# Reframing Effort to Improve Learners' Study Strategy Choices

Connie M. Fralick

Honors Thesis

Department of Psychology, The University of Texas at Austin

Faculty Advisor: Veronica Yan, Ph.D.

December 2021

**Abstract**

Even though the study strategy of mixing study topics ("interleaving") is better for learning than studying one topic at a time ("blocking"), people consistently underestimate the former and favor the latter. A barrier to interleaving is that it feels more difficult than blocking. People gauge the efficacy of study strategies by how easy they feel, inaccurately viewing lower levels of mental effort as a signal of good learning. However, reframing the study situation as something that the student *chooses* to invest in versus something that is *required* of them has been found to counter the view that easier learning is better learning. In the current study, I compared the effects of using this voluntary-versus-required framing on participants' decisions to interleave or block during studying. In the effort-as-voluntary condition, the more effortful that participants thought interleaving was compared to blocking, the more effective they perceived interleaving to be and more likely they were to choose to interleave in the future. This pattern was the opposite of that found in the effort-as-required condition. However, reframing effort did not significantly change the study strategy decisions made.

*Keywords*: interleaving, blocking, framing, misinterpreted effort hypothesis, heuristics

## Reframing Effort to Improve Learners' Study Strategy Choices

People generally want to learn in the most effective and efficient way possible, but prior findings indicate that they are not the best judges of what study strategies are most effective. They have difficulty discerning which strategies help them learn, and which strategies only make people *feel* like they help them learn. One of the most pervasive study strategy misconceptions involves interleaving versus blocking. "Blocking" is a study strategy in which items are studied by type or organized into similar groups (e.g., *aaabbbccc*), whereas "interleaving" is a strategy in which one studies items intermixed (e.g., *abcabcabc*). Students tend to choose blocking when given the choice, but many studies show that it is the less effective strategy (Yan et al., 2017; Carvalho et al., 2016; Tauber et al., 2013). Only one other study (Yan et al., 2016), however, has investigated the central question of this present study: how can one influence learners to choose interleaving over blocking, the study strategy more effective for their own learning?

### The Interleaving Effect: Generalizability, Mechanisms, and Boundary Conditions

Prior findings support interleaving's advantages over blocking and show that it is robust and generalizable (see Brunmair et al., 2019 for a meta-analytic review). Interleaved practice has been shown to be more effective than blocked practice for motor skills (e.g., surgical techniques, Goldin et al., 2014; piano melodies, Abushanab & Bishara, 2013; baseball batting practice, Hall et al. 1994), visual category learning (e.g., recognizing different artist styles, Kornell & Bjork, 2008; Kang & Pashler, 2012; Zulkiply & Burt, 2013; bird species, Kirk-Johnson et al., 2019, Wahlheim et al., 2012; butterfly species, Birnbaum et al., 2013), and educational and conceptual learning (e.g., mathematics, Ostrow et al., 2015, Taylor & Rohrer, 2010; statistics, Sana et al., 2017, 2018; science, Eglington & Kang, 2017; Sana & Yan, 2021).

Kornell & Bjork (2008) was the first to test the application of the interleaving effect in regards to recognizing various artist styles, which is the learning task used in the present study. In their Experiment 1b, participants viewed six different paintings for each of 12 artists. Some participants viewed these paintings blocked ("massed by artist") while others viewed the paintings interleaved (intermixed). After the initial exposure participants were shown new paintings by the 12 artists and were asked to determine from a list of the artist names which artist had created the painting in question. In line with previous findings on interleaving and those that followed, it was found that participants performed significantly better on the test if they had studied the paintings interleaved.

Although there has been a positive overall effect of interleaving across the literature, there are some exceptions and limitations. These boundary conditions reveal insights into why interleaving may be more effective. Specifically, interleaving facilitates between-category discrimination. However, in cases in which between-category discrimination is trivial, or in which within-category similarities are much more difficult to notice, blocking benefits may arise (e.g., Brunmair & Richter, 2019; Carvalho & Goldstone, 2014; Sorensen & Woltz, 2016). Perhaps as a result of these discrimination processes, Brunmair & Richter (2019) found that visual stimuli showed the strongest interleaving effects whereas word stimuli showed the weakest. Moreover, results from Yan & Sana (2021) suggest that when sequencing decisions can be made at different levels (e.g., concepts within a subject, multiple subjects), an intermediate amount of interleaving may be optimal: interleaving concepts but not subjects, or interleaving subjects but not concepts; interleaving at both the concept and subject level was no different from blocking at both levels.

**People Often Fail to Recognize the Interleaving Benefit**

Few learners know of interleaving's effectiveness as a study strategy (Kornell & Bjork, 2008; Kornell et al., 2010; Yan et al., 2017; Zulkiply et al., 2012). For example, McCabe (2010) found that students rated interleaving as one of the *least* effective on a list of various study strategies. This belief does not appear to be limited to students. Rohrer, Dedrick & Hartwig (2020) analyzed six of the most representative and popular mathematics textbooks used in the United States, with a total of 13,505 practice problems, and found only 9.7% of the problems were interleaved, and hence that 90.3% were not. This lack of exposure to interleaving could partially inform student beliefs about blocking and interleaving. Regardless, Rohrer, Dedrick, and Hartwig's study illustrates the collective misunderstanding on the part of both students and educators.

Beyond misinformed beliefs about blocking and interleaving, blocking in lieu of interleaving is also the main choice for learners (Tauber et al., 2013; Carvalho et al., 2016). Yan et al. (2017) found that people tend to choose blocking as opposed to interleaving when constructing their own study schedules. Participants did not think to interleave or to create a hybrid part-blocking, part-interleaving schedule when not prompted with the possibility. In other words, people incorrectly intuit interleaving as less helpful than blocking. Thus, convincing students to make better study choices requires a measure of intervention.

**How Can We Encourage Better Study Strategy Decision-Making?**

Effective interventions on decision-making appear to depend on the study strategies involved. For the study strategy of generation, in which people better remember material they helped produce rather than passively receiving information, DeWinstanley and Bjork (2004) found that a single experience of the benefits of generation regarding a text passage was enough for participants to apply what they had learned to a second text passage. Tullis et al. (2013) found

that for participants to appreciate the benefits of self-testing as a study strategy as opposed to restudying, it required not just an experience of the strategy, but also explicit guidance and feedback to show they performed better matching word pairs after self-testing. The only study that has investigated changing student beliefs about interleaving, Yan et al. (2016), required even more than this.

To counter student bias toward blocking, Yan et al. (2016) not only provided participants with an experience to compare the two strategies (experience-based debiasing) regarding learning artist style recognition in a paradigm and stimuli set similar to Kornell & Bjork (2008), but also introduced participants to theoretical knowledge about why interleaving was more effective (theory-based debiasing). It was found that individual experience and theory-based informational debiasing were each, on their own, unable to overcome participant belief that blocking was the better strategy. They then attempted a much more extensive debiasing intervention, which involved explaining why interleaving is more effective than blocking and establishing a norm that interleaving was better for 90% of learners, pointing out the connection between study schedule and learners' own performance, and highlighting the differences between the two study strategies by separating the blocked and interleaved schedules to better illustrate the students' personal performance improving with the interleaved schedule. Even so, only just over half of the students come to appreciate the value of interleaving; a large minority still denied that interleaving was better for them. Yan et al. (2016) concluded that learners resist recognizing interleaving is better for learning than blocking for a number of reasons: 1) pre-existing assumptions that blocking is better, 2) the fluency experienced when studying with blocking, and 3) learner exceptionalism (when informed of the broad findings on the most effective study strategies, they believe it does not apply to themselves).

