The Enactment Effect and Judgments of Learning:

Action Memory Meets Metamemory

Andy Jin

University of North Carolina at Chapel Hill

Spring 2015

A thesis submitted to the faculty at The University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the Bachelor of Science degree with Honors in Psychology.

Advisor _______________________
Dr. Neil Mulligan

Reader _______________________
Dr. Kelly Giovanello

Reader _______________________
Jonathan Susser
Acknowledgments

I would like to express my gratitude toward my advisor Dr. Neil Mulligan for all of his support and guidance throughout this project and my undergraduate career. I thank Dr. Kelly Giovanello for taking the time to participate in my defense. Finally, I am also sincerely grateful to Jonathan Susser for all of his continuous advice and assistance.
Abstract

Research on metamemory has largely used only verbal materials. However, analogous to the enactment effect in memory, metamemory accuracy has been shown to be different for subject-performed tasks (SPTs), experimenter-performed tasks (EPTs), and verbal tasks (VTs). Prior literature suggests that encoding differences create a strategic versus nonstrategic distinction that influences both memory and metamemory. The present study uses a between-subjects encoding manipulation (SPT vs. EPT vs. VT) to examine differences in judgments of learning (JOLs), memory performance, and metamemory accuracy. JOLs and recall percentage did not vary across condition, a result that is not unexpected due to the fact that such effects are commonly observed within-subjects. Relative accuracy, as measured by gamma correlations, were not significantly different from zero and also did not vary across conditions. Various explanations for this result are discussed, emphasizing methodological and statistical differences between the current study and prior research. Future implications for the study of metamemory and action memory are considered.
The Enactment Effect and Judgments of Learning: Action Memory Meets Metamemory

Metamemory refers to people’s beliefs and predictions about memory, and encompasses elements of monitoring and control. Whereas monitoring involves assessing one’s memory, control involves using those assessments to regulate cognitive activity. Investigating how people monitor and predict their memory is important because people often use these predictions to guide how they allocate cognitive resources for studying. Making effective study choices is crucial to successful learning. For example, consider a student studying for an exam. A student would likely study certain material until reaching a self-determined level of acceptable understanding (Kornell & Metcalfe, 2006). Gauging how well material is learned may help the student decide whether to continue studying a certain section or move on to new information.

Research on memory monitoring is often done by having subjects provide metacognitive judgments of memory strength, or judgments of learning (JOLs). JOLs are predictions of future memory performance and reflect the level of confidence in one’s memory. Typical metamemory experiments use the traditional study-test paradigm. A subject is presented with material such as a word list, and studies the material for a later memory test. During the study phase, the subject also gives JOLs about the probability of remembering the material on the later memory test. After a distractor task, the subject’s memory is tested using methods such as a recognition test or free-recall test. Researchers are often interested in the accuracy of JOLs and whether JOLs predict actual memory performance. JOL accuracy is important because it can determine whether future study decisions are adaptive or maladaptive (Kornell & Metcalfe, 2006).

There are two kinds of accuracy to consider when comparing predicted memory performance to actual memory performance. Relative accuracy is the degree to which one’s
JOLs predict the likelihood of correct performance on one item relative to another (Dunlosky & Metcalfe, 2009). High relative accuracy would mean that the items given high JOLs were actually better remembered than the items given low JOLs. On the other hand, absolute accuracy is the degree to which JOLs correspond to the specific level of performance. High absolute accuracy would mean that the overall magnitude of JOLs matched the percentage of items actually remembered. Whereas relative accuracy indicates whether a person can discriminate between differences in the memorability of items, absolute accuracy indicates whether a person can estimate the actual level of test performance (Dunlosky & Metcalfe, 2009). To assess absolute accuracy, judgments must be made on a scale that is directly comparable to the scale used for test performance (e.g., percentage predicted recall and percentage actual recall) (Dunlosky & Metcalfe, 2009). The majority of the metamemory literature, especially early research, has focused on relative accuracy measurements.

