Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

(Un)predictable false memories: the role of sensorimotor language prediction in episodic memory

Mara Stockner

mara.stockner@uniroma1.it

Sapienza University of Rome: Universita degli Studi di Roma La Sapienza https://orcid.org/0000-0002-2692-3458

Gianmarco Convertino

Sapienza University of Rome https://orcid.org/0000-0002-8612-9854

Giuliana Mazzoni

Sapienza University of Rome https://orcid.org/0000-0002-6377-0139

Research Article

Keywords: predictability, false memory, sensorimotor simulation, plausibility, metacognition

Posted Date: September 5th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10920568/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: The authors declare no competing interests.

Abstract

Humans are able to predict which word they are going to encounter next during language comprehension and this has also been shown to be an underlying mechanism of false memory creation, i.e. predicted words are not suppressed and subsequently erroneously remembered, even if they were never encoded. The present study aimed at investigating, for the first time, the sensorimotor component of this language prediction process and its potential role in episodic (false) memory. More specifically, we validated a stimulus set of predictable (e.g., "stirring coffee with a *spoon*"), totally unpredictable (e.g., "stirring coffee with *pliers*") and motor-similar unpredictable ("e.g., "stirring coffee with a *thermometer*") action-object sentences. The latter refer to objects that are semantically unpredictable but rely on similar motor affordances compared to the predictable object. These stimuli were administered to 118 participants in three encoding conditions: Action Observation, Motor Imagery and Mental Rehearsal, and 24 hours later, tested in a recognition memory task, also assessing memory confidence. Our results show that predictable action-object sentences had a higher probability of eliciting false recognitions compared to totally unpredictable sentences whilst motor-similar unpredictable sentences did not significantly differ from totally unpredictable sentences. However, motor-similar unpredictable sentences were shown to have beneficial effects on correct remembering and accurate metacognitive monitoring, if encoded through action observation. Our findings provide a first insight into the motor and semantic components of linguistic predictability in episodic memory.

Introduction

"Donna used the shampoo to wash her…?". If asked to complete this sentence, most people would easily insert the word "hair" (Matsuki et al., 2011). This example shows the well-known fact that humans can not only *interpret* language extremely rapidly, by analysing each word as they encounter it and integrate it with prior context, but are even able to *predict* which word they are going to encounter next (e.g., Pickering & Gambi, 2018). In this regard, predictability has been defined as the "conditional probability that a specific word will occur given the preceding word, sentential fragment, or discourse" (Matsuki et al., 2011 p. 19). The underlying mechanism is the predictive function of cognition (e.g., Clark, 2013) that continuously uses incoming information to predict how the world might be in future states, also including a comparison between these predictions and what we subsequently encounter (e.g., Pickering & Gambi, 2018). In this line, not only visual perception (e.g., Den Ouden et al., 2012; Ianì et al., 2023), and memory (e.g., Vecchi & Gatti, 2020) have been argued to be of predictive nature but also language processing and comprehension (Ryskin & Nieuwland, 2023).

One hypothesised underlying mechanism of this predictive ability is the "predict-by-production" hypothesis (Pickering & Gambi, 2018). According to this approach, language comprehenders first covertly imitate what they have comprehended so far. Based on this, the intention of the utterance is derived, taking into account the linguistic context provided, as well as aspects of background knowledge and other extra-linguistic information that the comprehender assumes are shared with the speaker (Pickering & Gambi, 2018). This leads to the emergence of mental representations which can be

activated at any linguistic level (e.g., form, syntax, semantics). In this regard, lower-level activations, such as the prediction of the nature of the first letter of a word (e.g., Delong et al., 2005), usually follow higher level activations (e.g., semantic concepts) (Pickering & Gambi, 2018).

Interestingly, these predictions have been shown to be relatively robust, so that even when expectations for a specific word are disconfirmed, predictable words do not become suppressed immediately (Rich & Harris, 2021; Rommers & Federmeier, 2018). This phenomenon contributes to the creation of false memories of never-encoded but predicted words. For instance, Hubbard et al. (2019) presented sentences with expected or unexpected endings (e.g., I added my name to the list vs. I added my name to the basket) to their participants and tested their memory in a recognition test in which both lure endings (e.g., "list" is the lure in case of encoding of "I added my name to the basket") and completely new endings were presented. The authors found higher false alarm rates for lures compared to new items, suggesting that prediction during language processing can indeed facilitate false memory creation. This is also supported by Haeuser and Kray (2022) who adopted a similar experimental procedure. They found that the level of plausibility (i.e., highly implausible vs. plausible) of the prediction-disconfirming words during encoding did not differently impact false memory rates for predictable (unpresented) words. This suggests that lure-predictable words are not suppressed even when alternative plausible words are presented, and highlights the robust and fairly automatic nature of this predictive mechanism (see also Hubbard & Federmeier, 2024).

In the field of embodied cognition, on the other hand, some literature has also started focusing on the potential embodiment of this language-related predictive cognitive mechanism. As widely shown, the comprehension of language can automatically activate a mental simulation in a quasi-automatic manner (e.g., Stockner et al., 2025). Seminal neurophysiological studies suggest that sensorimotor predictive mechanisms do not only apply to action observation (e.g., Ianì et al., 2024; XXXX, *under review*) and motor imagery (e.g., Rieger et al., 2024) but also to the linguistic domain (e.g., Grisoni et al., 2017; Lam et al., 2017). In this context, underlying mechanisms of language comprehension and prediction might also include sensorimotor information (e.g., see the above-mentioned extra-linguistic information that contributes to intention understanding). In their EEG study, Grisoni et al. (2017), for instance, showed that participants who listened to action sentences that were either hand or mouth related, pre-activated body-specific parts of the motor cortex involved in the action (if a high-constraining context was provided). Moreover, Lesage et al. (2012), in their TMS study, found that when stimulating (i.e., temporarily interfering with) the right cerebellum, participants took longer to fixate the predictable object (e.g., cake) when hearing the associated predictive sentence (e.g., They boy will eat the…). This further highlights the involvement of motor areas in predictive language processing (see also e.g., Lam et al., 2017).

The present study

Whilst literature has shown that the activation of predicted words can lead to the creation of false memories and, at the same time, this activation can also be rooted in (or at least associated with) the

sensorimotor system, no research has yet investigated the role of sensorimotor processes in language prediction underlying the creation of memory distortions. This represents a crucial gap as it would allow memory researchers to distinguish the semantic and the sensorimotor components underlying the mechanisms of episodic memory (distortions). Thus, the present study aimed to fill this gap by disentangling the semantic and sensorimotor nature of the linguistic prediction which can lead to the creation of false memories. Previous studies used stimuli that are plausible and/or predictable both from a semantic and also from a motor point of view (e.g., Hubbard et al., 2019; 2024). This makes it difficult to identify the characteristics or nature of the underlying activated (semantic and/or sensorimotor) mechanisms that lead to the creation of false memories. Here, we adopt, for the first time, a procedure that includes not only totally predictable and unpredictable sentences (from both a semantic and motor point of view), but also semantically non-predictable sentences that are predictable from a motor point of view. Specifically, we decided to include sentences that contain motor-similar objects that share sensorimotor activations with the highly predictable objects (for instance, "You are writing with a *chopstick*" instead of "You are writing with a *pencil*"). We tested the creation of false memories linked to these stimuli in a recognition memory task, also measuring confidence ratings associated with recognition responses, in order to provide a meta-cognitive perspective on the relationship between (sensorimotor-based) predictability and false memories. We expected fully predictable sentences to lead to the highest false memory rate, and completely unpredictable sentences to the lowest false memory rate. Finally, and crucially, based on the assumption that the motor system is crucially involved in language prediction (e.g., Lam et al., 2017), we expected motor-similar unpredictable sentences to show higher false memory rates compared to completely unpredictable sentences. However, as semantic predictability is expected to show the strongest contribution to the predicted representations (Pickering & Gambi, 2018), lower false memory rates are expected in the motor-similar condition compared to the totally predictable (semantically and motor) sentences. Regarding true memories (i.e., hits), in line with Haeuser and Kray (2026; Experiment 1), we expected (totally) unpredictable sentences to be remembered better compared to predictable sentences. As it has been argued that only the largest violation of expectedness (and thus, predictability) may be able to enhance memory (e.g., Hauser & Kray, 2026; Kuperberg et al., 2020), we maintained an exploratory hypothesis regarding motor-similar unpredictable sentences, which, from a motor point of view, should not lead to a great violation of predictability.

