*Journal of Articles in Support of the Null Hypothesis*
Vol. 12, No. 2

www.jasnh.com

# Memory for Five Novel Naturalistic Activities: No Memory Recall Advantage for Enactment over Observation or Pictorial Learning

Rul von Stülpnagel
**University of Freiburg, Germany**
Melanie C. Steffens
**University of Koblenz-Landau, Germany**
Janette C. Schult
**Friedrich-Schiller-University Jena, Germany**

"Learning by doing" promises to lead to more efficient acquisition than other learning strategies. Indeed, much research has established that enactment leads to better recognition and recall of simple verb-object phrases (e.g., "light the match", "touch your nose") than intentional learning without enactment. Only few studies have compared the acquisition of novel naturalistic activities (e.g., "to fold a paper frog") across different study conditions, and only a few different activities have been investigated overall. Two experiments tested whether five very different such activities can be carried out better after enactment learning than after observing a model or after pictorial learning instructions. No evidence of different performance across study conditions was obtained.

Keywords: Enactment, observation, pictorial learning, action sequences, novel naturalistic activities, memory recall, performance recall

Every day, people perform naturalistic activities that consist of well-known goal-oriented action sequences, such as making coffee, preparing cookie dough, or changing a light bulb. Occasionally, people also need to learn novel sequences of actions or novel naturalistic activities (NNAs, Gold & Park, 2009) such as "to build a bird feeder" or "to fold a paper frog". Many people seem to assume that in order to memorize such NNAs properly they have to perform all action steps involved themselves rather than observing someone else perform them or than simply reading instructions or seeing pictures of each action step. This conception is mirrored by action memory research where it is often claimed that "learning by doing" (i.e., the enactment of actions and activities) enables better memorization than "learning by viewing" (i.e., the observation of actions and activities) or than verbal learning, as enactment "provides optimal encoding" (Nilsson & Cohen, 1988, p.427) – the axiomatically named "enactment effect". Whereas memory for actions has typically been examined by presenting lists of simple verb-object phrases (e.g., "crack an egg", "clap your hands") to be retrieved in later memory tests, a number of recent studies examined more complex study materials: they referred to them as movement patterns (Helstrup, 2005), action sequences (Schult, von Stülpnagel, & Steffens, 2014; Steffens, 2007), or NNAs (von Stülpnagel, Schult, Richter, & Steffens, 2015). As a rule, recall of action sequences was not superior after enactment as compared to observation of the same sequence. However, the generalizability of this conclusion beyond a lab setting may be considered debatable. Most of the studies investigated only one NNA; only two of the studies featured two different NNAs, respectively (Steffens, 2007; von Stülpnagel et al., 2015). NNAs can differ in many features that could potentially benefit memory performance after enactment encoding. Thus, the aim of the present experiments was extending the evidence on memory for NNAs after "learning by doing" to five not-yet-examined NNAs from very different contexts.

## Memory performance after enactment, observation, and verbal learning

Memory for actions has been studied since the 1980's (e.g., Bäckman, Nilsson, & Chalom, 1986; Cohen, 1983; Zimmer & Engelkamp, 1984). In a prototypical experiment, participants study a list of unrelated, simple verb-object phrases (e.g., "clap your hands", "break the match", etc.). During this study phase, some participants enact the actions (with or without the actual objects). Other participants observe the experimenter demonstrate the actions, or they learn them verbally. In a subsequent test phase, memory is typically tested with verbal recall or recognition tests. Most researchers in the field agree that the enactment of actions generally improves memory over verbal learning (e.g., Earles, 1996; Engelkamp, 1998). When enactment is compared to observing a model perform the denoted actions, empirical findings are less clear. For example, studies comparing free recall performance after enactment and observation in between-list study designs[1] rarely found an enactment effect (e.g., Cohen, Peterson, & Mantini-Atkinson, 1987; Engelkamp, Jahn, & Seiler, 2003, but also see Engelkamp & Zimmer, 1983; Golly-Häring & Engelkamp, 2003; for a review, see Steffens, von Stülpnagel, & Schult, in press).

