



The Role of Mental Effort in Students' Perceptions of the Effectiveness of Interleaved and Blocked Study Strategies and Their Willingness to Use Them

Eva M. Janssen¹ · Tamara van Gog¹ · Laura van de Groep¹ · Anne Jóaia de Lange¹ · Roosmarijn L. Knopper¹ · Erdem Onan² · Wisnu Wiradhany³ · Anique B. H. de Bruin²

Accepted: 10 July 2023 / Published online: 15 August 2023
© The Author(s) 2023

Abstract

Students tend to avoid effective but effortful study strategies. One potential explanation could be that high-effort experiences may not give students an immediate feeling of learning, which may affect their perceptions of the strategy's effectiveness and their willingness to use it. In two experiments, we investigated the role of mental effort in students' considerations about a typically effortful and effective strategy (interleaved study) versus a typically less effortful and less effective strategy (blocked study), and investigated the effect of individual feedback about students' study experiences and learning outcomes on their considerations. Participants learned painting styles using both blocked and interleaved studying (within-subjects, Experiment 1, $N=150$) or either blocked or interleaved studying (between-subjects, Experiment 2, $N=299$), and reported their study experiences and considerations before, during, and after studying. Both experiments confirmed prior research that students reported higher effort investment and made lower judgments of learning during interleaved than during blocked studying. Furthermore, effort was negatively related to students' judgments of learning and (via these judgments) to the perceived effectiveness of the strategy and their willingness to use it. Interestingly, these relations were stronger in Experiment 1 than in Experiment 2, suggesting that effort might become a more influential cue when students can directly compare experiences with two strategies. Feedback positively affected students' considerations about interleaved studying, yet not to the extent that they considered it more effective and desirable than blocked studying. Our results provide evidence that students use effort as a cue for their study strategy decisions.

Keywords Mental effort · Judgments of learning · Study considerations · Effective study strategies · Interleaved study · Blocked study

Experiment 1 by Janssen, Van Gog, Van de Groep, De Lange, Onan, Wiradhany, and De Bruin, experiment 2 by Janssen, Van Gog, Knopper, Onan, Wiradhany, and De Bruin.

✉ Eva M. Janssen
e.m.janssen@uu.nl

Extended author information available on the last page of the article

For decades, instructional design research has focused on identifying effective study strategies, and with success. There is a large body of evidence, for instance, that for many types of learning materials, retrieval practice (i.e., retrieving studied information from memory) is more effective than restudy (Roediger & Butler, 2011); that spacing restudy over time is more effective than massing study into a single session (Cepeda et al., 2006); and that interleaved study, in which different types of exemplars are mixed (e.g., ACB BAC CBA), is more effective than blocked study (e.g., AAA BBB CCC; Brunmair & Richter, 2019). Note that ‘more effective’ means that these strategies will typically result in higher learning outcomes, better transfer performance, or better long-term retention, even if progress during the study session may be slower. For instance, blocked studying often yields faster performance improvements during the session than interleaved studying, whereas interleaved studying results in better performance on a later test (e.g., Schmidt & Bjork, 1992). Retrieval practice may yield lower performance than restudy on an immediate test at a 5 min. delay, whereas it yields higher performance than restudy after a two-day or one-week interval (e.g., Roediger & Karpicke, 2006).

Getting students to use those effective strategies, however, turns out to be a major challenge (e.g., Kornell & Vaughn, 2018; Tauber et al., 2013). This is a problem, because in many situations, especially in higher education, students have to self-regulate their learning process, and their learning outcomes may suffer if they use suboptimal study strategies. One potential explanation for why students might avoid the more effective study strategies, is that those are typically also more effortful to maintain and may not give students an immediate feeling of learning while they are studying, even if they do ultimately lead to better learning outcomes (i.e., higher test performance). That is, most students may not see the benefits of effective study strategies because they consider high effort as an indication that they are not learning well and fail to consider that—under some conditions—high effort results in better learning outcomes. However, there has been very little research that has tested this explanation. Therefore, the present study aims to investigate the role of mental effort in students’ perceptions of the effectiveness of study strategies and their willingness to use them, in the context of blocked and interleaved study strategies.

Interleaved Vs. Blocked Study

The benefits of interleaved over blocked study (or practice) schedules (also known as the ‘contextual interference effect’; Shea & Morgan, 1979) have been investigated for many different tasks. Many studies have shown these benefits for the acquisition of motor tasks (e.g., Cross et al., 2007; Guadagnoli & Lee, 2004; Magill & Hall, 1990; Shea & Morgan, 1979; Simon, 2007), but they have also been demonstrated for learning language and vocabulary (e.g., Jacoby, 1978; Schneider et al., 1998, 2002), procedural tasks (e.g., Carlson & Schneider, 1989; Carlson & Yaure, 1990), problem-solving and troubleshooting tasks (e.g., De Croock et al., 1998; Paas & Van Merriënboer, 1994; Rohrer & Taylor, 2007), logical rules (e.g., Schneider et al., 1995), complex judgment and decision-making tasks (e.g., Helsdingen et al.,

2011) and category learning tasks (e.g., Carvalho & Goldstone, 2017; Eglington & Kang, 2017).

Several (not mutually exclusive) hypotheses have been formulated to explain the benefits of interleaved studying. According to the elaboration hypothesis (Shea & Morgan, 1979), schema acquisition is fostered by the fact that interleaved studying evokes more elaborate processing of *intertask* similarities and differences. The reconstruction hypothesis (Lee & Magill, 1983) attributes the beneficial effects of interleaving to the fact that it also implies *spacing* study exemplars of a specific type over time (which is known to have beneficial effects compared to massed studying; Cepeda et al., 2006), due to which learners have forgotten some (aspects of) the task and need to engage in reconstructive processing. Finally, for some tasks, interleaved studying also implies that learners have to engage in *retrieval* processes (see Schmidt & Bjork, 1992), and (attempting to) retrieve previously learned information from memory is known to strengthen long-term retention (see also Roediger & Butler, 2011).

A common feature of these hypotheses is that these processes (elaboration, reconstruction, retrieval) evoked by interleaved studying initially seem to make it harder for learners (reflected in slower progress or worse performance during the study phase compared to blocked studying), yet lead to better learning outcomes. In other words, these processes impose ‘desirable difficulties’ for learners (Bjork & Bjork, 2020). Consequently, interleaved studying is considered to be more effortful than blocked studying because of the desirable difficulties it imposes. Interestingly, however, studies in which participants were asked to rate the amount of mental effort they invested, show inconsistent results. Several studies with problem-solving, troubleshooting, or reasoning tasks did not show significant differences in mental effort investment during the study phase between interleaved and blocked study (e.g., De Croock et al., 1998; Paas & Van Merriënboer, 1994; Van Peppen et al., 2021), whereas other studies with category learning tasks did (Kirk-Johnson et al., 2019; Onan et al., 2022). This difference in findings could perhaps be a consequence of differences in the nature of the tasks and the cognitive demands they impose, but it could also be a consequence of the research design: in the studies that did not find a significant difference in self-reported mental effort, blocked vs. interleaved study was a between-subjects factor, whereas in the category learning studies it was typically a within-subjects factor, meaning that learners experienced and could compare the two different methods.

An important question, however, is how we can get students to use those effective but effortful study strategies. As we will discuss in the next section, because an effective strategy like interleaved studying induces desirable difficulties and requires more effort than blocked studying (at least for category learning tasks, which are also used in the present study), this might negatively affect students’ judgments of learning and their perception of the effectiveness of the strategy. Consequently, they may not be inclined to use it.

Students' (Dis)Inclination to Use Effective but Effortful Study Strategies

Several studies have shown that students seem to underestimate what they learn from effective but effortful strategies and to overestimate what they learn from less effective and less effortful strategies. For instance, Zechmeister and Shaughnessy (1980) showed that students judged words to be more likely to be recalled after massed practice than after spaced practice, which was not the case. With a motor-learning task (key-stroke patterns) it was found that although participants who engaged in interleaved practice predicted their performance on a test the next day quite well, those given blocked practice were quite overconfident (Simon & Bjork, 2001). In a category learning task (learning to recognize artists' styles from their paintings), it was found that only 22% of the students thought they did better in the interleaved study condition than in the blocked study condition, whereas actually, 78% did better after interleaved than blocked study (Kornell & Bjork, 2008). One could argue that this is because students are unaware of the effectiveness of the strategy (cf. McCabe, 2011). Interestingly, however, Kornell and Bjork (2008) provided correctness feedback on the test, and the judgments of learning were made after the test, so students could have been aware that they performed better after interleaved studying. Moreover, it has been found that students who quite accurately rated study strategies in terms of being more effective versus less effective, still tended to use ineffective strategies (Blasiman et al., 2017).

It has been suggested that both students' judgments of how well they are learning with a specific strategy as well as their inclination to use that strategy are affected by their experiences during studying (Kornell & Bjork, 2007; Soderstrom & Bjork, 2015). However, these experiences are not necessarily predictive of actual learning outcomes. For instance, students often misinterpret feelings of fluency during studying (which will be higher during blocked studying) as an indication of learning (Koriat, 1997). The effort monitoring and regulation (EMR) framework (De Bruin et al., 2020) proposes that students use mental effort as a cue for monitoring their learning and for making study decisions. According to this framework, one reason why students are not inclined to use effective but effortful strategies might be that they misinterpret their experienced effort during studying. In other words, most students may not see the benefits of effective but effortful study strategies because they consider their experience that studying costs much effort as an indication that they are not learning well, and fail to consider that under some conditions (like interleaved studying) high effort investment may actually result in higher learning outcomes.

To date, however, not much is known about how students' monitoring of their effort investment relates to their monitoring of their learning process, their perceptions of the effectiveness of the study strategies, and their willingness to use those strategies. As mentioned earlier, relatively few studies have measured effort investment, and most of those did not measure judgments of learning or students' perceptions of the study strategies. We found only two studies that did look into these relationships.

First, in a study by Kirk-Johnson et al. (2019), participants studied bird families using both blocked and interleaved studying, reported their perceived mental effort, perceived learning effectiveness and, thereafter, chose which strategy they would use to study a new set of bird families. In their analyses, Kirk-Johnson et al. (2019) related participants' difference in experienced effort between the two strategies (i.e., perceived effort with interleaved minus perceived effort with blocked) to the difference in their experienced learning (i.e., perceived learning effectiveness with interleaved minus perceived learning effectiveness with blocked) and to participants' choice for interleaved studying. Results showed that learners who perceived interleaved studying as more effortful than blocked studying rated it as less effective for learning than blocked studying and were less likely to report that they would choose it for future study. Hence, these findings suggest that learners indeed use mental effort as a cue for their study decisions. Yet, based on these results we only know that a contrast in experience between the two strategies can predict study decisions. It remains unclear if and how mental effort would relate to study decisions when learners would experience high mental effort with both strategies or low mental effort with both strategies (i.e., both situations would be coded as no or low experienced contrast between the two strategies, even though they refer to very different mental effort perceptions). Moreover, it also remains unclear whether the relationship between experienced effort and study decisions would be the same for blocked and interleaved studying and how mental effort would relate to study decisions when learners would work with only one study strategy rather than comparing two strategies, which is more often the case in real study situations.