Whereas Yan and colleagues' (2016) intensive intervention focused on convincing students of interleaving's effectiveness, Do and Lee (2020) may suggest an alternate path to get students to make better decisions for their own learning. Their study used the same set up as the original Kornell and Bjork (2008) blocking versus interleaving study on artist painting recognition (12 artists, six paintings each for initial study), however, they also asked participants to make a judgment of learning about the percentage of questions they would get correct if tested on the paintings. Participants then took an interim test and were given a choice to study a new set of artists in either an interleaved or a blocked sequence. About two-thirds of participants chose to block their study of the second set; however, those who scored higher on the interim test and had higher estimates of their learning were more likely to choose to interleave at the second opportunity than those who performed more poorly and had lower estimates of learning. One way to interpret these findings is that giving learners success under interleaved practice is important for increasing the likelihood that they choose to interleave in the future. Do and Lee (2020) also suggested that the participants that chose to interleave may have been more confident in their learning overall. Another possibility, however, is that they merely continued with whatever strategy they were using if it appeared to work. This possibility should be explored further, for although Yan et al.'s (2016) findings illustrated that it takes a great deal of convincing for students to choose the best course of action for their learning, Do & Lee's (2020) findings indicate that students do not need to be convinced by others to choose interleaving. Although it would be preferable for students to understand the reasons they should choose certain study strategies over others, it is perhaps more important that they choose said strategies at all. Do and Lee's (2020) findings indicate that one could lead students to rely on alternative mental shortcuts

("heuristics") to influence student decision-making for the better instead of appealing to student's logic or engaging in extensive interventions.

**System 1 and System 2 Thinking**

Two types of thinking that Kahneman (2011) describes provide a useful framework to understand an alternative route to persuade students to make better learning choices. According to Kahneman, the vast majority of our thinking relies on "System 1" thinking; the unconscious, automatic, heuristic-based thinking that makes daily living easier. "System 2" thinking, meanwhile, is conscious, effortful, logical, and reasoning-based, the kind of thinking employed when carefully thinking through complex decisions. Whereas Yan et al. (2016) tried to convince students of the merits of interleaving by appealing to System 2 thinking, the present study investigates the possibility of using System 1 thinking to encourage students to choose more effective study strategies.

Reliance on System 1 thinking is a major reason why students tend to prefer blocking over interleaving. Interleaving is more effortful and feels more difficult than blocking; Yan and colleagues (2016) found that participants viewed their learning of the blocked artists as better than that of the interleaved artists throughout the study phase. These experiences of difficulty and effort often act as a System 1 signal that the strategy is not effective. Yan and colleagues' (2016) theory-based debiasing messages targeted toward System 2 thinking were unable to override participant System 1 reliance on ease as a cue for good learning. It seemed participants were relying on the heuristic that "if it's hard, it must not be good for learning," or what has elsewhere been referred to as the "misinterpreted effort" heuristic (Kirk-Johnson et al., 2019).

Kirk-Johnson and colleagues (2019) studied the misinterpreted effort hypothesis. In one experiment, they examined whether participants would choose to study photos of bird species via

blocking or interleaving after experiencing both study strategies in counterbalanced order. They found that a majority of participants (68%) chose to block and that this was connected to participant perception of the difficulty and efficacy of the study strategies. The more effortful participants found a strategy, the less effective they perceived it to be and the less likely they were to choose it for future study. In other words, the relationship between effort and strategy choice was mediated by perceived effectiveness.

**Leveraging Heuristics to Encourage Better Learning**

Situational framing can shift reliance away from the conflation of ease and efficacy that Kirk-Johnson et al. (2019) found that many people engage in. For example, Koriat et al. (2014) demonstrated that a subtle wording change to reframe effort can reverse the traditional "easy learning equals good learning" heuristic association. In their study, participants who were asked to imagine that they were studying for an exam were randomly assigned to one of two conditions, one in which study effort was framed as a choice and one in which study effort was framed as a requirement. In the choice condition, they were informed that people choose to invest greater effort and time to study certain areas over others. In the required condition, they were told that some topics necessitate greater study than others. Participants then studied four brief stories in a self-paced manner, as if for an upcoming exam. After reading the passages, participants in the effort-as-voluntary and effort-as-required conditions rated the effort they "chose to invest" in reading the passages or how much effort the passages "required," respectively. Participants in both conditions were then asked to estimate the likelihood of correctly answering questions about the passages, and then took a test on the passages.

Koriat et al. (2014) found that in the condition in which effort was framed as required, participants' judgments of learning showed a similar pattern as that which Kirk-Johnson and

colleagues (2019) described above. That is, the more effortful participants rated the study phase, the lower their judgments of learning. Participants seemed to be using a heuristic that effortful learning is not effective learning. However, in the condition which framed effort as a choice, participants' judgments of learning showed the exact opposite pattern: the more effortful they rated the study phase, the higher their judgments of learning. Participants seemed to be using a heuristic that effortful learning indicates effective learning.

In other words, when effort was framed as something required of a person, it implied the material was difficult, and more effortful experiences were perceived as leading to worse learning. However, if effort was framed as something the learner chose to invest into their studying, then more effortful experiences were perceived to lead to *better* learning. Koriat et al. (2014), however, did not examine whether this type of reframing can also affect how learners view the efficacy of study strategies. Nevertheless, the study suggests that the framing of mental effort in studying may be one way to leverage heuristics to reverse the misinterpreted effort hypothesis and potentially encourage learners to engage in more challenging and more effective learning strategies. Thus, the present study examined if this method of reframing effort to reverse the misinterpreted effort hypothesis would apply to the context of interleaving versus blocking study items, and if this new perspective would influence actual study strategy choices and make learners more likely to choose to interleave.

**Conclusion**

The prior literature shows that learners frequently do not choose the most optimal study strategies for their own learning and that these suboptimal decisions may stem from a mistaken assumption that "good" learning should feel easy. Blocking leads to an increased sense of fluency whereas interleaving feels disfluent (Kirk-Johnson et al., 2019; Yan et al., 2016). Using

the misinterpreted effort heuristic, learners rely on this feeling of ease or fluency to determine the effectiveness of a study strategy and how much they have learned. The perceived effectiveness of study strategies subsequently affects study strategy choice (Kirk-Johnson et al., 2019). Previous attempts to encourage interleaving such as Yan et al. (2016) have tried to use logical reasoning (System 2) to change learners' beliefs, but it may be more effective to leverage heuristics instead (System 1; Do & Lee, 2020; Koriat et al., 2014).

Although it may not be possible to change the experience of effort when interleaving, perhaps it could be possible to change how that effort is interpreted. In the present study, we examined how perceived effort, perceived effectiveness of study schedule, and study strategy choice are affected by an effort framing manipulation similar to that of Koriat et al. (2014). Specifically, by framing effort as a choice and an investment, this study tested whether shifting student focus away from the heuristic of "easy study equals good learning" and towards that of "effortful study equals good learning" could affect how perceived effort is related to perceived effectiveness, and whether it would cause more participants to interleave when given the choice.

We hypothesized that the likelihood students would choose to interleave for a hypothetical future test would depend on the effort framing of the experiment. Namely, that students in the effort framed as voluntary condition would be more likely to choose to interleave than students in the effort framed as required condition. Given previous findings on the tendency to heavily favor blocking over interleaving when presented with a choice between the two (Yan et al., 2017; Carvalho et al., 2016; Tauber et al., 2013), it was predicted that the majority would still choose to block. However, we predicted that significantly more students in the voluntary framing condition would choose to interleave students in the requirement framing condition.

## Method

### Experimental Design Overview

The present study was modeled after Kirk-Johnson et al. (2019) and administered via a Qualtrics survey. The study had a 2x2 mixed design, in which effort framing condition (effort-as-voluntary or effort-as-required) was manipulated between participants, and study schedule (blocked vs. interleaved) was manipulated within participants. The order in which participants experienced blocked and interleaved schedules was counterbalanced. Participants were introduced to both examples of paintings by the artists they would study and examples of interleaving and blocking with filler artists before moving onto the main study section. After each study session, participants answered questions regarding that session (framed in terms of either *choosing to invest* effort in the study task or the study task *requiring effort from* them). After studying using interleaving and blocking and answering their respective questionnaires, participants were asked to compare their experiences with the two study strategies and decide which strategy they would use to study similar material in the future. The experiment closed with a test over the studied material and several demographics items.

The primary independent variable of interest was effort framing—whether effort was phrased as something participants chose to invest in or something that was required of them. The dependent variables were perceptions of effort, perceptions of learning, and study strategy choice. We predicted that framing effort as voluntary would reverse the misinterpreted effort hypothesis, in which greater effort would be associated with greater learning and a higher likelihood of choosing interleaving. We predicted that interleaving would be perceived as more effortful across conditions, but that in the effort-as-voluntary condition, perceptions of learning

would be higher for more effortful strategies, and they would be more likely to choose to interleave than participants in the effort-as-required condition.