Accounts on the mechanisms underlying the so-called enactment converge on the distinction between item-specific (i.e., features of an individual action) and relational

1 Between-list study designs indicate that participants studied either by enactment or by observation only, which appears highly relevant for the question which study condition is recommendable in everyday life.

processing (i.e., conceptual or order relations between actions, Engelkamp & Dehn, 2000; Engelkamp, Seiler, & Zimmer, 2004; Schult et al., 2014; Steffens, 2007). It has been assumed that enactment draws attention to item-specific information, which is why enactment effects are particularly pronounced in recognition tests that rely on item-specific information (e.g., Dick, Kean, & Sands, 1989; Engelkamp & Dehn, 2000; Golly-Häring & Engelkamp, 2003; Hornstein & Mulligan, 2004; Manzi & Nigro, 2008; Mulligan & Hornstein, 2003; but see Feyereisen, 2006). Because of drawing the actor's focus of attention towards item-specific information, enactment typically does not provoke improved processing of relations among action phrases. In contrast, it has been found that observation facilitates the processing of order information more than enactment does (Engelkamp & Dehn, 2000). Free recall is based both on item-specific and on relational information (e.g., Einstein & Reed, 1980; Hunt & Mcdaniel, 1993). The efficient use of relational information during recall could be the source of similar net recall after observation encoding as compared to enactment encoding that uses item-specific information more efficiently.

We argue that goal-directed action sequences provide relational information, such as the hierarchical structure, "in-order-to", and "enable" relations (Lichtenstein & Brewer, 1980). Increased processing of item-specific information during enactment encoding should consume encoding resources necessary to understand such structural relations. Thus, during observation more encoding resources should be available to encode these relations. Relational information based on goals or on the outcomes of sets of actions could be used to reconstruct which action has to be done when. Thus, a good encoding of relations among action steps may provide efficient retrieval paths for recalling actions within action sequences. In line with this assumption, Steffens (2007) and Schult and colleagues (2014) found better organization of free recall protocols after observation than enactment of action sequences. In both studies, recall levels were comparable across encoding conditions (also see von Stülpnagel et al., 2015).

One problem with the existing evidence base is that only a few different NNAs have been investigated so far. At the same time, little is known about the features that make actions more or less memorable (Cohen et al., 1987), let alone NNAs. Such features may also interact with the encoding condition. For example, sequences differ much in their complexity and memorability (von Stülpnagel et al., 2015), and the goal structure of an NNAs may be quite obvious or not. The generalizability of existing findings on comparable memory recall after enactment and observation to different NNAs is therefore an open question.

**Aims of the present research**

In short, it is frequently assumed by both lay people and experts that "learning by doing" is the best way to encode NNAs as compared to other study conditions such as "learning by observation." Evidence to the contrary is so far limited to very few NNAs, and could thus be dismissed as being particular to the properties of the tested NNAs. Thus, the research at hand compared memory recall for five NNAs from diverse contexts studied by enactment versus observation in between-subjects designs, with the aim to provide a generalizable conclusion whether the frequently assumed enactment effect is found for NNAs. The two experiments we present used different conceptualizations of the observation condition. Additionally, we included a pictorial learning condition, which resembled a classic verbal learning condition as closely as possible; a pure verbal study condition was impossible because unambiguous verbal-only descriptions for the complex NNAs we examined could not be constructed (see von Stülpnagel et al., 2015, for the implementation of a similar study condition).

Our central hypothesis was that comparable memory (i.e., the ability to re-enact an NNA) after enactment encoding as compared to observing a model perform actions and as compared to pictorial learning is not limited to specific instances of NNAs, but is a generalizable and robust null-effect in several different contexts.

## Experiment 1

In Experiment 1, participants either studied the NNAs according to visualized instructions in a pictorial learning condition, or additionally carried out all of the visualized instructions in an enactment condition, or learned them by watching video clips in an observation condition, thus introducing typical differences in presentation modality between study conditions (see Engelkamp & Krumnacker, 1980).

### Method

*Participants*

Participants were 94 university students. Gender (18% men) and age (18-37 years, $M$ = 21.40, $SD$ = 3.18) were about equally distributed across study conditions. Due to technical problems of the apparatus and failures in the experimental procedure, data sets of 35 participants were partially incomplete. Details, actual Ns, and achieved power are presented in the results section.