**Experiments in Metamemory**

Early studies on metamemory yielded varied results on the accuracy of JOLs. Lovelace (1984) tested subjects’ memories for words and sentences, and subjects predicted their memory performance using a 1-5 Likert scale. Subjects showed high predictive accuracy for words and sentences, with the probability of recall matching predictions (Lovelace, 1984). On the other hand, other studies have shown that memory predictions can be quite inaccurate. Vesonder and Voss (1985) used a multiple-trial design where subjects had to recall the same items across several trials. Subjects predicted whether they would remember an item for the following trial using a “yes”/“no” response. For the first trial, subjects showed low predictive accuracy for items. In subsequent trials, subjects showed high predictive accuracy for previously recalled items, but low predictive accuracy for items that were not previously recalled. When predictive
accuracy was low, subjects exhibited an overconfidence bias by predicting correct recall more frequently than actually achieving correct recall (Vesonder & Voss, 1985).

Beyond simply observing and measuring metamemory, researchers have manipulated certain variables to influence JOLs, actual memory performance, or both. Prior literature has shown that JOLs can be manipulated independent of actual memory performance. By changing the instructions for when subjects should give their JOLs, Nelson and Dunlosky (1991) improved metamemory accuracy without affecting actual memory performance. Instead of having subjects provide JOLs immediately after the presentation of each item, subjects provided JOLs for all of the items in a block after the presentation of the entire block. These delayed JOLs were far more accurate than immediate JOLs, but the levels of actual recall were comparable across the two conditions. This manipulation showed that delayed JOLs have higher relative and absolute accuracy compared to immediate JOLs.

Alternatively, metamemory accuracy can be selectively influenced without affecting actual memory. Rhodes and Castel (2008) manipulated font size and predicted that words presented in a large font are more easily perceived than words presented in a small font. This perceptual fluency manipulation was designed to create a metamemory illusion such that more easily perceived words would be given higher JOLs. Indeed, subjects gave higher JOLs to larger words than smaller words, showing that perceptual fluency was used as a cue of memorability, even though actual recall was the same for the two fonts (Rhodes & Castel, 2008).

Finally, other studies have concurrently manipulated metamemory and actual memory performance (e.g., Mueller, Tauber, & Dunlosky, 2013). Laboratory experiments often use paired-associate learning to study memory. For this task, subjects studied pairs of words for a
later memory test. During the test phase, subjects were cued with the first word and had to recall the second word. Mueller et al. (2013) manipulated the relatedness of the word pairs such that some pairs were related (e.g., injury-hurt) and others were unrelated (e.g., cellar-hurt). Both recall performance and JOLs were significantly greater for related pairs than for unrelated pairs, showing that certain manipulations can improve both metamemory accuracy and actual memory performance.

Overall, metamemory research has elucidated circumstances where memory predictions are accurate and other situations where predictions are not accurate. One shortcoming of metamemory research, however, is that the large majority of studies have used only verbal materials such as words and sentences. This is somewhat unexpected, as memory for actions is a fairly well studied body of literature.

**Memory for Actions**

Although memory research is dominated by the use of verbal materials, memory for actions has received considerable attention. From a practical perspective, studying memory for actions is important because such memories constitute a large portion of everyday uses of memory. For example, one must be able to remember tasks that have already been completed: eating breakfast, filling the car up with gas, etc. Remembering these items can generally be done with ease and accuracy. From a theoretical standpoint, memory for actions has generated noticeable interest due to the enactment effect. The enactment effect is defined by improved memory for *self-performed tasks* (SPTs) compared to memory for *experimenter-performed tasks* (EPTs) or *verbal tasks* (VTs). Subjects generally have better memory for a list of action phrases after acting out list items compared to reading the items or watching someone act out the items.
A typical research paradigm investigating action memory includes the following characteristics. During the study phase, subjects are presented with a list of action phrases (e.g., *drop the pencil*) to remember for a later memory test. Subjects either listen to the phrases (VTs), watch an experimenter act out the phrases (EPTs), or act out the phrases themselves (SPTs). This encoding manipulation can vary between-subjects or within-subjects, but an enactment effect is traditionally found only with the within-subjects manipulation (McDaniel & Bugg, 2008). Following a distractor task, memory is tested through various methods such as a recognition test or free-recall test. Numerous studies have observed the superiority of SPTs, or enacted items, over EPTs and VTs (see Engelkamp, 1998 for an overview).

