What is the best strategy for retaining gestures in working memory?

Guillaume Gimenes\textsuperscript{1}, Valérie Pennequin\textsuperscript{1}, Tom Mercer\textsuperscript{2}

\textsuperscript{1} EA 2114, PAV, University of Tours, France

\textsuperscript{2} Institute of Psychology, University of Wolverhampton, UK

Correspondence should be addressed to:

Dr. Guillaume GIMENES

EA 2114, Psychologie des Ages de la Vie \textsuperscript{1}

Université François-Rabelais de Tours

3, rue des tanneurs BP 4103

37 041 Tours Cedex 1, France

Telephone: +33 247 36 67 21

E-Mail: guillaume.gimenes@univ-tours.fr

Pr. Valérie PENNEQUIN

EA 2114 Psychologie des Ages de la Vie

Université François-Rabelais de Tours

3, rue des tanneurs BP 4103

37 041 Tours Cedex 1, France

Telephone: +33 247 36 67 21

E-mail: valerie.pennequin@univ-tours.fr
STRATEGIES TO RETAIN GESTURES IN WM

Dr. Tom MERCER

Institute of Psychology,

University of Wolverhampton,

Wulfruna Street,

Wolverhampton, WV1 1LY, UK

Telephone: +44 1902 321368

E-mail: t.mercer2@wlv.ac.uk
What is the best strategy for retaining gestures in working memory?

The present study aimed to determine whether the recall of gestures in working memory could be enhanced by verbal or gestural strategies. We also attempted to examine whether these strategies could help resist verbal or gestural interference. Fifty-four participants were divided into three groups according to the content of the training session. This included a control group, a verbal strategy group (where gestures were associated with labels) and a gestural strategy group (where participants repeated gestures and were told to imagine reproducing the movements). During the experiment, the participants had to reproduce series of gestures under three conditions: “no interference”, gestural interference (gestural suppression) and verbal interference (articulatory suppression). The results showed that task performance was enhanced in the verbal strategy group, but there was no significant difference between the gestural strategy and control groups. Moreover, compared to the “no interference” condition, performance decreased in the presence of gestural interference, except within the verbal strategy group. Finally, verbal interference hindered performance in all groups. The discussion focuses on the use of labels to recall gestures and differentiates induced strategies from self-initiated strategies.

Keywords: working memory; gestures; strategy

Introduction

The present study focuses on the retention of gestures within working memory, and specifically examines whether the recall of gestures can be improved. This experiment was influenced by Baddeley and Hitch’s (1974) multi-component model of working memory, which has generated extensive research. This model has been updated several times over the past 40 years (Baddeley, 1986, 2000; Baddeley, Allen, & Hitch, 2011), with the addition of components and a rearrangement of the way in which they interact. In its current form, the model postulates the existence of four components, including the visuo-spatial sketchpad, the phonological loop, the episodic buffer and the central executive. The visuo-spatial sketchpad processes visual and spatial material, along with haptic information (tactile and kinaesthetic; Baddeley et al., 2011). Conversely, the phonological loop – which consists of a passive store
STRATEGIES TO RETAIN GESTURES IN WM

and an articulatory loop to refresh stored information – processes speech, music and sound.

According to Wilson and Emmorey’s (Wilson & Emmorey, 1997a, 1997b, 1998, 2003; Wilson, 2001) work on deaf people, the phonological loop also handles sign language and lip reading (Baddeley et al., 2011). The third component – the episodic buffer – is connected to both the visuo-spatial sketchpad and the phonological loop. It consists of a temporary multimodal store that binds different types of information from the other components. The episodic buffer also processes olfactory and gustatory material. The fourth and final component is the central executive, which manages attentional resources.