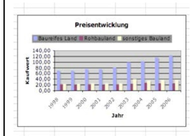
*Materials*

We included five very different NNAs from different contexts (building a Lego airplane, folding a napkin into a hat, binding a knot with a rope to a karabiner, a wood-and-string puzzle, and creating a computer graph; see Table 1). For the enactment condition and the pictorial learning condition, we prepared illustrated stepwise instructions to be presented on a computer screen, which demonstrated each action step of each NNA. Some steps needed more than one illustration to demonstrate the action, resulting in instructions including between 15 and 43 illustrated steps. Hands were visible in the illustrations if necessary. Iconographies (e.g., arrows) were added to some illustrations in order to increase comprehensibility. For the observation condition, video clips were prepared that resembled the presentations in the other conditions as much as possible. No additional iconographies were included in the video clips. The video clips were adjusted to the same duration as in the other conditions with frozen images after the demonstration of each step.

Seven pretest participants performed all activities with the illustrated instructions at their own pace with a subsequent, unlimited performance recall for each activity. The average time the pre-testers needed to complete each action step was used to set the presentation rate in the main experiment, where the stepwise presentations ran automatically. The fixed presentation rate was implemented to keep study time constant across the experimental conditions. A further extension of the presentation rate would have imposed fewer demands on participants in the enactment condition, but would have further strained the concentration and patience of participants in the other conditions, who already experienced the presentation rate as very slow.

Performance recalls in the pretest were used as an indicator of a recall time limit in the main experiment. This limit was set with the aim that only participants with complete understanding of the activity would be able to finish.

***Table 1:*** *Illustrations of Completed NNAs, the Number of Action Steps Necessary to Complete each NNA, the Overall Study Time, and the Recall Time Limit, Separately for Each NNA.*

| | Lego | Napkin | Knot | Puzzle | Graph |
|---|---|---|---|---|---|
| |  |  |  |  |  |
| Action Steps | 30 | 12 | 10 | 9 | 20 |
| Study Time | 230s | 123s | 137s | 125s | 371s |
| Recall Time Limit | 180s | 60s | 90s | 90s | 240s |

*Procedure*

Up to four participants at a time were assigned to the same study condition. Participants were told that they would study several activities, each followed by a time-limited memory test. In the study phase, participants in the enactment condition received the materials of the respective NNA with the instruction to execute the steps as shown in the step-wise instructions displayed with a beamer. If participants made mistakes or failed to catch up with the presentation, the experimenter told them just to watch the presentation (see Results & Discussion for further details). In the pictorial learning condition, participants were asked to carefully watch the same instructions without additional enactment. In the observation condition, participants watched the video clip demonstrating the respective NNA instead of the stepwise presentation.

After the study phase of each activity, participants worked on unrelated questionnaires for two minutes as distractor tasks. In the subsequent recall phase, participants in all conditions received the respective material in its initial state and were asked to complete the activity as far as possible within the given time limit. Pictures (respectively screenshots for the graph) of the final state were taken and subsequently analyzed for recognizable action steps. For each activity the number of correctly completed action steps was counted by a rater blind to the experimental condition. If participants disassembled an object and repeated action steps these steps were counted only once. Due to variations in the precision with which action steps were executed, an action step was not counted if it was not unambiguously recognizable as correctly performed. Recall proportions were computed and used in the respective analyses.

This procedure was repeated for all five activities, which were presented in a counterbalanced order. Finally, demographic data as well as familiarity with the activities in general (with 7-point Likert scales) were collected. Participants were thanked and debriefed. The whole experiment lasted about 60 minutes.

### *Design*

The independent variable was study condition (enactment vs. observation vs. pictorial learning), manipulated between subjects. Dependent variables were recall proportions of the five activities.

## Results and Discussion

Participants reported comparable and intermediate familiarity with the different activities in general (ranging from $M$ = 3.48, $SD$ = 1.85, for using spreadsheet software to $M$ = 5.03, $SD$ = 1.59, for building with Lego bricks). However, the study phase of each of the NNAs was not completed by about 15% of the participants, respectively (see top line of Table 2 for the remaining number of valid data sets per NNA). Scores from the respective NNAs were excluded from further analyses. The missing data resulted mostly from participants in the enactment condition who did not manage to keep pace with the timed instructions or made errors during the execution of the displayed action steps[2]. Taken together, there were 31 participants with one or two corrupted NNAs and four participants who did not complete three or four of the five NNAs. As those four participants were obviously overstrained with their task, they were excluded from all further analyses.