Moreover, research has shown that the encoding modality in memory tasks can influence recognition performance. As research on linguistic predictability and false memories has so far only used amodal encoding modalities, we decided to administer the memory test via three encoding conditions: 1) an Action Observation (AO), 2) a Motor Imagery (MI) and 3) a Mental Rehearsal (MR) encoding condition. This makes it possible to provide a more nuanced understanding of the impact of predictability on false memories. In this regard we expected, in line with the existing literature (e.g., for MI: see Foley et al., 2009; Oliver et al., 2016; for AO see: Bouazzaouui et al., 2024; Senkfor et al., 2002), both mental simulation conditions (i.e., AO and MI) to lead to higher recognition accuracy (i.e., more hits and less false memories) compared to the control group (MR).

Finally, regarding the potential interaction between sensorimotor encoding and predictability levels of sentences, we hypothesized that higher false memory rates for predictable (or motor-similar predictable) sentences vs. unpredictable sentences are predominantly encountered in the control group (MR). This is based on the assumption that an activation of sensorimotor information (as in AO and MI) might facilitate and strengthen the creation of the memory trace that has been actually encoded (e.g., Ianì, 2019). This would lead to an increased suppression of the (linguistically and automatically) activated predictable (but not presented) words. In this regard, MI-based encoding might show a stronger benefit (i.e., less false memories compared to AO), as it involves more higher-level cognitive processes compared to AO (e.g., Stockner et al., 2025), which, compared to bottom-up processes involved in AO, might enhance both source monitoring (e.g., Foley et al., 2009) and activate the suppression of the predictable (not presented) words. Regarding false memory creation for motor-similar predictable sentences across the three encoding conditions, we adopted an exploratory approach due to the lack of relevant literature.

Methods

Experimental design

A mixed design was adopted with Action Type (Predictable / Motor-similar unpredictable / Totally unpredictable) as a within-subject factor and Group (Action Observation (AO) / Motor Imagery (MI) / Mental Rehearsal (MR)) as a between-subject factor. The sensorimotor encoding condition (i.e., Group) was manipulated between-subjects to avoid making participants aware of other encoding conditions, controlling for potential carry-over effects. The control condition was the mental rehearsal group (e.g., see also Marre et al., 2021), as verbally and silently rehearsing a linguistic stimulus is not expected to trigger any additional mental simulation compared to the mere sensorimotor activation, which is (almost) automatically triggered by language comprehension.

Participants

In order to determine the necessary sample size, we computed an a-priori power-analysis with G-Power (G*Power 3.1.9.7). We assumed a small to medium effect size and ran an approximate power analysis for mixed-design ANOVAs including an effect size of .15, an alpha of .05, a power of .80, with 3 groups (between-subject: AO / MI / MR) and 3 measurements (within-subject: predictable / motor-similar unpredictable / unpredictable). At least 102 participants were required to detect the expected effect, i.e. 34 participants per group (see Hubbard et al., 2019 for a similar effect size regarding linguistic prediction and false memories). 118 right-handed Italian young adults (M_{age} = 27.10 years, *SD* = 6.83; 74 female, 1 participant preferred not to indicate their gender) took part in the main experiment. All participants were Italian native speakers and did not report any psychiatric and/or neurological disorders, and they all had normal (or corrected-to-normal) eyesight. Participants were randomly assigned to the three experimental groups: AO: *N* = 40, MI: *N* = 39, MR: *N* = 39.

Materials.

Linguistic stimuli development.

For the aim of this study, initially, the authors created a set of 22 predictable linguistic stimuli, consisting of Italian action-object phrases (e.g., You are stirring the coffee with a teaspoon). From a grammar point of view, all sentences were created using the second-person singular form (in line with Brunye et al., 2009; Ditman et al., 2010) in order to induce a sense of engagement also in the control condition. All stimuli had the same syntactic structure, that is, subject (You) + action verb + object + *with* + target object. Neither fixed objects nor target objects nor action verbs were present more than once in the stimulus set. Moreover, for each predictable sentence, a motor-similar unpredictable sentence was created. In order to do so, the target object used to perform the action (e.g., the spoon) was substituted with an object that elicits similar motor affordances when grasping it (e.g., You are stirring the coffee with a *thermometer*). Objects were chosen based on similar grasping and physical properties (i.e., size and shape of the objects) compared to the predictable target objects. Finally, a completely unpredictable sentence was created by substituting the target object with a completely unrelated object (i.e., both semantically and motorwise) (e.g., You are stirring the coffee with *pliers*). For all three categories, an interrater agreement between the first two authors and an external rater, naive to the aim of the study, was obtained before proceeding with the validation study (see Supplementary Materials for all 66 sentences).

Linguistic stimuli validation.

Nineteen university students, different from those who took part in the main experiment, participated in the validation study (M_{age} = 25.20 years, *SD* = 5.50; 15 female). All participants were native Italian speakers and did not report any neurological or psychiatric disorders. The validation study consisted of two parts: 1) the completion of predictable action sentences and 2) the rating of their plausibility. After signing the informed consent and providing socio-demographic information, the participants were presented, one at a time and in completely randomized order, with the predictable sentences but without the final target object (e.g., “You are stirring the coffee with”), and were asked to complete the sentence with the most suitable object that came to their mind (i.e., the so-called Cloze procedure, e.g., Staub et al., 2015; Taylor, 1953). The instructions specified that there were no right or wrong answers and they should reply as spontaneously as they could. After performing the cloze-procedure for all 22 predictable stimuli, participants were presented with the second part of the validation study which aimed at additionally validating the motor-similar unpredictable and totally unpredictable stimuli. More specifically, in this second task, participants were asked to rate, on a Likert scale ranging from 1 (= not at all) to 7 (= very much) how *plausible* they considered each of the 66 action-object phrases to be. This approach was adopted based on existing studies and the intertwined nature of the concepts of plausibility and predictability (e.g., Matsuki et al., 2011). More specifically, we considered both *semantic* plausibility (how plausible is it *that an individual uses* a certain object to perform a said action? e.g., Lam et al., 2017) and, tailored to our motor-similar unpredictable stimuli, also *motor* plausibility of the stimuli

(how plausible is it that *an individual is able to perform* a said action with a certain target object?). Thus, participants evaluated each stimulus twice alongside two additional attention check questions “Please select the number X”. The whole procedure was carried out on Qualtrics.

Definition of predictable sentences (i.e., Cloze Procedure). Written text answers were grouped and frequencies for each object were computed. Subsequently, binomial tests were performed assessing whether the most frequent object response was significantly more frequent than chance-level ($p < .05$), indicating predictability of said object. Fifteen stimuli passed this threshold, see Supplementary Materials for the final stimuli set.

