THE QUARTERLY JOURNAL OF EXPERIMENTAL PSYCHOLOGY, 2013
Vol. 66, No. 10, 2060–2072, http://dx.doi.org/10.1080/17470218.2013.777751

Routledge
Taylor & Francis Group

Separation of encoding fluency and item difficulty effects on judgements of learning

Monika Undorf and Edgar Erdfelder

Department of Psychology, School of Social Sciences, University of Mannheim, Mannheim, Germany

The fluency of information encoding has frequently been discussed as a major determinant of predicted memory performance indicated by judgements of learning (JOLs). Previous studies established encoding fluency effects on JOLs. However, it is largely unknown whether fluency takes effect above and beyond the effects of item difficulty. We therefore tested whether encoding fluency still affects JOLs when numerous additional cues indicating the difficulty of an item are available as well. In three experiments, participants made JOLs for another participant while observing his or her self-paced study phase. However, study times were swapped in one experimental condition, so that items with short study times (indicating high fluency) were presented for long durations, whereas items with long study times (indicating low fluency) were presented for short durations. Results showed that both item difficulty and encoding fluency affected JOLs. Thus, encoding fluency in itself is indeed an important cue for JOLs that does not become redundant when difficulty information is available in addition. This observation lends considerable support to the ease-of-processing hypothesis.

Keywords: Judgements of learning; Encoding fluency; Ease of processing; Metacognitive heuristics; Metamemory.

Successful learning critically depends on people's ability to monitor and control their cognitive processes (e.g., Dunlosky & Metcalfe, 2008; Kornell & Bjork, 2007; Sitzmann & Ely, 2011). The impact of metacognition on self-regulated learning can, for example, be seen from the fact that judgements of learning (JOLs) are systematically related to learning behaviour. JOLs are numerical predictions of the likelihood that recently studied information will be retrieved in future. Both people's choices concerning the relearning of to-be-studied information and people's allocation of study time to different pieces of information are known to depend on metacognitive monitoring as reflected in JOLs (Metcalfe & Finn, 2008; Son & Metcalfe, 2000).

JOLs, in turn, are generally agreed to be inferences based upon a variety of cues and heuristics (e.g., Koriat, 1997; Schwartz, 1994). Among the heuristics that have attracted particular attention in recent years is the ease-of-processing hypothesis (Dunlosky & Metcalfe, 2008).[1] According to the ease-of-processing hypothesis, JOLs are governed by the ease with which items are committed to

Correspondence should be addressed to Monika Undorf, Department of Psychology, School of Social Sciences, University of Mannheim, Schloss, Ehrenhof-Ost, 68131 Mannheim, Germany. E-mail: undorf@uni-mannheim.de

Experiment 1 is based in part on Monika Undorf's doctoral thesis.

[1] We stick to a narrow definition proposed by Dunlosky and Metcalfe (2008), according to which the ease-of-processing hypothesis encompasses the influence of encoding fluency on JOLs exclusively. The ease-of-processing hypothesis more broadly defined (e.g., Kornell, Rhodes, Castel, & Tauber, 2011) also includes influences of retrieval fluency and perceptual fluency on JOLs (for a discussion of different aspects of processing fluency, see Undorf & Erdfelder, 2011).

memory during study—that is, by *encoding fluency* (e.g., Begg, Duft, Lalonde, Melnick, & Sanvito, 1989; Hertzog, Dunlosky, Robinson, & Kidder, 2003; Koriat & Ma'ayan, 2005; Kornell, Rhodes, Castel, & Tauber, 2011). Over the years, numerous findings have accumulated that are compatible with the ease-of-processing hypothesis (Begg et al., 1989; Hertzog et al., 2003; Koriat, 2008; Koriat, Ackerman, Lockl, & Schneider, 2009a, 2009b; Koriat, Ma'ayan, & Nussinson, 2006; Miele, Finn, & Molden, 2011). More specifically, it was shown that JOLs are positively correlated with measures of encoding fluency: JOLs are related to the speed with which an interactive image of paired associates is formed (Hertzog et al., 2003), to the amount of self-paced study time spent for studying word pairs (Koriat, 2008; Koriat et al., 2009b; Koriat & Ma'ayan, 2005; Koriat et al., 2006), and to the number of trials needed for a correct recall of word pairs (Koriat, 2008; Koriat et al., 2009a). These results are certainly in line with the ease-of-processing hypothesis, but they do not provide conclusive evidence for it. Instead, differences in item difficulty might have acted as a common cause for the obtained relationships between JOLs and encoding fluency in each of the studies (cf. Undorf & Erdfelder, 2011).

Recently, Undorf and Erdfelder (2011) suggested a method to assess encoding fluency effects in the absence of item difficulty information. This method is based on a variant of the learner–observer–judge method (Vesonder & Voss, 1985; for similar methods cf. Koriat & Ackerman, 2010; Matvey, Dunlosky, & Guttentag, 2001) in which participants observe the study phase of another participant. Participants were shown each item as long as a yoked participant had chosen to study it and made JOLs for the yoked participant. In this *observing-without-difficulty* task, JOLs could not rely on item difficulty, because the to-be-studied word pairs were concealed by strings of symbols. In line with previous results, we found that JOLs from the observing-without-difficulty task were negatively correlated with study time when participants themselves had studied and recalled word pairs before. In contrast to previous findings, however, this result not only is consistent with the ease-of-processing hypothesis, but also provides strong support for it, because the relationship between encoding fluency (indicated by short study times) and JOLs cannot possibly stem from a confound between encoding fluency and item difficulty.