Onan et al. (2022) did study the relationship between mental effort and study decisions separately for a blocked and an interleaved strategy. Participants used both blocked and interleaved strategies to study painting styles. Across multiple short study blocks, learners rated their invested effort and made concurrent judgments of learning for each strategy. In their analyses, however, Onan et al. (2022) tested the relation of *temporal change* in ratings of effort across study blocks (i.e., of the increase or decrease in invested effort over time) with *temporal change* in judgments of learning (i.e., of the increase or decrease in judgments of learning over time) and learners' choice for interleaved studying. Results showed that over time (i.e., over study blocks), effort investment in interleaved study decreased, which makes sense, as exemplars have been repeated several times then and learners start to recognize the categories (i.e., the effort demands of interleaved practice decreased with increased ability). Furthermore, judgments of learning increased over time, and learners' choice for interleaved studying increased from 13 to 40%. Moreover, these changes in effort (i.e., decrease) and judgments of learning (i.e., increase) over time predicted learners' decisions to use interleaved studying (no such relationship was found for blocked studying). Hence, in line with Kirk-Johnson et al. (2019), these results also show that learners use invested mental effort as cue for their study decisions and, additionally, that the relation between invested effort and study decisions might be different for blocked and interleaved studying. However, similar to the study by Kirk-Johnson et al. (2019), because of the relative measure, it remains unclear from these findings if and how *the amount* of invested effort is related to study decisions. For example, a learner who reports high mental effort across all

study blocks and a learner who reports low mental effort across all study blocks would be coded the same in the Onan study (i.e., no change in experienced effort), even though both situations refer to a very different overall effort experience. Moreover, as in the study by Kirk-Johnson et al. (2019), study strategy was manipulated within subjects and, therefore, it remains unclear if and how students use mental effort as a cue when using only blocked or only interleaved studying. The goal of the present study was to gain further insight into how learners use invested mental effort as a cue for their study decisions. Therefore, we investigated for each of the strategies separately the extent to which the overall amount of experienced effort negatively affects students' judgments of learning, their perception of the effectiveness of the used strategy, and consequently, their willingness to use it. We tested these relationships in two experiments: in Experiment 1 study strategy was manipulated within subjects and in Experiment 2 it was manipulated between subjects.

Another interesting finding from the Onan et al. study was that even though changes in effort predicted the willingness to use interleaved studying, the number of students who actually chose to use interleaved studying for the second study phase was still rather low (40%). This raises the question: How can we change students' willingness to use effective but effortful strategies? Based on the proposition from the EMR framework (De Bruin et al., 2020) that students use mental effort as a cue for monitoring their learning and making study decisions, we expect that presenting students with information about their effort experiences in relation to their actual learning outcomes might change their study decisions. In other words, we hypothesize that students' perceptions of the effectiveness of interleaved studying and their willingness to use it in the future would become more positive if students received visual feedback after the test, displaying their self-reported effort, judgments of learning, and their actual test score following interleaved and blocked studying. This feedback might make them aware of their misconception that high effort would mean they are not learning, by showing them that they actually learned more under the more effortful study strategy (i.e., based on prior research, we can assume this to be the case for the majority of students), which might change their considerations.

The Present Study

The present study aims to investigate the role of mental effort and feedback in students' perceptions of the effectiveness of blocked and interleaved study strategies and their willingness to use those strategies using a category learning task (learning to recognize artists' styles from their paintings). These strategies are very well suited to address this question, as prior research has shown that (at least in a within-subjects design) blocked studying is typically associated with less mental effort investment during the study phase than interleaved studying, but also leads to lower learning outcomes than interleaved studying (Kirk-Johnson et al., 2019; Onan et al., 2022). We expect that students use their mental effort during a particular study session as a cue when judging the effectiveness of that study strategy and their willingness to use that strategy. We also expect that visual feedback on their ratings of invested mental effort and judgments of

learning in relation to their actual test performance (provided after a test), would alter students' perceptions of the effectiveness of the strategies and their willingness to use those strategies.

In Experiment 1, we address these questions using study strategy as a within-subjects factor (cf. Kirk-Johnson et al., 2019; Onan et al., 2022) and feedback as a between-subjects factor. The fact that in a within-subjects design, students are able to compare their experiences with the two different study strategies (e.g., that interleaved studying requires more effort than blocked studying), yields relevant information, yet also raises the question to what extent the findings regarding the role of effort and students' strategy perceptions would be colored by this comparison. Thus, an important and interesting question is to what extent findings would hold when students only engage with one of the strategies. Therefore, we conducted a second experiment with study strategy as between-subjects factor.

Hypotheses

We had similar hypotheses for Experiment 1 and Experiment 2. First, we expected to replicate findings from prior research that students' experiences while studying and their learning outcomes do not match: Interleaved studying is expected to lead to higher mental effort investment and lower judgments of learning (Kirk-Johnson et al., 2019; Onan et al., 2022), yet is also expected to yield higher test performance than blocked studying (Kirk-Johnson et al., 2019; Kornell & Bjork, 2008; Onan et al., 2022). Second, we expect that students use their invested mental effort during studying as a cue when judging the effectiveness of blocked and interleaved studying and deciding on their willingness to use each strategy (i.e., we expect a negative association between invested effort and perceived effectiveness/willingness to use a strategy). Therefore, we tested per study strategy, the direct and indirect correlational effects of students' mental effort on perceived strategy effectiveness (via judgments of learning), and on their willingness to use that strategy (via perceived effectiveness) after studying. Perceived effectiveness of the study strategies and willingness to use the strategies was measured at three time points: before the study phase, after the study phase, and after the test. We expected that immediately after studying, students would consider blocked studying (in which we expected them to invest less effort) to be more effective than interleaved studying, and would be more willing to use blocked studying. In addition, we expected that the students who received feedback on their actual learning outcomes (test performance) after the test, would become more positive about the effectiveness of interleaved studying than students who did not receive feedback, and would be more willing to use it in the future.

Experiment 1

Method

This study was approved by the faculty's ethics committee of the first author. The study data and analyses scripts are stored on an Open Science Framework (OSF) page for this project, see <https://osf.io/5uzbq/>.

Participants and Design

Experiment 1 had a $2 \times 3 \times 2$ mixed design with study strategy (blocked vs. interleaved) and measurement moment (before study phase, after study phase, and after test) as within-subjects factors, and feedback intervention (yes/no) as between-subjects factor.

To estimate the required sample size, we first calculated how much participants were needed to detect the interleaving effect. A meta-analysis (Brunmair & Richter, 2019) showed that the interleaving effect for studying paintings is medium-sized (Hedge's $g=0.67$). An a priori power analysis in G*Power (Faul et al., 2007) for ANOVA with one within-subjects factor (study strategy), a medium effect size (Cohen's $f=0.25$), an alpha level of 0.05, a power of 95%, a correlation of 0.30 among repeated measures, and a non-sphericity correction of 1, showed that 75 participants would be sufficient to detect the interleaving effect. For detecting the feedback intervention effect for each study strategy, an a priori analysis in G*Power with a within-between interaction effect (measurement moment \times feedback condition) and the same settings revealed that 76 participants would be required. However, we used path analyses to test the direct and indirect correlational effects of students' mental effort on perceived strategy effectiveness (via prospective judgments of learning), and on their willingness to use that strategy (via perceived effectiveness; see [Analyses](#) section.). Power analyses for path models are quite complex, but there is a general consensus that 10 participants per estimated parameter is sufficient (Schreiber et al., 2006). As our path models estimated a total 15 parameters (see Fig. 6) we would need 150 participants.

Hence, a total of 150 Dutch students were recruited on Prolific Academic (www.prolific.ac) and paid £3.75 for their participation. Most participants were from a research university ($n=99$) or a university of applied sciences ($n=45$), three participants were enrolled in secondary vocational education, and three indicated "other." The sample consisted of 66 females, 83 males, and one participant who indicated "other." and their average age was 22.47 years ($SD=3.93$).

All participants engaged in a study session in which they had to learn to recognize ten different painters by their paintings. They studied five painters via blocked studying (five study units, each unit contained five paintings by the same painter, see Fig. 1) and the other five painters via interleaved studying (five study units, each unit contained five paintings by the five different painters, see Fig. 1). Participants were

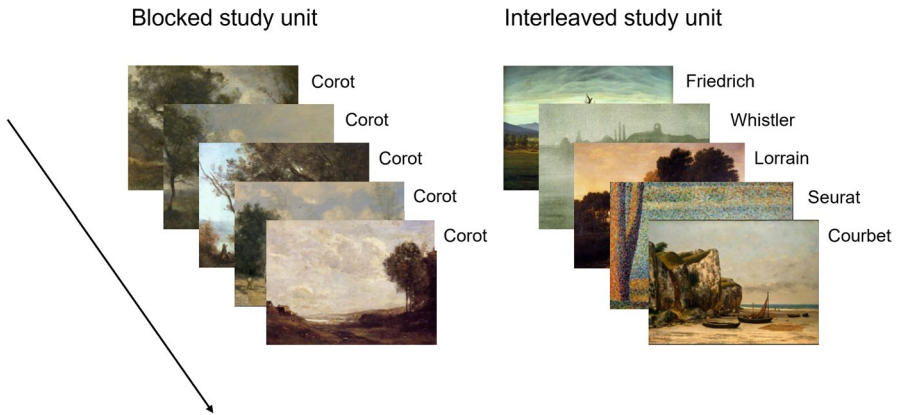


Fig. 1 Examples of a blocked study unit and an interleaved study unit

asked to report their opinion about the strategies before the study phase, after the study phase, and after a test. After the test, participants were randomly assigned to a feedback condition ($n=75$) or no-feedback condition ($n=75$). Participants in the feedback condition received visual feedback on their self-reported study experiences and on their test performance.

Materials

Prior Knowledge Test

Participants completed a short prior knowledge test that consisted of three multiple-choice questions on which participants had to select the correct art movement belonging to paintings they saw. As expected, prior knowledge was low and did not differ between feedback conditions, $M=0.96$ questions correct, $SD=1.05$, $t(148)=-0.46$, $p=0.644$, Cohen's $d=-0.08$.

Learning Materials

Participants had to learn to recognize the work of ten different painters (Corot, Church, Constable, Turner, Cézanne, Friedrich, Whistler, Lorrain, Seurat, and Courbet). All paintings were obtained from the Painting-91 database (Khan et al., 2014). The selected paintings were landscapes or skyscapes and resized to 500×375 pixels.

Mental Effort

Participants rated their mental effort (“You have just practiced learning to recognize the work of a painter [for blocked units]/painters [for interleaved units]. How much effort did you invest?”) on a 9-point rating scale (Paas, 1992) ranging from 1 (very, very little mental effort) to 9 (very, very much mental effort).

Concurrent Judgments of Learning

Concurrent judgments of learning were measured with the question “You have just practiced learning to recognize the work of a painter/painters. How well do you think you would be able to recognize the work of this painter/these painters now?” on a 9-point rating scale ranging from 1 (very, very poorly) to 9 (very, very well).

Prospective Judgments of Learning

Prospective judgments of learning were measured with the following question: “How well will you be able to recognize the work of the painters that you just studied with the blocked/interleaved study strategy on the test? I expect that I will correctly recognize the painter in _____ out of the 10 paintings that will be tested.”

Learning Outcomes

Learning outcomes were measured with a test that consisted of two unseen paintings by each of the ten studied painters (20 in total). Below each painting, participants could select the name of the painter from a drop-down menu including all painters' names. Participants received one point for each correct answer. Reliability of the test scores was acceptable, blocked test: Cronbach's $\alpha = .60$; interleaved test: Cronbach's $\alpha = .63$.