The remaining (listwise) $N$ = 59 with complete data sets was not sufficient to detect even a large effect of study condition ($f$ = .4) with adequate statistical power (Cohen, 1977; Faul, Erdfelder, Lang, & Buchner, 2007). Thus, we computed five one-way ANOVAs for study condition (enactment vs. observation vs. pictorial learning, between subjects), separately for each NNA, with a Bonferroni-Holm-corrected Type-I-error. Descriptive data and statistics are presented in Table 2. We found no indications that enactment leads to generally superior memory performance (i.e., performing the same steps again) as compared to observation learning. Surprisingly, performance after pictorial learning was not inferior to both other study conditions − rather on the contrary. However, due to the unexpected high dropout rate, a test power sufficient to accept the null-hypothesis with $1 - \beta > .80$ was achieved for only two of the five NNAs. Furthermore, we could not compute potential interaction effects between study condition and NNA. Thus, we aimed to replicate our findings with a larger data set in Experiment 2.

## Experiment 2

In addition to the large number of incomplete data sets, the validity of Experiment 1 potentially suffered from a presentation bias because participants in the observation condition watched video clips rather than the stepwise illustrations presented to the other study conditions. Furthermore, studying pictorial learning and enactment along illustrated instructions may have been more familiar and closer to everyday life situations of many participants than studying video clips was to observing an actual person demonstrating a NNA. Thus, in Experiment 2, participants in the observation condition studied the same stepwise presentations as the other conditions and additionally observed the experimenter performing the activities.

2 Anticipating this problem, we had collected more data for the enactment condition than for the other experimental conditions.

***Table 2:*** *Recall Percentages (and standard deviations) and Statistical Test Results for each NNA in Experiment 1, Separately for Study Conditions.*

| | Lego ($n$ = 80) | Napkin ($n$ = 79) | Knot ($n$ = 79) | Puzzle ($n$ = 79) | Graph ($n$ = 82) |
|---|---|---|---|---|---|
| Enactment | .41 (.23) | .56 (.30) | .40 (.28) | .64 (.45) | .45 (.22) |
| Observation | .25 (.14) | .41 (.22) | .35 (.24) | .60 (.45) | .42 (.27) |
| Pictorial learning | .34 (.24) | .52 (.29) | .50 (.28) | .75 (.43) | .47 (.26) |
| One-way ANOVA statistics | $F(1,77) = 4.20$, $\eta_p^2 = .10$, $p = .02$ | $F(1,76) = 2.21$, $\eta_p^2 = .06$, $p = .12$ | $F(1,76) = 1.79$, $\eta_p^2 = .05$, $p = .17$ | $F(1,76) < 1$, $\eta_p^2 = .02$, $p = .45$ | $F(1,79) < 1$, $\eta_p^2 = .01$, $p = .70$ |
| Bonferroni-Holm-corrected Type-I-error | $\alpha = .01$ | $\alpha = .013$ | $\alpha = .016$ | $\alpha = .025$ | $\alpha = .05$ |
| Achieved statistical power to detect a large effect ($f$ = .4) | $1 - \beta = .72$ | $1 - \beta = .75$ | $1 - \beta = .77$ | $1 - \beta = .82$ | $1 - \beta = .90$ |

## Method

*Participants & Design*

Participants were 91 university students. Gender (20% men) and age (18-39 years, $M = 22.03$, $SD = 3.71$) were about equally distributed over study conditions. Design and hypotheses were identical to Experiment 1.

*Materials & Procedure*

Materials and procedure corresponded to Experiment 1 with the following exceptions. Minor adjustments to single illustrations and the presentation rate were made in order to increase comprehensibility of the instructions. Participants in the observation condition watched the same stepwise presentation as the other study conditions and additionally observed the experimenter, who executed all steps in accordance with the presentation. Care was taken that all participants could see the presentation as well as the actions of the experimenter. Familiarity with the activities in general was assessed (with 12-point Likert scales).