Completely withholding information on item difficulty provides an excellent means of disentangling effects of self-paced study time from effects of item difficulty. However, it cannot address the question of whether study time is used as a cue for JOLs when other potentially useful cues reflecting differences in item difficulty are available as well. Unlike in the artificial observing-without-difficulty task, JOLs might reflect a variety of influences in study situations that are more typical for learning contexts in everyday life. Presumably, standard study situations provide even more cues than people can reasonably be expected to use (cf. Oppenheimer, 2008). Thus, Undorf and Erdfelder's (2011) finding does not necessarily imply that study time is used in typical study situations characterized by a variety of cues associated with memory performance. To resolve this issue, we conducted three experiments in which study time was not confounded with item difficulty, and, in addition, information on both study time and item difficulty was available to participants.

Specifically, we used an *observing-with-difficulty* task, in which participants were shown intact rather than concealed word pairs whose presentation times corresponded to the study times of a yoked participant. Resembling the observing-without-difficulty task, participants were instructed to indicate probable JOLs of the observed participant. At the same time, study times were swapped between some of the items for one group of participants, so that word pairs with below-median study times were presented for long durations whereas word pairs with above-median study times were presented for short durations (swapped study time condition). For the other group of participants, presentation times always corresponded to the actual study times of the yoked participant (original study time condition). JOLs could thus reflect both study duration and item difficulty (as indicated by word pair content). However, swapping study times in one of the two experimental conditions disentangled effects of study duration and item difficulty.

Making item difficulty information available to participants results in a multitude of cues on which JOLs could be based in principle, including word frequency, concreteness, imagery, and the associative relatedness between the members of paired associates, among others. In principle, all these cues could affect JOLs directly through the deliberate application of a theory or belief about memory—for example, the belief that related paired associates are more likely to be remembered than unrelated paired associates. They could, however, also influence JOLs indirectly, mediated by their effects on mnemonic cues such as encoding fluency (Koriat, 1997). Importantly, swapping study times eliminates confounds between presentation time and any other cue that may indicate item difficulty.

Earlier work has shown that study time is used as a cue for JOLs for other people only when participants had previously gained experiences with the learning situation (Koriat & Ackerman, 2010; Undorf & Erdfelder, 2011). Therefore, in the experiments reported here, all participants completed a study phase with immediate JOLs and a cued-recall test prior to the observing-with-difficulty task. However, whereas participants worked on the very same word pairs both in the learning task and in the observing-with-item-difficulty task of Experiment 1 and Experiment 3, participants worked on two different sets of word pairs in Experiment 2. Previous studies have shown that studying exactly the same set of items repeatedly may be crucial for JOL effects. For example, JOLs shift from overconfidence in a first study–test–trial to underconfidence in further trials only when exactly the same word list is repeatedly used (Koriat, 1997). Moreover, study–test practice has been shown to reduce metacognitive illusions for the list of items that was practised, but not for a similar new list (Koriat & Bjork, 2006). It seems thus important to investigate whether study time affects JOLs irrespective of whether the same or a different word list has been studied before.

The main predictions for the observing-with-difficulty task are identical for all three experiments. Because people spend more time studying more difficult items (e.g., Son & Metcalfe, 2000), study time is positively correlated with item difficulty in the original study condition. Correspondingly, study time is negatively correlated with item difficulty in the swapped study time condition. Study time thus acts in concert with further item difficulty cues in the original study time condition, whereas the two types of cues act in opposition in the swapped study time condition (for a similar dissociation logic, see Jacoby, 1991). Thus, if study time is used as a cue for JOLs in the presence of other cues indicating item difficulty, correlations between actual item difficulty and JOLs should be reduced in the swapped study time condition (where the two types of cues act in opposition) compared to the original study time condition (where they act in concert). If, however, correlations between actual item difficulty and JOLs do not differ between the two conditions, it follows that study time is not used as a cue for JOLs. Analogously, the pattern of correlations between study time and JOLs reveals whether JOLs are affected by item difficulty information over and above study time. If participants base their JOLs on study time exclusively, study time–JOL correlations should be generally negative and should not differ between conditions. In contrast, if item difficulty information other than study time is used to make JOLs, this will counteract the negative study time–JOL correlation in the swapped study time condition. Thus, the study time–JOL correlation should be less pronounced (i.e., less negative or even positive) in the swapped study time condition than in the original study condition.

EXPERIMENT 1

Method

Participants

Participants were 80 German-speaking psychology undergraduates at the University of Mannheim.

They were randomly assigned to either the original or the swapped study time condition ($n = 40$ each).

Design

We manipulated whether presentation times in the observing-with-difficulty task always corresponded to the actual study times of the yoked participant or were swapped between some of the items (original vs. swapped study time condition). Dependent variables were the within-subject gamma correlations between (a) observed study time and JOLs and (b) JOLs and item difficulty in the observing-with-difficulty task. For control purposes, we also analysed the within-subject gamma correlations between study time, recall, and JOLs in the learning task.

Materials

The study list consisted of 30 related and 30 unrelated German word pairs. The associative strengths of related word pairs ranged between .01 and .67 ($M = .15$, $SD = .16$) according to German word-association norms (Hager & Hasselhorn, 1994), whereas the words of the remaining pairs were unrelated both according to word-association norms and according to intuitive judgement. Three apparently related and three apparently unrelated buffer word pairs were added to the list. Stimuli were presented on 19-in. CRT monitors using E-Prime (Schneider, Eschman, & Zuccolotto, 2002).

Procedure

In the first part of the experiment, both groups were instructed to study 66 word pairs for a later memory test. Word pairs were displayed one by one in the centre of the screen. Participants were told that the first word of each pair would be presented in the test phase and that their task would then be to recall the second word. They were asked to choose their study times so that they could recall the second word in the test phase while spending as little study time as possible. Participants indicated by mouse click that they had finished studying a word pair. Immediately after studying each word pair, participants made a JOL for that item in which they estimated the probability of recalling it in the test phase on a percentage scale ranging from 0% to 100%. As soon as the JOL was made, the next item was presented for study. After all items were studied and judged, participants worked on an unrelated filler task for 90 seconds (easy mathematical problems) and then completed the test phase. In this phase, the first words of all 66 pairs were presented, and participants were asked to type in the second word or—when they could not recall it—“xxx”. Responses that were very similar to the target word (e.g., *arrow* instead of *arrows*) were scored as correct responses. The order of word pairs in the study and test phases was randomly determined for each participant. However, buffer word pairs were always presented first, and word pairs presented in the first half of the study phase were part of the first half of the test phase.