Perceived Effectiveness

Participants' perception of the effectiveness of each study strategy (“Blocked /Interleaved studying is an effective study strategy for the current learning task, namely, learning to recognize the work of painters by their paintings”) were assessed with a 9-point rating scale ranging from 1 (completely disagree) to 9 (completely agree).

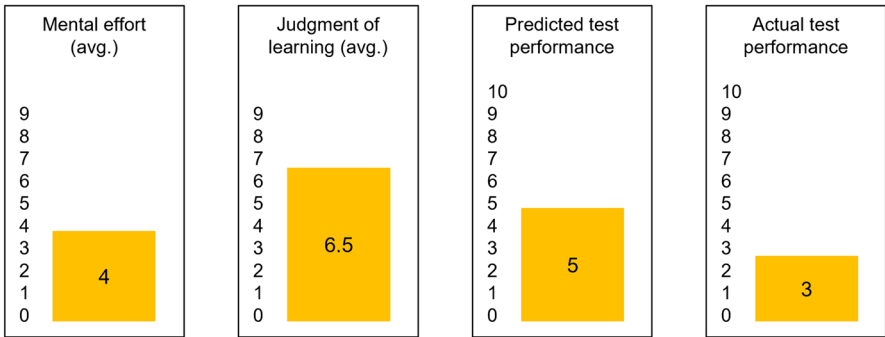
Willingness to Use

Participants' willingness to use each strategy (“I would like to use blocked/interleaved studying as a study strategy for learning tasks that are similar to the current learning task, namely, learning to recognize the work of painters from their paintings”) was also measured with on a 9-point rating scale ranging from 1 (completely disagree) to 9 (completely agree).

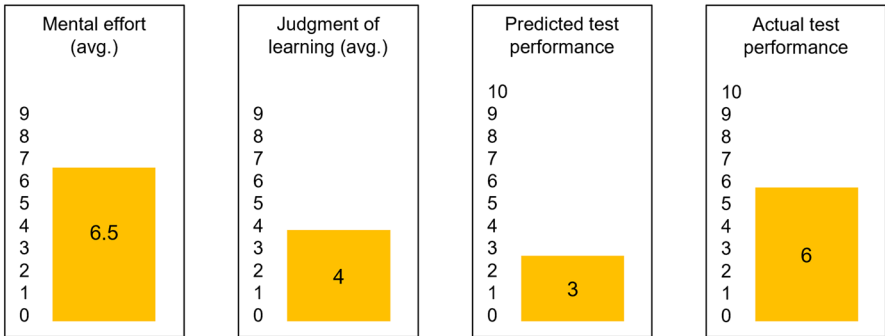
Feedback Intervention

In the feedback condition a personalized visual overview (see Fig. 2) was provided of the participant's self-reported experiences (mental effort, concurrent and prospective judgments of learning) and actual learning outcomes (test performance), per study strategy. To encourage processing of the feedback, participants were prompted to interpret the feedback by writing down what they could conclude from the visual overview. In the no-feedback condition participants did not receive a visual

Your results for blocked studying



Your results for interleaved studying



Briefly describe what you can conclude from the results above

Fig. 2 Example of the personal feedback that was shown to participants in the feedback condition (translated from Dutch)

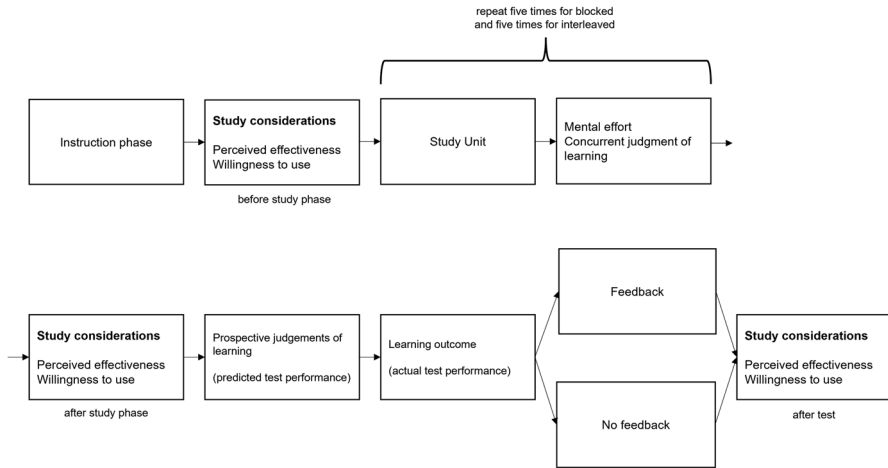


Fig. 3 Overview of the procedure of experiment 1

overview, but were prompted to take some time to reflect on the two study strategies that they just had used.

Procedure

Figure 3 shows an overview of the procedure. First, an instruction phase: (1) informed participants about the study's procedure; (2) explained the difference between a blocked and an interleaved study strategy; and (3) showed three sample test questions, so that participants knew what to expect on the test they would complete at the end. Next, participants rated their study strategy considerations (perceived effectiveness of and willingness to use each strategy) for the first time (out of three). The order of these two questions was counterbalanced between participants. Half of the participants first rated perceived effectiveness of the study strategies and then their willingness to use them; for the other half, it was the other way around. Question order did not affect participants' ratings at any of the three measurement moments for any of the two study strategies ($p_s \geq 0.091$).

Hereafter, participants completed the prior knowledge test, and then started with the study phase in which they had to learn to recognize the painting styles of the ten painters through the blocked and interleaved study units. In line with Kornell and Bjork (2008), participants studied the paintings in study units following this study sequence: B-I-I-B-B-I-I-B-B-I, where B refers to a blocked study unit and I refers to an interleaved study unit. Each painting was presented as follows: first, a fixation cross appeared at the center of the screen for 1 s, and then the painting, with the painter's last name above for 3 s. The order of the paintings presented within each unit was randomized per participant. We counterbalanced between participants which painters were studied by which study strategy. Half of the participants studied Corot, Church, Constable, Turner, and Cézanne with a blocked study strategy and

Friedrich, Whistler, Lorrain, Courbet, and Seurat with an interleaved study strategy; for the other half of the participants, it was the other way around.

At the end of each study unit, participants reported their invested mental effort and made concurrent judgments of learning. The order of these two questions was counterbalanced between participants; half of the participants first rated invested mental effort and then made concurrent judgments of learning; for the other half, it was the other way around. Question order did not affect participants' concurrent judgments of learning for either of the two study strategies ($p \geq 0.091$). It did not affect their mental effort ratings for interleaved studying ($p = 0.646$) but there was an order effect for blocked studying: Averaged mental effort ratings for blocked studying were significantly higher ($M = 4.33$) when participants first judged their learning than when they first rated their mental effort ($M = 3.60$), $t(148) = 4.51$, $p < 0.001$. After completing all study units, participants engaged in a distractor task to clear working memory ("Henk buys 50 apples of two different varieties for 14 euros. We know that Henk paid € 4.00 for 10 apples of type A. How much do the apples of type B cost each?"). Next, participants rated their study strategy considerations for the second time (perceived effectiveness of and willingness to use each strategy).

Hereafter, participants were informed that the test was about to start and were asked to make a prediction about their performance on the test (prospective judgments of learning), for the painters studied via blocked studying and for the painters studied via interleaved studying. Thereafter, participants completed the test. The paintings on this test were presented on separate screens in a randomized order.

After the test, participants reflected on the two study strategies they just had used, with or without personalized feedback depending on their assigned condition, and then rated their study strategy considerations (perceived effectiveness and willingness to use) for the last time and briefly explained their considerations. Finally, participants were debriefed on the precise goal of the study and informed about the effectiveness of interleaved studying over blocked studying in category learning.

Analyses

Statistical analyses were conducted with R Studio 4.0.0 (R Core Team, 2020). The R code including the full list of used packages is available on our OSF project page. First, we conducted paired t-tests to check whether we would replicate findings from prior research that interleaved studying leads to higher mental effort, lower concurrent and prospective judgments of learning, yet yields higher learning outcomes (actual test performance) than blocked studying. Second, we used path analyses to investigate the direct and indirect correlational effects of students' mental effort on perceived effectiveness of each study strategy (via prospective judgments of learning), and on their willingness to use that strategy (via perceived effectiveness). Third, we constructed linear mixed models to evaluate how participants' study considerations changed over time and whether they were affected by the feedback intervention. For both study strategies, we created two models: One with perceived effectiveness and one with willingness to use as the outcome measure. In all models, the fixed effects were measurement moment (before study phase, after study

phase, and after test), feedback intervention (no/yes), and their interaction. Measurement moment was dummy-coded with the second moment (after study phase) as the reference category. Participants were specified as a random effect. We checked for multivariate outliers in each model: cases with a standardized residual at a distance greater than 2.5 standard deviations from 0 were considered an outlier—because all effects were still replicated after omitting these cases, results are reported on the complete dataset.

Results

Mental Effort and Concurrent Judgments of Learning

Figure 4 (left panel) shows the average of the mental effort ratings and concurrent judgments of learning after each study unit. Overall, in line with our expectations, students reported blocked studying as being less effortful ($M=3.97$, $SD=1.06$) than interleaved studying ($M=6.26$, $SD=1.19$), $t(149)=-21.93$, $p<0.001$, Cohen's $d=-1.79$, and felt they learned more during blocked studying ($M=6.18$, $SD=0.86$) than during interleaved studying ($M=3.96$, $SD=1.20$), $t(149)=20.64$, $p<0.001$, Cohen's $d=1.69$.

Prospective Judgments of Learning and Learning Outcomes

Figure 5 (left panel) shows, per study strategy, students' prospective judgments of learning (predicted test performance) and their learning outcomes (actual test performance) after the study phase. To test for significant differences, we used paired t -tests with a Bonferroni corrected significance level of $p<0.025$ (i.e., $0.05/2$). As expected, prospective judgments of learning were higher for blocked studying ($M=5.31$, $SD=1.83$) than for interleaved studying ($M=3.52$, $SD=1.92$), $t(149)=11.05$, $p<0.001$, Cohen's $d=0.90$, but actual learning outcomes were higher for interleaved studying ($M=4.83$, $SD=2.21$) than for blocked studying ($M=3.70$, $SD=2.11$), $t(149)=-6.36$, $p<0.001$, Cohen's $d=-0.52$.

Mental Effort as Cue for Study Considerations

Figure 6 (upper panel) graphically displays, per study strategy, the results of the path model analyses that tested whether students' mental effort during studying was a direct and/or indirect predictor of their considerations after studying about the effectiveness of the study strategies and their willingness to use them. Note that both models we tested were saturated and, therefore, there is no point in reporting any model fit statistics.

