. The formal test then began.

Condition 5：Action observation-verbal recall condition

The experimenter first introduced the task requirement by saying, ‘I will play the video clips on my computer. Each instruction contains a series of action-object pairs and there is a silent pause after each one. During this time, you just need to *watch the video clips carefully*. At the end of the instruction, you will hear the reminding tone *du*, and you need to *speak out loud* all the action-object pairs of the instruction in order.’

The experimenter then demonstrated the procedure, and said, ‘I will demonstrate, and you must observe how I perform it.’ The experimenter played the sample video clip and watched it, and *repeated* the entire instruction after hearing the *du* sound. After the demonstration, the experimenter said, ‘This time, I will play the video clip, and you do the task as I just demonstrated. When you hear the *du* sound, say out loud all the action-object pairs in order. Let’s practice.’ The experimenter played two practice trials, and observed and corrected any rule-breaking behaviors until the participants could perform the task as required.

After the experimenter had ensured the participants could understand and perform the task correctly, the experimenter started the formal test.

Condition 6：Action observation-enacted recall condition

The experimenter first introduced the task requirement by saying, ‘I will play the video clips on my computer. Each instruction contains a series of action-object pairs and there is a silent pause after each one. During this time, you just need to *watch the video clips carefully*. At the end of the instruction, you will hear the reminding tone *du*, and you need to *perform* all the action-object pairs of the instruction in order. Please be aware that you can use one hand or both hands to complete the task, but do not use both hands at the same time on the same object.'

The experimenter then demonstrated the procedure, and said, 'I will demonstrate, and you must observe how I perform it.' The experimenter played the sample video clip and watched it, and *performed* the entire instruction sequentially after the *du* sound. After the demonstration, the experimenter said, 'This time, I will play the video clip, and you do the task as I just demonstrated. When you hear the *du* sound, perform all the action-object pairs in order. Let's practice.' The practice procedure was similar to that in Condition 5. After practice, the formal test began.

Condition 7：Self-enactment-verbal recall condition

The experimenter first introduced the task requirement by saying, 'I will play the spoken instructions on my computer. Each instruction contains a series of action-object pairs and there is a silent pause after each one. During this time, you need to you need to *perform the action-object pair and place the object back immediately*. At the end of the instruction, you will hear the reminding tone *du*, and you need to *speak out* loud all the action-object pairs of the instruction in order.'

The experimenter then demonstrated the procedure, and said, 'I will demonstrate, and you must observe how I perform it.' The experimenter played the sample recording, *performed each action-object pair and placed that object back immediately*, and *repeated* the entire instruction again after hearing the *du* sound. After the demonstration, the experimenter said, 'This time, I will play the instructions, and you do the task as I just demonstrated. After each action, you must perform it immediately and place the object back, and when you hear the *du* sound, say out loud all the action-object pairs in order. Let's practice.' The experimenter played two practice trials, and observed and corrected any rule-breaking behaviors until the participants could perform the task as required.

After the experimenter had ensured the participants could understand and perform the task correctly, the experimenter started the formal test.

Condition 8：Self-enactment-enacted recall condition

The experimenter first introduced the task requirement by saying, 'I will play the spoken instructions on my computer. Each instruction contains a series of action-object pairs and there is a silent pause after each one. During this time, you need to *perform the action-object pairs and place the object back*. At the end of the instruction, you will hear the reminding tone *du*, and you need to *perform* all the action-object pairs of the instruction in order. Please be aware that you can use one hand or both hands to complete the task, but do not use both hands at the same time on the same object.'

The experimenter then demonstrated the procedure, and said, 'I will demonstrate, and you must observe how I perform it.' The experimenter played the sample recording, *performed each action-object pair and placed the object back immediately*, and *performed* the entire instruction again after hearing the *du* sound. After the demonstration, the experimenter said, 'This time, I will play the instructions, and you do the task as I just demonstrated. After each action, you must perform it immediately and place the object back, and when you hear the *du* sound, perform all the action-object pairs in order. Let's practice.' The practice procedure was similar to that in Condition 7. After practice, the formal test began.

5. Instruction of formal test

The experimenter said, 'The official game is about to begin, please put your hands on your lap. Are you ready?' The experimenter said 'ready' at the beginning of each instruction trial, and then played the audio recordings or the video clips based on the conditions.