Research on the enactment effect has emphasized the motoric component as being critical for producing the benefit, not the visual or tactile feedback (Mulligan, 2013). Kormi-Nouri (2000; reported in Mulligan, 2013) reported that blind subjects demonstrate an enactment effect, suggesting that visual feedback is not critical for the enactment effect. In a similar way, Engelkamp and Zimmer (1997) argued against the role of tactile feedback by demonstrating that the enactment effect is equally strong for SPTs that use real objects and SPTs that do not. Furthermore, the enactment effect is a robust effect that is found under many different conditions (e.g., phrases vs. whole sentences, body-related items vs. object-related items, short vs. long lists; Engelkamp, 1998). The effect is also generalizable across adult and child populations, including individuals with memory disorders such as Alzheimer’s disease (Mulligan, 2013).

The robustness and generalizability of the enactment effect has led some researchers to suggest a qualitative difference between memory for actions and memory for verbal material. Cohen (1981) argued that action memory is fundamentally distinct from verbal memory, citing that SPTs are not subject to the same memory phenomena as VTs. Results from Cohen (1981)
suggested that memory for SPTs does not show a primacy effect or a typical levels-of-processing effect. However, later research demonstrated at least some effect of levels-of-processing for enacted items; furthermore, the distinction between memory for SPTs and memory for VTs fades under deep encoding (see Mulligan, 2013 for an overview). Yet other research has highlighted meaningful similarities between memory for actions and memory for verbal materials. Memory for both materials are subject to similar developmental and age effects. Also similar to verbal memory, memory for enacted items suffers under divided attention. Overall, the pattern of differences does not completely support that actions are governed by different memory principles than those commonly applied to verbal materials (Mulligan, 2013).

The Current Study

Although memory for actions is not a qualitatively different type of memory, the superiority of memory for enacted events is still a robust and widely found effect. Given the prominence of the enactment effect in memory, one may wonder whether enactment also influences metamemory. Previous research has shown that metamemory can be influenced by physical characteristics such as weight and motoric fluency (Alban & Kelley, 2013; Susser & Mulligan, 2014). Similarly, the process of physically enacting phrases may influence metamemory.

The intersection of metamemory and action memory has not been well studied because existing research on metamemory has largely used verbal materials. Given the differences between memory for actions and memory for verbal materials, it would be interesting to examine memory monitoring for actions and verbal materials. Previous research on metamemory for actions has suggested some preliminary differences. Cohen (1983) asked subjects to predict
memory for events, though he did not specifically use the term “judgments of learning”.

Subjects used a 4-point Likert scale to rate how likely they would recall each event on a future recall test. Ratings for events presented as either auditory or visual words showed good predictive power, but ratings for SPTs were non-predictive (Cohen, 1983).

Certain researchers have focused on the distinction between strategic and nonstrategic encoding to explain why metamemory for actions is less accurate than metamemory for verbal material. Some prior research has suggested that whereas SPTs are encoded nonstrategically, VTs are encoded strategically (e.g., Cohen, 1983). As mentioned before, SPT recall does not show the typical primacy effect seen in VT recall. This pattern suggests that actions may be encoded nonstrategically, automatically, or without controlled rehearsal. With SPTs, subjects may simply enact the action phrases and assume that encoding is complete. In contrast, VTs have been shown to be encoded strategically with rehearsal (e.g., Cohen, 1983). The presence or absence of controlled rehearsal is relevant for metamemory because learners are sensitive to how effectively they have engaged in controlled rehearsal (Dunlosky & Metcalfe, 2009). This sensitivity to rehearsal may give learners an insight into their memory for verbal materials, but would not be present in assessing memory for actions. Thus, enactment may lead SPTs to produce worse metamemory accuracy compared to EPTs and VTs.