Plausibility of action-object sentences. In order to evaluate plausibility ratings between the three different types of sentences (predictable vs. motor-similar unpredictable vs. totally unpredictable), Likert-scale responses were treated as ordinal variables, and cumulative linear mixed models were computed by applying the *ordinal* R package (Christensen, 2018). The following model was applied: Rating ~ Action Type * Plausibility Type + (1|Participant) + (1|Item), including both participants as well as stimuli as random effects. Plausibility type referred to semantic and motor plausibility. Only the 15 selected stimuli (and their associated motor-similar unpredictable and totally unpredictable sentences) were included in the analysis and one participant was excluded as they failed both attention checks. The model showed a significant main effect of Action Type ($\chi^2 = 218.91$, $df = 2$, $p < .0001$) and Plausibility Type ($\chi^2 = 19.96$, $df = 1$, $p < .0001$). More precisely, predictable sentences were overall rated as more plausible compared to motor-similar unpredictable sentences ($b = 9.60$, $SE = .85$, $z_{ratio} = 12.85$, $p < .0001$) and totally unpredictable sentences ($b = 11.52$, $SE = .78$, $z_{ratio} = 154.77$, $p < .0001$), and motor-similar unpredictable were evaluated as more plausible compared to totally unpredictable sentences ($b = 1.92$ $SE = .59$, $z_{ratio} = 3.28$, $p = .003$). Moreover, motor plausibility was overall higher compared to semantic plausibility. Finally, the significant interaction ($\chi^2 = 12.06$, $df = 2$, $p = .002$) revealed that there was no difference in plausibility ratings (semantic vs. motor) for predictable sentences ($b = .74$, $SE = .54$, $z_{ratio} = 1.37$, $p = .172$) but higher motor plausibility (compared to semantic plausibility) was encountered both for motor-similar unpredictable ($b = - .79$, $SE = .21$, $z_{ratio} = -3.34$, $p = .0008$) and totally unpredictable ($b = -1.40$, $SE = .31$, $z_{ratio} = -4.50$, $p < .0001$) sentences. Taken together, these results confirm that the motor-similar unpredictable stimuli differed, as expected, from predictable and totally unpredictable stimuli.

Finally, stimuli were divided into three encoding protocols which consisted, respectively, of 15 stimuli (five predictable, five motor-similar unpredictable and five totally unpredictable) and participants were randomly assigned to one of the encoding protocols.

Visual stimuli

All videos used for the Action Observation group lasted 4.5 seconds and depicted the validated action-object phrases. All actions were filmed in first-person and showed a young male actor performing the

action (see Fig. 1). More specifically, at the start of each video, both hands of the actor were placed still on a white table and the objects were visible. The actor then performed the action with his right hand.

Procedure of the main experiment

After signing the informed consent and providing socio-demographic information, participants were presented with the encoding instructions (see Supplementary Materials for the specific instructions in all three groups). In all three groups, each encoding trial started with the oral presentation of the action-object sentence, see also Fig. 2A. Subsequently, participants heard a "beep" sound which indicated the start of the second part of the trial (different for each group – see below) and after 4.5 seconds, a second "beep" sound indicated the end of this second part (i.e., imagery, action observation or mental rehearsal), followed by a final fixation cross (1000 ms). More specifically, the second part of the trial consisted, in the AO group, of the presentation of the video depicting the action heard in the first part of the trial whilst the MI group was instructed to imagine performing the action in their mind in first person perspective and with the instruction to feel the associated bodily and motor sensations. The mental rehearsal group was instructed to mentally repeat in their mind the sentence they just heard. After 24 hours (*M* = 24.20 hours; *SD* = 0.91), participants completed the recognition task, see Fig. 2B. More precisely, participants were presented, one at a time, with all 45 linguistic stimuli (15 old and 30 new), that were shown written on the screen, and were asked to indicate orally whether they remembered each sentence from the first part of the task (by saying out loud "yes") or not (by saying out loud "no"). They had 5 seconds to give their answer for each trial, after which they were prompted to indicate their confidence in their response on a Likert scale ranging from 1 = not at all, to 7 = very confident. Sentences were presented in a completely randomized order. Finally, participants were debriefed and thanked for their participation. The experiment was built on PsychoPy (Version 2024.2.5; Peirce et al., 2019) and both parts of the procedure were carried out by participants on the platform Pavlovia on their personal computers.

Figure 2.

A). Example encoding trial of the Action Observation (AO) group, B). Example trial of the recognition task (same for all groups)

Data analysis

Recognition responses were transcribed and each "yes" answer was coded as 1 (= old) and each "no" answer was coded as 0 (= new). Trials with missing answers were removed (5.93% of all trials) and outliers (trials) based on Response Times were identified by calculating the first (Q1) and third quartile (Q3). Q1—1.5*IQR (interquartile range) presented the lower limit and Q3 + 1.5*IQR the upper limit (e.g., Stockner et al., 2025). Trials below and above these limits were removed (0.57% of all trials).

A mixed effects model approach was adopted, and a Generalized Linear Mixed Model (GLMM) was run on recognition answers on the old items and the new items datasets, respectively (see e.g., Gatti et al., 2022). The main model included Group (AO / MI / MR) and Action Type (Predictable / Motor-similar Unpredictable / Totally unpredictable) as predictors, and both participants and items were included as random effects. Thus, the following model structure was adopted: Response ~ Group * Action Type + (1 | Participant) + (1|Stimulus), implemented with the glmer() function from the *lme4* package (Bates et al., 2015) and tested with the Anova() function. Response times (RT) from the verbally provided recognition responses were computed by measuring the time from stimulus onset to the onset of the participants' verbal response. In order to do so, the Matlab script provided by Limata et al. (2023) was applied, which calculates the medium noise of the entire 5 seconds during which participants could respond, and sets a threshold for each trial. At the specific moment the noise exceeded the threshold, the response time was calculated. Finally, RTs were log-transformed in order to account for skewness (e.g., Gatti et al., 2023). The same model structure was applied, and linear mixed models (LMM) were run on both old and new items, respectively.

Finally, the same model structure was also applied on confidence responses which were treated as ordinal variables. Thus, cumulative linear mixed models (clmm) were run using the ordinal package (Christensen, 2018). Post hoc comparisons were carried out, on all three types of mixed models, with the emmeans() function (Lenth et al., 2019).

Results

Recognition responses

The GLMM on recognition responses of the new items (i.e., probability to falsely recognize a new item) revealed a main effect of both Group (χ^2 = 17.21, *df* = 2, *p* = .0002) and Action Type (χ^2 = 29.14, *df* = 2, *p* < .0001). Post hoc comparisons on the Group effect show that participants in the MR group had significantly more false alarms compared to the AO group (*b* = -1.04, *SE* = .25, z_{ratio} = -4.14, *p* = .0001) (see Fig. 3A) whilst the other group differences were not significant (*p*s > .092).

Table 1
False memory rates (SD)

		Action Type			**Overall *M***
		Predictable sentences	**Motor-similar unpredictable sentences**	**Unpredictable sentences**	
	AO	.16 (*.19*)	.13 (*.17*)	.09 (*.14*)	.13 (*.17*)
Group	MI	.24 (*.21*)	.17 (*.16*)	.16 (*.22*)	.19 (*.20*)
	MR	.33 (*.25*)	.26 (*.20*)	.21 (*.18*)	.27 (*.21*)
Overall *M*		.24 (*.23*)	.19 (*.18*)	.15 (*.18*)	

Notes. Group: AO = Action Observation; MI = Motor Imagery; MR = Mental Rehearsal

Regarding Action Type, predictable actions elicited higher false alarms compared to totally unpredictable sentences (b = .64, SE = .12, z_{ratio} = 5.3220 p < .0001), and compared to motor-similar unpredictable sentences (b = .36, SE = .12, z_{ratio} = 3.08, p = .006). The comparison between motor-similar unpredictable and totally unpredictable sentences did not reach significance (b = .28, SE = .13, z_{ratio} = 2.19, p = .073). The interaction between Group and Action Type was not significant (χ^2 = 1.81, df = 4, p = .771).