The amount of mental effort students reported during blocked studying was not directly related to how effective they perceived blocked studying to be. The model only showed a small indirect negative relationship of mental effort with perceived effectiveness (via prospective judgments of learning). The more mental effort

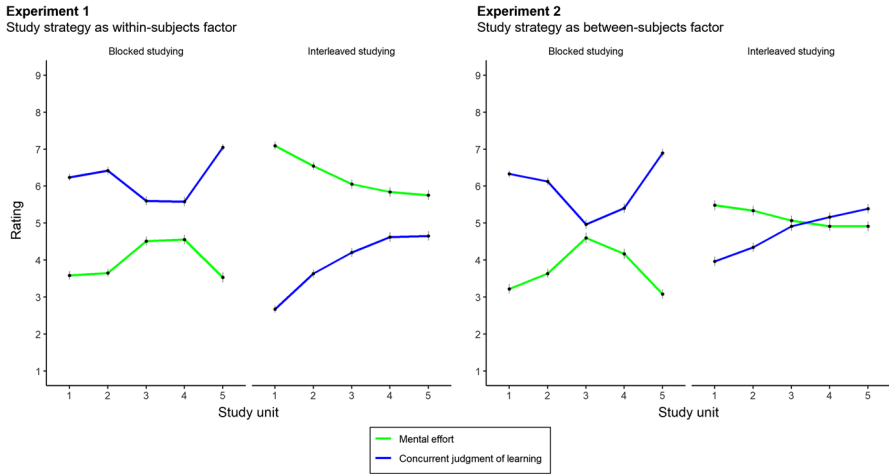


Fig. 4 Mental effort and concurrent judgments of learning. *Note.* Error bars represent standard errors

students had to invest during blocked studying, the lower their prospective judgments of learning, and the less effective they perceived this study strategy to be. Furthermore, students’ mental effort was not directly or indirectly (via perceived effectiveness) related to their willingness to use blocked studying. For interleaved studying, on the other hand, the amount of mental effort students reported was both

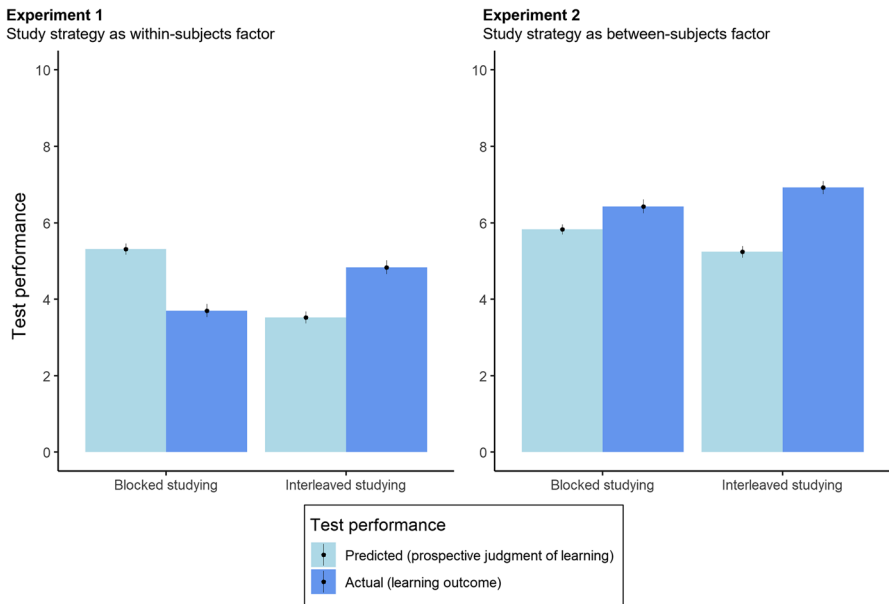


Fig. 5 Prospective judgments of learning (predicted test performance) and learning outcomes (actual test performance). *Note.* Error bars represent standard errors

directly and indirectly (via prospective judgments of learning) negatively related to how effective they perceived interleaved studying to be. In other words, even when controlling for the indirect effect via prospective judgments of learning, mental effort was still a negative predictor of perceived effectiveness. Moreover, mental effort was also indirectly (via perceived effectiveness) negatively related to students' willingness to use interleaved studying: The more effort students had to invest during interleaved studying, the less effective they perceived this study strategy to be and the less willing they were to use this strategy again.¹ Thus, in line with our hypotheses, perceived mental effort indeed played a role in students' study considerations, and this role seemed more influential for interleaved studying than for blocked studying.

Effects of Feedback on Study Considerations

Figure 7 (upper panel) displays how effective students perceived the study strategies to be and how willing they were to use them before the study phase, immediately after the study phase, and after the test. We used paired t-tests with a Bonferroni corrected significance level of $p < 0.013$ (i.e., $0.05/4$) to compare perceived effectiveness of and the willingness to use blocked vs. interleaved studying at the different time points. To test our hypothesis that feedback would affect students' study considerations over time, we conducted linear mixed model analyses on perceived effectiveness and willingness to use the study strategies (the results of the analyses can be found in Table 1). Here, we summarize the main findings.

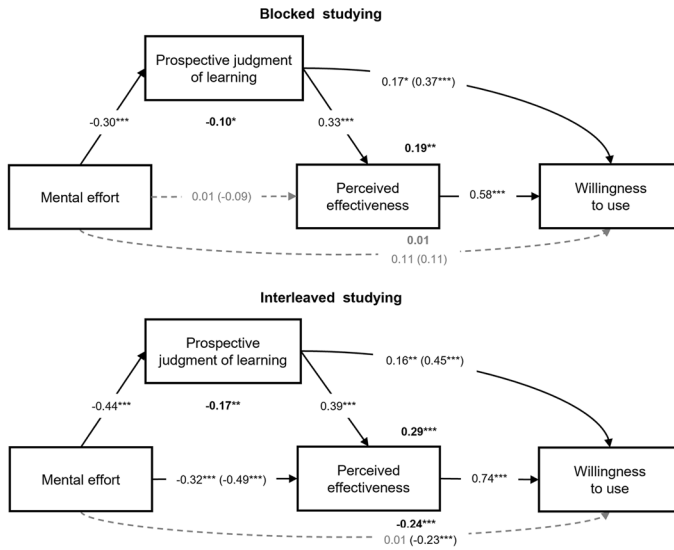
Prior to the study phase, students considered blocked studying to be more effective ($M = 6.75$, $SD = 1.48$) than interleaved studying ($M = 5.41$, $SD = 1.91$), $t(149) = 5.69$, $p < 0.001$, Cohen's $d = 0.46$. They were also more willing to use blocked studying ($M = 6.76$, $SD = 1.77$) than interleaved studying ($M = 5.13$, $SD = 1.99$), $t(149) = 5.92$, $p < 0.001$, Cohen's $d = 0.48$.

Immediately after the study phase (i.e., after having experienced the two strategies), these convictions were even stronger: The perceived effectiveness of blocked studying and the willingness to use it had increased significantly, whereas the perceived effectiveness of interleaved studying and willingness to use it had decreased significantly. Interestingly, completing the test seemed to have negatively affected students' considerations about blocked studying: Both perceived effectiveness of blocked studying and willingness to use it decreased significantly after the test, and this was not affected by the feedback intervention. The feedback intervention did affect students' considerations about interleaved studying in a positive direction: Those who received feedback after the test became significantly more positive about the effectiveness of interleaved studying, $\beta = 0.24$,

¹ Note that the non-significant relationships between mental effort and willingness to use a strategy do not imply that the two are not related to each other, but point to a statistical artefact, caused by strong correlations between willingness to use and perceived effectiveness. When testing for (in)direct effects of mental effort on willingness to use a strategy (via judgment of learning), we also find significant effects. This applied to both models (on blocked and interleaved studying).

Experiment 1

Study strategy as within-subjects factor



Experiment 2

Study strategy as between-subjects factor

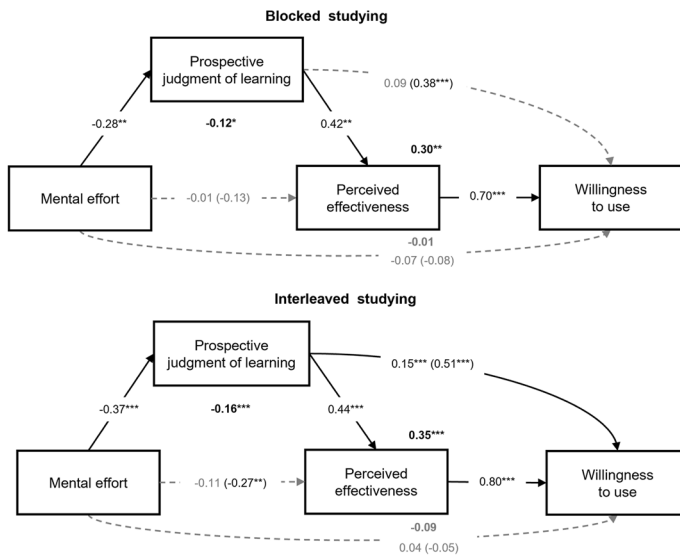


Fig. 6 Path analyses. *Note.* Beta values are standardized. Values in boldface represent indirect effects. Values in parentheses represent total effects. Grey values and dashed lines are non-significant effects. * $p < .05$, ** $p < .01$, *** $p < .001$. For completeness, we additionally tested these models with concurrent judgments of learning. Direction of all effects remained the same. Yet, in line with previous studies (for a meta-analysis, see Baars et al., 2020), the effect size of the correlation between mental effort ratings and the concurrent judgments of learning was larger than that of the correlation between mental effort ratings and the prospective judgments of learning, exp. 1 blocked: $\beta = -0.51, p < .001$; exp. 1 interleaved: $\beta = -0.68, p < .001$; exp. 2 blocked: $\beta = -0.53, p < .001$; exp. 2 interleaved: $\beta = -0.63, p < .001$

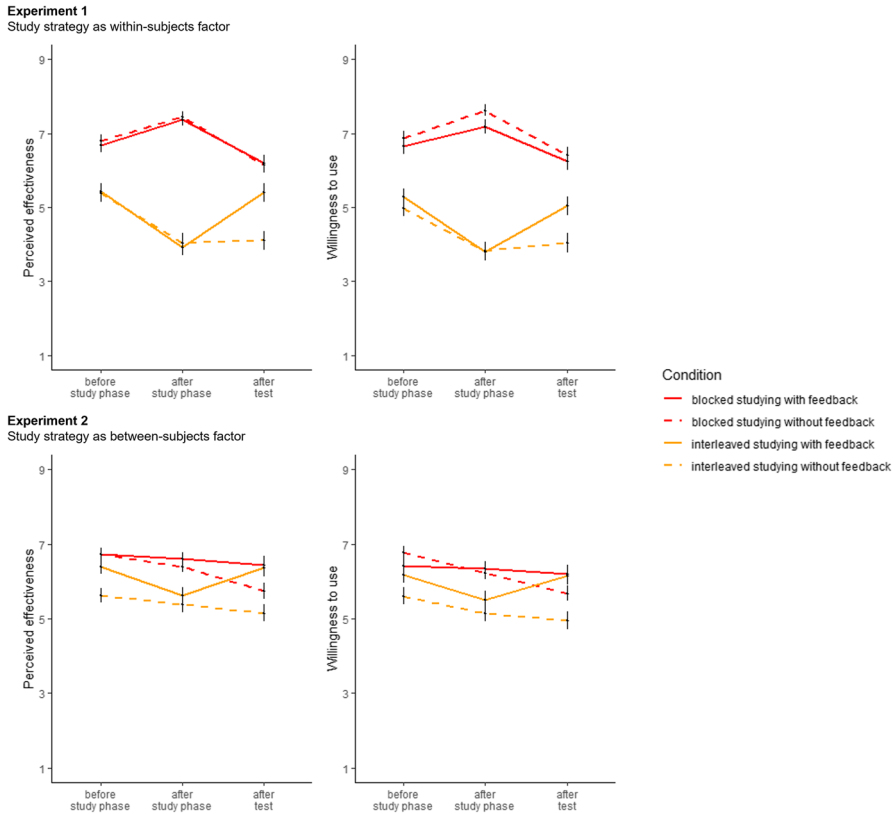


Fig. 7 Students' study strategy considerations about blocked and interleaved studying across measurement moments. *Note.* Error bars represent standard errors

$t(296) = 3.82, p < 0.001$, and were more willing to use it in the future than those who had not received feedback, $\beta = 0.17, t(296) = 2.56, p = 0.011$.