The present study seeks to replicate the existing results of metamemory for actions, but with stricter methodological rigor. Several previous studies on metamemory and actions lack an EPT condition (e.g., Cohen, 1988; Cohen & Bryant, 1991; Cohen, Sandler, & Keglevich, 1991; McDonald-Misczczak, Hubley, & Hultsch, 1996). Though VTs are used as a control condition to compare against SPTs, verbal tasks alone are not a sufficient control. EPTs more closely mirror SPTs than VTs in that EPTs also feature a visual component, thereby better isolating the motoric
element of performing a task. The current experiment uses both VTs and EPTs as control groups to compare against the SPTs experimental group.

Also, much of the existing literature used Likert scales to measure JOLs, and different studies used arbitrarily different scales (e.g., Cohen, 1983; Cohen & Bryant, 1991). The varying use of JOL scales obscures a clear interpretation of the results. The use of Likert scales also makes it impossible to interpret absolute accuracy, because there is no way to objectively correspond JOL ratings to percentages of recall. The current experiment uses a 0-100 scale to report JOLs, which is more consistent with current research on JOLs (e.g., Alban & Kelley, 2013). This 100-point scale allows for a fine-grained analysis of predictive accuracy, especially compared to studies such as Cohen (1983) that used a two-point scale for analyzing predictive accuracy. Using a JOL scale that is directly comparable to the scale used for test performance will also allow for greater statistical interpretation of the results, namely absolute accuracy.

Finally, several other methodological shortcomings exist in previous research. Cohen (1983) did not maintain a constant presentation time across different types of events. SPTs were presented at a rate of eight seconds per event, EPTs were presented at a rate of six seconds per event, and VTs were presented at four seconds per event (Cohen, 1983). Varying presentation time creates different amounts encoding between the three conditions, which can affect memory performance and metamemory judgments. Cohen (1983) also did not clearly report statistics about recall performance, further limiting the interpretation of results.

In the present experiment, a between-subjects encoding manipulation is used to examine memory monitoring for SPTs, EPTs, and VTs. Because encoding condition is being manipulated between-subjects, the conditions may not show statistically significant differences in percent
recall or JOL ratings. As mentioned above, the enactment affect is usually found in mixed-list designs and if often not seen in pure lists (McDaniel & Bugg, 2008). Manipulations that influence JOLs are also largely seen in mixed-list designs (Susser, Mulligan, & Besken, 2013). The main focus of this experiment is the metamemory accuracy of JOL ratings. Subjects in the SPT condition are predicted to have less, or no, relative accuracy compared to subjects in the VT condition. All conditions are expected to show poor absolute accuracy due to overconfidence. Overall, this experiment seeks to replicate previous results with greater statistical and methodological thoroughness to allow for a clearer understanding of whether JOLs are less accurate for SPTs than control conditions (i.e., EPTs and VTs).

Method

Participants

Sixty undergraduates from the University of North Carolina at Chapel Hill participated in this experiment for course credit.

Materials and Design

The study list consisted of 32 critical action phrases involving objects, adapted from Peterson and Mulligan (2010) (see the Appendix). Each action consisted of a short phrase involving a verb and an object (e.g., drop the pencil). Two additional items were placed at the beginning and end of the study list to serve as primacy and recency buffers. The 36 action phrases were sequenced in the same fixed random order for all conditions, and all of the 36 actions and 36 verbs were unique. All action phrases were recorded and presented to the subjects over computer speakers.
The type of encoding was manipulated between subjects with three conditions (SPT vs. EPT vs. VT). In the SPT condition, subjects performed each action phrase with the relevant object. In the EPT condition, subjects watched the experimenter perform each action phase with the relevant object. In the VT condition, subjects listened to each action phrase while looking at the relevant object.