In the second part of the experiment, the same participants completed an observing-with-difficulty task. They were told that they would observe the study phase of another participant and were asked to indicate his or her probable JOLs. They were subsequently informed that each item would be presented to them as long as the other participant chose to study it and that they themselves were not supposed to study the words again. In the observing task, participants were presented with the same list of words as the list that they had previously processed in the learning task. The observing task was otherwise identical to the learning task except that participants did not complete a cued recall test and that they were not allowed to self-pace their study. Instead, with the exception of the word pairs with swapped study times (see below), each word pair was presented to participants for as long as one particular participant from an earlier experiment (Undorf, 2010, Experiments 3 and 4, total $N = 80$) had studied it under identical conditions. In order to ensure that swapping study times between items resulted in unusual combinations of study time and item difficulty, we selected the yoked participant and his or her study times according to the following criteria. First, individual study times for each item had to be highly correlated with mean study times, and, second, individual JOLs for each item also had to be highly correlated with mean JOLs. We additionally ensured that median study times were not above average to keep the duration of the experiment as short as possible. The study times of the selected participant ranged from 1.30 to 12.10 s ($Mdn = 5.17$) and correlated $r = .68$ with mean study times; the yoked participant's JOLs correlated $r = .79$ with mean JOLs for each item.

In the original study time condition, all 66 word pairs were presented as long as the yoked participant had chosen to study them. In contrast, study times for 30 of the 66 word pairs were swapped in the swapped study time condition. In order to ensure that participants in this condition experienced the yoked participant as sufficiently similar to themselves, both the 6 buffer items and the 10 word pairs with highest associative strengths kept their original study times. The remaining 50 word pairs were ranked from 1 to 50 according to the actual study times of the yoked participant. Then, the two word pairs with the most extreme study times (i.e., Word Pair 1 with the shortest study time and Word Pair 50 with the longest study time) were preselected for possible exchanges of their study times. The same was done for Word Pairs 2 and 49, 3 and 48, and so on, up to Word Pairs 20 and 31. The 40 word pairs that were preselected for swapping their study times consisted of two related word pairs in 10% of cases, two unrelated word pairs in 25% of cases, and one related and one unrelated word pair in the remaining 65% of cases. Importantly, not all of the corresponding study times were actually exchanged. Rather, for each participant, study times of a random sample of 30 of these 40 word pairs were swapped, whereas the other 10 pairs kept their original study times. The remaining Word Pairs 21 to 30 whose study times differed by less than 1 s also kept their original study times. In the swapped study time condition, therefore, original study times remained preserved for 36 of the 66 items.

Results and discussion

Learning task

Because the two experimental conditions did not differ in the learning task, we expected no group differences in this task. Within-subject gamma correlations were calculated using study times, JOLs, and cued recall performance for all word pairs in the list except for buffer items (i.e., 60 word pairs in total). The mean within-subjects gamma correlation between study time and cued recall performance was $G = -.16$ ($SD = .24$) in the original study time condition and $G = -.18$ ($SD = .26$) in the swapped study time condition. Both correlations were significantly smaller than zero: $t(38) = 4.13$, $p < .001$, $d = 0.66$,[2] for the original study time condition, and $t(39) = 4.32$, $p < .001$, $d = 0.68$, for the swapped study time condition. The same holds for the correlations between study time and JOLs, $G = -.30$ ($SD = .26$), $t(39) = 7.28$, $p < .001$, $d = 1.15$, and $G = -.30$ ($SD = .22$), $t(39) = 8.84$, $p < .001$, $d = 1.40$, for the original study time condition and the swapped study time condition, respectively. In contrast, the correlations between JOLs and recall performance were significantly positive both in the original study time condition, $G = .56$ ($SD = .26$), $t(38) = 13.57$, $p < .001$, $d = 2.17$, and in the swapped study time condition, $G = .54$ ($SD = .29$), $t(39) = 11.83$, $p < .001$, $d = 1.87$. We thus found that JOLs from the learning phase were predictive of memory performance and that both JOLs and recall performance were inversely related to study time. As expected, none of these gamma correlations differed between the two conditions: $t(77) = 0.35$, $p = .730$, $d = 0.08$, for the study time–cued recall correlation; $t(78) = 0.05$, $p = .962$, $d < 0.01$, for the study time–JOL correlation; and $t(77) = 0.32$, $p = .754$, $d = 0.07$, for the JOL–cued recall correlation.

[2] Analyses of gamma correlations between cued recall performance and both study time and JOLs included data from 39 participants in the original study time condition only. Data from one participant had to be excluded from these analyses, because he or she correctly recalled all items, so that neither the study time–cued recall correlation nor the JOL–cued recall correlation was defined for this participant.

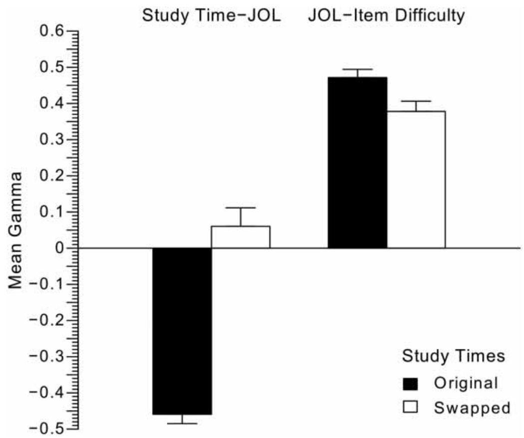

Figure 1. *Mean within-subjects gamma correlations between study time and judgements of learning (JOLs) and between JOLs and item difficulty in the observing-with-difficulty task for the original and the swapped study time condition in Experiment 1. The error bars represent one standard error of the mean.*

Observing-with-difficulty task

In the observing-with-difficulty task, within-subject gamma correlations were based on the 40 word pairs whose study times could differ in principle between the original study time condition and the swapped study time condition.[3] In this task, the mean correlation between study time and JOLs (see Figure 1) was significantly negative for the original study time condition, $t(39) = 17.78$, $p < .001$, $d = 2.81$, whereas it did not significantly differ from zero in the swapped study time condition, $t(39) = 1.18$, $p = .245$, $d = 0.19$. Mean study time–JOL correlations differed significantly between groups, $t(78) = 9.06$, $p < .001$, $d = 2.03$. Swapping study times thus essentially eliminated the correlation between study time and JOLs.