Moreover, the GLMM on the old items showed a significant main effect of Action Type (χ^2 = 8.28, df = 2, p = .016), see Table 2 for accuracy means (i.e., hit rates). More precisely, post hoc comparisons revealed higher probability of remembering for both motor-similar unpredictable (b = − .41, SE = .15, z_{ratio} = -2.78, p = .015) and totally unpredictable (b = − .36, SE = .15, z_{ratio} = -2.42, p = .041) sentences compared to predictable sentences. The comparison between motor-similar unpredictable and totally unpredictable sentences was not statistically significant (p = .93). Finally, a significant interaction between Action Type and Group was encountered (χ^2 = 10.39, df = 4, p = .034). More specifically, motor-similar predictable sentences were shown to be recognized more accurately compared to predictable sentences in the AO group (b = -1.02, SE = .26, z_{ratio} = -3.95, p = .0002), whilst no other post hoc comparisons were significant (all ps > .11).

Figure 3.

False memory rates based on A) Group and B) Action Type

Notes. Group: AO = Action Observation; MI = Motor Imagery; MR = Mental Rehearsal. Action Type: 1 = Predictable; 2 = Motor-similar unpredictable; 3 = Totally unpredictable sentences

Table 2
Hit rates (SD)

		Action Type			**Overall *M***
		Predictable sentences	**Motor-similar unpredictable sentences**	**Totally unpredictable sentences**	
Group	AO	.65 (*.28*)	.82 (*.22*)	.73 (*.25*)	.73 (*.26*)
	MI	.69 (*.22*)	.71 (*.29*)	.77 (*.28*)	.72 (*.26*)
	MR	.66 (*.29*)	.68 (*.26*)	.69 (*.27*)	.68 (*.27*)
Overall *M*		.66 (*.26*)	.73 (*.26*)	.73 (*.27*)	

Notes. Group: AO = Action Observation; MI = Motor Imagery; MR = Mental Rehearsal

Confidence

Regarding confidence, the first CLMM on new items revealed significant group differences (χ^2 = 18.69, *df* = 2, *p* < .0001) between AO and MI (*b* = .85, *SE* = .33, *z* = 2.61, *p* = .025) and AO and MR (*b* = 1.40, *SE* = .33, *z* = 4.30, *p* = .0001), indicating higher confidence for participants in the AO condition compared to both other conditions (Fig. 4A). No other significant effects or interaction effects were encountered (all *p*s > .059).

Notes. Group: AO = Action Observation; MI = Motor Imagery; MR = Mental Rehearsal. Action Type: 1 = Predictable; 2 = Motor-similar unpredictable; 3 = Totally unpredictable sentences

The second CLMM on confidence for old items, revealed a significant group difference (χ^2 = 24.10, *df* = 2, *p* < .0001), with both AO (*b* = 1.21, *SE* = .25, z_{ratio} = 4.79, *p* < .0001) and MI (*b* = .92, *SE* = .25, z_{ratio} = 3.67, *p* = .0007) showing higher confidence values compared to the MR group. Moreover, the significant effect of Action Type (χ^2 = 31.47, *df* = 2, *p* < .0001) showed that, overall, predictable sentences were associated with lower confidence compared to both motor-similar unpredictable (*b* = − .76, *SE* = .14, z_{ratio}= -5.42, *p* < .0001) and totally unpredictable (*b* = − .59, *SE* = .14, z_{ratio} = -4.26, *p* = .0001) sentences .

Finally, a significant interaction between Group and Action Type was found (χ^2 = 11.74, *df* = 4, *p* = .019). More specifically, predictable sentences were associated with lower confidence compared to motor-similar unpredictable sentences in all three groups (AO: *b* = − .84, *SE* = .24, z_{ratio} = -3.53, *p* = .001; MI: *b* = − .90, *SE* = .22, z_{ratio} = -4.02, *p* = .0002; MR: *b* = − .54, *SE* = .20, z_{ratio} = -2.68, *p* = .020). Interestingly, however, when compared to totally unpredictable sentences, predictable sentences elicited lower confidence only in the MI group (*b* = -1.09, *SE* = .23, z_{ratio} = -4.83, *p* < 0001). Finally, on the other hand, motor-similar unpredictable sentences were associated with lower confidence compared to totally

unpredictable sentences only in the AO group (b = .62, SE = .24, z_{ratio} = 2.61, p = .024). No other comparisons were significant (all ps > .060).

Response Times

Finally, regarding RTs, the first LMM on new items showed a significant effect of Action Type (χ^2 = 7.30, df = 2, p = .026), with responses to totally unpredictable sentences significantly faster (M = 2.68 s; SD = .44), compared to predictable sentences (M = 2.82 s, SD = .83), (b = .05, SE = .02, z_{ratio} = 2.67, p = .021). No other post hoc comparisons or predictor effects were significant (all ps > .21). The LMM on RTs regarding old items did not yield any significant effect (all ps > .424).

Discussion

The present study investigated, for the first time, the role of sensorimotor language prediction in episodic (false) memory. By manipulating both sentence predictability (i.e., predictable, motor-similar unpredictable and totally unpredictable sentences) and sensorimotor encoding modalities (i.e., action observation, motor imagery and mental rehearsal), we aimed at disentangling the potential contribution of sensorimotor and semantic mechanisms of language prediction in memory distortions.

False memories

First, our findings showed that, overall, predictable (vs. totally unpredictable) action-object sentences had a higher probability of eliciting false recognitions. That is, participants remembered the predictable target object (e.g., stirring coffee with a teaspoon) even when the object was not encoded. This is in line with existing literature showing that automatically activated language prediction is not suppressed (e.g., Rich & Harris, 2021) and, thus, false memories of the predicted words can be formed, even if they were never overtly encoded (e.g., Haeuser & Kray, 2022; 2026; Hubbard et al., 2019). This increased false memory tendency was also reflected in slower Response Times, possibly indicating a higher level of uncertainty in recognition memory of predictable actions (e.g., De Boeck & Jeon, 2019). However, it was not reflected by confidence levels, in line with literature showing that memory distortions and accuracy are not always accompanied by adequate metacognitive monitoring (e.g., Convertino et al., 2024; Hanczawoski et al., 2021). On the other hand, this is not in line with Hauser and Kray (2026) who found that predictable false memories were also associated with high confidence levels, suggesting a rather complex accuracy-confidence relationship in predictability and memory processes.

Crucially, we remind that the effect of action predictability could be either purely semantic in nature, or also occur with the contribution of the sensorimotor component of the action. Thus, motor-similar and yet unpredictable sentences (e.g., you are stirring the coffee with a thermometer) were introduced in the study in an attempt to disentangle these two components. Importantly, and central to this novel aim, in our results, motor-similar unpredictable action sentences elicited significantly less false memories compared to predictable sentences. This confirms that the predictability of the target object of an action (e.g., the object with which we typically stir coffee) plays a major role in determining what people

remember. In the sentence-predictable condition participants tended to complete the sentence with the predictable target object (e.g., stirring coffee with a teaspoon) also when that object had not been encoded. When the action was performed with an object that was not typical (e.g., stirring coffee with a thermometer), this effect occurred significantly less even if the movement of the hand (grasping) was similar to that used for the target object. This is in line with Haeuser and Kray (2022) who showed that independently of the level of (im)plausibility of the alternative stimuli, the (highly) predictable target word is not suppressed.