In addition to these components, studies by Smyth, Pearson and Pendleton (1988) and Gimenes, Pennequin and Sorel (2013) suggest that a dedicated component is responsible for gestures. In five experiments, Smyth et al. (1988) asked their participants to learn patterned movements (neither gestures or mimes), words and spatial patterns. There were also three different interference conditions: articulatory suppression (repeating “one, two, three, four, five”), gestural suppression (tapping a sequence on the top of the head, the shoulders and the hips with both hands) and spatial suppression (tapping patterns on blocks). By examining the impact of different types of interference, Smyth et al. reported multiple dissociations. Specifically, memory for certain stimuli was only affected by the corresponding sensory task (e.g. the spatial patterns were only influenced by spatial suppression). However, the double dissociation between verbal and gestural learning was not fully realised, as gestural learning was hindered by both gestural and articulatory suppression. The authors subsequently modified their paradigm to facilitate gestural learning and decrease the effect of articulatory suppression. To do this, they eased the attentional demands of the task by repeating each of the gestures before the main experiment. This resulted in a simple dissociation as gestural learning was still damaged by gestural suppression, but not by articulatory suppression.

Gimenes et al. (2013) further examined the possible dissociation between gestural and verbal
learning, but their method differed from that employed by Smyth et al. Gimenes et al. presented sequences of three gestures, and each gesture included a combination of three components (movements, hand shapes, and locations; *ibid* Wilson & Fox, 2007). Verbal labelling of the gestures was minimised by using short encoding times (1 s) and limited repetitions of each gesture. The study revealed that gestural learning was only hindered by gestural suppression, and not by articulatory suppression.

Whilst the findings of Gimenes et al. (2013) showed that articulatory suppression does not affect working memory for gestures, similar experiments have found that a verbalisation strategy can be used to enhance performance. Frencham, Fox and Maybery (2004) asked their participants to recall hand movements from the KHMT (the Kaufman Hand Movement Test). In addition to a baseline condition, Frencham et al.’s study used a verbal shadowing task during stimulus presentation. This task required participants to repeat words that either were congruent or incongruent with the hand movements. The congruent words were shown to facilitate recall, in comparison to the baseline, but incongruent words had the opposite effect. Miyahara (2007) also used the KHMT in an experiment testing both young and old adults. Utilising a verbal labelling strategy improved performance on the KHMT in comparison to an initial attempt that lacked this strategy. However, Miyahara noted that all of the young adult participants spontaneously labelled the hand gestures in their initial attempt, although they may have used ineffective labels. Miyahara et al. (2013) further studied the effects of verbal labelling in a neuroimaging study. They examined the neural substrates involved in recalling hand movements by comparing participants who spontaneously used verbal labelling with those who did not. The use of verbal labelling seemed to reduce the cost of cortical activation, resulting in faster recall.

These earlier studies confirm that verbalisation can be helpful when retaining gestures, highlighting some involvement of verbal labelling. Working memory for gestures may
therefore be optimised through the use of verbal encoding strategies. Yet verbalisation may also make the gestural representation more vulnerable to verbal interference. Specifically, the involvement of words via the phonological loop can make memory traces susceptible to articulatory suppression. Frencham et al. (2003) found that performance on the KHMT was impaired by both gestural and articulatory suppression, compared to a baseline condition. This result is congruent with the report that 77% of Frencham et al.'s (2003) participants spontaneously used verbal labels, thereby involving the phonological loop. The authors suggested that articulatory suppression prevented the representation from being verbally recoded. Moreover, this spontaneous verbalisation is not as efficient as induced verbalisation because, according to Frencham et al. (2004), spontaneous labels may be less specific and consistent.

Smyth et al. (1988) also found that gestural and articulatory suppression hindered gestural learning. However, when they instructed participants to perform the gestures prior to the learning task, verbal suppression no longer affected performance. This could be due to the initial instructional phase acting as a training session, which facilitated the encoding of the gestures by reducing cognitive load. Preliminary training might allow participants to spontaneously verbalise gestures, as observed by Frencham et al. (2003), and thereby counteract verbal interference.