As a measure of item difficulty, the mean probability that an item was successfully recalled in the learning task was used (i.e., the prototypical recall performance). Mean gamma correlations between JOLs and prototypical recall performance were significantly positive for both conditions: original study time condition, $t(39) = 20.74$, $p < .001$, $d = 3.28$; swapped study time condition, $t(39) = 13.38$, $p < .001$, $d = 2.12$. As predicted, gamma correlations were significantly higher for the original study time condition than for the swapped study time condition, $t(78) = 2.58$, $p = .012$, $d = 0.58$. Thus, swapping study times reduced the correlation between JOLs and item difficulty.

In sum, Experiment 1 revealed that JOLs reflect study time even in the presence of other cues disclosing item difficulty and when confounds between study time and these other cues are controlled. Experiment 2 was conducted to replicate this finding and to assess its robustness against procedural changes.

EXPERIMENT 2

Experiment 1 showed that study time affects JOLs when other cues reflecting differences in item difficulty are available as well. In contrast to previous studies, it was ensured that study time was not confounded with other cues that indicate item difficulty. Because Experiment 1 is the first demonstration of this effect, we tried to establish that it is replicable and does not depend on procedural details. As outlined in the introduction, one detail of potential importance is the nature of the word list used in the learning task preceding the observing-with-difficulty task. Do people rely on study time in making JOLs for other people only if they had studied the same word list before? Or does reliance on study time generalize to new word lists? If the learning task serves to establish study time as a cue for other JOLs in similar (but not necessarily identical) learning contexts (cf. Undorf & Erdfelder, 2011), and if the ease-of-processing hypothesis holds, then generalization must be expected. To test this hypothesis, we designed Experiment 2 as a conceptual replication of Experiment 1. More precisely, Experiment 2 is identical to Experiment 1 with the exception that participants worked on different lists of word pairs in the learning task and in the observing-with-difficulty task.

[3] The pattern of results is identical, however, when all items are included in the analyses. The same is true for Experiment 2 and Experiment 3.

Method

Participants

Participants were 95 native German-speaking psychology undergraduates at the University of Mannheim. They were randomly assigned to either the original ($n = 47$) or the swapped study time condition ($n = 48$).

Materials and procedure

All participants worked on the study list from Experiment 1 in the observing-with-difficulty task. However, a different study list was used in the preceding learning task. Just like the study list from Experiment 1, it was composed of 30 unrelated and 30 related German word pairs with associative strengths between .01 and .67 ($M = .15$, $SD = .15$; Hager & Hasselhorn, 1994) and was supplemented with six buffer word pairs (three apparently related and three apparently unrelated). The procedure was identical to that of Experiment 1 except that participants were presented with the new list of words in the learning task. Instructions for the observing-with-difficulty task thus emphasized that the other participant had completed the same experiment, but had studied different word pairs.

Results and discussion

Learning task

The mean within-subject gamma correlation between study time and cued recall performance was significantly smaller than zero in both conditions: $G = -.21$ ($SD = .23$), $t(46) = 6.31$, $p < .001$, $d =$ 1.12, and $G = -.19$ ($SD = .21$), $t(46) = 6.13$, $p < .001$, $d = 0.89$,[4] in the original study time condition and in the swapped study time condition, respectively. The same was true for the study time–JOL correlation, $G = -.30$ ($SD = .22$), $t(47) =$ 9.27, $p < .001$, $d = 1.35$, and $G = -.30$ ($SD = .20$), $t(47) = 10.02$, $p < .001$, $d = 1.45$, for the original study time condition and for the swapped study time condition, respectively. As expected, the correlations between JOLs and recall performance were significantly positive in both conditions: original study time condition, $G = .61$ ($SD = .24$), $t(46) =$ 17.61, $p < .001$, $d = 2.57$; swapped study time condition, $G = .60$ ($SD = .26$), $t(46) = 15.80$, $p < .001$, $d = 2.30$. As in Experiment 1, none of these correlations differed between conditions: study time–cued recall correlation, $t(92) = 0.58$, $p = .564$, $d = 0.09$; study time–JOL correlation, $t(93) = 0.23$, $p = .818$, $d < 0.01$; JOL–cued recall correlation, $t(92) = 0.22$, $p = .830$, $d = 0.04$.

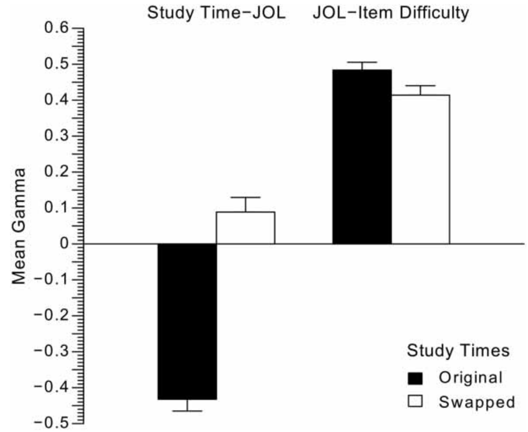

Figure 2. *Mean within-subjects gamma correlations between study time and judgements of learning (JOLs) and between JOLs and item difficulty in the observing-with-difficulty task for the original and the swapped study time condition in Experiment 2. The error bars represent one standard error of the mean.*

Observing-with-difficulty task

In the observing-with-difficulty task, the mean correlation between study time and JOLs (see Figure 2) was significantly negative for the original study time condition, $t(46) = 13.17$, $p < .001$, $d =$ 1.92, and significantly positive for the swapped study time condition, $t(47) = 2.17$, $p = .035$, $d =$ 0.31. Mean study time–JOL correlations differed significantly between groups, $t(93) = 9.93$, $p < .001$, $d = 2.04$. Swapping study times thus reversed the sign of the correlation between study time and JOLs.