## Results and Discussion

Participants reported about comparable and intermediate familiarity with the different activities in general (ranging from $M = 4.28$, $SD = 2.80$, for using spreadsheet software to $M = 7.24$, $SD = 2.94$, for building with Lego bricks).

The adjusted instructions decreased the number of dropouts. There were 21 participants with one or two corrupted NNAs and one participant not completing three of the five NNAs, who was excluded from all further analyses. The remaining ns for each NNA are presented in Table 3. Data sets of 70 participants were complete, which was sufficient to detect a large effect of study condition ($f$ = .4) (Cohen, 1977) with a Type-I-error of $\alpha = .05$. and a statistical power of $1 - \beta = .84$ (Faul et al., 2007). A 3 (study condition) × 5 (NNA) ANOVA with repeated measures on the second factor showed that there were general differences in the recall level between the NNAs, $F(4,268) = 10.65$, $p < .001$, $\eta_p^2 = .14$[3]. However, there was no main effect of study condition (enactment: $M = .47$, $SE = .04$; observation: $M = .39$, $SE = .05$; and pictorial learning: $M = .43$, $SE = .03$) and no interaction effect, both $F$s < 1.33, ns[4]. Taken together, Experiment 2 confirmed the impression of Experiment 1 regarding a null-effect for the effects of encoding NNAs via different study conditions with a higher statistical power.

## General Discussion

A central assumption of action memory research is that the enactment of actions leads to superior recall performance as compared to other study conditions such as verbal learning or observation. Whereas many findings confirm this assumption for an enactment effect over verbal learning, many other studies report that memory performance after enactment encoding is frequently comparable to memory performance after observation encoding (see Schult et al., 2014; Steffens et al., submitted for publication, for reviews). However, the vast majority of studies in this field have been conducted under rather artificial

***Table 3:*** *Recall Percentages (and standard deviations) and Statistical Test Results for each NNA in Experiment 2, Separately for Study Conditions.*

| | Lego ($n$ = 89) | Napkin ($n$ = 90) | Knot ($n$ = 81) | Puzzle ($n$ = 89) | Graph ($n$ = 87) |
|---|---|---|---|---|---|
| Enactment: | .35 (.20) | .56 (.36) | .47 (.35) | .64 (.41) | .32 (.22) |
| Observation: | .27 (.20) | .51 (.32) | .34 (.26) | .62 (.45) | .29 (.18) |
| Pictorial learning: | .28 (.17) | .46 (.29) | .42 (.24) | .54 (.43) | .45 (.19) |

# General Discussion

3 Recall differences between the five NNAs were not the focus of the research at hand and were to be expected due to the different number of to be performed action steps. Thus, we refrain further analyses inflating the type-I-error probability.

4 The sample size of Experiment 2 only allowed conclusions assuming a large statistical effect. Given the nearly identical setup of both experiments, we ran a supplementary joint analysis of Experiments 1-2, which was sufficient to detect an about medium-sized effect of encoding condition ($f$ = .28), based on $N$ = 129, $\alpha = .05$, and $1 - \beta = .80$ (Faul et al., 2007). This 3 (study condition) × 5 (NNA) × 2 (experiment) ANOVA showed no main effects or interactions regarding the factor "experiment" (all $F$s < 1.16, all $p$s > .32), thus corroborating the comparability of the data sets. Furthermore, the analysis yielded a main effect of NNA, $F(4,488) = 17.29$, $p < .001$, $\eta_p^2 = .12$, again demonstrating the differing difficulty of the NNAs. Concerning the study conditions, enactment ($M = .47$, $SE = .03$) and pictorial learning ($M = .47$, $SE = .03$) showed descriptively a slightly enhanced memory performance as compared to observation ($M = .39$, $SE = .03$), but this effect was not significant, $F(2,122) = 2.56$, $p = .08$, $\eta_p^2 = .04$. There were no other significant main or interaction effects, all $F$s < 1.16, all $p$s > .32. Thus, the joint analysis of Experiments 1-2 corroborated our conclusions.

laboratory conditions. In contrast to lists of unrelated and easy-to-verbalize action phrases recalled verbally, most real-life activities do not consist of such easy-to-verbalize actions, and the aim of real-life activities normally is to perform them again, not to recall them verbally. Thus, it is possible that an enactment advantage over observation is obtained in an experimental setting closer to everyday life. Only a few studies from this field investigated such study materials and reported comparable memory performance after enactment and observation, but this evidence is limited to a small number of different NNAs (Steffens, 2007; von Stülpnagel et al., 2015). The aim of the present paper was to extend the evidence on this issue on memorizing and re-performing five very different goal-directed and close-to-real-life activities. In two experiments the present study found no evidence that enactment during study leads to better memory as compared to observation or learning from pictorial instructions.