In summary, verbalisation may be a crucial strategy for retaining gestures, whether it is spontaneous (i.e. self-initiated) or controlled (i.e. experimenter-initiated). Spontaneous verbalisation may be a component of baseline performance, as gestures are hindered by articulatory suppression. But this can be counteracted by introducing a training session that involves gestural learning (Smyth et al., 1988). Controlled verbalisation can improve performance compared to a control condition, but it is unclear whether this form of verbalisation can counteract different forms of interference.
The present experiment therefore aimed to study whether induced verbal and gestural encoding strategies enhanced working memory for gestures. The study addressed an important gap in the working memory literature since, to our knowledge, no prior study has evaluated the effects of verbal and gestural strategies on learning gestures within the same experiment. More precisely, two objectives were pursued. The first objective was to evaluate the effectiveness of verbal and gestural strategies on performance. The second objective was to determine the best strategy for counteracting the deleterious effect of interfering tasks.

Method

Participants

Fifty-four students were recruited from the François-Rabelais University of Tours, France. They were aged between 18 and 26 ($M = 21.39$, $SD = 2.34$). None of the participants were familiar with sign language. They were divided into three experimental groups: a control group, a “verbal strategy” group, and a “gestural strategy” group. Each group contained 18 participants, including nine males and nine females. Each participant also took part in all three interference conditions (no interference, verbal interference, and gestural interference). No financial remuneration was given to participants.

Materials

The materials were inspired by Wilson and Fox (2007) and Gimenes et al. (2013), and used the same sets of gestures. Participants were shown video sequences comprising a succession of gestures. The gestures included three components, and each had three possible variations: hand shape (pointing index finger, fist, or five fingers spread), location (forehead, chin/mouth, or chest), and movement (away from the body, circular, or across the body). The original protocol by Wilson and Fox (2007) involved 27 different sequences of three gestures. In the present experiment, only 12 gestures were used and these were chosen randomly (see Table 1
for descriptions of the 12 gestures and Figure 1 for three examples). The use of a smaller number of gestures was intended to limit the length of the training session, whilst still retaining enough variation within the stimulus set to minimise repetition. During training, each gesture lasted 1.5 s and was followed by a blank screen for 3 s.

In each condition, participants were shown 12 video sequences of gestures. The first sequence comprised two gestures, and one gesture was added after every three sequences. Gestures were randomly selected to create the sequences, although successive repetition of gestures was prevented (cf. Figure 1). During experimental trials, each gesture lasted 1.5 s and was followed by a blank screen for 120 ms (3 frames).

**Design**

The experiment employed a 3 (strategy group) x 3 (interference condition) mixed design, with strategy type as the between groups variable and the interference condition as the within groups variable. The order of the interference conditions was balanced between participants. The condition with which participants began the experiment was randomly determined, but the experimental sequence had a fixed order: no interference condition, verbal interference condition and gestural interference condition.

**Procedure**

Each participant was given oral instructions about the procedure, and a demonstration of the different components of each gesture. In the training session, they watched four series of each of the 12 gestures. Gestures were randomly ordered and preceded by a 3 s unfilled interval. However, each group of participants received specific instructions during this preliminary
session. In the control group, participants were told to watch the gestures without doing anything else (neither talking or moving). In the verbal strategy group, participants were given a word or idea associated with an upcoming gesture. For example, “kiss” was stated before the first gesture (cf. Figure 1). Participants had to repeat the word during the 3 s interval to ensure they understood, and they were instructed to associate the word with the gesture. The labels used in this condition were generated from two pre-tests. In the first pre-test, participants were asked to write down words they thought were associated with the gestures. In the second pre-test they had to evaluate the words on a 10-point scale, varying from 0 (no association) to 10 (perfect match). The words and ideas that were selected had mean scores equal to or exceeding 8. Finally, participants in the gestural strategy group were asked to watch the gesture on the screen and imagine producing it. During the 3 s interval, they had to reproduce the gesture twice and as accurately as possible. We did not control whether participants’ eyes were closed when imagining the gesture.

Before beginning the experimental session, all participants were instructed to reproduce the gestures after each sequence in the exact order and as precisely as possible. They were told to remain silent and motionless whilst watching the sequences. Additionally, each group received specific instructions regarding the strategy they had to use during the experiment. Participants in the verbal strategy group were asked to use the words presented in the training session, but they were not permitted to talk aloud. Participants in the gestural strategy group were told to imagine themselves performing the gestures.