[4] Analyses of gamma correlations between cued recall performance and both study time and JOLs included data from 47 participants in the swapped study time condition only, because one participant correctly recalled all items.

Unlike in Experiment 1, recall performance in the learning task could not be used as a measure of item difficulty in the current experiment, because participants studied a different list of word pairs in the learning task. Instead, we again used the mean probability of successfully recalling an item in the learning task of Experiment 1 as a measure of item difficulty. In line with the results of Experiment 1, prototypical recall performance was positively correlated with JOLs in both conditions: original study time condition, $t(46) = 22.14$, $p < .001$, $d = 3.23$; swapped study time condition, $t(47) = 15.75$, $p < .001$, $d = 2.27$. Correlations were significantly higher for the original study time condition than for the swapped study time condition, $t(93) = 2.03$, $p = .045$, $d = 0.42$. Swapping study times thus reduced the correlation between JOLs and item difficulty, as it did in Experiment 1.

Taken together, Experiment 2 replicated the results observed in Experiment 1 using different word pairs in the learning and the observing-with-difficulty task. Results from both experiments demonstrate that study time is used as a cue for JOLs in the presence of other cues disclosing item difficulty. As predicted by the ease-of-processing hypothesis and shown in Experiment 2, this effect does not depend on using the same word list in the learning task and the observing-with-difficulty task. Adding to previous studies showing that experience with the study situation is required to promote reliance on study time when making JOLs for other people (Koriat & Ackerman, 2010; Undorf & Erdfelder, 2011), Experiment 2 thus revealed that experience with the study list is not necessary.

EXPERIMENT 3

In Experiment 1 and Experiment 2, participants made JOLs for paired associates. JOLs in paired-associate tasks are known to be strongly influenced by the degree of associative relatedness between the members of a word pair (e.g., Rabinowitz, Ackerman, Craik, & Hinchley, 1982; for a review see Mueller, Tauber, & Dunlosky, 2012). Although this task is used in the majority of JOL studies, we were interested in assessing whether our findings generalize to a task in which item difficulty is determined primarily by cues other than associative relatedness. According to the ease-of-processing hypothesis, using encoding fluency as a cue for JOLs should not be limited to specific encoding contexts or specific types of item difficulty. To test this prediction, Experiment 3 examined whether study time influences JOLs over and above item difficulty in the learning of foreign-language vocabulary, too.

Learning foreign-language translations consists of mapping novel pattern of sounds and letters to known concepts (e.g., Ellis & Beaton, 1993). The ease of learning translation pairs is mainly influenced by the overlap between the two words in orthography and phonology (e.g., De Groot & Keijzer, 2000; Ellis & Beaton, 1993). This stands in sharp contrast to paired associates whose difficulty is governed by the similarity of cue and target with regard to meaning (Mueller et al., 2012). Hence, foreign language vocabulary is well suited to investigating whether the findings from Experiment 1 and Experiment 2 generalize to learning tasks other than the study of paired associates.

In Experiment 3, JOL instructions for the observing-with-difficulty task were slightly modified: Participants were asked to predict the other person's likelihood of recall (cf. Koriat & Ackerman, 2010; Matvey et al., 2001) rather than the other person's JOLs (cf. Experiment 1 and Experiment 2; Undorf & Erdfelder, 2011). Because the two procedures were shown to produce very similar results in the observing-without-difficulty task (cf. Undorf & Erdfelder, 2011), we did not expect it to influence the results of the current study. However, given that framing effects may occur with JOLs (e.g., Finn, 2008; Koriat, Bjork, Sheffer, & Bar, 2004), it seems worthwhile to examine this issue.

Method

Participants

Participants were 65 German-speaking psychology undergraduates at the University of Mannheim who were unfamiliar with the Polish language. They were randomly assigned to either the original ($n = 32$) or the swapped study time condition ($n = 33$).

Materials and procedure

Participants studied 36 German–Polish translation pairs, of which 30 were of medium difficulty, and the remaining 6 were either easy or difficult. German and Polish words from easy pairs were heterographic cognates (e.g., Vulkan–wulkan). For medium pairs, a synonym of the German word was a cognate of the Polish word (e.g., Boden–grunt). In contrast, we could not find any link between German and Polish words from difficult pairs (e.g., Blüte–kwiat). Four additional easy or difficult translation pairs served as buffer items.

The procedure was identical to that of Experiment 1 with two important exceptions. First, in the learning task, participants studied German–Polish translation pairs and judged the probability that they would be able to recall the Polish words in the test phase. Second, in the observing-with-difficulty task, participants were asked to predict the other person's likelihood of recall rather than the other person's JOLs. Translation pairs used in the learning task were identical to those used in the subsequent observing task.