Yet, despite this significant increase, students who received the feedback still seemed to perceive blocked studying ($M = 6.20, SD = 1.79$) to be more effective than interleaved studying ($M = 5.40, SD = 2.22$), but this difference was not statistically significant (after correction for multiple testing), $t(74) = 2.04, p = 0.045$, Cohen's $d = 0.24$. Furthermore, they were still more willing to use blocked studying ($M = 6.24, SD = 1.95$) than interleaved studying ($M = 5.04, SD = 2.26$), $t(74) = 2.82, p = 0.006$, Cohen's $d = 0.33$.

Subgroup

The preference for blocked studying after feedback disappears, however, if we only look at subsample of students who actually performed better with interleaved studying ($n = 94$ out of 150; see Fig. 8). These exploratory analyses on those 94 students, showed a similar pattern of results over time as found in the whole sample, except

Table 1 Mixed models on students' perceived effectiveness of blocked and interleaved studying across measurement moments (experiment 1: study strategy as within-subjects factor)

Models	<i>B</i>	<i>S.E</i>	<i>df</i>	<i>t</i>	<i>p</i>	β
Models for blocked studying						
Perceived effectiveness~						
(Intercept)	7.45	0.18	413.61	41.66	<0.001	0.00
Measurement moment (T1 minus T2)	-0.64	0.23	296.00	-2.81	0.005	-0.19
Measurement moment (T3 minus T2)	-1.29	0.23	296.00	-5.69	<0.001	-0.38
Feedback (yes minus no)	-0.08	0.25	413.61	-0.32	0.752	-0.02
Measurement moment (T1 minus T2)×Feedback	-0.05	0.32	296.00	-0.17	0.868	-0.01
Measurement moment (T3 minus T2)×Feedback	0.12	0.32	296.00	0.37	0.709	0.03
Willingness to use~						
(Intercept)	7.63	0.20	396.06	37.84	<0.001	0.00
Measurement moment (T1 minus T2)	-0.76	0.25	296.00	-3.07	0.002	-0.20
Measurement moment (T3 minus T2)	-1.21	0.25	296.00	-4.90	<0.001	-0.32
Feedback (yes minus no)	-0.44	0.29	396.06	-1.54	0.123	-0.12
Measurement moment (T1 minus T2)×Feedback	0.23	0.35	296.00	0.65	0.517	0.05
Measurement moment (T3 minus T2)×Feedback	0.27	0.35	296.00	0.76	0.447	0.06
Models for interleaved studying						
Perceived effectiveness~						
(Intercept)	4.05	0.24	336.19	16.98	<0.001	0.00
Measurement moment (T1 minus T2)	1.33	0.26	296.00	5.10	<0.001	0.29
Measurement moment (T3 minus T2)	0.05	0.26	296.00	0.20	0.838	0.01
Feedback (yes minus no)	-0.12	0.34	336.19	-0.36	0.722	-0.03
Measurement moment (T1 minus T2)×Feedback	0.17	0.37	296.00	0.47	0.640	0.03
Measurement moment (T3 minus T2)×Feedback	1.41	0.37	296.00	3.82	<0.001	0.24
Willingness to use~						
(Intercept)	3.83	0.25	354.04	15.53	<0.001	0.00
Measurement moment (T1 minus T2)	1.16	0.28	296.00	4.15	<0.001	0.25
Measurement moment (T3 minus T2)	0.23	0.28	296.00	0.81	0.418	0.05
Feedback (yes minus no)	-0.03	0.35	354.04	-0.08	0.939	-0.01
Measurement moment (T1 minus T2)×Feedback	0.32	0.40	296.00	0.81	0.419	0.05
Measurement moment (T3 minus T2)×Feedback	1.01	0.40	296.00	2.56	0.011	0.17

T1 = before study phase, T2 = after study phase, T3 = after test

that the effect of feedback after the test is even stronger. Students who received feedback that they performed better with interleaved studying than with blocked studying ($n=42$ out of 75) did not consider blocked studying ($M=5.48$, $SD=1.76$) to be significantly more effective than interleaved studying ($M=6.33$, $SD=1.88$), $t(41)=-1.94$, $p=0.060$, Cohen's $d=-0.30$, and were not significantly more willing to use blocked studying ($M=5.43$, $SD=1.88$) than interleaved studying ($M=6.14$, $SD=1.95$), $t(41)=-1.47$, $p=0.148$, Cohen's $d=-0.23$. Yet, interestingly, they still did not significantly favor interleaved studying over blocked studying either.

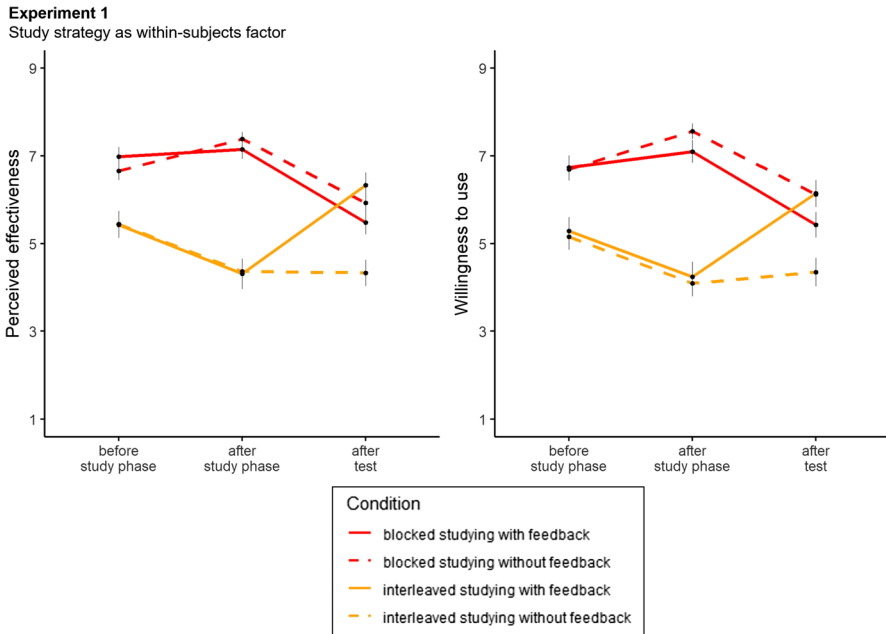


Fig. 8 Study strategy considerations of blocked and interleaved studying across measurement moments for those students who performed better with interleaved studying than with blocked studying ($n=94$). *Note.* Error bars represent standard errors

Discussion

In sum, we replicated prior findings showing that students typically report investing less mental effort and make higher judgments of learning with blocked studying, yet attain higher actual learning outcomes with interleaved studying. In addition, our path analyses suggest that students indeed use mental effort as a cue for their considerations about the effectiveness of each study strategy and for their willingness to use them again. Interestingly, mental effort played a more influential role in considerations about interleaved than about blocked studying. Furthermore, students' study strategy considerations changed over time. Considerations about blocked studying were not additionally affected by personal feedback on self-reported study experiences and actual learning outcomes. Feedback did positively affect students' considerations about interleaved studying, yet not to such an extent that students considered this strategy better than blocked studying. An important and interesting question is to what extent these findings will hold when students are not able to directly compare their experiences with the two different study strategies. Therefore, we conducted a second experiment with study strategy as between-subjects factor.

Experiment 2

With Experiment 2, we aimed to test the robustness of the findings from Experiment 1 and applied study strategy as between-subject factor instead of as within-subjects factor. In contrast to Experiment 1, participants in Experiment 2 could not directly compare their experiences with the two study strategies to each other. Thus, Experiment 2 will show to what extent findings regarding the role of mental effort from Experiment 1 will hold or might have been due to relative comparison (i.e., one strategy requiring more effort than the other).

This study was approved by the faculty's ethics committee of the first author. The study data and analyses scripts are stored on an Open Science Framework (OSF) page for this project, see <https://osf.io/5uzbq/>.

Method

Participants and Design

Experiment 2 had a $2 \times 2 \times 3$ mixed design with study strategy (blocked vs. interleaved) and feedback intervention (yes/no) as between-subjects factors, and measurement moment (before study phase, after study phase, and after test) as within-subjects factor.

Because of the path analyses, we again needed 150 participants, but this time per study strategy condition, so we recruited 300 participants. This was also more than sufficient to detect the interleaving effect ($N=210$) and the measurement-moment \times feedback condition within-between interaction for each study strategy ($N=152$) according to G*Power analyses with the same settings as in Experiment 1 (except that study strategy was a between-subjects factor here). In total, 299 Dutch students were recruited on Prolific Academic (www.prolific.ac) and paid £2.50 for their participation (one participant was lost to due to a technical error). Again, most participants were from a research university ($n=155$) or a university of applied sciences ($n=106$), 23 participants were enrolled in secondary vocational education, and 15 participants indicated "other". The sample consisted of 130 females, 165 males, and four participants who indicated "other". The average age was 22.63 years ($SD=3.81$).

Instead of studying ten different painters using both blocked and interleaved studying, each group of participants now studied five different painters, either using blocked studying or using interleaved studying. Again, half of the participants received feedback on their self-reported study experiences and their actual test performance (feedback condition), whereas the other half did not (no-feedback condition). Participants were randomly assigned to the one of the four experimental conditions: blocked studying with feedback ($n=76$); blocked studying without feedback ($n=73$); interleaved studying with feedback ($n=74$); and interleaved studying without feedback ($n=76$). All participants reported their opinion about the used strategy at three measurement moments: before the study phase, immediately after the study phase, and after the test.

Materials

Materials were the same as in Experiment 1; therefore, we only describe the differences in implementation compared to Experiment 1 below.

Learning Materials

In the study phase participants had to learn to recognize the work of five painters (Corot, Church, Constable, Turner, Cézanne) in five study units (BBBBB or IIIII).

Learning Outcomes

Learning outcomes were measured with a test that consisted of two unseen paintings by each of the five studied painters (10 in total). Reliability of the test score was acceptable, Cronbach's $\alpha = .62$.

Feedback Intervention

Participants in the feedback condition saw a personalized overview of their self-reported experiences (mental effort, concurrent and prospective judgments of learning) and their actual learning outcomes (test performance) for the strategy they used. As in Experiment 1, participants were encouraged to process the feedback by prompting them to write down what they could conclude based on the visual overview. Note that, in contrast to Experiment 1, participants could not compare their study experiences or learning outcomes as they only studied with one strategy in Experiment 2). Participants in the no-feedback condition were again prompted to take some time to reflect on the two study strategies that they just had used.

Choosing Between Blocked and Interleaved Studying

Once participants had reported their final considerations regarding the study strategy they used, we explained there was another study strategy that they could have used for the same learning task (i.e., blocked studying for those in the interleaved condition and interleaved studying for those in the blocked condition), and how this other strategy was different from the one that they just had used (but participants received no further information on what strategy was more effective). Then, we asked them which of these two study strategies they thought would be most effective for the current learning task and which of these two study strategies they would want to use for comparable learning tasks. To make sure they answered these questions attentively, we also asked for a brief explanation of their choices.