**Procedure**

The experiment used a traditional study-test paradigm. During the study phase, subjects were instructed to study action phrases for a later memory test. The particular class of event (SPT, EPT, or VT) that was used for the testing session was also described. In the SPT condition, subjects sat at a table, half of which was screened from view to conceal the objects required for performing the actions phrases. At the beginning of a trial, the experimenter removed the relevant object from behind the screen and placed it on the table. An action phrase was then presented over the computer speakers (e.g., *drop the pencil*). Following the phrase, subjects were given six seconds to carry out the action with the relevant object. Immediately following the completion of the action phrase, the experimenter retrieved the object and placed it behind the screen. Following each item, subjects were asked to rate how confident they were in their ability to remember the phrase on a later memory rest. JOL ratings were made on a scale from 0 (not confident at all) to 100 (extremely confident). Subjects were encouraged to use the entire scale, and orally reported their JOLs to the experimenter.

The EPT condition involved a similar manner of presentation. However, instead of the subject enacting the action phrases, the subjects watched as the experimenter had six seconds to
enact the phrases. The VT condition was also similar, but the subjects only listened to the action phrases and looked at the relevant object for six seconds.

After the study phase, subjects completed a three-minute distractor task consisting of mental arithmetic problems. This was followed by a free recall test in which subjects were asked to write down as many action phrases as they could remember from the first part of the experiment. Subjects were instructed to recall the phrases in any order and to try to retrieve the entire action phrase. Subjects were also told that if they could not recall the entire phrase, they should write down as much of the phrase as they could remember. A maximum of five minutes was allotted for the test phase, and subjects who attempted to complete the task before the two minute mark were instructed to keep trying.

Results

Recall performance was scored in two ways. For stringent scoring, both the noun and verb of the action phrase must have been recalled together to count as correct (e.g. *drop* and *pencil* for *drop the pencil*). For lenient scoring, either the noun or verb of the phrase must have been recalled to count the phrase as having been recalled. All analyses were conducted with both types of scores and yielded the same pattern of results. Thus, only the lenient recall scores are reported.

A one-way ANOVA was conducted to examine percent recall performance. There was no significant effect of encoding condition on recall performance, \( F(2, 57) = 0.58, p = .565 \), such that SPTs \((M = 53.44, SD = 9.23)\), EPTs \((M = 50.00, SD = 12.21)\), and VTs \((M = 50.47, SD = 11.26)\) exhibited similar levels of memory performance. This suggests that recall performance was not influenced by encoding condition, and that no enactment effect occurred.
Another one-way ANOVA was conducted to examine the effect of encoding condition on JOLs, and did not reveal a significant effect, $F(2, 57) = 0.24, p = .786$. Subjects reported similar JOLs for the SPT ($M = 72.31, SD = 17.94$), EPT ($M = 71.50, SD = 22.36$), and VT ($M = 68.32, SD = 16.76$) conditions. As with recall performance, JOL ratings also did not vary by condition.

The relative accuracy of JOLs was assessed by calculating within-subject gamma correlations between JOL ratings and recall percentage. Gamma correlations are the standard measure for determining the association between JOLs and recall, and reflect the accuracy of the subject’s JOLs in predicting actual memory performance (Nelson, 1984). A one-way ANOVA was conducted and did not reveal a significant effect of encoding condition on JOL accuracy, $F(2, 57) = 1.68, p = .196$. Relative accuracy was similar for SPTs ($M = .032, SD = .285$), EPTs ($M = .029, SD = .277$), and VTs ($M = -.106, SD = .251$). Subjects showed similar levels of predictive accuracy, regardless of encoding condition. Separate $t$-tests were conducted to assess whether the gamma correlations were significantly different from zero, and revealed that no condition had significant relative accuracy: SPT [$t(19) = 0.50, p = .625$], EPT [$t(19) = 0.47, p = .644$], VT [$t(19) = -1.89, p = .075$].