Presentation times in the observing task corresponded to the study times of one particular participant from a pilot study ($N = 27$). The study times of the selected participant ranged from 1.90 to 47.51 s ($Mdn = 8.32$) and correlated $r = .71$ with mean study times; the corresponding JOLs correlated $r = .68$ with mean JOLs. In the original study time condition, all 40 translation pairs were presented as long as the yoked participant had chosen to study them. In contrast, study times for 20 of the 40 translation pairs were exchanged in the swapped study time condition. In the swapped study time condition, study times were exchanged only between the translation pairs of medium difficulty; all easy and difficult translation pairs (including the four buffer items) kept their original study times. As before, the translation pair with the shortest study time (i.e., Translation Pair 1) and the translation pair with the longest study time (i.e., Translation Pair 30) were preselected for possible exchanges of their study times, as were Translation Pairs 2 and 29, 3 and 28, and so on, up to Translation Pairs 15 and 16. For each participant, study times of randomly selected 20 of these 30 translation pairs were swapped, whereas the other 10 kept their original study times.

Results and discussion

Learning task

The mean gamma correlation between study time and cued recall performance was $G = -.20$ ($SD = .22$) and $G = -.24$ ($SD = .24$) in the original study time condition and in the swapped study time condition, respectively. Both correlations were significantly smaller than zero: original study time condition, $t(31) = 5.08$, $p < .001$, $d = 0.90$; swapped study time condition, $t(32) = 5.59$, $p < .001$, $d = 0.97$, as were the study time–JOL correlations: original study time condition, $G = -.25$ ($SD = .25$), $t(31) = 5.68$, $p < .001$, $d = 1.00$; swapped study time condition, $G = -.18$ ($SD = .22$), $t(32) = 4.66$, $p < .001$, $d = 0.81$. As expected, the mean correlation between JOLs and recall performance was significantly positive in both conditions: original study time condition, $G = .52$ ($SD = .23$), $t(31) = 12.69$, $p < .001$, $d = 2.24$; swapped study time condition, $G = .48$ ($SD = .22$), $t(32) = 12.75$, $p < .001$, $d = 2.22$. None of these correlations differed between conditions: study time–cued recall correlation, $t(63) = 0.70$, $p = .486$, $d = 0.17$; study time–JOL correlation, $t(63) = 1.23$, $p = .223$, $d = 0.30$; JOL–cued recall correlation, $t(63) = 0.76$, $p = .452$, $d = 0.18$.

Observing-with-difficulty task

In the observing-with-difficulty task, the mean correlation between study time and JOLs (see Figure 3) was significantly negative for the original study time condition, $t(31) = 8.80$, $p < .001$, $d = 1.55$, but not for the swapped study time condition, $t(32) = 0.64$, $p = .530$, $d = 0.11$. Mean study time–JOL correlations differed significantly between groups, $t(63) = 5.62$, $p < .001$, $d = 1.40$. Swapping study times thus essentially eliminated the correlation between study time and JOLs.

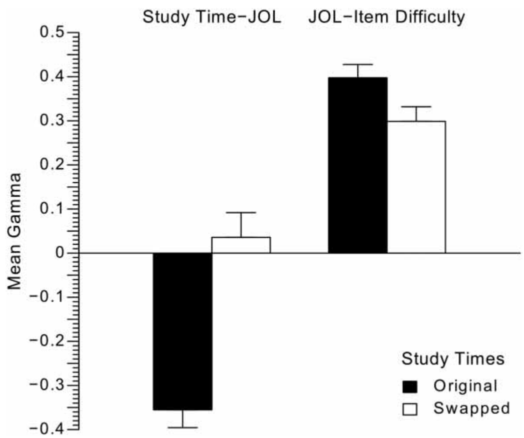

Figure 3. *Mean within-subjects gamma correlations between study time and judgements of learning (JOLs) and between JOLs and item difficulty in the observing-with-difficulty task for the original and the swapped study time condition in Experiment 3. The error bars represent one standard error of the mean.*

The mean probability that an item was successfully recalled in the learning task was used as a measure of item difficulty. Mean gamma correlations between JOLs and item difficulty were significantly positive for both conditions: original study time condition, $t(31) = 13.20$, $p < .001$, $d = 2.33$; swapped study time condition, $t(32) = 9.02$, $p < .001$, $d = 1.57$. As predicted, gamma correlations were significantly higher for the original study time condition than for the swapped study time condition, $t(63) = 2.20$, $p = .031$, $d = 0.55$. Swapping study times thus reduced the correlation between JOLs and item difficulty.

In summary, Experiment 3 replicates Experiment 1 and Experiment 2 in showing that JOLs are based on study time even when other cues are available as well. It thus shows that using encoding fluency as a cue for JOLs does not depend on specific JOL instructions and generalizes to learning tasks other than the study of paired associates.

GENERAL DISCUSSION

The present experiments investigated whether encoding fluency affects JOLs when numerous other cues pointing to the difficulty of an item are available as well. We made sure that study time was not confounded with any other cue that may reveal item difficulty by using an observing-with-difficulty task. In this task, participants observed how long another person had studied word pairs in a self-paced study phase and made JOLs for this person. Item difficulty as revealed by word pair content was disentangled from study time by swapping study times between items in one experimental condition. Experiment 1 showed that swapping study times between paired associates reduced the correlation between JOLs and item difficulty. It thus revealed that JOLs reflect study time in the presence of other cues disclosing item difficulty. Experiment 2 replicated this finding using a list that was new to participants in the observing-with-difficulty task; it suggested that the findings from Experiment 1 are both reliable and robust across procedural variations. Finally, Experiment 3 demonstrated that Experiment 1 and Experiment 2 findings extend to the learning of foreign-language translation equivalents. Taken together, the reported experiments suggest that JOLs are indeed based on self-paced study time and do thus decisively support the presumption that encoding fluency is an important cue for metacognitive judgements (e.g., Begg et al., 1989; Hertzog et al., 2003; Koriat & Ma'ayan, 2005; Kornell et al., 2011).