## Participants

A total of 173 undergraduates (99 females, 1 declined to disclose; age range: 18-62, $M_{age}$= 19.66, *SD* = 3.55) were recruited from a research participant pool of students enrolled in introductory psychology courses at The University of Texas at Austin. They received course credit for their participation. The sample size was determined using G*Power analysis using the method of Gelman (2018). Thus, 100 participants total was the minimum number of participants we aimed to recruit. However, the study remained open to participants through the end of the academic semester, and participant count easily surpassed the minimum target. Participants provided their informed consent and were randomly assigned to either the requirement framing condition (*n* = 86) or the voluntary framing condition (*n* = 87).

## Materials and Measures

### *Learning Stimuli*

The study stimuli consisted of paintings by 12 different artists, including four artists used for practice and eight artists used for the critical study and test phases. Four of these artists (Yie Mei, Edward Cross, Georges Braque, and Ciprian Stratulat) were used to illustrate the concepts of interleaving and blocking. Three paintings were displayed per artist during a practice phase for a total of 12 paintings shown in the practice phase. The main study stimuli included realistic paintings by four artists (Philip Juras, George Wexler, Gerald Schwartz, Takeyce Walter) and stylistic paintings by four artists (Karen Margulis, Toni Grote, Judy Hawkins, Marilyn Mylrea). See Figure 1 for an example of a painting from each set.

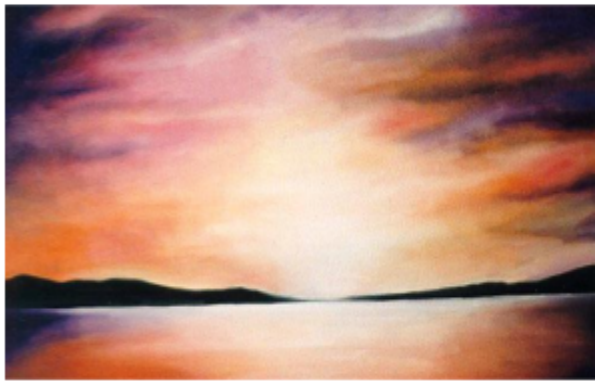
For each of the eight critical artists, one painting was presented during the initial preparation phase (to communicate the difficulty of the study material), eight paintings were presented during the study phase, and four paintings were presented during the test phase. The order of the paintings presented in each stage were randomized once and then applied to all participants. The paintings stimuli were taken from those used in prior studies (Kornell & Bjork, 2008; Yan et al., 2015; 2017). All paintings were resized to be as close to 400 x 500 pixels for portrait paintings or 500 x 500 for square-shaped paintings as possible while maintaining the original aspect ratio.

**Figure 1**

*Realistic and Stylistic Paintings Study Stimuli Examples*

A.

B.



*Note*. A) Painting at left by George Wexler, an artist in the realistic (Set A) artist group.

B) Painting at right by Judy Hawkins, an artist in the stylistic (Set B) artist group.

***Interim Questionnaire: Perception of Study Strategy***

After studying a set of artists, participants answered four questions about the perceived effort of the study strategy used. These questions differed by framing condition (see Table 1), and all questions were rated on a 6-point Likert scale (1 = very little; 6 = a great deal). The four effort perception questions formed a reliable scale, both when participants were responding about

the experience of blocked learning and the experience of interleaved learning (see Table 2). The responses to these items were averaged to create a scale score.

In the same interim questionnaires, participants were asked four questions about their perceived learning using each study strategy. These were the same across framing conditions: 1) How likely are you to be able to distinguish between the different artists by their paintings? 2) How good do you think your memory for the different artist styles/paintings will be? 3) How effective was this exercise in helping you to distinguish between different artists' paintings? 4) How well did you learn to distinguish between different artists' paintings? These four items formed a reliable scale, both for blocked and interleaved learning (see Table 2). The responses to the four items were averaged to create a scale score.

**Table 1**

*Effort Questions by Framing Condition*

| Questionnaire | Requirement Framing | Voluntary Framing |
|---|---|---|
| Interim | How much mental effort did the last exercise require? | How much mental effort did you choose to invest in the last exercise? |
| | How hard was the last exercise for you? | How hard did you choose to work on the last exercise? |
| | How much work did the last exercise require? | How much work did you choose to put into the last exercise? |
| | How much concentration did the last exercise require? | How much did you choose to concentrate on the last exercise? |
| Comparison | Which strategy required more mental effort? | Which strategy did you choose to put more mental effort into? |
| | Which strategy would require more mental effort for other people? | Which strategy would other people choose to put more mental effort into? |
| | Which strategy was harder for you? | Which strategy did you choose to work harder on? |
| | Which strategy would be harder for other people? | Which strategy would other people choose to work harder on? |
| | Which strategy required greater concentration from you? | Which strategy did you choose to concentrate more on? |
| | Which strategy would require greater concentration from other people? | Which strategy would other people choose to concentrate more on? |

*Note*. All questions are rated on a 6-point Likert scale (1 = very little; 6 = a great deal, for the interim questionnaire; 1 = Grouped; 6 = Mixed, for the comparison questionnaire).

**Table 2**

*M (SD) and Reliability of Effort and Learning Scales*

| Questionnaire | Measure | Cronbach's $\alpha$ | M (SD) |
|---|---|---|---|
| Interim | Perceived effort: Blocked | .91 | 4.28 (1.12) |
| | Perceived effort: Interleaved | .93 | 4.59 (1.22) |
| | Perceived learning: Blocked | .92 | 3.52 (1.17) |
| | Perceived learning: Interleaved | .93 | 2.81 (1.21) |
| Comparison | Perceived effort | .94 | 4.21 (1.53) |
| | Perceived learning | .87 | 2.18 (1.32) |

*Note*. All questions are rated on a 6-point Likert scale (1 = very little; 6 = a great deal, for the interim questionnaire; 1 = Grouped; 6 = Mixed, for the comparison questionnaire).

***Study Strategy Comparison Questionnaire***

A study strategy comparison questionnaire was given at the end of the study, after participants studied the paintings of all eight critical artists (four in a blocked manner; four in an interleaved manner). Participants were asked to compare their experience of effort under each strategy and to compare the effectiveness of the two strategies. They were also asked to choose which strategy they would want to use for future study. All the questions were scored on a 6-point Likert scale where 1 = Grouped to 6 = Mixed (i.e., a higher number reflected that the interleaved condition was more effortful, more effective, their choice).

Participants were asked three questions directly comparing the perceived effort of the two strategies and three analogous questions about what they thought others would say about the two strategies (Table 1). These six questions formed a reliable scale ($\alpha$ = .94), so the responses were averaged to create a "comparison experience of difficulty" scale score.

Next, another three questions were geared towards assessing perceived effectiveness for learning: 1) Which strategy's set of artists do you think you'll remember better? 2) Which do you think is a more effective learning strategy for you? 3) Which do you think is a more effective learning strategy for the average person?

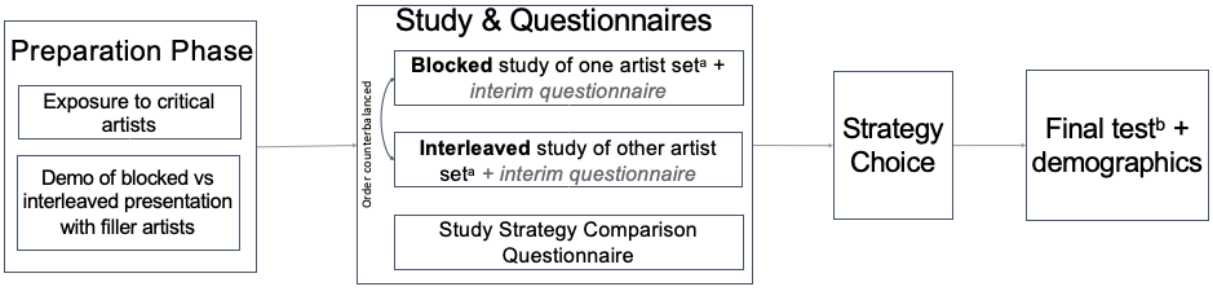
Finally, participants were asked a single choice question about which strategy they would use if they were to study more artist paintings ("Imagine that you had to study more artist paintings like you did today. Which strategy would you use to study the paintings, so you would be able to do well on a test on them later?") Participants answered using a 6-point Likert scale in which 1 = Grouped and 6 = Mixed, and the intervening numbers indicated combinations of the two that lean either more towards grouped or mixed study.