In the no interference condition, participants simply had to watch the sequence and reproduce it afterwards. In the gestural interference condition, participants performed a simple repetitive gesture whilst watching the sequence. This gesture had to be repeated from the moment participants pressed the key to launch the video until the sequence ended. They then had to reproduce the gestures observed in the sequence. The repetitive gesture was inspired by
the interference task used by Wilson and Fox (2007). It involved creating an “L” shape with the thumb and middle finger of each hand. A square-like shape was then created with both hands by touching the thumb of one hand with the middle finger of the other hand. The dominant hand had to repeatedly circle around the other hand and touch the fingers again. A training session was carried out before this condition to allow participants to practice the interference activity. This ensured that it could be carried out sufficiently quickly to be repeated more than twice per second. The verbal interference task (articulatory suppression) involved repeating (in French) the words “One, two, three”. It was performed from the moment participants pressed the key to start the video until the end of the sequence.

Each correctly reproduced component of a gesture received one point. Thus, each gesture could score a maximum of three points, with a total score ranging from 0 to 126 for each condition. To score a point, the components had to be executed as accurately as possible. We used the same reliable scoring technique as Gimenes et al. (2013). One point was given for the gesture starting in the right place, including: 1) the forehead, comprising the area from the top of the head to the nose (with the bottom of the palm); 2) the mouth, with the palm placed between the mouth (finger on the nose) and the neck; 3) the chest, from above the torso (fingers at the beginning of the neck) to the abdomen. Movements across and away from the body could be made either in a straight line or with a slight curve, but it had to be in the correct place. Circular movements could be executed clockwise or counterclockwise, as long as circles began at the top and in the correct area. If the hand shape changed during the gesture, it was considered incorrect. Participants were video recorded and responses were coded afterwards. Scores were then converted into percentages.

Results

A 3x3 mixed ANOVA was used to analyse the data, with strategy as the between groups variable (control group, verbal strategy group, gestural strategy group) and interference
condition as the within groups variable (no interference, verbal interference, gestural interference). Newman-Keuls’ post-hoc tests were employed throughout the analysis to limit type II errors, in agreement with our objective of measuring the effect of strategies on interference.

The ANOVA revealed a significant main effect of strategy, \( F(2, 51) = 7.16, \) \( MSE = .014, p < .01, \eta_p^2 = .22 \) (cf. Figure 2). Subsequent Newman-Keuls’ post-hoc tests showed that scores in the control and gestural strategy groups were not reliably different \( (p = .93) \), but the verbal strategy group showed a significant increase in performance when compared with the other two groups \( (p < .001) \).

A significant main effect of interference condition was also found, \( F(2, 102) = 18.69, \) \( MSE = .0035, p < .001, \eta_p^2 = .27 \). Newman-Keuls’ post-hoc tests revealed that gestural interference led to significantly lower performance than the no interference condition \( (p < .001) \). Additionally, performance in the verbal interference condition was significantly poorer than the gestural interference condition \( (p < .05) \). As such, verbal interference led to significantly worse performance than the no interference condition \( (p < .001) \).

Lastly, the interaction was significant, \( F(4, 102) = 5.45, MSE = .0035, p < .001, \eta_p^2 = .18 \). To further explore this, Newman-Keuls’ post-hoc test was used to contrast each interference condition within the three strategy groups. For the control and the gestural strategy groups, a significant decrease in accuracy was observed for both interference conditions \( (p = .05) \), compared with no interference. However, performance within the two interference conditions was not significantly different in both the control and gestural strategy groups. In the verbal strategy group, the no interference condition did not differ from the gestural interference condition \( (p = .84) \), but verbal interference significantly lowered performance compared with the other two conditions \( (p < .001) \).
When the interference conditions were compared across groups, performance in the no interference and gestural interference conditions within the verbal strategy group was significantly higher than the corresponding conditions in the other two groups ($p < .05$). Conversely, performance in the verbal interference condition was not significantly different across strategy groups.