Our conclusions rely on JOLs that were made for another person. Such other JOLs have been used in a couple of studies before (Koriat & Ackerman, 2010; Matvey et al., 2001; Undorf & Erdfelder, 2011), and although it has been argued that metacognitive judgements for oneself are fundamentally different from metacognitive judgements for other persons (cf. Koriat & Ackerman, 2010), neither we (Undorf & Erdfelder, 2011) nor others (e.g., Kelley, 1999; Kelley & Jacoby, 1996; Matvey et al., 2001; Thomas & Jacoby, 2012) found crucial differences so far. However, previous studies revealed that using study time as a cue for other JOLs requires individual learning experiences in the relevant study context (cf. Undorf & Erdfelder, 2011). Therefore, all participants were given the opportunity to gain experiences prior to completing the observing-with-difficulty task in the present experiments.

Our experimental paradigm also allows assessing of the relative importance of study time for JOLs as compared to other cues disclosing item difficulty. We found that study time–JOL correlations differed markedly between the original study time condition and the swapped study time condition of each experiment: The correlation was negative for the original study time condition, whereas it was essentially zero or even positive for the swapped study time condition. The large difference between the two correlations indicates that JOLs heavily relied on item difficulty as revealed by item content. This result is in line with previous findings and current theorizing in metacognition research (e.g., Koriat, 1997). Importantly, however, we additionally found in all three experiments that the correlation between JOLs and item difficulty was significantly reduced in the swapped study time condition. Hence, as predicted by the ease-of-processing hypothesis (e.g., Dunlosky & Metcalfe, 2008), encoding fluency also influences JOLs when further information on item difficulty is available.

The effect of swapping study times on the correlations between JOLs and item difficulty was rather small. However, it is important to keep in mind that the influence of encoding fluency on JOLs is likely to be underestimated in the current experiments. In principle, encoding fluency as measured by study time can influence JOLs both through analytic processes (e.g., the deliberate application of the rule that easily learned information will be easily remembered, cf. Koriat, 2008; Koriat et al., 2009a) and through nonanalytic influences (e.g., subjective feelings of mastery of a vague origin). Nonanalytic influences of study time on other JOLs should have been severely attenuated in our swapped study time condition: It seems implausible that study time elicited strong feelings of whether an item was mastered with ease or difficulty when counteracted by item content (see also Koriat & Ackerman, 2010). In contrast, it is reasonable to assume that study time influenced other JOLs through analytic processes. After all, study time was the only cue revealing how difficult an item was to the other person. In sum, in our experiments the influence of encoding fluency on JOLs was presumably mediated by analytical processes only and was thus underestimated (for a detailed discussion of the contributions of analytic and nonanalytic processes to the relatedness effect, see Mueller et al., 2012).

Our results lend strong support to the ease-of-processing hypothesis and other theories assuming that encoding fluency is an important cue for JOLs (e.g., Koriat, 1997, 2008; Koriat et al., 2006). More generally, our findings support the idea that fluency of processing is a major determinant of JOLs (e.g., Benjamin, Bjork, & Schwartz, 1998; Busey, Tunnicliff, Loftus, & Loftus, 2000; Kornell et al., 2011; Matvey et al., 2001; Rhodes & Castel, 2008, 2009; Serra & Dunlosky, 2005). With respect to the underlying cognitive mechanisms, our results thus emphasize similarities between JOLs and other metacognitive judgements including judgements of familiarity, liking, and confidence (e.g., Ackerman & Zalmanov, 2012; Alter & Oppenheimer, 2009).

Original manuscript received 9 July 2012
Accepted revision received 13 November 2012
First published online 10 April 2013

REFERENCES

Ackerman, R., & Zalmanov, H. (2012). The persistence of the fluency–confidence association in problem solving. *Psychonomic Bulletin & Review*, *19*(6), 1187–1192. doi:10.3758/s13423-012-0305-z

Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review*, *13*(3), 219–235. doi:10.1177/1088868309341564

Begg, I., Duft, S., Lalonde, P., Melnick, R., & Sanvito, J. (1989). Memory predictions are based on ease of processing. *Journal of Memory and Language*, *28*(5), 610–632. doi:10.1016/0749-596X(89)90016-8

Benjamin, A. S., Bjork, R. A., & Schwartz, B. L. (1998). The mismeasure of memory: When retrieval fluency is misleading as a metamnemonic index. *Journal of Experimental Psychology: General*, *127*(1), 55–68. doi:10.1037/0096-3445.127.1.55

Busey, T. A., Tunnicliff, J., Loftus, G. R., & Loftus, E. F. (2000). Accounts of the confidence accuracy relation in recognition memory. *Psychonomic Bulletin & Review*, *7*, 26–48. doi:10.3758/BF03210724

De Groot, A. M. B., & Keijzer, R. (2000). What is hard to learn is easy to forget: The roles of word concreteness, cognate status, and word frequency in foreign-language vocabulary learning and forgetting. *Language Learning*, *50*(1), 1–56. doi:10.1111/0023-8333.00110

Dunlosky, J., & Metcalfe, J. (2008). *Metacognition*. Los Angeles, CA: Sage Publications.

Ellis, N. C., & Beaton, A. (1993). Psycholinguistic determinants of foreign language vocabulary learning. *Language Learning*, *43*(4), 559–617. doi:10.1111/j.1467-1770.1993.tb00627.x

Finn, B. (2008). Framing effects on metacognitive monitoring and control. *Memory & Cognition*, *36* (4), 813–821. doi:10.3758/MC.36.4.813

Hager, W., & Hasselhorn, M. (1994). *Handbuch deutschsprachiger Wortnormen* [Handbook of German word norms]. Göttingen: Hogrefe.

Hertzog, C., Dunlosky, J., Robinson, A. E., & Kidder, D. P. (2003). Encoding fluency is a cue used for judgments about learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*(1), 22–34. doi:10.1037/0278-7393.29.1.22

Jacoby, L. L. (1991). A process dissociation framework: Separating automatic from intentional uses of memory. *Journal of Memory and Language*, *30*(5), 513–514. doi:10.1016/0749-596X(91)90025-F

Kelley, C. M. (1999). Subjective experience as basis of "objective" judgments: Effects of past experience on judgments of difficulty. In D. Gopher & A. Koriat (Eds.), *Attention and Performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 515–536). Cambridge. MA: MIT Press.