Procedure

The procedure was the same as the procedure of Experiment 1 with the only exception that participants used only one study strategy and that we asked them to make a final choice between blocked and interleaved studying. The order of mental effort and concurrent judgment of learning ratings was again counterbalanced between participants. None of the ratings were affected by question order, $ps \geq 0.503$. Likewise, the order of the perceived effectiveness and willingness to use questions was again counterbalanced between participants. One of these ratings was affected by question order. Perceived effectiveness of blocked studying at the third measurement moment was significantly higher ($M=6.47$) when participants first rated perceived effectiveness of blocked studying than when they first rated their willingness to use it ($M=5.75$), $t(147)=2.32$, $p=0.021$. None of the other ratings were affected by question order, $ps \geq 0.061$.

Analyses

The analyses were also highly similar to those used in Experiment 1, with the main difference being that study strategy was a between-subjects factor now (instead of a within-subjects factor). In addition, we conducted chi-square difference tests to identify pattern differences in participants' final choices between blocked and interleaved studying. We again checked for multivariate outlier cases in each mixed model—two non-significant effects became significant after omitting the multivariate outlier cases from the model. Results are reported on the complete dataset and we additionally reported the two deviant findings in a table note.

Results

Mental Effort and Concurrent Judgments of Learning

Figure 4 (right panel) shows the average self-reported mental effort and concurrent judgment of learning after each study unit. Overall, in line with our expectations and consistent with Experiment 1, students who used blocked studying perceived the study phase as being less effortful ($M=3.74$, $SD=1.22$) than students who used interleaved studying ($M=5.14$, $SD=1.43$), $t(297)=-9.10$, $p<0.001$, Cohen's $d=-1.05$. They also made higher concurrent judgments of learning ($M=5.94$, $SD=0.97$) than students who used interleaved studying ($M=4.75$, $SD=1.32$), Welch's $t(272.8)=8.87$, $p<0.001$, Cohen's $d=1.03$.

Prospective Judgments of Learning and Learning Outcomes

Figure 5 (right panel) shows, per study strategy, students' prospective judgments of learning (predicted test performance) and their learning outcomes (actual test performance) after the study phase. To test for significant differences, we used independent *t*-tests with a Bonferroni corrected significance level of $p < 0.025$ (i.e., $0.05/2$). As expected and again consistent with Experiment 1, students who had used blocked studying made higher prospective judgments of learning ($M = 5.83$, $SD = 1.60$) than students who had used interleaved studying ($M = 5.24$, $SD = 1.82$), $t(297) = 2.95$, $p = 0.003$, Cohen's $d = 0.34$. Numerically, the actual learning outcomes were higher for the interleaved study condition ($M = 6.92$, $SD = 2.06$) than for the blocked study condition ($M = 5.83$, $SD = 1.60$). However, in contrast to our expectations, this difference was not statistically significant, $t(297) = -1.99$, $p = 0.048$, Cohen's $d = -0.23$.

Mental Effort as Cue for Study Considerations

Figure 6 (bottom panel) graphically displays, per study strategy, the results of the path model analyses that tested whether students' mental effort during studying was a direct and/or indirect predictor of their considerations regarding the effectiveness of the study strategies and their willingness to use them. Note (again) that both models we tested were saturated and, therefore, there is no point in reporting any model fit statistics.

As in Experiment 1, the amount of mental effort students reported during blocked studying was only indirectly (not directly) negatively related to how effective they perceived blocked studying to be: The more mental effort students had to invest during blocked studying, the lower their prospective judgments of learning and the less effective they perceived this study strategy to be. Furthermore, as in Experiment 1, students' mental effort was not significantly related (directly, or indirectly via perceived effectiveness) to their willingness to use blocked studying. For interleaved studying, the amount of mental effort students reported was only indirectly (via prospective judgments of learning) negatively related to how effective they perceived interleaved studying to be. This differs from the findings from Experiment 1, where there was both a direct and indirect negative relation. Also, in contrast to Experiment 1, mental effort was not directly or indirectly (via perceived effectiveness) related to students' willingness to use interleaved studying.²

² Again, the non-significant relationships between mental effort and willingness to use a strategy do not imply that the two are not related to each other, but point to a statistical artefact, caused by strong correlations between willingness to use and perceived effectiveness. When testing for (in)direct effects of mental effort on willingness to use a strategy (via judgment of learning), we do find significant relationships between the two variables. This applied to both models (on blocked and interleaved studying).

Table 2 Mixed models on students' perceived effectiveness of blocked and interleaved studying across measurement moments (experiment 2: study strategy as between-subjects factor)

Models	<i>B</i>	<i>S.E</i>	<i>df</i>	<i>t</i>	<i>p</i>	β
Models for blocked studying						
Perceived effectiveness ~						
(Intercept)	6.40	0.19	338.80	33.78	<0.001	0.00
Measurement moment (T1 minus T2)	0.34	0.21	294.00	1.64	0.103	0.10
Measurement moment (T3 minus T2)	-0.64	0.21	294.00	-3.07	0.002	-0.18
Feedback (yes minus no)	0.21	0.27	338.80	0.78	0.433	0.06
Measurement moment (T1 minus T2)×Feedback	-0.21	0.29	294.00	-0.72	0.473	-0.04
Measurement moment (T3 minus T2)×Feedback	0.49	0.29	294.00	1.66	0.099	0.11
Willingness to use ~						
(Intercept)	6.23	0.20	282.17	31.04	<0.001	0.00
Measurement moment (T1 minus T2)	0.53	0.19	294.00	2.75	0.006	0.15
Measurement moment (T3 minus T2)	-0.55	0.19	294.00	-2.82	0.005	-0.15
Feedback (yes minus no)	0.12	0.28	282.17	0.44	0.664	0.04
Measurement moment (T1 minus T2)×Feedback ^a	-0.48	0.27	294.00	-1.77	0.078	-0.10
Measurement moment (T3 minus T2)×Feedback	0.40	0.27	294.00	1.48	0.140	0.09
Models for interleaved studying						
Perceived effectiveness ~						
(Intercept)	5.38	0.22	285.61	25.01	<0.001	0.00
Measurement moment (T1 minus T2)	0.24	0.21	296.00	1.13	0.259	0.06
Measurement moment (T3 minus T2)	-0.22	0.21	296.00	-1.07	0.286	-0.05
Feedback (yes minus no)	0.24	0.31	285.61	0.78	0.434	0.06
Measurement moment (T1 minus T2)×Feedback ^a	0.55	0.30	296.00	1.84	0.068	0.11
Measurement moment (T3 minus T2)×Feedback	0.97	0.30	296.00	3.24	0.001	0.19
Willingness to use ~						
(Intercept)	5.16	0.23	271.41	22.34	<0.001	0.00
Measurement moment (T1 minus T2)	0.45	0.22	296.00	2.08	0.039	0.10
Measurement moment (T3 minus T2)	-0.20	0.22	296.00	-0.92	0.361	-0.05
Feedback (yes minus no)	0.36	0.33	271.41	1.08	0.280	0.09
Measurement moment (T1 minus T2)×Feedback	0.21	0.31	296.00	0.70	0.485	0.04
Measurement moment (T3 minus T2)×Feedback	0.85	0.31	296.00	2.76	0.006	0.15

T1=before study phase, T2=after study phase, T3=after test. ^a This effect became significant after omitting the multivariate outlier cases

Effects of Feedback on Study Strategy Considerations

Figure 7 (bottom panel) displays how effective students perceived the study strategies to be and how willing they were to use them before the study phase, immediately after the study phase, and after the test. We used independent *t*-tests with a Bonferroni corrected significance level of $p < 0.013$ (i.e., $0.05/4$) to compare the perceived effectiveness of blocked vs. interleaved studying and the willingness to use these strategies at the different measurement moments. To test our hypothesis that

Table 3 Percentages of students' choosing for blocked or interleaved studying

	Blocked study condition		Interleaved study condition	
	No feedback (<i>n</i> = 73)	Feedback (<i>n</i> = 76)	No feedback (<i>n</i> = 76)	Feedback (<i>n</i> = 74)
Perceived effectiveness				
Choice for blocked studying (%)	61.64	73.68	64.47	41.89
Choice for interleaved studying (%)	38.36	26.32	35.5	58.11
	100.00	100.00	100.00	100.00
Willingness to use				
Choice for blocked studying (%)	61.64	75.00	64.47	39.19
Choice for interleaved studying (%)	38.36	25.00	35.53	60.81
	100.00	100.00	100.00	100.00

feedback would affect students' study considerations, we conducted linear mixed model analyses on perceived effectiveness and willingness to use the study strategies over time (the results of the analyses can be found in Table 2). Here, we summarize the main findings.

Prior to the study phase, students in the blocked study condition considered their strategy as more effective ($M=6.74$, $SD=1.44$) than students in the interleaved study condition, ($M=6.01$, $SD=1.76$), Welch's $t(286.32)=3.93$, $p<0.001$, Cohen's $d=0.45$. Students in the blocked study condition were also more willing to use their strategy ($M=6.58$, $SD=1.57$) than students in the interleaved condition ($M=5.89$, $SD=1.97$), Welch's $t(283.63)=3.38$, $p<0.001$, Cohen's $d=0.39$. Immediately after the study phase (i.e., after having experienced the strategy), students' convictions did not seem to change that much. To both study conditions applied that perceived effectiveness did not change significantly after the study phase. Willingness to use the strategy decreased significantly in both study conditions, but only very slightly.

Interestingly, completing the test seemed to negatively affect students' considerations in the blocked study condition: As in Experiment 1, both perceived effectiveness of blocked studying and willingness to use it decreased significantly after the test, and this was not affected by the feedback intervention. The feedback intervention did positively affect students' considerations in the interleaved study condition: those who received feedback after the test became significantly more positive about the effectiveness of interleaved studying, $\beta=0.19$, $t(296)=3.24$, $p=0.001$, and were more willing to use it than those who had not received feedback, $\beta=0.15$, $t(296)=2.76$, $p=0.006$. Yet, despite this significant increase, students who received the feedback in the interleaved condition did not consider their strategy to be significantly more effective ($M=6.36$, $SD=2.02$) than students in the blocked condition who received feedback ($M=6.45$, $SD=2.00$), $t(148)=0.25$, $p=0.802$, Cohen's $d=0.04$, nor were they significantly more willing to use their strategy again ($M=6.16$, $SD=2.15$) than students in blocked condition who received feedback ($M=6.21$, $SD=2.11$), $t(148)=0.14$, $p=0.890$, Cohen's $d=0.02$.

Choosing Between Blocked and Interleaved Studying

At the end of Experiment 2, we gave participants information on the study strategy they had not used and asked them to choose between blocked and interleaved studying in terms of perceived effectiveness and willingness to use the strategy. Table 3 shows an overview of students' choices per condition. Response patterns were highly similar for both questions. Students' choices in the blocked study condition were not affected by the feedback intervention, perceived effectiveness: $\chi^2(1)=1.95$, $p=0.162$; willingness to use: $\chi^2(1)=2.49$, $p=0.115$. Irrespective of whether they received feedback, the majority of students in the blocked study condition chose for blocked studying. In contrast, students' choices in the interleaved study condition were affected by the feedback intervention, perceived effectiveness: $\chi^2(1)=6.80$, $p=0.009$; willingness to use: $\chi^2(1)=8.62$, $p=0.003$. The majority of those who had received feedback chose for interleaved studying, whereas the majority of those who had not received feedback chose for blocked studying.