**Procedure**

After providing informed consent, participants completed the present study via Qualtrics. The procedure is illustrated in Figure 2.

**Figure 2**

*Illustration of Study Procedure*




*Note.* (a) Each interim questionnaire only focused on the study strategy that preceded it (as opposed to comparing the two strategies). (b) The final test was a composite of both sets of artists.

***Preparation Phase***

In a preparation phase, participants were shown one painting from each of the eight critical artists they would study in the main study portion of the experiment. Participants were asked to identify the correct artist from a list of artists from the same style that would appear concurrently with it during the study session. For example, *Juras, Wexler, Walter, or Schwartz* for the realistic set vs *Mylrea, Hawkins, Grote, or Margulis* for the stylistic set. Participants were given feedback after answering (e.g., "*The correct answer is* **Walter**") before the survey advanced to the next example. The purpose of this exposure was to illustrate the type and difficulty level of the material that the participants would study.

The demonstration of blocked and interleaved paintings was intended to introduce participants to the study strategies they would be using, with one practice section for blocking and interleaving each (described as "grouped together by type" and "mixed together," respectively). The artists used in the practice section (Stratulat, Braque, Mei, and Cross) were filler artists unconnected to the content of the main study portion and the final test. The demonstration was only meant to familiarize participants with the study strategies they would use during the experiment. Participants were shown three paintings from each artist with the artist's surname labeled above it for three seconds. For this demonstration, participants were only shown paintings from two artists per study strategy. In other words, participants studied paintings in the order of Braque-Braque-Braque then Stratulat-Stratulat-Stratulat to demonstrate blocking and then Mei-Cross-Mei-Cross-Mei-Cross to demonstrate interleaving. At the end of this preparation phase, the participants moved on to the primary study task.

### *Main Study Phase*

The main study phase consisted of two study sets divided by style (either realistic or stylistic). In one study set, the paintings were presented in a blocked sequence; in the other study set, the paintings were presented in an interleaved sequence. Whether participants experienced blocking or interleaving first was counterbalanced, as was the set (realistic or stylistic) that was interleaved, so that there was an equal chance of receiving any possible combination of study strategy and artist set.

The blocked study portion consisted of studying all eight paintings by an artist consecutively before moving onto the next artist. The order of the artists was randomized. Comparatively, the interleaved study portion intermingled paintings from the four artists in the set, randomized so that there were equal probabilities of each artist appearing in the first and second halves of the sequence. After each study set, participants answered the interim questions regarding the perceived effort and perceived effectiveness of the study strategy they just experienced.

After completion of both study sets, participants then answered the study strategy comparison questionnaire, where they directly compared blocking and interleaving (studying "grouped" versus "mixed") in terms of perceived effort and effectiveness.

### *Strategy Choice, Final Test, and Demographics*

After completing all study sections and questionnaires, participants were asked a final question regarding what study strategy they would choose to use in the future: interleaving or blocking.

Then, participants were tested over the material they studied during the experiment. They were shown four paintings from each critical artist, for a total of 32 questions, presented one at a

time. The presentation order of paintings was randomized by sections, with two paintings of each artist in each of two sections, so that each artist had equal presentation in both the first and second halves of the test. For each painting, participants were presented with a list of four possible names. The four options were always the four names from the same artist set. Thus, when the test item was a realistic painting, the multiple-choice answer options only included artists from the realistic set; when the test item was a stylistic painting, the multiple-choice answer options only included artists from the stylistic set. This final test was self-paced.

Finally, participants completed an 8-item demographic questionnaire. This included information about participant age, gender, year in college, ethnicity, and native language.

**Statistical Analysis**

Cronbach's alpha calculations indicated that the perceived effort and perceived learning questionnaire items each formed reliable scales, respectively. With the data from the interim questionnaires, a 2x2 mixed ANOVA evaluated the effects of effort framing and study strategy on perceived effort, and another evaluated the effects of this on perceived learning. With the data from the study strategy comparison questionnaire, a between-subjects t-test compared the effect of strategy on perceived effort, and another between-subjects t-test examined the effect of strategy on perceived learning. A third t-test compared study strategy choice by framing condition.

Kirk-Johnson et al. (2019) found that the relationship between perceived effort and strategy choice was fully mediated by perceived efficacy of the strategies. Specifically, they found that the more effortful participants rated a strategy, the less effective they thought it was and the less likely they were to choose to use it for future study. Using the responses from the

study strategy comparison questionnaire, we tested the same mediation model in our study. We also tested whether this mediated effect was moderated by the framing condition.

Two mediation models were tested, one for each framing condition. These examined the relationships between perceived effort, perceived learning, and study strategy choice. Then the difference in indirect and direct effects between the two conditions was tested using the *mediation* package in R.

## Results

### Interleaving Benefit

The first thing that was examined was whether there was an interleaving benefit: if students performed better on the final test on the material they had studied using interleaving as opposed to blocking. A within-subjects *t*-test showed that there was indeed an interleaving benefit. Participants were significantly more likely to correctly identify paintings by the interleaved artists ($M$ = .40, $SD$ = .18), than paintings by the blocked artists ($M$ = .36, $SD$ = .18), $t(171) = 2.84$, $p$ = .005, Cohen's $d$ = .22.

### Strategy Perceptions: Effort and Learning

We evaluated experiences of effort and perceived learning in two ways: first, through self-report items on the interim questionnaire directly after participants experienced each strategy, and then through self-report items after all study sections were completed, when participants compared interleaving and blocking in the study strategy comparison questionnaire. See Table 2 for the Cronbach's alpha reliability scores for each scale. If reframing effort did successfully lead participants to see interleaving as a more effective strategy, the data would show the effort-as-voluntary condition resulting in higher ratings of efficacy *without* affecting the experience of effort.

#### *Perceived Effort*

**Interim perceived effort.** On average, the interleaved strategy ($M$ = 4.59, $SD$ = 1.22) was experienced as more effortful than the blocked strategy ($M$ = 4.28, $SD$ = 1.12). The interim effort judgments by strategy and framing condition are depicted in the left half of Figure 3. A 2x2 mixed ANOVA examined how strategy and framing affected participants' experience of effort. As expected, there was a main effect of strategy, $F(1,171) = 17.22$, $MSE$ = 0.50, $p$ <.001,

$\eta_p^2$ = .09, which revealed that participants rated interleaving as significantly more effortful than blocking. This aligns with previous findings that learners tend to perceive interleaving as more difficult than blocking (Kirk-Johnson et al., 2019; Yan et al., 2016). There was no main effect of framing, $F(1,171) = 0.64$, $MSE = 2.16$, $p = .425$, $\eta_p^2 = .004$. Overall, then, participants in the effort-as-voluntary condition did not experience the study tasks to be significantly more or less effortful than participants in the effort-as-required condition.