Finally, comparing false recognitions in the two action-unpredictable conditions, revealed that there were only tendentially more false memories for the motor-similar unpredictable sentences (p = .07) compared to totally unpredictable sentences. Thus, our preliminary findings are only partially in line with our initial hypotheses of a possible important sensorimotor contribution to false recognition of predictability. While we expected motor-similar but unpredictable sentences to elicit less false memories compared to totally predictable sentences (as the semantic component is expected to contribute majorly to predictability effects, e.g., Pickering & Gambi, 2018), we also expected motor-similar unpredictable sentences to elicit more false memories compared to totally unpredictable sentences (due to the hypothesized sensorimotor activations involved in language prediction; e.g., Lam et al., 2017), but this difference was not significant. Whilst our findings can not completely rule out the presence of a sensorimotor component in linguistic prediction associated with false memory creation, several reasons may account for our not-significant result. On one hand, it is possible that our experimental stimuli were not sufficiently specific regarding the similarity of motor-activations (e.g., grasping of “spoon” vs. “thermometer”) and further validation of the actual (similar) motor involvement/affordances of the objects (compared to the predictable objects) would be needed to verify this. In this regard, it would also be crucial not only to consider grasping similarities based on the similarities of the objects’ physical properties but also the associated functions and use of the objects. Our decision to not include this as a criterion was based on the fact that often similar functions or similar use of objects are also associated with increasing semantic plausibility (i.e., while stirring coffee with a thermometer is not semantically plausible, stirring coffee with a fork, on the other hand, may be more plausible also from a semantic point of view). Future research should however, try to further disentangle this.

On the other hand, it is also possible to hypothesize, that despite the potential presence of a sensorimotor component, the semantic aspects of the language-based predictions may be of much more important influence, overriding the sensorimotor contribution which may, in comparison, have less relevance, potentially leading to a very small effect which we were not able to (significantly) detect. Indeed, considering the results of our validation study, it is worth noting that not only motor-similar unpredictable but also totally unpredictable sentences yielded higher motor (compared to semantic) plausibility ratings (even if motor-similar unpredictable had higher plausibility ratings compared to totally unpredictable). Thus, future research is needed to replicate and further clarify our preliminary findings. One way of partially addressing these open questions would be by disentangling not only the motor from the semantic component (as we aimed at doing in the current study) but also the semantic from the motor component. Whilst, in the present study, we aimed at eliminating semantic (but not motor)

predictability in our motor-similar unpredictable action stimuli, it would be crucial to apply a complementary paradigm in order to investigate a potential double dissociation: that is, removing motor (but not semantic) predictability by applying, for instance, abstract concepts that are not “graspable” (e.g., Borghi et al., 2017), that could be “You are stirring coffee with love”. In this regard, it is worth repeating, however, that even for totally unpredictable sentences we encountered higher motor (vs. semantic) plausibility. Whilst an extensive discussion on this finding exceeds the aim of the present paper, these results do raise the question of the different underlying cognitive mechanisms of motor vs. semantic plausibility: whilst semantic plausibility may rely exclusively on previous knowledge and semantic nodes in order to make logical sense of the world (e.g., Connell & Keane, 2006), motor plausibility might elicit more bottom-up cognitive processes (e.g., if you are able to think about or imagine an action, this might elicit motor but not semantic plausibility). Future research should further investigate these different implications, also considering the well-established role of plausibility in our memory (e.g., Mazzoni, 2007; Scoboria et al., 2004).

Interestingly, we did not encounter any interaction effects between action predictability and the different (sensorimotor) encoding groups (i.e., action observation vs. motor imagery, vs. mental rehearsal) on false recognitions. This suggests that the linguistic predictability effect on false memory creation is a solid phenomenon which does not only occur in verbal encoding modalities (as reading in Haeuser & Kray, 2022; Hubbard et al., 2019) but also via sensorimotor encoding, which did not protect from predictable false memories in our study. However, it is worth noting that, independently from predictability, our findings suggest lower false alarms in the action observation (compared to the mental rehearsal) group, also associated with higher confidence (and lower confidence in both the mental rehearsal and motor imagery groups). This is in line with the literature suggesting that sensorimotor encoding can protect from false memories (e.g., Robin & Garnier, 2022). Moreover, it is in line with Stockner et al. (2026) who found a beneficial effect of action observation compared to other sensorimotor modalities. In line with the latter study, our findings also show that imagery did not significantly reduce false memories. This suggests that the protective role, usually attributed to mental imagery (e.g., Foley et al., 2009), might be mainly encountered in studies that do not include a comparison with other sensorimotor conditions (e.g., Foley, 2012; Marre et al., 2021; 2024).

True memories

Regarding the memory of actually encoded action-object sentences, we encountered better memory for both motor-similar unpredictable and totally unpredictable compared to predictable sentences. This can theoretically be explained by the violation of expectedness (e.g., Kuperberg et al., 2020) (i.e., predictabilityt) that can enhance memory and is in line with literature showing that implausible events are better remembered, also related to a “surprise” effect (e.g., de Bruine et al., 2025). This finding was reflected in confidence ratings which were higher for unpredictable sentences (both motor unpredictable and totally unpredictable) compared to predictable sentences. This suggests metacognitive awareness of one’s memory for actually encoded (un)predictable action sentences. Moreover, interestingly, this beneficial effect of unpredictable sentences on memory was especially found in the action observation

group, in which there was better memory for motor-similar unpredictable (compared to totally predictable) sentences. For correct recognition of encoded actions, this result seems to suggest that the beneficial “surprise” effect on memory works particularly well when stimuli are encoded through action observation. A potential explanation could be the integration between the “surprise” effect, leading implausible events to be remembered better (e.g., Ranganath & Rainer, 2003) and, on the other hand, the creation of a rich sensorimotor memory trace. Whilst motor resonance in totally unpredictable action-object sentences may have been limited due to the implausible nature of the actions (e.g., Brass et al., 2007), it is possible to hypothesize that the sensorimotor plausibility of the motor-predictable sentences may have elicited motor-resonance and thus, the creation of a richer memory trace and better memory for those stimuli.

This finding was also supported by our confidence results. Whilst the significant interaction between action type and group revealed lower confidence for predictable (vs. motor-predictable) sentences in all encoding groups, only in the action observation group confidence of previously encoded totally unpredictable sentences was lower compared to motor-similar unpredictable sentences. This again confirms the above hypothesized sensitivity of action observation regarding motor-similar but unpredictable content, suggesting adequate metacognitive monitoring of this type of content (see also Haeuser & Kray, 2026).

Finally, only in the motor imagery group higher confidence for totally unpredictable (compared to predictable) sentences was encountered. As totally unpredictable sentences were remembered better across all encoding groups, this finding suggests that imagery might have a particular impact on metacognition regarding very implausible stimuli or events. Indeed, on one hand, confidence can reflect the level of strength of a memory, or at least, the subjective sense of a strong memory elicited by this type of encoding (e.g., Roediger & DeSoto, 2014). On the other hand, imagining an action performed with an unpredictable object may have required greater cognitive effort, which can be reflected in memory monitoring (e.g., Mazzoni & Nelson, 1995). Thus, the imagination of such implausible events during encoding might have led to the creation of memory traces that were accurately (metacognitively) monitored. This may have potentially been facilitated by the involvement of higher-order top-down processes involved in imagery (e.g., Stockner et al., 2025).