[Insert Figure 2 about here]

**Discussion**

The first objective of the present experiment was to study the effectiveness of different encoding strategies on the retention of gestures in working memory. We compared the performance of three experimental groups, including a control group without specific instructions, a “gestural strategy” group and a “verbal strategy” group. The type of strategy was important, as participants in the verbal group performed better than participants in the control and gestural strategy groups. The latter two groups did not differ. The (induced) verbal strategy may have allowed participants to use their phonological loop, and retain gestures as words. This possibility was also suggested by Frencham et al. (2004). Whilst this does not eliminate a role for gestural encoding, it does signify that verbal encoding can be used to support the learning of gestures.

The absence of any difference between the control and gestural strategy groups can be interpreted in two ways. Firstly, participants in the control group may also have used a motor imagery strategy, but spontaneously. If so, repeating gestures during training does not improve performance. Alternatively, the induced motor imagery strategy may have been ineffective or not applied. If motor imagery had been used spontaneously, it would explain why performance in the control and gestural strategy groups was similar. A spontaneous gestural strategy could be highly efficient, so an induced gestural strategy would not further
improve performance. This may correspond with Allen and Waterman’s (in press) findings. They reported that adults created spontaneous motoric codes when expecting to enact sequences of actions at recall. Thus, our results could be interpreted as a ceiling effect for gestural encoding. Consequently performance may only be enhanced through the use of other encoding strategies (e.g. verbal labelling). However, it is possible that participants simply did not use the gestural strategy, although it seems unlikely that they did not deliberately follow the instructions.

A third plausible explanation proposes that the gestural strategy may have been too difficult to successfully implement. The gestural strategy group had to watch gestures and imagine reproducing the action, so it is possible that the two activities were difficult to perform simultaneously. Participants may have prioritised the information presented on the screen over attempts at visualisation, resulting in the same processing as the control group. Future studies could investigate whether longer training and repeated practice could make the gestural strategy easier to implement (as shown in the fourth experiment of Smyth et al., 1988).

Despite the limited impact of the gestural strategy, the comparison between the no interference conditions showed that the induced verbal strategy was effective. It increased performance from a mean of 74.33% in the control group to a mean of 82.01% in the verbal strategy group. In line with Frencham et al.’s (2004) hypothesis, this improvement can be explained by the recruitment of the phonological loop, which stores labels corresponding to gestures. The phonological loop is not affected by gestural interference, as already demonstrated in previous experiments (Gimenes et al., 2013; Smyth et al., 1988; Wilson & Fox, 2007). The present results offer further support for this interpretation, as gestural suppression did not affect performance for participants within the verbal strategy group. However, verbal interference (repetition of “one, two, three”) significantly decreased recall
scores compared to the no interference condition, regardless of strategy. Post-hoc analyses also indicated that performance in the verbal interference condition did not differ between groups. These results suggest that verbalisation was used in each group. This would involve spontaneous verbalisation in the control and gestural strategy groups, and induced verbalisation in the verbal strategy group. Nonetheless, verbal interference prevents both induced and spontaneous verbalisation, yet gestural encoding is still possible and allows participants to remain efficient: indeed, performance persisted at a high level (approximately 70%). Our interpretation therefore proposes a parallel encoding mechanism, in which gestural and verbal traces are both used in the formation of gestural representations.

The results from the no interference conditions attests that any spontaneous verbalisation in the control and gestural strategy groups was not as efficient as induced verbalisation. This could be due to the heightened attentional demands of spontaneous verbalisation, or it may simply be less effective at characterising the gesture. Furthermore, our labels were congruent with the gestures, which could have helped to improve the retention of the to-be-remembered stimuli by allowing deeper semantic encoding (Craik & Lockhart, 1972). As such, it would be interesting to examine the effects of incongruent words, and how they interact with the different interference conditions.