The absolute accuracy of JOLs was examined with a 2 (measure: JOL vs. Recall performance) x 3 (encoding condition: SPT vs. EPT vs. VT) mixed ANOVA. There was no main effect of encoding condition, $F(2, 57) = 0.43, p = .650$. There was a main effect of measure, $F(1, 57) = 55.36, p < .001$, such that subjects’ JOLs ($M = 70.71, SD = 18.93$) were significantly higher than their recall percentage ($M = 51.30, SD = 10.90$). Furthermore, this overconfidence effect did not vary by condition, $F(2, 57) = 0.17, p = .841$, suggesting that subjects showed similar amounts of overconfidence in each of the three conditions.
Discussion

The present study used a between-subjects encoding manipulation to examine metamemory accuracy for actions and verbal material. JOL ratings and recall percentage were not significantly different across the three conditions: SPT, EPT, and VT. Regardless of encoding condition, subjects showed similar metamemory predictions and similar levels of memory performance. This result is not unexpected, as these differences are traditionally seen when encoding is manipulated within-subjects. The data are consistent with prior research showing that the enactment effect and influences on JOLs are usually found in mixed-list designs (e.g., McDaniel & Bugg, 2008; Susser et al., 2013). Absolute accuracy was examined and revealed a consistent overconfidence effect across all conditions. Regardless of encoding condition, subjects tended to give JOLs higher than their respective percentage of actual recall. This result is consistent with the large amount of literature showing an overconfidence bias in metamemory judgments (for a review see Dunlosky & Metcalfe, 2009).

Relative accuracy, as measured by gamma correlations, was not significantly different across the three conditions. Furthermore, none of the conditions showed a relative accuracy significantly different from zero. Subjects across all three conditions were generally unable to predict which action phrases they would remember and which they would not. The fact that the SPT condition showed poor relative accuracy is consistent with a large portion of existing literature (e.g., Cohen, 1983). On the other hand, the poor relative accuracy in the EPT and VT conditions is a more unexpected result. The majority of the following discussion will be directed towards interpreting this result and its implications.
One factor that can partially explain this discrepancy in results is the difference in study materials used. The majority of previous studies used different materials for the verbal (VT, and where appropriate, EPT) and SPT conditions (e.g., Cohen, 1983; Cohen, 1988; Cohen et al., 1991). In many of these studies, subjects studied a list of actions for the SPT condition, and then studied a separate list of words for the verbal conditions. Thus, the pattern of relative accuracy differences found by prior studies may not be due to enactment, but could be based on the nature of the study materials (sentences vs. words). This disparity in study materials could impact relative accuracy in a number of ways.

Single words may allow for more flexible encoding because individuals can choose to encode items in different ways (McDonald-Miszczak et al., 1996). Greater emphasis is placed on active rehearsal and controlled memorization strategies, which allows for more metamemorial insight. However, with sentences, controlled processes may play less of a role because the individual is focused on sentence comprehension, sentence structure, etc. This would allow for potentially less metamemorial insight and result in lower accuracy of metacognitive judgments.

Other possible ways that sentences and words could differently affect gammas are suggested by various theories about the nature of JOLs. According to the direct-access hypothesis, individuals directly monitor memory trace strength and use this information to make metacognitive judgments (King, Zechmeister, & Shaughnessy, 1980; reported in Koriat, 1997). Under this hypothesis, the ability to directly monitor trace strength could differ for single words compared to whole sentences, leading to differing levels of JOL accuracy. Alternatively, the cue-utilization view posits that individuals use different types of cues (e.g., internal mnemonic indicators, characteristics of the study item) to predict subsequent memory performance (Koriat, 1997). Individuals use these cues to diagnose future memory performance and form JOLs.
Applying these cues and heuristics may be easier to do when studying words compared to sentences because it is easier to make these sort of judgments for single words than groups of words. The cues themselves may also suggest that single words will be more likely remembered than groups of words.

Besides differences in study material, other factors may help account for the varied results between the current study and previous research. The method for scoring free recall has been inconsistent throughout the literature. The present study used the strict and lenient scoring method described above, which is the current standard practice. Several previous studies did not specify how free recall was scored (e.g., Cohen, 1988; Cohen et al., 1991; McDonald-Miszczak et al., 1996). Other past studies used a very lenient scoring criterion; a recall response was scored as correct if it enabled the unequivocal identification of a list item (Cohen, 1983; Cohen & Bryant, 1991). This difference in scoring practices could greatly impact the reported recall percentages, which would subsequently influence gamma correlations.