To sum up the results, this is the first study to investigate the role of sensorimotor language prediction in episodic memory. We introduced motor-similar yet unpredictable action sentences in an attempt to disentangle semantic from motor components in predictability. While we need to take into account the preliminary nature of our findings, our results confirm existing evidence that linguistic predictability can lead to false memories, independently of the type of (sensorimotor or verbal) encoding. Whilst motor-similar but unpredictable sentences did not elicit more false memories compared to totally unpredictable sentences, suggesting a predominant role of semantic components in linguistic predictability, we encountered beneficial effects on episodic memory of actually encoded (motor-similar and totally) unpredictable contents, which were also associated with adequate metacognitive monitoring. Overall, the present study provides an important first contribution to the role of sensorimotor

mechanisms in linguistic predictability impacting episodic (false) memory and raises the question on how to fully disentangle motor and semantic predictability.

Declarations

Ethics approval.

This study was performed in line with the principles of the Declaration of Helsinki. Approval was granted by the departmental Ethics Committee of Sapienza University of Rome.

Consent to participate

and to publish. Written informed consent was obtained by all participants regarding both their participation and the publication of their anonymized data.

Competing interests.

The authors declare that they have no competing interests.

Funding.

This research was funded by PRIN 2022 nr. 20225ECXPP – “When the body tricks the mind: movement, action and touch as determinants of memory illusions”.

Availability of data and materials.

All data, statistical analysis scripts and materials are available at: https://osf.io/n7fxe/overview?view_only=a04e985b2ab749aa9e29247650f4d75c

References

1. Bates D, Maechler M, Bolker B, Walker S, Christensen RHB, Singmann H, Bolker MB (2015) Package ‘lme4’. *convergence*, *12*(1), 2
2. Borghi AM, Binkofski F, Castelfranchi C, Cimatti F, Scorolli C, Tummolini L (2017) The challenge of abstract concepts. Psychol Bull 143(3):263
3. Bouazzaoui B, Angel L, Moutoussamy I, Villatte J, Isingrini M, Bidet-Ildei C (2024) Kinematic observation reduces effect of aging on episodic memory performance. Acta Psychol 243:104130
4. Brunyé TT, Ditman T, Mahoney CR, Augustyn JS, Taylor HA (2009) When you and I share perspectives: Pronouns modulate perspective taking during narrative comprehension. Psychol Sci 20(1):27–32

5. Christensen RHB (2018) Cumulative link models for ordinal regression with the R package ordinal. Submitted J Stat Softw 35:1–46
6. Clark A (2013) Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behav Brain Sci 36(3):181–204
7. Connell L, Keane MT (2006) A model of plausibility. Cogn Sci 30(1):95–120
8. De Boeck P, Jeon M (2019) An overview of models for response times and processes in cognitive tests. Front Psychol 10:102
9. de Bruïne A, Vel Tromp M, Koornneef A, Brod G, Jolles D (2025) The interactive effects of surprise and plausibility on memory. J Experimental Psychology: Learn Memory Cognition 51(6):954
10. DeLong KA, Urbach TP, Kutas M (2005) Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nat Neurosci 8(8):1117–1121
11. Den Ouden HE, Kok P, De Lange FP (2012) How prediction errors shape perception, attention, and motivation. Front Psychol 3:548
12. Ditman T, Brunyé TT, Mahoney CR, Taylor HA (2010) Simulating an enactment effect: Pronouns guide action simulation during narrative comprehension. Cognition 115(1):172–178
13. Foley MA, Hughes K, Librot H, Paysnick AMY (2009) Imagery encoding effects on memory in the DRM paradigm: A test of competing predictions. Appl Cogn Psychology: Official J Soc Appl Res Memory Cognition 23(6):828–848
14. Gatti D, Marelli M, Rinaldi L (2023) Out-of-vocabulary but not meaningless: Evidence for semantic-priming effects in pseudoword processing. J Exp Psychol Gen 152(3):851
15. Gatti D, Rinaldi L, Marelli M, Mazzoni G, Vecchi T (2022) Decomposing the semantic processes underpinning veridical and false memories. J Exp Psychol Gen 151(2):363
16. Grisoni L, Miller TM, Pulvermüller F (2017) Neural correlates of semantic prediction and resolution in sentence processing. J Neurosci 37(18):4848–4858
17. Haeuser KI, Kray J (2026) Do you remember? Effects of lexical predictability on false and true recognition memory. *Journal of experimental psychology. Learning, memory, and cognition*, 10.1037/xlm0001592. Advance online publication. https://doi.org/10.1037/xlm0001592
18. Haeuser K, Kray J (2022) Uninvited and unwanted: False memories for words predicted but not seen. In *Proceedings of the Annual Meeting of the Cognitive Science Society* (Vol. 44, No. 44)
19. Hanczakowski M, Butowska E, Beaman CP, Jones DM, Zawadzka K (2021) The dissociations of confidence from accuracy in forced-choice recognition judgments. J Mem Lang 117:104189
20. Hubbard RJ, Federmeier KD (2024) The impact of linguistic prediction violations on downstream recognition memory and sentence recall. J Cogn Neurosci 36(1):1–23
21. Hubbard RJ, Rommers J, Jacobs CL, Federmeier KD (2019) Downstream behavioral and electrophysiological consequences of word prediction on recognition memory. Front Hum Neurosci 13:291

22. Ianì F (2019) Embodied memories: Reviewing the role of the body in memory processes. Psychon Bull Rev 26(6):1747–1766
23. Ianì F, Limata T, Bucciarelli M, Mazzoni G (2023) The implicit effect of action mental simulation on action evaluation. Q J Experimental Psychol 76(2):257–270
24. Ianì F, Limata T, Ras IN, Bucciarelli M (2024) Forward effects from action observation: the role of attentional focus. Psychol Res 88(3):773–785
25. Kuperberg GR, Brothers T, Wlotko EW (2020) A tale of two positivities and the N400: Distinct neural signatures are evoked by confirmed and violated predictions at different levels of representation. J Cogn Neurosci 32(1):12–35
26. Lam KJ, Bastiaansen MC, Dijkstra T, Rueschemeyer SA (2017) Making sense: Motor activation and action plausibility during sentence processing. Lang Cognition Neurosci 32(5):590–600
27. Lenth R, Singmann H, Love J, Buerkner P, Herve M (2019) Package 'emmeans'. R package version 1(32):1199
28. Lesage E, Morgan BE, Olson AC, Meyer AS, Miall RC (2012) Cerebellar rTMS disrupts predictive language processing. Curr Biol 22(18):R794–R795
29. Limata T, Bucciarelli M, Schmidt S, Tinti C, Ras IN, Ianì F (2023) Action and posture influence the retrieval of memory for objects. Memory 31(5):652–664
30. Marre Q, Huet N, Labeye E (2021) Embodied mental imagery improves memory. Q J Experimental Psychol 74(8):1396–1405
31. Marre Q, Huet N, Labeye E (2024) Does context matter for memory? Testing the effectiveness of learning by imagining situated interactions with objects. Memory 32(4):502–514
32. Matsuki K, Chow T, Hare M, Elman JL, Scheepers C, McRae K (2011) Event-based plausibility immediately influences on-line language comprehension. J Experimental Psychology: Learn Memory Cognition 37(4):913
33. Mazzoni G (2007) Did you witness demonic possession? A response time analysis of the relationship between event plausibility and autobiographical beliefs. Psychon Bull Rev 14(2):277–281
34. Mazzoni G, Nelson TO (1995) Judgments of learning are affected by the kind of encoding in ways that cannot be attributed to the level of recall. J Experimental Psychology: Learn Memory Cognition 21(5):1263–1274. https://doi.org/10.1037/0278-7393.21.5.1263
35. Oliver MC, Bays RB, Zabrucky KM (2016) False memories and the DRM paradigm: Effects of imagery, list, and test type. J Gen Psychol 143(1):33–48
36. Peirce J, Gray JR, Simpson S, MacAskill M, Höchenberger R, Sogo H, Lindeløv JK (2019) PsychoPy2: Experiments in behavior made easy. Behav Res Methods 51(1):195–203
37. Pickering MJ, Gambi C (2018) Predicting while comprehending language: A theory and review. Psychol Bull 144(10):1002