These findings converge with research from related domains. Foley and colleagues found not advantage of learning dancing figures by performing them as compared to observing them being performed (Foley, Bouffard, Raag, & DiSanto-Rose, 1991). Stull and Mayer (2007) set up an experiment where a scientific text was either studied with participants creating their own illustrations (resembling an enactment condition), or with preset illustrations (resembling an observation condition). The authors reported comparable recall for both study conditions and even better transfer performance for participants who studied preset illustrations than participants who created their own graphs. Somewhat contrasting evidence was reported by Dijkstra, MacMahon, and Misirlisoy (2008), who found superior memory for golf and everyday items as compared to other study conditions. However, they applied a within-subject design, which has been repeatedly shown to bolster enactment encoding (e.g., Engelkamp & Dehn, 2000).

It was uncertain whether there would be an enactment effect compared to the visual learning condition, as previous research agrees on an enactment effect as compared to verbal learning. However, visual learning is superior to verbal learning (Maisto & Queen, 1992; Paivio, 1986). Neither Experiment 1 nor Experiment 2 showed evidence that visual learning of NNAs led to a substantial disadvantage as compared to the other study conditions. Similar findings were reported in another recent experiment (von Stülpnagel et al., 2015, Exp. 2). In our interpretation, participants (knowing about the subsequent performance recalls) were able to transfer the visual instructions into an imagery of the displayed activity. This was sufficient to generate a mnemonic model and to enable a decent level of recall. Thus, in contrast to verbal learning, visual learning does not necessarily lead to a disadvantage compared to other study conditions (see also Bird, Osman, Saggerson, & Heyes, 2005).

A possible explanation why re-enactment did not benefit from enactment is the extra effort that enactment brings with it (von Stülpnagel et al., 2015). Whereas "learning by doing" (i.e., enactment) may provide richer encoding and memorization of the tested activities, it also increases cognitive load, reducing memory performance to the level of "learning by viewing" (i.e., observation and visual learning); for further discussions of this issue, see Steffens and colleagues (submitted for publication).

A limitation regarding the interpretation of the present null finding concerns the longevity of the observed lack of effect. In this as in most previous experiments, recall followed shortly after encoding (but also see, for example, Manzi & Nigro, 2008). One may argue that even though there was no enactment effect in the short run, enactment may have left deeper memory traces that would provide a memory advantage to enacting participants in a delayed recall (Knopf & Neidhardt, 1989). Further data are needed to resolve this issue.

A potential bias in the present experiments results from the large number of participants who were not able to enact the NNAs correctly or quickly enough during the study phase to keep up with the instructions. However, time limits were necessary to keep the exposition constant across experimental conditions. Future studies approaching this issue should take great care to adapt the instructions accordingly. Extending the available time further could enable more participants to enact the displayed activities properly, but this may result in rapidly decreasing motivation and attention of participants in the other study conditions. This difference in the time span people are willing to spend on encoding a single step may be one reason for memory differences outside of the lab. Effects of self-paced studying on encoding time and memory performance differences between study conditions appear to be a worthwhile topic for future research. Concerning the dropout in the studies at hand, we argue that this potential confound may have actually led to an *overestimation* of performance after enactment encoding, as only those participants were included in the analyses who were able to process and execute all action steps properly and quickly. We thus believe that our conclusions are not questioned by this issue.

To conclude: We attempted to extend the enactment effect to real-life activities. Participants learned by enactment, observation, or visual instruction, and were tested with a performance recall. There was no enactment effect. On the contrary: Even visual learning led to a similar level of recall as enactment. Thus, an enactment benefit does not appear to extend to real-life activities. All in all, the widespread assumption that "learning by doing" is always the best way to remember an action or activity was not supported by this research.