Discussion

In sum, blocked studying still resulted in lower mental effort ratings and higher judgments of learning even when students could not directly contrast their used study strategy to another strategy (i.e., Experiment 2). Yet, the size of the effects was smaller than when students did directly contrast the two study strategies (Experiment 1). Furthermore, our path analyses indicated that, also when students did not directly contrast their used study strategy directly to another strategy, mental effort was an indirect predictor of students' study considerations about their used strategies. Interestingly, the role of mental effort seemed approximately equally influential in both study strategies, which was in contrast with findings in Experiment 1 showing that experienced effort played a more (negative) influential role in considerations about interleaved studying than about blocked studying. Similar to Experiment 1, students' study strategy considerations changed over time and the feedback intervention only affected students in the interleaved condition (not in the blocked condition). Students who received the feedback after interleaved studying perceived this strategy as more effective and were more willing to use it than students who did not receive feedback. In addition, these students were also more likely to choose interleaved studying over blocked studying when asked at the end of the experiment.

General Discussion

Getting students to use effective but effortful study strategies is a challenge. One potential explanation for why students might avoid such strategies, is the fact that they are effortful. A high effort experience may not give students the feeling that they are learning while they are studying, which may affect their perceptions of the effectiveness of the strategy and their willingness to use it. In this case, providing

students with feedback about their experiences and learning outcomes might help to change their perceptions of the strategies. Therefore, the goal of the present study was to investigate the role of invested mental effort in students' considerations about a study strategy that is typically found to be more effortful yet also more effective (interleaved) versus a strategy that is less effortful and less effective (blocked), and the effect of personal feedback about their (subjective) study experiences and learning outcomes on their considerations.

To this end, we measured students' study experiences and considerations before, during, and after using interleaved *and* blocked studying (Experiment 1), and before, during, and after using interleaved *or* blocked studying (Experiment 2). Importantly for addressing the main questions regarding the role of effort and effect of feedback, both experiments replicated findings from prior research that, overall, at the group level, students reported higher effort investment and made lower judgments of learning during interleaved studying than during blocked studying (Kirk-Johnson et al., 2019; Onan et al., 2022). Yet, we only replicated the finding that students actually learned significantly more from interleaved studying than from blocked studying (as evidenced by their test performance) in Experiment 1. In Experiment 2, the difference in learning outcome, although numerically in the hypothesized direction, was not statistically significant, which was surprising given that many previous studies demonstrated the interleaving effect in both within-subjects and between-subjects designs (for a meta-analysis, see Brunmair & Richter, 2019).

In line with our expectations, our results indicated that students indeed seem to use mental effort as a cue in forming their perceptions about the strategies that they used. In both experiments and for both study strategies we found that mental effort during studying was negatively related to students' prospective judgments of learning and, via these judgments, to their perceived effectiveness of the strategy. Moreover, perceived effectiveness of the strategy correlated positively and strongly with students' willingness to use that strategy. These findings are in line with those of the study by Kirk-Johnson et al. (2019) and provide even stronger evidence for role of mental effort in students' perceptions about effective study strategies. That is, Kirk-Johnson et al. (2019) used the contrast in mental effort that students experienced during blocked and interleaved studying (i.e., difference scores) as predictor of students' decisions about what strategy to use. In our experiments, we used the actual mental effort ratings for each strategy as predictor for students' study considerations, and we also tested the role of effort when students could not directly compare the strategies (in Experiment 2). Thus, overall, our findings demonstrate that mental effort plays a role in students' study strategy perceptions, and the findings from Experiment 2 show that this is not simply due to the comparison that learners make when they contrast two strategies against each other.

Yet, despite the fact that we observed very similar data patterns in Experiment 1 and 2, the strength of relations was smaller in the second experiment. Differences in perceptions about interleaved studying versus blocked studying were smaller in Experiment 2, where students could not directly compare their experiences to another strategy. Also, in Experiment 1, where they could compare, we found that mental effort played a more influential role in students' considerations about

interleaved studying than about blocked studying: Students' judgments of learning only partially mediated the effect of mental effort on perceived effectiveness of interleaved studying, implying that mental effort also directly affected perceived effectiveness above and beyond students' judgments of learning in a negative direction. However, this direct correlational effect did not occur in Experiment 2. Thus, it could be that mental effort becomes a more influential cue for judging the effectiveness of a strategy (and willingness to use it) when students have an immediate other option that is less effortful. When not presented with this other option, students seem to be somewhat less negatively affected by mental effort.

Interestingly, perceptions of required effort might also be affected by the comparison, as illustrated by the fact that the mean ratings of invested effort during interleaved studying were lower in Experiment 2 than in Experiment 1. While we cannot rule out that the difference in the total amount of study materials (10 painters in Experiment 1, 5 in Experiment 2) might also at least partly explain this, Fig. 4 shows that a large difference between experiments is already present in the effort ratings of the first interleaved study unit. Since students in Experiment 1 always started with blocked studying, it seems likely that the switch to interleaved studying in the second unit, and thus, the contrast between blocked and interleaved studying, exacerbated their feelings of the amount of effort required.

As for the effects of feedback, we expected that presenting students with an overview of their (subjective) study experiences (mental effort and judgments of learning) next to their actual learning outcomes (test performance), would positively affect their perceptions of the effectiveness of interleaved studying and their willingness to use it in the future. As expected, the feedback did indeed positively affect students' perceptions about interleaved studying. Interestingly, however, it did not negatively affect perceptions about blocked studying. Moreover, even though students became more positive about interleaved studying due to the feedback, they still did not favor it over blocked studying, and this even applied when looking only at the subsample of students who attained a higher test performance with interleaved studying than blocked studying in Experiment 1. The fact that being confronted with the feedback that interleaved studying had a learning benefit for them, and that their concurrent judgments of learning and prediction of test performance (prospective judgments of learning) were incorrect, was still not sufficient to convince students to favor interleaved studying over blocked studying, underlines what a major challenge it is to get students to use effective but effortful study strategies. On the other hand, when forced to make a choice between blocked and interleaved studying, we did find that the majority of the students (60.8%) who received feedback after interleaved studying (Experiment 2) chose for interleaved studying rather than blocked studying, although it remains to be seen if they would actually use it in practice, when given the choice.

Limitations and Future Research

This study has some limitations that need to be taken in to account. First, surprisingly, the difference in learning outcome between blocked and interleaved studying,

although numerically in the hypothesized direction, was not statistically significant in Experiment 2. It is unlikely that we had insufficient power to detect this effect, as the meta-analysis by Brunmair and Richter (2019) showed that within-participants versus between-participants designs did not affect the size of the interleaving effect, and we had more than sufficient participants to detect a medium-sized effect (which, according to their meta-analysis, can be expected with our materials). A potential explanation could be that we used too few categories (i.e., participants studied five painters compared to a total ten in Experiment 1) or that there were too few test items to detect the interleaving effect (cf. Onan et al., 2022, in which a similar concern was raised based on pilot study results). However, the learning outcomes were numerically in the expected direction and the rest of the data pattern including mental effort and judgments of learning fully supported the assumptions on which we based our main hypotheses regarding the role of effort in students' experiences and study strategy considerations.

Second, in Experiment 1, participants always started with a blocked study unit followed by an interleaved study unit, which might have caused some sort of anchoring bias towards blocked studying: Participants might have perceived interleaved studying as even more effortful than they would have done if they had engaged in interleaved studying prior to blocked studying. Indeed, the results of Experiment 2 suggest that while mental effort is still higher for interleaved than for blocked studying when participants could not compare their experiences across study strategies, the effects seem to be smaller. Hence, the effects in Experiment 1 might have been smaller if participants had started with interleaved studying.

Third, the goal of Experiment 2 was to see whether we would replicate the findings from Experiment 1 when treating study strategy as a between-subjects factor. The fact that Experiment 2 largely replicated the findings from Experiment 1, shows the robustness of the effects. Note, however, that we cannot make any statistical comparisons or causal claims about the small differences in results between the two experiments since we did not randomly assign participants to Experiment 1 or Experiment 2. Nevertheless, because both participant samples were recruited on Prolific, using the same selection criteria, we do feel this descriptive comparison of findings between experiments is informative.

Fourth, we counterbalanced the order of the mental effort ratings and the concurrent judgments of learning and of the questions on perceived effectiveness and willingness to use the study strategies. These ratings were, by and large, not affected by question order. We consider the two significant effects we did find to be negligible given the many statistical comparisons across the two experiments; moreover, it is hard to meaningfully interpret them as they did not reveal a clear response pattern.

Fifth, we measured students' study strategy considerations in terms of the perceived effectiveness of a strategy and their willingness to use that strategy. However, our results indicate that students answered both questions almost identically, causing a statistical artefact in our path model analyses due to the high intercorrelations. Conceptually, it seems to imply that students are only willing to use strategies that they perceive as effective, which makes sense. However, we do not know what students' own definition is of an effective strategy (e.g., do they also think about

long-term and transfer effects?). It would be interesting in future research to use think aloud protocols or other verbal reporting methods to get more insight into this issue.

Another option would be to give students an actual choice over which strategy to use or even to allow them to alternate and study when they decide to do so. For instance, Tauber et al. (2013) showed that most students opted for blocked study when given a choice. Kornell and Vaughn (2018) found that the majority of participants judged blocked studying to more effective than interleaved studying when asked to choose between the two; interestingly, however, when asked to judge if blocked studying or a combination of blocked and interleaved studying would be more effective, the majority of participants chose the combination. Moreover, in their actual study choices, they also opted mainly for blocked trials, but did switch to interleaved studying sometimes. Lu et al. (2021) showed that students were inclined towards blocked studying but mainly decide to switch between categories (i.e., use interleaved study) when categories are highly similar, which is also when interleaved studying presumably is most effective to help students to learn to discriminate between them. It would be interesting to study whether and how students use effort as a cue for deciding when to use which strategy in future research (see also Onan et al., 2022). Using verbal report methods or choice paradigms combined with effort measures would also be a way to obtain more direct evidence of whether and how students use invested effort as a cue (and whether this differs in within and between-subjects designs), because another limitation of the present study is that we only have indirect evidence of the use of effort as a cue, via the path models.

Finally, another relevant next step would be to replicate our findings with other types of learning materials with which the effectiveness of interleaved study has also been demonstrated (e.g., math problem solving tasks; Rohrer & Taylor, 2007), as well as to replicate these findings with other types of effective but effortful study strategies (e.g., massed versus spaced studying, Cepeda et al., 2006; or retrieval practice versus restudy, Roediger & Butler, 2011). Such studies could provide further evidence for the proposed Effort-Monitoring-and-Regulation-framework (De Bruin et al., 2020).

Conclusion

Overall, our results provide further evidence that students use mental effort as a cue for their judgments of learning and, in turn, for their decisions to use a particular study strategy or not. This is relevant information to take into account when aiming to promote students' use for more effective and effortful study strategies. Interventions could focus on how to teach students to better interpret their effort experiences during studying (i.e., to help them understand when more effort is actually helpful to achieve better learning outcomes). Both of our experiments indicated that showing students their own mental effort experiences combined with their learning outcomes already positively affected considerations of interleaved studying, yet was not sufficient to fully convince students to actually prefer this strategy over blocked studying. Future research should focus on what type of interventions are more effective for improving students' interpretation of their effort experiences and for promoting their choices for effective study strategies.