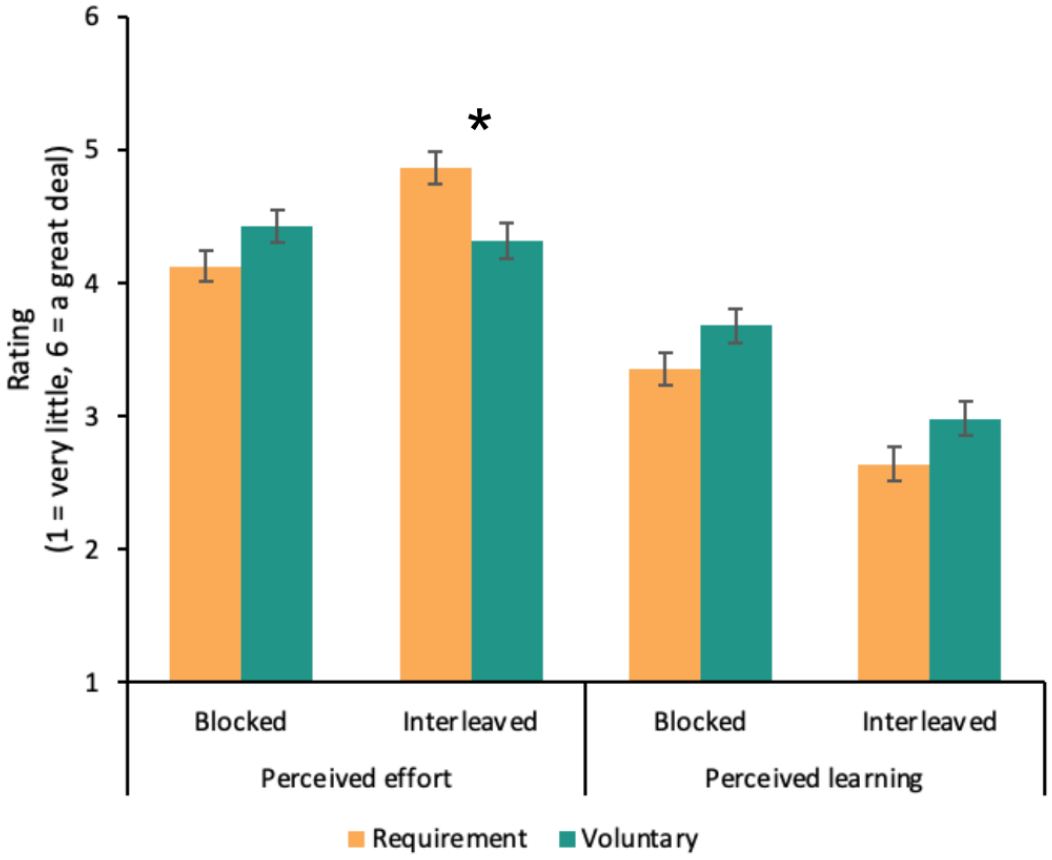
However, there was a significant interaction effect between framing condition and strategy $F(1,171) = 31.60$, $MSE = 0.50$, $p < .001$, $\eta_p^2 = .16$. Post hoc comparisons using Tukey tests were conducted to examine this interaction. For the interleaved strategy, effort ratings *were* affected by framing: experience of effort was significantly higher in the requirement framing condition ($M = 4.87$, $SD = 1.10$) than in the voluntary framing condition ($M = 4.32$, $SD = 1.28$), $t(167.76) = -3.04$, $p = .010$, Cohen's $d = .46$. For the blocked strategy, on the other hand, effort ratings were not affected by framing. Participants rated blocking to be moderately difficult whether in the requirement framing condition ($M = 4.13$, $SD = 1.11$) or the voluntary framing condition ($M = 4.43$, $SD = 1.11$), $t(170.98) = 1.78$, $p = .077$, Cohen's $d = .27$. It appears that framing had a greater effect on perceptions of interleaving than on perceptions of blocking due to the strong natural bias towards blocking, consistent with the findings of Yan et al. (2017).

**Comparison perceived effort.** The comparison experience of difficulty scale score yielded a $M$ of 4.21 ($SD = 1.53$), indicating that participants, on average, found the interleaved study to be more effortful. The average effort judgment by condition is depicted by the left pair of bars in Figure 4. A between-subjects t-test comparing the two framing conditions revealed that people in the requirement framing condition were significantly more likely to rate interleaving as effortful ($M = 5.04$, $SD = 1.05$) compared to those in the voluntary framing condition ($M = 3.40$,

*SD* = 1.50), $t(153.73) = 8.32$, $p < .001$, Cohen's $d = 1.29$. In fact, the average response of those in the voluntary framing condition was not significantly different from that of the mid-point of responses, hence showing no bias toward either strategy, $t(86) = -0.63$, $p = 0.530$, $d = .08$.

**Figure 3**

*Study Strategy Interim Evaluations: Perceptions of Effort and Learning*




*Note.* Participants in the effort-as-voluntary condition perceived interleaving to be less effortful on the interim questionnaires than participants in the effort-as-required condition. Bars represent the M, error bars represent ±1 SE (* $p < .05$).

***Perceived Learning***

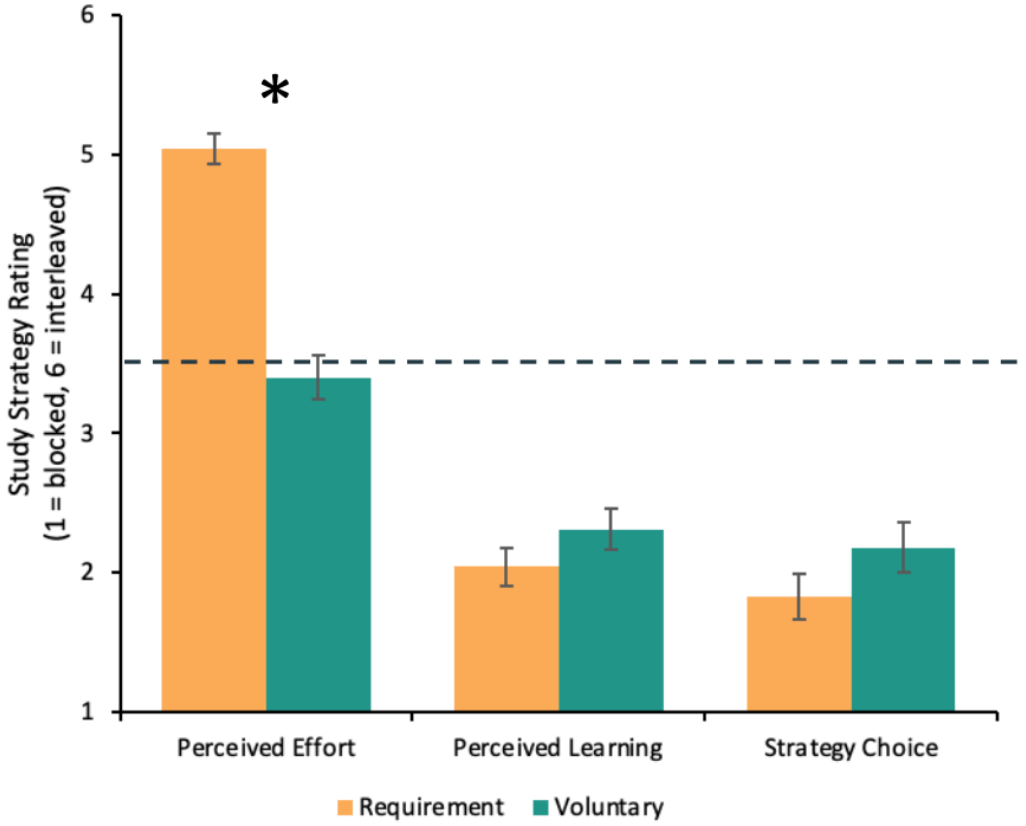
**Interim perceived learning.** The interim learning judgments by strategy and framing condition are depicted in the right half of Figure 4. Another 2 (strategy) x 2 (framing) mixed

ANOVA was conducted, this time examining how strategy and framing affect participants' perceived learning. This analysis revealed a main effect of strategy, $F(1,171) = 60.59$, $MSE = 0.72$, $p < .001$, $\eta_p^2 = .26$. On average, participants believed they had learned more when blocking ($M = 3.52$, $SD = 1.17$) than when interleaving ($M = 2.81$, $SD = 1.21$). There was also a main effect of framing condition, $F(1,171) = 4.51$, $MSE = 2.08$, $p = .035$, $\eta_p^2 = .03$. Participants in the effort-as-voluntary condition ($M = 3.33$, $SD = 1.02$) perceived that they had learned more than those in the effort-as-required condition ($M = 3.00$, $SD = 1.02$). Unlike with perceptions of effort, there was no strategy x framing interaction for perceived learning, $F(1, 171) = 0.02$, $MSE = 0.72$, $p = .901$.

**Comparison perceived learning.** The average perception of learning by condition is depicted by the middle pair of bars of Figure 4. A between-subjects t-test comparing the two framing conditions revealed no significant difference between the requirement framing condition ($M = 2.04$, $SD = 1.27$) compared to those in the voluntary framing condition ($M = 2.31$, $SD = 1.37$) on the comparison questionnaire, $t(170.24) = 1.33$, $p = 0.185$. Overall, the mean score was 2.18 ($SD = 1.32$), indicating a strong tendency for participants to judge blocking to be better for learning than interleaving.

**Figure 4**

*Study Strategy Comparison: Effort, Learning, and Study Strategy Choice*




*Note.* In the study strategy comparison questionnaire, as in the interim questionnaire, participants in the effort-as-voluntary condition rated interleaving as less effortful than participants in the effort-as-required condition. Regardless of framing condition, participants believed they learned better with blocking and chose to block more often. The dotted line represents the midpoint, meaning that at that point, participants' responses, on average, did not show a bias toward either blocking or interleaving (* $p$ <.05).