One final factor that may have influenced the differing results is JOL variability. The JOLs obtained in this study are of relatively greater magnitude compared to some other studies that assessed JOLs for action phrases (e.g., Koriat, Sheffer, & Ma’ayan, 2002). This increase in magnitude toward ceiling could contribute to a decrease in the variability of JOLs, subsequently decreasing the quality of gamma correlations. Gamma correlations cannot accurately capture the relationship between JOL ratings and recall percentage if too many JOLs are identical. This line of thinking is reflected in Cohen (1983), where several subjects are removed for giving unrealistically high or unrealistically low metacognitive judgments.
The three factors outlined above could account for the differing data patterns of the current study and prior literature: varied study material for VTs, inconsistent free recall scoring, and decreased JOL variability. These key differences can help explain why the present study did not show high relative accuracy for VTs. Beyond interpreting the current results, these factors also highlight future directions for research.

The present study was designed to replicate Cohen (1983) with more rigorous methodology. A logical follow-up experiment would have a similar design but use a within subjects manipulation. Manipulating encoding within-subjects would be more likely to produce an enactment effect and an effect on JOLs. It would then be interesting to see if a pattern similar to the one found in this study would be maintained after the dependent measures are above floor. That is, would SPTs and VTs continue to show similar levels of relative accuracy? Or would VTs begin to show greater relative accuracy as predicted by certain theories? Once a clear pattern is established, research can elucidate the underlying mechanism by manipulating SPTs to increase the relative accuracy of their JOLs. For example, if metacognitive sensitivity to controlled rehearsal is fundamental to accurate JOLs, then it should be possible to induce rehearsal of SPTs and measure changes in relative accuracy.

Other future directions could seek to clarify the circumstances under which memory and metamemory are different for actions and verbal material, and how patterns differ under changing factors such as age and characteristics of the study material. The present study used a between-subjects manipulation of enactment to examine the metamemory accuracy of SPTs, EPTs, and VTs. By improving statistical and methodological practices, this study provides new information relevant to the discussion of memory for actions versus memory for verbal material, and whether such a distinction extends to metamemory.
References


http://dx.doi.org.libproxy.lib.unc.edu/10.1016/0749-596X(85)90034-8
### Appendix

**Actions Used in the Experiment**

<table>
<thead>
<tr>
<th>Action Used in the Experiment</th>
<th>Action Used in the Experiment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Crumble the plastic bag</td>
<td>Roll the dice</td>
</tr>
<tr>
<td>Turn over the clipboard</td>
<td>Play with the ball and paddle</td>
</tr>
<tr>
<td>Dial your telephone number</td>
<td>Wipe the plate</td>
</tr>
<tr>
<td>Place the marker upright</td>
<td>Twist off the lid</td>
</tr>
<tr>
<td>Toss the jack in the air</td>
<td>Shoot the toy gun</td>
</tr>
<tr>
<td>Press the stapler</td>
<td>Remove a piece of tape</td>
</tr>
<tr>
<td>Wave the handkerchief</td>
<td>Insert the thumb tack</td>
</tr>
<tr>
<td>Erase the mark</td>
<td>Balance the man on the skateboard</td>
</tr>
<tr>
<td>Switch the light on and off</td>
<td>Slide the quarter into the piggy bank</td>
</tr>
<tr>
<td>Lift the paperclip</td>
<td>Write your first name</td>
</tr>
<tr>
<td>Pick up the napkin</td>
<td>Bounce the ball</td>
</tr>
<tr>
<td>Pull off a post-it</td>
<td>Flip the coin</td>
</tr>
<tr>
<td>Take the cap off the pen</td>
<td>Brush the dog</td>
</tr>
<tr>
<td>Push the car</td>
<td>Shuffle the cards</td>
</tr>
<tr>
<td>Stack the pieces</td>
<td>Drop the pencil</td>
</tr>
<tr>
<td>Fold the index card</td>
<td>Tear a ticket</td>
</tr>
<tr>
<td>Spray the bottle</td>
<td>Pour the water</td>
</tr>
<tr>
<td>Draw a circle</td>
<td></td>
</tr>
<tr>
<td>Squeeze the hole punch</td>
<td></td>
</tr>
</tbody>
</table>