**Authors' Notes**

Rul von Stülpnagel (Abteilung Kognitionswissenschaft, Universität Freiburg, Friedrichstr. 50, D-79098 Freiburg, Germany); Melanie C. Steffens, Faculty of Psychology, Social and Economic Psychology, Fortstr. 7, D-76829 Landau, Germany; Janette C. Schult (Institut für Psychologie, Friedrich-Schiller-University Jena, Am Steiger 3, Haus 1, D-07743 Jena, Germany).
The current research was supported by grants from the Deutsche Forschungsgemeinschaft to the second author (German Science Foundation, Ste 938/7-1, Ste 938/7-2). The authors declare to have no conflict of interest regarding this research. We thank Sascha Poppitz for valuable support with the preparation and organization of Experiment 2, Dave Cromm, Nele Fischer, Linda Köhler, Claudia Richter, and Linda Seidemann for support with the data collection, and Anne Berthold for valuable comments on previous drafts of this article.
Please address correspondence to the first author (rul.von.stuelpnagel@cognition.uni-freiburg.de).

## References


Bäckman, L., Nilsson, L.-G., & Chalom, D. (1986). New evidence on the nature of the encoding of action events. *Memory & Cognition, 14*(4), 339–346.

Bird, G., Osman, M., Saggerson, A., & Heyes, C. (2005). Sequence learning by action, observation and action observation. *British Journal of Psychology, 96*(3), 371–388. http://doi.org/10.1348/000712605X47440

Cohen. (1977). *Statistical power analysis for the behavioral sciences* (Vol. Rev). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Cohen, R. L. (1983). The effect of encoding variables on the free recall of words and action events. *Memory & Cognition, 11*(6), 575–582.

Cohen, R. L., Peterson, M., & Mantini-Atkinson, T. (1987). Interevent differences in event memory: Why are some events more recallable than others? *Memory & Cognition, 15*(2), 109–118.

Dick, M. B., Kean, M.-L., & Sands, D. (1989). Memory for action events in Alzheimer-type dementia: Further evidence of an encoding failure. *Brain and Cognition, 9*(1), 71–87.

Dijkstra, K., MacMahon, C., & Misirlisoy, M. (2008). The effects of golf expertise and presentation modality on memory for golf and everyday items. *Acta Psychologica, 128*(2), 298–303. http://doi.org/10.1016/j.actpsy.2008.03.001

Earles, J. L. (1996). Adult age differences in recall of performed and nonperformed items. *Psychology and Aging, 11*(4), 638.

Einstein, G. O., & Reed, R. (1980). Levels of processing and organization: Additive effects of individual-item and relational processing. *Journal of Experimental Psychology: Human Learning and Memory, 6*(5), 588–598. http://doi.org/10.1037/0278-7393.6.5.588

Engelkamp, J. (1998). *Memory for actions*. Hove, U.K.: Psychology Press.

Engelkamp, J., & Dehn, D. M. (2000). Item and order information in subject-performed tasks and experimenter-performed tasks. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(3), 671–682. http://doi.org/10.1037//0278-7393.26.3.671

Engelkamp, J., Jahn, P., & Seiler, K. H. (2003). The item-order hypothesis reconsidered: The role of order information in free recall. *Psychological Research, 67*(4), 280–290. http://doi.org/10.1007/s00426-002-0118-1

Engelkamp, J., & Krumnacker, H. (1980). Imaginale und motorische Prozesse beim Behalten verbalen Materials. *Zeitschrift Für Experimentelle Und Angewandte Psychologie, 27*, 511–533.

Engelkamp, J., Seiler, K. H., & Zimmer, H. D. (2004). Memory for actions: Item and relational information in categorized lists. *Psychological Research Psychologische Forschung, 69*(1-2), 1–10. http://doi.org/10.1007/s00426-003-0160-7

Engelkamp, J., & Zimmer, H. D. (1983). Zum Einfluss von Wahrnehmen und Tun auf das Behalten von Verb-Objekt-Phrasen. *Sprache & Kognition, 2*(2), 117–127.

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behav Res Methods, 39*(2), 175–91.

Feyereisen, P. (2006). Further investigation on the mnemonic effect of gestures: Their meaning matters. European *Journal of Cognitive Psychology, 18*(2), 185–205. http://doi.org/10.1080/09541440540000158

Foley, M. A., Bouffard, V., Raag, T., & DiSanto-Rose, M. (1991). The effects of enactive encoding, type of movement, and imagined perspective on memory of dance. *Psychological Research, 53*(3), 251–259.