**Strategy Choice**

Students' mean study strategy choices for future study are depicted by condition by the rightmost pair of bars of Figure 4. As this figure shows, regardless of framing condition,

participants overwhelmingly indicated that they would choose the blocked schedule if they were to study similar material in the future ($M = 2.01$, $SD = 1.58$). The responses did not differ by framing condition, $t(169.7) = 1.49$, $p = 0.137$.

**Moderated Mediation Analysis**

In the effort-as-required condition, the effect of perceived effort on strategy choice was fully mediated via the perceived learning of the strategy. As panel A (left) of Figure 5 illustrates, perceived effort was negatively related to strategy choice. Perceived effort was negatively related to perceived learning, and in turn, perceived learning was positively related to strategy choice. In other words, the more effortful participants perceived interleaving to be, the less likely they were to think they learned well with that strategy. The worse they perceived their learning to be for interleaving, the less likely they were to choose to interleave in the future. All relationships were significant. The indirect effect was (-.37)*(.33) = -.12. The significance of this indirect effect was tested using bootstrapping procedures. Unstandardized indirect effects were computed for each of 1,000 bootstrapped samples, and the 95% confidence interval was computed by determining the indirect effects at the 2.5th and 97.5th percentiles. The bootstrapped unstandardized indirect effect was -.12, and the 95% confidence interval ranged from -.17 to -.07. Thus, the indirect effect was statistically significant ($p < .001$).

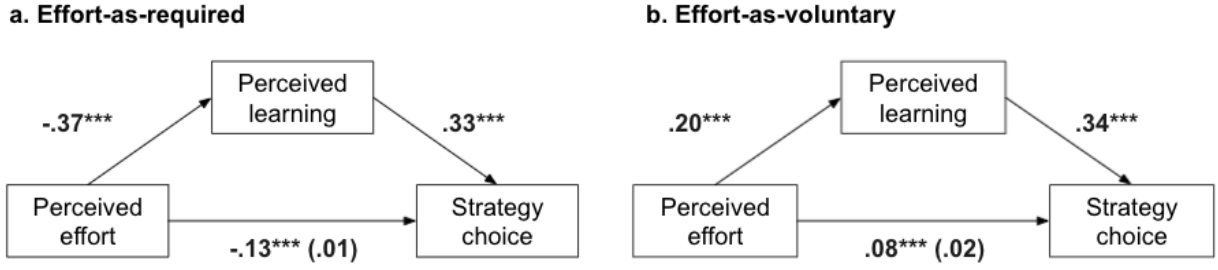
In the effort-as-voluntary condition, the effect of perceived effort on strategy choice was fully mediated via the perceived learning of the strategy. As panel A (left) of Figure 5 illustrates, the relationships between perceived effort and strategy choice and between perceived effort and perceived learning were positive (i.e., the opposite direction compared to the effort-as-required condition). Perceived learning was positively related to strategy choice. Thus, the more effortful participants perceived interleaving to be, the more likely they were to perceive they learned well

from interleaving, and the more likely they were to choose to interleave in the future. All relationships were significant. The indirect effect was (.20)*(.34) = .07. The bootstrapped unstandardized indirect effect was .07, and the 95% confidence interval ranged from .03 to .10. Thus, the indirect effect was statistically significant ($p$ <.001).

We then tested whether the bootstrapped direct and indirect effects were significantly different between these two mediation models. The indirect effects were significantly different from each other, M difference = .19 [95% CI: .13, .25], $p$ < .001. The direct effects were not significantly different from each other, M difference = .03 [95% CI: -.01, .06], $p$ < .108.

**Figure 5**

*Moderated Mediation Analysis*




*Note.* Variables from the study strategy comparison questionnaire, in which ratings were provided on a 6-point Likert scale (1 = blocked, 6 = interleaved).

**Interleaving Benefit Size**

As aforementioned, examining the performance of the participants on the final test yielded an interleaving benefit, in which participants performed better evaluating the group of artists they had studied using interleaved study. However, even though there was an overall benefit of interleaved study over blocked study, there was also substantial individual variation. Some participants experienced better performance after blocked study, some participants experienced only a small interleaved benefit, and others experienced a large interleaved benefit.

We thus examined whether the size of the interleaving benefit mattered. Since there was such variation in how participants performed via blocked or interleaved study (see the spread of data along the x-axis in Figure 6), we predicted that those who experienced greater differences in performance might also be more sensitive to recognizing the interleaving benefit. We examined this for both (a) effort ratings (to see if the size of the interleaving benefit affected people's ratings of strategy effort), and (b) perceived learning (to see if the size of the interleaving benefit affected people's ratings of strategy effectiveness).

We conducted regression analyses, predicting these dependent variables from the interaction between framing condition and interleaving benefit, calculated as the raw difference in proportion between the interleaved and blocked conditions. The results of the regression analysis predicting effort ratings are shown in Table 3, the results of the regression analysis predicting learning judgments are shown in Table 4.

**Table 3**

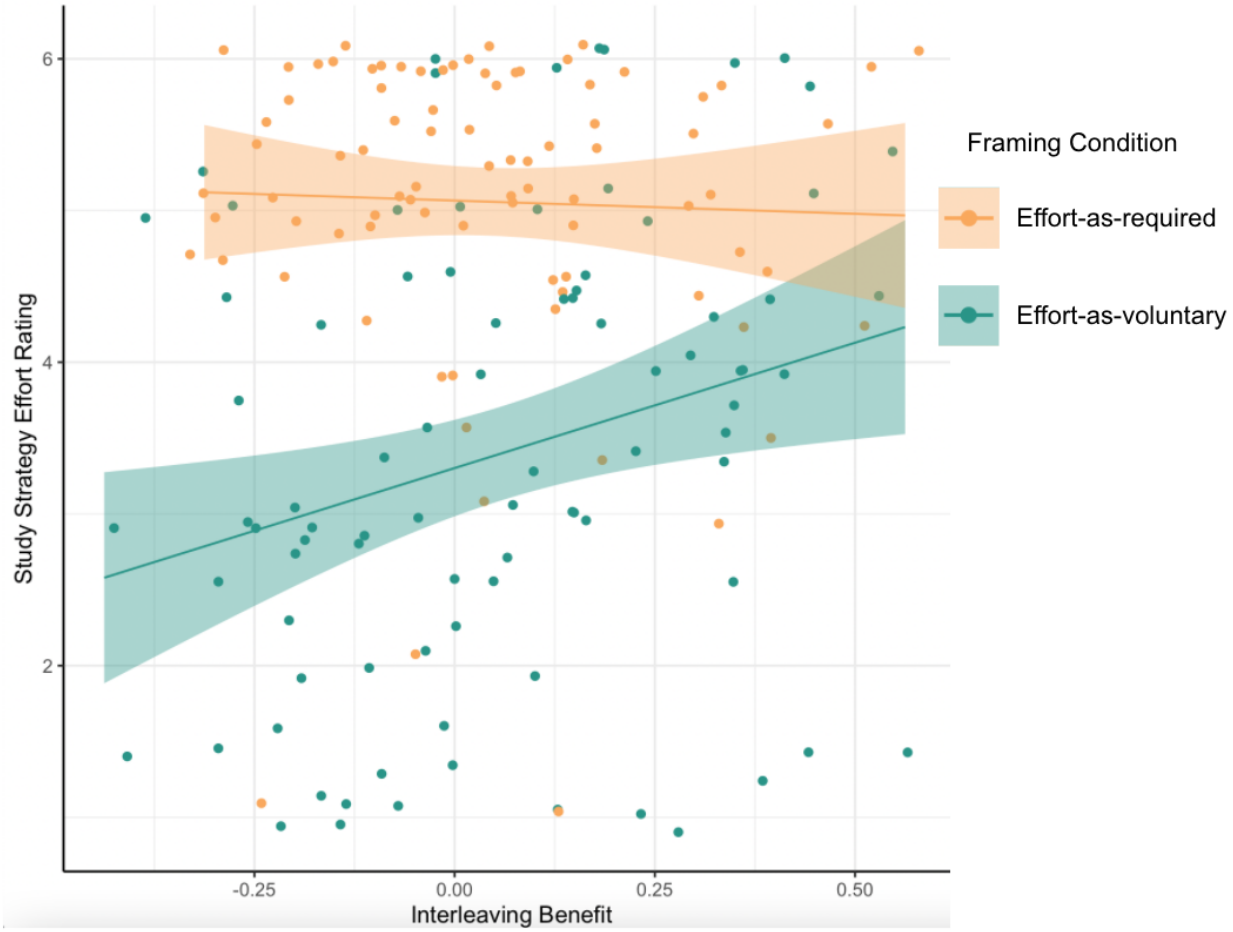
*Perception of Effort and Interleaving Benefit*

| | Unstandardized *B* | *SE* | *t*-value | *p*-value |
|---|---|---|---|---|
| Intercept | 3.30 | .14 | 23.72 | <.001*** |
| Interleaving Benefit | 1.65 | .55 | 3.00 | .003** |
| Framing Condition: Effort-as-required | 1.76 | 0.20 | 8.94 | <.001*** |
| Interleaving Benefit x Effort-as-required | -1.83 | .86 | -2.12 | .036* |