38. Ranganath C, Rainer G (2003) Neural mechanisms for detecting and remembering novel events. Nat Rev Neurosci 4(3):193–202
39. Rich S, Harris J (2021) Unexpected guests: When disconfirmed predictions linger. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 43, No. 43)
40. Rieger M, Boe SG, Ingram TG, Bart VK, Dahm SF (2024) A theoretical perspective on action consequences in action imagery: Internal prediction as an essential mechanism to detect errors. Psychol Res 88(6):1849–1858
41. Robin F, Garnier C (2022) The impact of enactment and imagery encoding on false memory. Psychol Appl Developments VIII, 285–293
42. Roediger III, H. L., DeSoto KA (2014) Understanding the relation between confidence and accuracy in reports from memory. Remembering. Psychology, pp 371–391
43. Rommers J, Federmeier KD (2018) Predictability's aftermath: Downstream consequences of word predictability as revealed by repetition effects. Cortex 101:16–30
44. Ryskin R, Nieuwland MS (2023) Prediction during language comprehension: what is next? Trends Cogn Sci 27(11):1032–1052
45. Scoboria A, Mazzoni G, Kirsch I, Relyea M (2004) Plausibility and belief in autobiographical memory. Appl Cogn Psychology: Official J Soc Appl Res Memory Cognition 18(7):791–807
46. Senkfor AJ, Petten CV, Kutas M (2002) Episodic action memory for real objects: An ERP investigation with perform, watch, and imagine action encoding tasks versus a non-action encoding task. J Cogn Neurosci 14(3):402–419
47. Staub A, Grant M, Astheimer L, Cohen A (2015) The influence of cloze probability and item constraint on cloze task response time. J Mem Lang 82:1–17
48. Stockner M, Mazzoni G, Ianì F (2025) Are fluent letter dyads really fluent? An update on objective and subjective motor fluency in an Italian student population. Cogn Research: Principles Implications 10(1):42
49. Stockner M, Mazzoni G, Enrici I, Ianì F (2026) Embodied simulations improve memory, but which is the most effective? The crucial role of action observation and its dynamic interplay with motor imagery ability. Front Psychol 17:1889010
50. Stockner M, Mazzoni G, Perrin D, Ianì F (2025) The same label, different processes: What lies behind the term mental simulation in the embodied cognition literature? Brain Cogn 188:106335
51. Taylor WL (1953) Cloze procedure: A new tool for measuring readability. Journalism Q 30(4):415–433
52. Vecchi T, Gatti D (2020) Memory as prediction: From looking back to looking forward. MIT Press

Figures

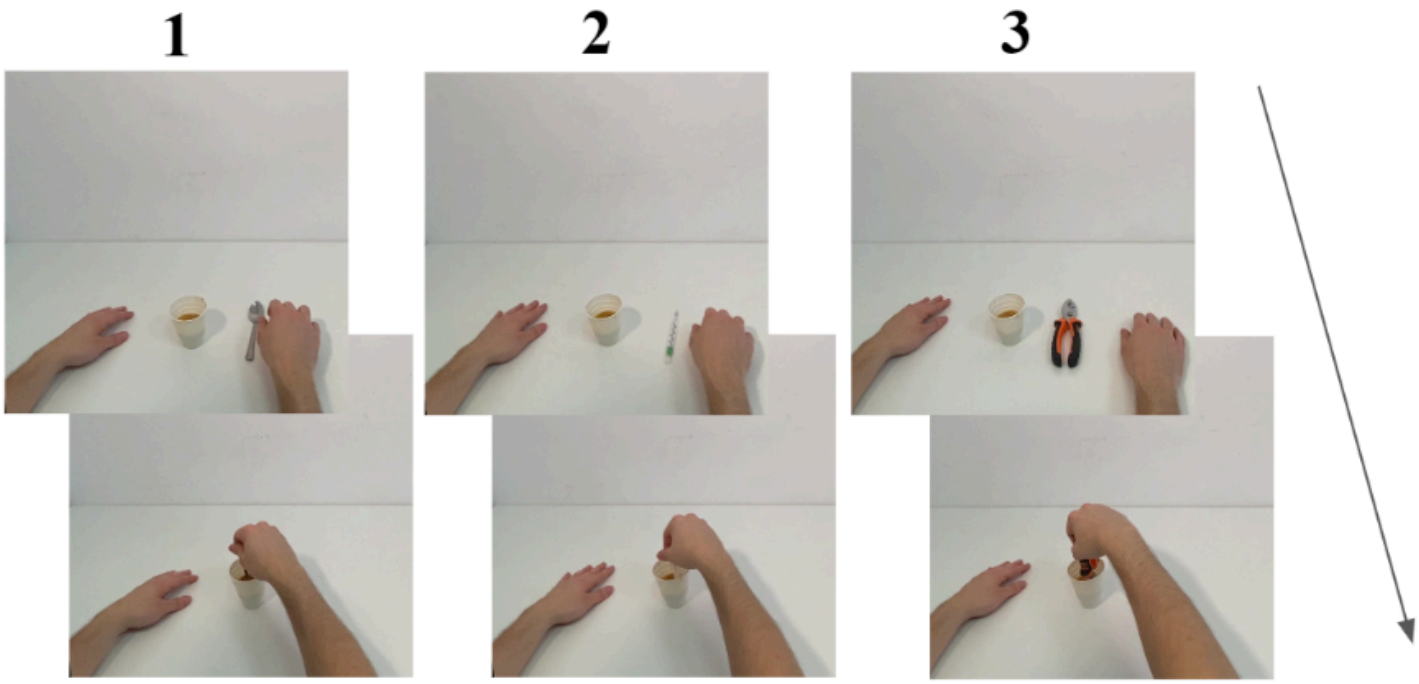


Figure 1

First (first row) and last (second row) frames of the three types of stimuli "You are stirring the coffee with" 1) a spoon (predictable), 2) a thermometer (motor-similar unpredictable) and 3) pliers (totally unpredictable)