Acknowledgements We thank members of the EARLI Emerging Field Group (EFG) “Monitoring and Regulation of Effort” for comments and suggestions on this study.

Funding This work was supported by the Jacobs Foundation.

Data Availability The data and the analyses scripts are stored on an Open Science Framework (OSF) page for this project, see <https://osf.io/5uzbq/>

Declarations

Conflict of Interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Baars, M., Wijnia, L., De Bruin, A., & Paas, F. (2020). The relation between students’ effort and monitoring judgments during learning: A meta-analysis. *Educational Psychology Review*, 32, 979–1002. <https://doi.org/10.1007/s10648-020-09569-3>
- Bjork, R. A., & Bjork, E. L. (2020). Desirable difficulties in theory and practice. *Journal of Applied Research in Memory and Cognition*, 9, 475–479. <https://doi.org/10.1016/j.jarmac.2020.09.003>
- Blasiman, R. N., Dunlosky, J., & Rawson, K. A. (2017). The what, how much, and when of study strategies: Comparing intended versus actual study behaviour. *Memory*, 25(6), 784–792. <https://doi.org/10.1080/09658211.2016.1221974>
- Brunmair, M., & Richter, T. (2019). Similarity matters: A meta-analysis of interleaved learning and its moderators. *Psychological Bulletin*, 145, 1029–1052. <https://doi.org/10.1037/bul0000209>
- Carlson, R. A., & Schneider, W. (1989). Acquisition context and the use of causal rules. *Memory & Cognition*, 17, 240–248. <https://doi.org/10.3758/BF03198462>
- Carlson, R. A., & Yaure, R. G. (1990). Practice schedules and the use of component skills in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 16, 484–496. <https://doi.org/10.1037/0278-7393.16.3.484>
- Carvalho, P. F., & Goldstone, R. L. (2017). The sequence of study changes what information is attended to, encoded, and remembered during category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 43, 1699–1719. <https://doi.org/10.1037/xlm0000406>
- Cepeda, N. J., Pashler, H., Vul, E., Wixted, J. T., & Rohrer, D. (2006). Distributed practice in verbal recall tasks: A review and quantitative synthesis. *Psychological Bulletin*, 132, 354–380. <https://doi.org/10.1037/0033-2909.132.3.354>
- Cross, E. S., Schmitt, P. J., & Grafton, S. T. (2007). Neural substrates of contextual interference during motor learning support a model of active preparation. *Journal of Cognitive Neuroscience*, 19, 1854–1871. <https://doi.org/10.1162/jocn.2007.19.11.1854>
- De Bruin, A. B. H., Roelle, J., Carpenter, S. K., & Baars, M. (2020). Synthesizing cognitive load and self-regulation theory: A theoretical framework and research agenda. *Educational Psychology Review*, 32, 903–915. <https://doi.org/10.1007/s10648-020-09576-4>
- De Croock, M. B. M., Van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). High versus low contextual interference in simulation-based training of troubleshooting skills: Effects on transfer performance and invested mental effort. *Computers in Human Behavior*, 14, 249–267. [https://doi.org/10.1016/S0747-5632\(98\)00005-3](https://doi.org/10.1016/S0747-5632(98)00005-3)

- Eglington, L. G., & Kang, S. H. K. (2017). Interleaved presentation benefits science category learning. *Journal of Applied Research in Memory and Cognition*, 6, 475–485. <https://doi.org/10.1016/j.jar-mac.2017.07.005>
- 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. *Behavior Research Methods*, 39(2), 175–191. <https://doi.org/10.3758/BF03193146>
- Guadagnoli, M. A., & Lee, T. D. (2004). Challenge Point: A framework for conceptualizing the effects of various practice conditions in motor learning. *Journal of Motor Behavior*, 36, 212–224. <https://doi.org/10.3200/JMBR.36.2.212-224>
- Helsdingen, A. S., Van Gog, T., & Van Merriënboer, J. J. G. (2011). The effects of practice schedule on learning a complex judgment task. *Learning and Instruction*, 21(1), 126–136. <https://doi.org/10.1016/j.learninstruc.2009.12.001>
- Jacoby, L. L. (1978). On interpreting the effects of repetition: Solving a problem versus remembering a solution. *Journal of Verbal Learning and Verbal Behavior*, 17, 649–667. [https://doi.org/10.1016/S0022-5371\(78\)90393-6](https://doi.org/10.1016/S0022-5371(78)90393-6)
- Khan, F. S., Beigpour, S., Van de Weijer, J., & Felsberg, M. (2014). Painting-91: A large scale database for computational painting categorization. *Machine Vision and Applications*, 25, 1385–1397. <https://doi.org/10.1007/s00138-014-0621-6>
- Kirk-Johnson, A., Galla, B. M., & Fraundorf, S. F. (2019). Perceiving effort as poor learning: The misinterpreted-effort hypothesis of how experienced effort and perceived learning relate to study strategy choice. *Cognitive Psychology Journal*, 115, 101237. <https://doi.org/10.1016/j.cogpsych.2019.101237>
- Koriat, A. (1997). Monitoring one's own knowledge during study: A cue-utilization approach to judgments of learning. *Journal of Experimental Psychology: General*, 126, 349–370. <https://doi.org/10.1037/0096-3445.126.4.349>
- Kornell, N., & Bjork, R. A. (2007). The promise and perils of self-regulated study. *Psychonomic Bulletin & Review*, 14, 219–224. <https://doi.org/10.3758/BF03194055>
- Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the “enemy of induction”? *Psychological Science*, 19, 585–592. <https://doi.org/10.1111/j.1467-9280.2008.02127.x>
- Kornell, N., & Vaughn, K. E. (2018). In inductive category learning, people simultaneously block and space their studying using a strategy of being thorough and fair. *Archives of Scientific Psychology*, 6, 138–147. <https://doi.org/10.1037/arc0000042>
- Lee, T. D., & Magill, R. A. (1983). The locus of contextual interference in motor-skill acquisition. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 9, 730–746. <https://doi.org/10.1037/0278-7393.9.4.730>
- Lu, X., Penney, T. B., & Kang, S. H. K. (2021). Category similarity affects study choices in self-regulated learning. *Memory & Cognition*, 49, 67–82. <https://doi.org/10.3758/s13421-020-01074-w>
- Magill, R. A., & Hall, K. G. (1990). A review of the contextual interference effect in motor skill acquisition. *Human Movement Science*, 9, 241–289. [https://doi.org/10.1016/0167-9457\(90\)90005-X](https://doi.org/10.1016/0167-9457(90)90005-X)
- McCabe, J. (2011). Metacognitive awareness of learning strategies in undergraduates. *Memory & Cognition*, 39, 462–476. <https://doi.org/10.3758/s13421-010-0035-2>
- Onan, E., Wiradhany, W., Biwer, F., Janssen, E. M., & De Bruin, A. B. H. (2022). Growing out of the experience: How subjective experiences of effort and learning influence the use of interleaved practice. *Educational Psychology Review*. <https://doi.org/10.1007/s10648-022-09692-3>
- Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: A cognitive-load approach. *Journal of Educational Psychology*, 84, 429–434. <https://doi.org/10.1037/0022-0663.84.4.429>
- Paas, F. G. W. C., & Van Merriënboer, J. J. G. (1994). Variability of worked examples and transfer of geometrical problem-solving skills: A cognitive-load approach. *Journal of Educational Psychology*, 86, 122–133. <https://doi.org/10.1037/0022-0663.86.1.122>
- R Core Team. (2020). *R: A Language and environment for statistical computing*. <https://www.R-project.org/>
- Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, 15, 20–27. <https://doi.org/10.1016/j.tics.2010.09.003>
- Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, 17, 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>

- Rohrer, D., & Taylor, K. (2007). The shuffling of mathematics problems improves learning. *Instructional Science*, 35, 481–498. <https://doi.org/10.1007/s11251-007-9015-8>
- Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science*, 3, 207–218. <https://doi.org/10.1111/j.1467-9280.1992.tb00029.x>
- Schneider, V. I., Healy, A. F., Ericsson, K. A., & Bourne, L. E. (1995). The effects of contextual interference on the acquisition and retention of logical rules. In A. F. Healy & L. E. Bourne (Eds.), *Learning and memory of knowledge and skills: Durability and specificity* (pp. 95–131). Sage.
- Schneider, V. I., Healy, A. F., & Bourne, L. E. (1998). Contextual interference effects in foreign language vocabulary acquisition and retention. In A. F. Healy & L. E. Bourne (Eds.), *Foreign language learning: Psycholinguistic studies on training and retention* (pp. 78–90). Erlbaum.
- Schneider, V. I., Healy, A. F., & Bourne, L. E. (2002). What is learned under difficult conditions is hard to forget: Contextual interference effects in foreign vocabulary acquisition, retention, and transfer. *Journal of Memory and Language*, 46, 419–440. <https://doi.org/10.1006/jmla.2001.2813>
- Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting structural equation modeling and confirmatory factor analysis results: A review. *The Journal of Educational Research*, 99(6), 323–338. <https://doi.org/10.3200/JOER.99.6.323-338>
- Shea, J. B., & Morgan, R. L. (1979). Contextual interference effects on the acquisition, retention, and transfer of a motor skill. *Journal of Experimental Psychology: Human Learning and Memory*, 5, 179–187. <https://doi.org/10.1037/0278-7393.5.2.179>
- Simon, D. A. (2007). Contextual interference effects with two tasks. *Perceptual and Motor Skills*, 105, 177–183. <https://doi.org/10.2466/pms.105.1.177-183>
- Simon, D. A., & Bjork, R. A. (2001). Metacognition in motor learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(4), 907–912. <https://doi.org/10.1037/0278-7393.27.4.907>
- Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative review. *Perspectives on Psychological Science*, 10, 176–199. <https://doi.org/10.1177/1745691615569000>
- Tauber, S. K., Dunlosky, J., Rawson, K. A., Wahlheim, C. N., & Jacoby, L. L. (2013). Self-regulated learning of a natural category: Do people interleave or block exemplars during study? *Psychonomic Bulletin & Review*, 20, 356–363. <https://doi.org/10.3758/s13423-012-0319-6>
- Van Peppen, L. M., Verkoeijen, P. P. J. L., Kolenbrander, S. V., Heijltjes, A. E. G., Janssen, E. M., & Van Gog, T. (2021). Learning to avoid biased reasoning: Effects of interleaved practice and worked examples. *Journal of Cognitive Psychology*, 33, 304–326. <https://doi.org/10.1080/20445911.2021.1890092>
- Zechmeister, E. B., & Shaughnessy, J. J. (1980). When you know that you know and when you think that you know but you don't. *Bulletin of the Psychonomic Society*, 15, 41–44. <https://doi.org/10.3758/BF03329756>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Eva M. Janssen¹  · Tamara van Gog¹  · Laura van de Groep¹ · Anne Jóia de Lange¹ · Roosmarijn L. Knopper¹ · Erdem Onan²  · Wisnu Wiradhany³  · Anique B. H. de Bruin² 

¹ Department of Education, Utrecht University, Heidelberglaan 1, 3584 CS Utrecht, the Netherlands

² Department of Educational Development and Research, School of Health Professions Education, Maastricht University, Maastricht, the Netherlands

³ Faculty of Psychology, Atma Jaya Catholic University of Indonesia, Jakarta, Indonesia