*Note*. 0 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' ', 1

**Figure 6**

*Effort Framing Moderates the Perceptions of Effort and Interleaving Benefit Relationship*




*Note.* There was a significant interaction between framing conditions and perceptions of effort as the interleaving benefit increased. Thus, the greater the interleaving benefit that participants experienced in the effort framed as voluntary condition, the more effortful they rated interleaving. Meanwhile, the effort-as-required condition was insensitive to this interleaving benefit. Study strategy effort was rated on a 6-pt scale (1 = blocked was a lot more effortful and 6 = interleaved was a lot more effortful). The interleaving benefit on the x-axis represents the difference in test score (proportion accurate) between the two sequences (interleaved performance – blocked performance). Shaded regions represent 95% confidence intervals.

**Table 4**

*Perception of Learning and Interleaving Benefit*

| | Unstandardized *B* | *SE* | *t*-value | *p*-value |
|---|---|---|---|---|
| Intercept | 2.30 | .15 | 15.75 | <.001*** |
| Interleaving Benefit | .09 | .58 | .15 | .883 |
| Framing Condition: Effort-as-required | -.30 | .21 | -1.43 | .154 |
| Interleaving Benefit x Effort-as-required | .59 | .91 | .65 | .52 |

*Note*. 0 '***', 0.001 '**', 0.01 '*', 0.05 '.', 0.1 ' ', 1

While the effort framing manipulation changed how people perceived effort, it did not affect perceptions of learning. In this case, both conditions were insensitive to the interleaving benefit. Contrary to our predictions, participants who performed far better using interleaving rated their learning using interleaving as about the same as those who experienced smaller interleaving benefits or even those who performed better with blocking.

## Discussion

Participants in the effort-as-required condition conformed to the misinterpreted effort pattern described in Koriat et al. (2014) and in Kirk-Johnson et al. (2019). The more effortful participants judged interleaving to be, the less they perceived that they had learned using the study strategy, and the less likely they were to choose to use it moving forward. However, in the effort-as-voluntary condition, this finding was reversed—the greater their estimations of effort for interleaving, the greater their estimations of their learning using the study strategy, and the more likely they were to choose to interleave in the future. This supported our predictions.

However, despite this finding, the effort-framing manipulation did not significantly affect perceptions of learning or study strategy choice. The participants in the voluntary condition had lower estimates of effort for interleaving than in the effort-as-required condition, rating effort as about the same when comparing interleaving and blocking. Participants in both conditions thought that they learned better using blocking, and overwhelmingly indicated that they would choose to use blocking to study similar material in the future. These findings did not support the hypothesis. Findings that would have supported the hypothesis would have shown no change in the perceptions of effort—with participants in the effort-as-voluntary condition continuing to view interleaving as far more difficult than blocking, as the participants in the effort-as-required condition did. Instead, there would have been greater perceptions of learning under interleaving in the effort-as-voluntary condition and a higher number of participants choosing to interleave over block for the future study strategy choice question.

It is possible that instead of changing the experience of effort, the framing manipulation only changed the source that the judgements of effort were based on. Instead of viewing effort as something located purely in the individual and invested into the task, or solely as something

intrinsic to the difficulty of the task, people are likely to hold and act on a mixture of these views. In other words, when considering one's effort in a task, people likely implicitly believe there is both a part that they can control (attention, time and concentration dedicated to the task) and a part that they cannot control (the difficulty of the items being studied or the task to be accomplished). Framing effort as voluntary may have increased the voluntary portion of effort so that participants increased their reflections on how much effort they had put into blocking.

There was a significant interaction regarding participant perceptions of effort and the interleaving benefit, with participants in the effort-as-voluntary condition increasing their estimations of effort the more they benefited from interleaved study. However, this degree to which participants performed better using interleaved study was unrelated to participant perception of learning using interleaving. Thus, even the participants that benefited more from interleaved study judged blocking to be far more effective for learning.

**Is the Interleaving Effect *Too* Counterintuitive?**

As the great lengths Yan et al. (2016) went to in order to counter student bias towards blocking suggest, perhaps there is something particular to the counterintuitive interleaving-blocking dynamic that makes it less feasible to alter student perceptions and decision-making. This potentiality begs the question of whether the same effort framing manipulation as used in the study might be more effective when applied to a different set of study strategies—for example, rereading versus retrieving information. There is a large body of work on strategies that are "desirably difficult" (Bjork, 1994; Yan, Clark, & Bjork, 2016). Each of these strategies increase the experience of effort while promoting better long-term retention. Interleaving is one strategy that falls under the "desirable difficulties" umbrella, but there are others, too, such as the retrieval practice (Roediger & Karpicke, 2006), distributed practice

(Carpenter, 2017), and the pre-testing (Richland, Kornell, & Kao, 2009). Each of these strategies tend to be underappreciated and under-utilized. But the use of these strategies may also be more amenable to an effort-reframing intervention. For example, while retrieval is experienced as more effortful and often judged to lead to worse learning, there is also likely less inherent doubt about the potential for retrieval to be beneficial (e.g., given the popularity of retrieval-based learning tools, such as Quizlet and flashcards). Perhaps an effort framing manipulation may have more impact when applied to these more intuitive desirably difficult study strategies.

**Other Possible Approaches to Improving Study Choices**

Alternatively, perhaps the effort framing manipulation is insufficient on its own to alter student decision-making. Future studies may investigate the potential of pairing this manipulation with more explicit statements discussing how oftentimes the most effective study strategies are those that subjectively feel more difficult. Another possibility would be to make use of other heuristics by activating the concept of a "good student" and connecting that to a student who chooses more desirably difficult and more effective study strategies. It would also be interesting to investigate whether showing participants their scores on the final test prior to their study strategy choice for future study would impact their choice. If their scores were presented to them divided by study strategy (i.e. if they received separate scores to show how they performed on the items they studied using blocking versus those they studied using interleaving), it is possible that more participants would choose to interleave in the final question. However, it is also likely that the difference would not be pronounced enough to change their minds.

**Concluding Comments**

This study reaffirms the challenge to encouraging appreciation of interleaving described in Yan et al. (2016). It replicated and extended the findings in Koriat et. al (2014) within the context of interleaving and blocking; framing effort as a choice and an investment instead of a requirement reversed the traditional association of easy learning with good learning to an association of *effortful* learning with good learning. However, it also exposed the limits of this manipulation by demonstrating that this effort-framing manipulation was insufficient to affect the choice of future study strategy use itself. Participants still overwhelmingly chose blocking over interleaving.

Regardless of what precise future study directions are undertaken, the matters of student decision making and addressing latent learning-related biases are vital ones. Further investigation could shed light on the mechanisms and processes by which people make study strategy decisions for themselves, as well as the potential to improve learners' study strategy choices.

## References

Abushanab, B., & Bishara, A. (2013). Memory and metacognition for piano melodies: Illusory advantages of fixed- over random-order practice. *Memory & Cognition, 41*. https://doi.org/10.3758/s13421-013-0311-z.

Birnbaum, M. S., Kornell, N., Bjork, E. L., & Bjork, R. A. (2013). Why interleaving enhances inductive learning: The roles of discrimination and retrieval. *Memory & Cognition, 41*(3), 392-402. https://doi.org/10.3758/s13421-012-0272-7.