Kelley, C. M., & Jacoby, L. L. (1996). Adult egocentrism: Subjective experience versus analytic bases for judgment. *Journal of Memory and Language*, *35*(2), 157–175. doi:10.1006/jmla.1996.0009

Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, *126*(4), 349–370. doi:10.1037/0096-3445.126.4.349

Koriat, A. (2008). Easy comes, easy goes? The link between learning and remembering and its exploitation in metacognition. *Memory & Cognition*, *36*(2), 416–428. doi:10.3758/MC.36.2.416

Koriat, A., & Ackerman, R. (2010). Metacognition and mindreading: Judgments of learning for self and other during self-paced study. *Consciousness and Cognition*, *19*(1), 251–264. doi:10.1016/j.concog.2009.12.010

Koriat, A., Ackerman, R., Lockl, K., & Schneider, W. (2009a). The easily learned, easily remembered heuristic in children. *Cognitive Development*, *24*(2), 169–182. doi:10.1016/j.cogdev.2009.01.001

Koriat, A., Ackerman, R., Lockl, K., & Schneider, W. (2009b). The memorizing effort heuristic in judgments of learning: A developmental perspective. *Journal of Experimental Child Psychology*, *102*(3), 265–279. doi:10.1016/j.jecp.2008.10.005

Koriat, A., & Bjork, R. A. (2006). Mending metacognitive illusions: A comparison of mnemonic-based and theory-based procedures. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*(5), 1133–1345. doi:10.1037/0278-7393.32.5.1133

Koriat, A., Bjork, R. A., Sheffer, L., & Bar, S. K. (2004). Predicting one's own forgetting: The role of experience-based and theory-based processes. *Journal of Experimental Psychology: General*, *133*(4), 643–656. doi:10.1037/0096-3445.133.4.643

Koriat, A., & Ma'ayan, H. (2005). The effects of encoding fluency and retrieval fluency on judgments of learning. *Journal of Memory and Language*, *52*(4), 478–492. doi:10.1016/j.jml.2005.01.001

Koriat, A., Ma'ayan, H., & Nussinson, R. (2006). The intricate relationships between monitoring and control in metacognition: Lessons for the cause-and-effect relation between subjective experience and behavior. *Journal of Experimental Psychology: General*, *135*(1), 36–69. doi:10.1037/0096-3445.135.1.36

Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, *14*(2), 219–224. doi:10.3758/BF03194055

Kornell, N., Rhodes, M. G., Castel, A. D., & Tauber, S.K. (2011). The ease-of-processing heuristic and the stability bias: Dissociating memory, memory beliefs, and memory judgments. *Psychological Science*, *22*(6), 787–794. doi:10.1177/0956797611407929

Matvey, G., Dunlosky, J., & Guttentag, R. (2001). Fluency of retrieval at study affects judgments of learning (JOLs): An analytic or nonanalytic basis for JOLs? *Memory & Cognition*, *29*(2), 222–233. doi:10.3758/BF03194916

Metcalfe, J., & Finn, B. (2008). Evidence that judgments of learning are causally related to study choice. *Psychonomic Bulletin & Review*, *15*(1), 174–179. doi:10.3758/PBR.15.1.174

Miele, D. B., Finn, B., & Molden, D. C. (2011). Does easily learned mean easily remembered? It depends on your beliefs about intelligence. *Psychological Science*, *22*(3), 320–324. doi:10.1177/0956797610397954

Mueller, M. L., Tauber, S. K., & Dunlosky, J. (2012). Contributions of beliefs and processing fluency to the effect of relatedness on judgments of learning. *Psychonomic Bulletin & Review.* Advance online publication. doi:10.3758/s13423-012-0343-6

Oppenheimer, D. M. (2008). The secret life of fluency. *Trends in Cognitive Sciences*, *12*(6), 237–241. doi:10.1016/j.tics.2008.02.014

Rabinowitz, J. C., Ackerman, B. P., Craik, F. I. M., & Hinchley, J. L. (1982). Aging and metamemory: The roles of relatedness and imagery. *Journal of Gerontology*, *37*(6), 688–695. doi:10.1093/geronj/37.6.688

Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology: General*, *137*(4), 615–625. doi:10.1037/a0013684

Rhodes, M. G., & Castel, A. D. (2009). Metacognitive illusions for auditory information: Effects on monitoring and control. *Psychonomic Bulletin & Review*, *16*(3), 550–554. doi:10.3758/PBR.16.3.550

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-prime user's guide*. Pittsburgh, PA: Psychology Software Tools.

Schwartz, B. L. (1994). Sources of information in metamemory: Judgments of learning and feelings of knowing. *Psychonomic Bulletin & Review*, *1*(3), 357–375. doi:10.3758/BF03213977

Serra, M. J., & Dunlosky, J. (2005). Does retrieval fluency contribute to the underconfidence-with-practice effect? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(6), 1258–1266. doi:10.1037/0278-7393.31.6.1258

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin*, *137*(3), 421–442. doi:10.1037/a0022777

Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*(1), 204–221. doi:10.1037//0278-7393

Thomas, R. C., & Jacoby, L. L. (2012). Diminishing adult egocentrism when estimating what others know. *Journal of Experimental Psychology: Learning, Memory, and Cognition.* Advance online publication. doi:10.1037/a0028883

Undorf, M. (2010). *Metagedächtnis und Lernen* [Metamemory and learning]. Berlin, Germany: Logos.

Undorf, M., & Erdfelder, E. (2011). Judgments of learning reflect encoding fluency: Conclusive evidence for the ease-of-processing hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(5), 1264–1269. doi:10.1037/a0023719

Vesonder, G., & Voss, J. (1985). On the ability to predict one's own responses while learning. *Journal of Memory and Language*, *24*(3), 363–376. doi:10.1016/0749-596X(85)90034-8