Gold, D. A., & Park, N. W. (2009). The effects of dividing attention on the encoding and performance of novel naturalistic actions. *Psychological Research Psychologische Forschung, 73*(3), 336–349. http://doi.org/10.1007/s00426-008-0148-4

Golly-Häring, C., & Engelkamp, J. (2003). Categorical-relational and order-relational information in memory for subject-performed and experimenter-performed actions. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 29*(5), 965–975.

Helstrup, T. (2005). In search of a motor element in memory for enacted events. *European Journal of Cognitive Psychology, 17*(3), 389–403.

Hornstein, S. L., & Mulligan, N. W. (2004). Memory for actions: Enactment and source memory. *Psychonomic Bulletin & Review, 11*(2), 367–372.

Hunt, R. R., & Mcdaniel, M. A. (1993). The Enigma of Organization and Distinctiveness. *Journal of Memory and Language, 32*(4), 421–445. http://doi.org/10.1006/jmla.1993.1023

Knopf, M., & Neidhardt, E. (1989). Aging and memory for action events: The role of familiarity. *Developmental Psychology, 25*(5), 780.

Lichtenstein, E. H., & Brewer, W. F. (1980). Memory for goal-directed events. *Cognitive Psychology, 12*(3), 412–445. http://doi.org/10.1016/0010-0285(80)90015-8

Maisto, A. A., & Queen, D. E. (1992). Memory for Pictorial Information and the Picture Superiority Effect. *Educational Gerontology, 18*(2), 213–223. http://doi.org/10.1080/0360127920180207

Manzi, A., & Nigro, G. (2008). Long-term memory for performed and observed actions: retrieval awareness and source monitoring. *Memory (Hove, England), 16*(6), 595–603. http://doi.org/10.1080/09658210802070749

Mulligan, N. W., & Hornstein, S. L. (2003). Memory for actions: Self-performed tasks and the reenactment effect. *Memory & Cognition, 31*(3), 412–421.

Nilsson, L.-G., & Cohen, R. L. (1988). Enrichment and generation in the recall of enacted and non-enacted instructions. In M. M. Gruneberg, P. E. Morris, & R. N. Sykes (Eds.), *Practical aspects of memory: Current research and issues* (pp. 427–432). Chinchester: John Wiley & Sons.

Paivio, A. (1986). *Mental Representation: A dual coding approach*. Oxford, England: Oxford University Press.

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. http://doi.org/10.1371/journal.pone.0099985

Steffens, M. C. (2007). Memory for goal-directed sequences of actions: Is doing better than seeing? *Psychonomic Bulletin & Review, 14*(6), 1194–1198.

Steffens, M. C., von Stülpnagel, R., & Schult, J. (in press). Memory recall after 'learning by doing' and 'learning by viewing': A review of the evidence. *Frontiers in Psychology*. doi:10.3389/fpsyg.2015.01907

Stull, A. T., & Mayer, R. E. (2007). Learning by doing versus learning by viewing: Three experimental comparisons of learner-generated versus author-provided graphic organizers. *Journal of Educational Psychology, 99*(4), 808–820. http://doi.org/10.1037/0022-0663.99.4.808

von Stülpnagel, R., Schult, J. C., Richter, C., & Steffens, M. C. (2015). Cognitive Costs of Encoding Novel Natural Activities: Can 'Learning by Doing' Be Distracting and Deceptive? *The Quarterly Journal of Experimental Psychology*, 1–19. (Advance online publication) http://doi.org/10.1080/17470218.2015.1087581

Zimmer, H., & Engelkamp, J. (1984). Planungs- und Ausführungsanteile motorischer Gedächtniskomponenten und ihre Wirkung auf das Behalten ihrer verbalen Bezeichnungen. [Planning and execution elements of motor memory components and their effect on the retention of their verbal descriptions.]. *Zeitschrift Für Psychologie Mit Zeitschrift Für Angewandte Psychologie, 192*(4), 379–402.

*Received: 9.20.2015*
*Revised: 11.25.2015*
*Accepted: 11.26.2015*