Smyth et al. (1988) also observed lower performance when participants performed a verbal interference task while learning gestures. Yet their interpretation focused on the attentional load involved in learning gestures when interference was present. In reference to previous experiments, which demonstrated the involvement of spontaneous verbal encoding when memorising gestures (Frencham et al., 2003; Miyahara, 2007; Miyahara et al., 2013), it is possible that verbal interference hinders spontaneous verbalisation. In the present case, it seems plausible that performance could reflect the remaining capacity to recall gestures.
without the help of verbalisation. It would be useful if future studies explicitly measured spontaneous verbalisation, compared to induced verbalisation, in this particular protocol.

The innovative result of the present experiment was the interaction between the strategy group and the type of interference. As in previous studies (Gimenes et al., 2013; Smyth et al., 1988; Wilson & Fox, 2007), we found the classical decrease in recall scores when participants were exposed to gestural interference. In the control group, gestural and verbal interference decreased performance, whereas only verbal interference damaged recall in the verbal strategy group. The labels alone may be sufficient to allow participants to reproduce the gestures, so they may have relied solely on verbal traces. Indeed, these results suggest that participants do not need gestural traces to reproduce gestures, as the verbal strategy counteracted the deleterious effect of gestural suppression.

In conclusion, the present results showed that an induced verbal strategy improved performance compared to an induced gestural strategy and a hypothesised spontaneous verbal strategy in the control group (see Frencham et al., 2004). Furthermore, the verbal strategy not only enhanced performance, but counteracted the deleterious effect of gestural interference. From an applied perspective, it is possible that verbalisation could help individuals retain sequences of movements or gestures in procedures that do not require complex motoric skills. For example, gesture verbalisation could help in the early stages of learning dance choreography in dancing school. Individuals beginning to learn a sequence of movement may benefit from naming the action, instead of recreating the movement without labels.

References


Table 1. Descriptions of the 12 gestures with: the word/idea associated with the gesture in the verbal strategy group; the hand shape (index finger, fist, fingers spread); the location (forehead, chin/mouth, chest); the movement (away, circular, across the body).

<table>
<thead>
<tr>
<th>Word/Idea</th>
<th>Hand shape</th>
<th>Location</th>
<th>Movement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Move over</td>
<td>Fingers spread</td>
<td>Chest</td>
<td>Across</td>
</tr>
<tr>
<td>Halo</td>
<td>Index finger</td>
<td>Forehead</td>
<td>Circular</td>
</tr>
<tr>
<td>Kiss</td>
<td>Fingers spread</td>
<td>Chin/Mouth</td>
<td>Away</td>
</tr>
<tr>
<td>Door</td>
<td>Fist</td>
<td>Chest</td>
<td>Across</td>
</tr>
<tr>
<td>Idea</td>
<td>Index finger</td>
<td>Forehead</td>
<td>Away</td>
</tr>
<tr>
<td>Delicious</td>
<td>Fingers spread</td>
<td>Chest</td>
<td>Circular</td>
</tr>
<tr>
<td>Windscreen wiper</td>
<td>Index finger</td>
<td>Chin/Mouth</td>
<td>Across</td>
</tr>
<tr>
<td>Hat</td>
<td>Fist</td>
<td>Forehead</td>
<td>Away</td>
</tr>
<tr>
<td>To wash one's face</td>
<td>Fingers spread</td>
<td>Chin/Mouth</td>
<td>Circular</td>
</tr>
<tr>
<td>You</td>
<td>Index finger</td>
<td>Chest</td>
<td>Away</td>
</tr>
<tr>
<td>Curtain</td>
<td>Fist</td>
<td>Forehead</td>
<td>Across</td>
</tr>
<tr>
<td>Generous</td>
<td>Fingers spread</td>
<td>Chest</td>
<td>Away</td>
</tr>
</tbody>
</table>
Figure 1. Presentation of three gestures used in the experiment: “Halo”, “Kiss” and “Door”.

- Figure 1(a) Halo
- Figure 1(b) Kiss
- Figure 1(c) Door

Figure 2. Percentage of correct movements as a function of strategies and experimental conditions (error bars represent standard errors).
Figure 1.
Figure 2.