Bjork, R. A. (1994). Memory and metamemory considerations in the. *Metacognition: Knowing about knowing, 185*, 7-2.

Brunmair, M., & Richter, T. (2019). Similarity matters: A meta-analysis of interleaved learning and its moderators. *Psychological Bulletin, 145*(11), 1029–1052. https://doi.org/10.1037/bul0000209.

Carpenter, S. K. (2017). Spacing effects in learning and memory. In J. T. Wixted, & J. H. Byrne (Eds.), *Learning and memory: A comprehensive reference* (pp. 465–485). Academic Press.

Carvalho, P. F., Braithwaite, D. W., de Leeuw, J. R., Motz, B. A., & Goldstone, R. L. (2016). An in vivo study of self-regulated study sequencing in introductory psychology courses. *PloS one, 11*(3), e0152115.

DeWinstanley, P.A., Bjork, E.L. (2004) Processing strategies and the generation effect: Implications for making a better reader. *Memory & Cognition 32*, 945–955. https://doi.org/10.3758/BF03196872.

Dweck, C.S., Chiu, C. and Hong, Y. (1995) Implicit theories and their role in judgments and reactions: a word from two perspectives. *Psychological Inquiry; An International Journal for the Advancement of Psychological Theory 6*(4): 267–285.

Eglington, L.G. and Kang, S.H. (2017) Interleaved presentation benefits science category learning. *Journal of Applied Research in Memory and Cognition 6*(4): 475–485.

Fisher, O., & Oyserman, D. (2017). Development and validation of interpretation of experienced ease and difficulty scales. *Long Beach, CA: Society for Personality and Social Psychology*.

Gelman, A. (2018, March 15). *You need 16 times the sample size to estimate an interaction than to estimate a main effect.* Statistical Modeling, Causal Inference, and Social Science. https://statmodeling.stat.columbia.edu/2018/03/15/need-16-times-sample-size-estimate-interaction-estimate-main-effect/.

Goldin, S. B., Horn, G. T., Schnaus, M. J., Grichanik, M., Ducey, A. J., Nofsinger, C., Hernandez, D. J., Shames, M. L., Singh, R. P., & Brannick, M. T. (2014). FLS Skill Acquisition: A Comparison of Blocked vs Interleaved Practice. *Journal of Surgical Education, 71*(4), 506–512. https://doi.org/10.1016/j.jsurg.2014.01.001.

Do, L.A. & Lee, H. (2020). Why They Chose What They Chose: Exploring Effects of Test Performance and Metacognitive Judgments on Learners' Selection of Interleaving Schedule. *The Korean Journal of Cognitive and Biological Psychology, 32*(2), 169-188. https://doi.org/10.22172/COGBIO.2020.32.2.004.

Hall, K. G., Domingues, D. A., & Cavazos, R. (1994). Contextual interference effects with skilled baseball players. *Perceptual and Motor Skills, 78,* 835-841.

Kahneman, D. (2011). *Thinking, fast and slow.* Farrar, Straus and Giroux.

Kang, S. H., & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology, 26*(1), 97-103.

Kirk-Johnson, A., Galla, B. M., & Fraundorf, S. H. (2019). Perceiving effort as poor learning: The misinterpreted-effort hypothesis of how experienced effort and perceived learning relate to study strategy choice. *Cognitive Psychology, 115*, 101237. https://doi.org/10.1016/j.cogpsych.2019.101237.

Koriat, A., Nussinson, R., & Ackerman, R. (2014). Judgments of learning depend on how learners interpret study effort. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40*(6), 1624-1637. https://doi.org/10.1037/xlm0000009.

Kornell, N., & Bjork, R. A. (2008). Learning concepts and categories: Is spacing the "enemy of induction"? *Psychological Science, 19*, 585-592. doi:10.1111/j.1467-9280.2008.02127.x.

Kornell, N., Castel, A. D., Eich, T. S., & Bjork, R. A. (2010). Spacing as the friend of both memory and induction in young and older adults. *Psychology and Aging, 25*, 498-503.

Kühl, T., Eitel, A. Effects of disfluency on cognitive and metacognitive processes and outcomes. *Metacognition Learning* 11, 1–13 (2016). https://doi.org/10.1007/s11409-016-9154-x.

McCabe, J. (2010). Metacognitive awareness of learning strategies in undergraduates. *Memory & Cognition, 39*, 462-476.

Ostrow, K., Heffernan, N., Heffernan, C., & Peterson, Z. (2015). Blocking Vs. Interleaving: Examining Single-Session Effects Within Middle School Math Homework. *Artificial Intelligence in Education*, 338–347. https://doi.org/10.1007/978-3-319-19773-9_34.

Richland, L. E., Kornell, N., & Kao, L. S. (2009). The pretesting effect: Do unsuccessful retrieval attempts enhance learning? *Journal of Experimental Psychology: Applied, 15*(3), 243.

Roediger III, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological science, 17*(3), 249-255.

Rohrer, D., Dedrick, R. F., & Hartwig, M. K. (2020). The Scarcity of Interleaved Practice in Mathematics Textbooks. *Educational Psychology Review, 32*(3), 873-883. https://doi.org/10.1007/s10648-020-09516-2.

Sana, F. & Yan, V. X. (in press). Interleaving retrieval practice promotes science learning. *Psychological Science.* https://doi.org/10.1177/09567976211057507.

Sana, F., Yan, V. X., & Kim, J. A. (2017). Study sequence matters for the inductive learning of cognitive concepts. *Journal of Educational Psychology, 109,* 84-98. http://dx.doi.org/10.1037/edu0000119.

Sana, F., Yan, V. X., Kim, J. A., Bjork, R. A., Bjork, E. L. (2018). Do learners with lower working memory capacity benefit from interleaving? *Journal of Applied Research in Memory and Cognition, 7,* 361-369. https://doi.org/10.1016/j.jarmac.2018.05.005.

Sorensen, L.J., Woltz, D.J. (2016). Blocking as a friend of induction in verbal category learning. *Memory & Cognition, 44*, 1000–1013. https://doi.org/10.3758/s13421-016-0615-x.

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.

Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology, 24*(6), 837-848.

Tullis, J.G., Finley, J.R. & Benjamin, A.S. (2013). Metacognition of the testing effect: Guiding learners to predict the benefits of retrieval. *Memory & Cognition, 41*, 429–442. https://doi.org/10.3758/s13421-012-0274-5.

Wahlheim, C., Finn, B., Jacoby, L. (2012). Metacognitive judgments of repetition and variability effects in natural concept learning: Evidence for variability neglect. *Memory & Cognition, 40*(5), 703-16. doi:10.3758/s13421-011-0180-2.

Yan, V. X., Bjork, E. L., & Bjork, R. A. (2016). On the difficulty of mending metacognitive illusions: A priori theories, fluency effects, and misattributions of the interleaving benefit. *Journal of Experimental Psychology: General, 145*, 918-933.

Yan, V. X., Clark, C. M., & Bjork, R. A. (2016). Memory and metamemory considerations in the instruction of human beings revisited: Implications for optimizing online learning. In *From the Laboratory to the Classroom* (pp. 73-90). Routledge.

Yan, V. X., Sana, F., Kim, J., Bjork, R. A., & Bjork, E. L. (2015, Nov). *Why does interleaving the exemplars of to-be-learned categories enhance concept learning? Testing the discriminative-contrast hypothesis.* Poster presented at the 56th Annual Scientific Meeting of the Psychonomic Society, Chicago, IL.

Yan, V. X., & Sana, F. (2021). Does the interleaving effect extend to unrelated concepts? Learners' beliefs versus empirical evidence. *Journal of Educational Psychology, 113*(1), 125–137. https://doi.org/10.1037/edu0000470.

Yan, V. X., Soderstrom, N. C., Seneviratna, G. S., Bjork, E.L., & Bjork, R. A. (2017). How should exemplars be sequenced in inductive learning? Empirical evidence versus learners' opinions. *Journal of Experimental Psychology: Applied, 23*, 403-416.

Zulkiply, N., & Burt, J. S. (2013). The exemplar interleaving effect in inductive learning: Moderation by the difficulty of category discriminations. *Memory & Cognition, 41*(1), 16-27.

Zulkiply, N., McLean, J., Burt, J. S., & Bath, D. (2012). Spacing and induction: Application to exemplars presented as auditory and visual text. *Learning and Instruction, 22*, 215-221.