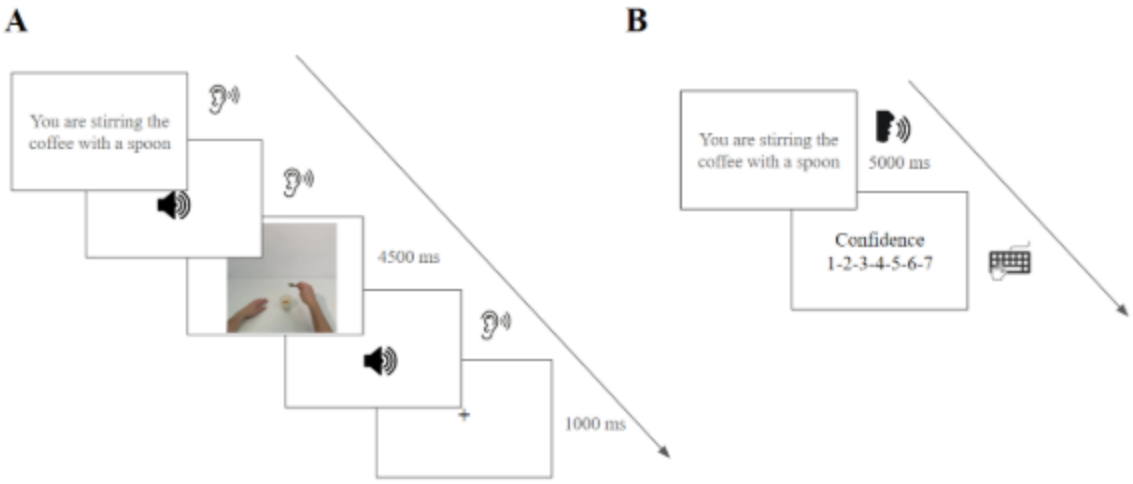


Figure 2

A). Example encoding trial of the Action Observation (AO) group, B). Example trial of the recognition task (same for all groups)

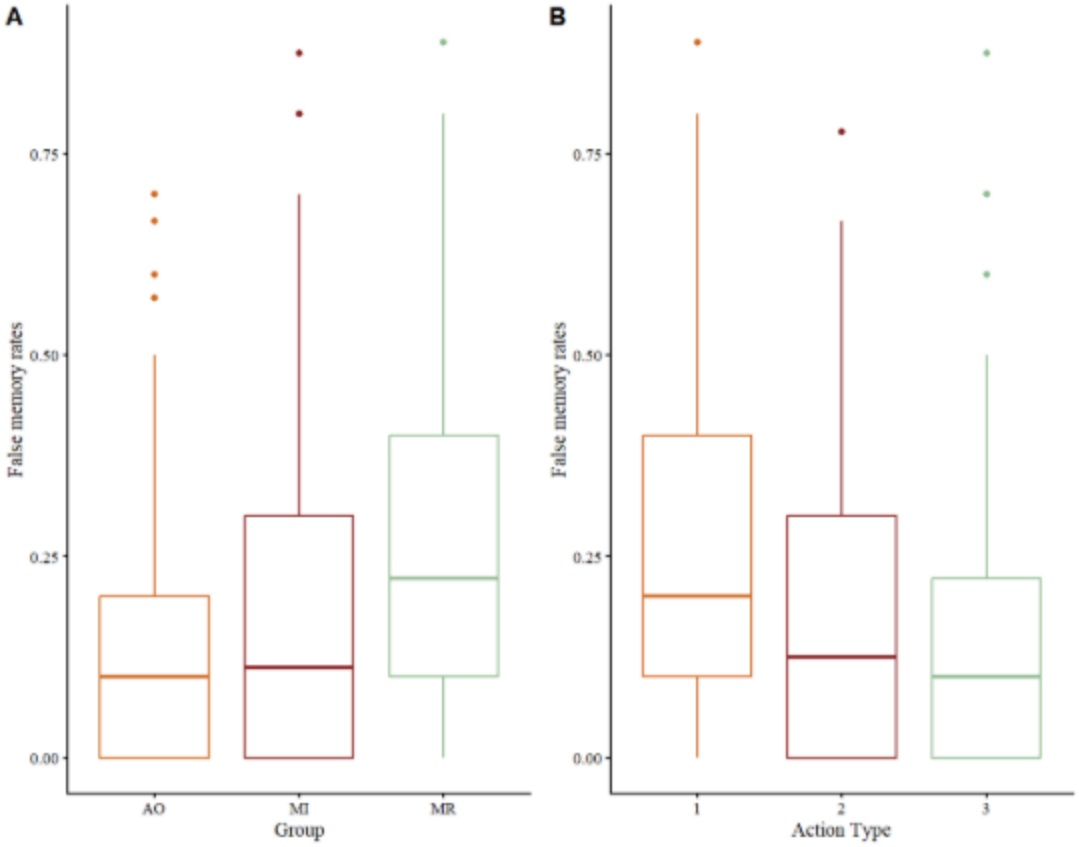


Figure 3

False memory rates based on A) Group and B) Action Type

Notes. Group: AO = Action Observation; MI = Motor Imagery; MR = Mental Rehearsal. Action Type: 1 = Predictable; 2= Motor-similar unpredictable; 3 = Totally unpredictable sentences

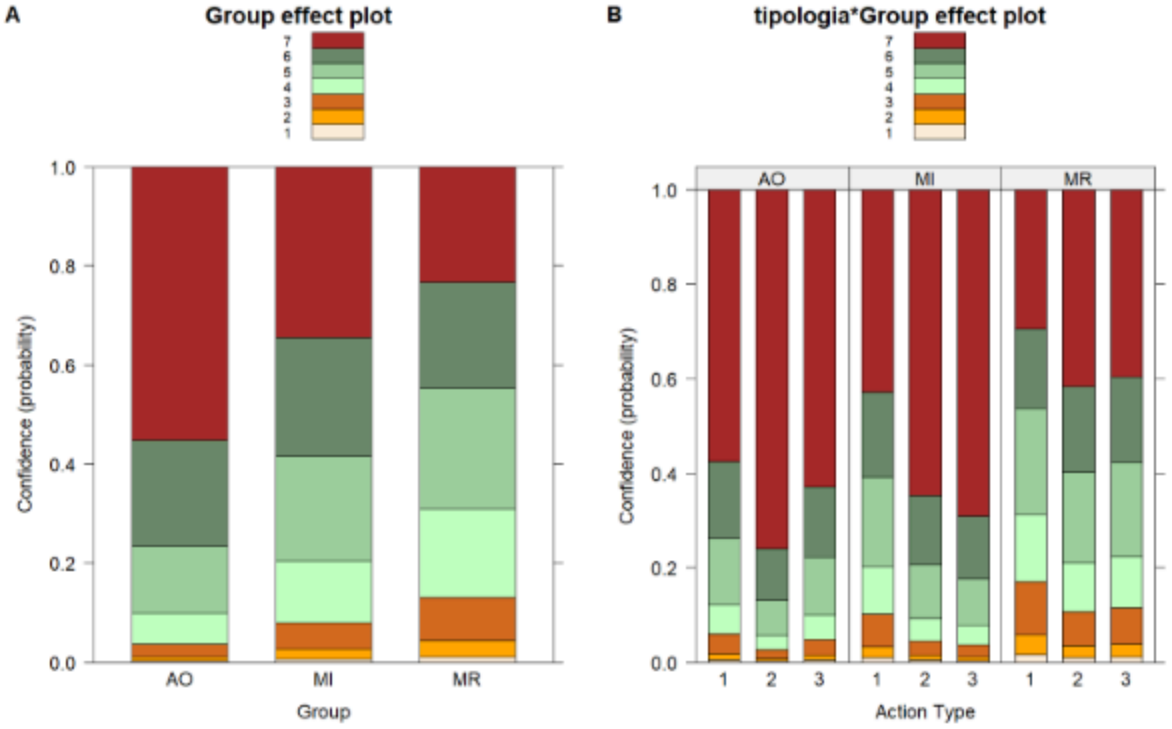


Figure 4

Confidence (probability) based on A) new items related to Group differences and B) old items related to the interaction between Action Type and Group

Notes. Group: AO = Action Observation; MI = Motor Imagery; MR = Mental Rehearsal. Action Type: 1 = Predictable; 2= Motor-similar unpredictable; 3 = Totally unpredictable sentences