Cognitive Psychology 115 (2019) 101237

Contents lists available at ScienceDirect

Cognitive Psychology

journal homepage: www.elsevier.com/locate/cogpsych

Perceiving effort as poor learning: The misinterpreted-effort hypothesis of how experienced effort and perceived learning relate to study strategy choice☆

Afton Kirk-Johnson[a], Brian M. Galla[a], Scott H. Fraundorf[b,*]

[a] Learning Research and Development Center and Department of Applied Developmental Psychology, University of Pittsburgh, United States
[b] Learning Research and Development Center and Department of Psychology, University of Pittsburgh, United States

ARTICLE INFO

Keywords:
Mediation
Metacognition
Self-regulated learning

ABSTRACT

How do learners make decisions about how, what, and when to study, and why are their decisions sometimes ineffective for learning? In three studies, learners experienced a pair of contrasting study strategies (Study 1: interleaved vs. blocked schedule; Studies 2 & 3: retrieval practice vs. restudy) and rated their perceptions of each strategy before choosing one for future use. In all three studies, mediation analysis revealed that participants who perceived a strategy as more effortful rated it as less effective for learning and, in turn, were less likely to choose it for future study. Further, choosing the *more* effortful strategy was associated with better long-term retention (Study 3), contrary to participants' judgments. A final fourth study suggested that these relationships were not driven by the mere act of providing ratings. Our results thus support a misinterpreted-effort hypothesis in which the mental effort associated with many normatively effective learning strategies (*desirable difficulties*; Bjork & Bjork, 1992) leads learners to misinterpret them as ineffective for learning and consequently not to employ them in self- regulated learning.

1. Introduction

Learners often study with no explicit knowledge of the science of human learning and memory (Geller et al., 2018; Hartwig & Dunlosky, 2012; Kornell & Bjork, 2007; McCabe, 2011). Nevertheless, they must make decisions about what, when, and how to study: Should they re-read an assigned textbook chapter or use the questions at the back of the chapter to quiz themselves? Is their time better spent distributing study over a week or cramming the night before an exam? Self-regulated learning theory holds that controlling and choosing study strategies is critical to effective learning (Bjork, Dunlosky, & Kornell, 2013; Dunlosky & Metcalfe, 2009; Finley, Tullis & Benjamin, 2011; Winne & Hadwin, 1998; Zimmerman, 2000). Indeed, such skills become increasingly important in formal education as students progress to higher grades and begin to study more independently and without study materials created for them (Poropat, 2009; Yeager et al., 2014). Unfortunately, learners do not always make effective choices (e.g., Chi, Bassok, Lewis,

☆ We thank Rachel Peters, Eva Bacas, Rodica Constantine, Cody Gilbert, Sriteja Gummadi, Allison Hill, Robert Hines, Alan Lu, Sophia Lu, Brenna Mauro, and Genevieve O'Brien for assistance with data collection and scoring, and Kristin Divis, Jason Finley, Kathleen Hourihan, and Jonathan Tullis for comments and suggestions. The work was supported by an internal award from the Learning Research and Development Center to Scott Fraundorf and Brian Galla.

* Corresponding author at: Learning Research and Development Center, University of Pittsburgh, 3939 O'Hara Street, Pittsburgh, PA 15260, United States.

E-mail address: sfraundo@pitt.edu (S.H. Fraundorf).

https://doi.org/10.1016/j.cogpsych.2019.101237
Received 11 January 2019; Received in revised form 5 July 2019; Accepted 18 August 2019
Available online 27 August 2019

Reimann, & Glaser, 1989; Karpicke, 2009; Kornell & Bjork, 2007, 2008b; Tullis & Benjamin, 2011; Yan, Bjork, & Bjork, 2016). Thus, there is a critical need to understand the reasons and processes behind learners' choices in order to empower learners to make better self-regulated learning decisions.

Here, we develop and test an account of self-regulated learning decisions grounded in both cognitive and motivational science: Learners often do not have direct access to their objective level of performance or learning during study. Thus, their perceptions of how well they feel they are learning the material may instead be linked to the visceral experience of effort required by a particular study strategy. In particular, learners may interpret the effort required by certain strategies as a sign of poor learning and consequently avoid them in their self-regulated study—even when those strategies might have been quite efficacious. This chain of influences is suggested by several lines of existing experimental evidence; however, few studies have simultaneously assessed all of these links within the same participants at the same time.

In four studies, we directly test this misinterpreted-effort hypothesis by giving learners experience with different study strategies and then assessing both their perceptions of those strategies and their intent to use them in the future. We test the mediation hypothesis that the mental effort participants experience while engaging in study predicts how much participants feel they learn from a given study strategy, which in turn predicts the choices they make about using those study strategies in the future. And, we contrast the predictions of this hypothesis with those of alternative accounts in which poor self-regulated learning decisions reflect a direct avoidance of mental effort or a reliance on other metacognitive bases entirely.

1.1. *Choices are required for studying*

Learners need to make decisions about what material to study, in what order to study that material, how to study that material, how long to study, when to study, and when to terminate study (Finley, Tullis, & Benjamin, 2011). For instance, given a set of statistical tests to master for an exam, learners must choose whether to practice all of the problems of one type (e.g., *t*-tests) before moving onto the next type (e.g., correlations) or to do the problems mixed together. And, the learners must decide whether to practice each principle by working problems for themselves or reviewing already worked examples.

These decisions can have consequences for learning because not all study strategies are equally effective (for review, Bjork et al., 2013; Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Kornell & Bjork, 2007; Pashler, Rohrer, Cepeda, & Carpenter; Rohrer & Pashler, 2010). For instance, meta-analysis (Rowland, 2014) indicates that *retrieval practice* (studying material once and then taking a quiz on it) leads to better long-term retention than *restudying* (studying material one time and then studying it again) in many contexts (Carpenter, Pashler, Wixted, & Vul, 2008; Karpicke & Roediger, 2008; Karpicke, 2012; Roediger & Karpicke, 2006). Thus, the more participants employ retrieval practice, the more they learn, as reflected both in laboratory performance (Karpicke, 2009) and in college GPA (Geller et al., 2018; Hartwig & Dunlosky, 2012). Other lab studies have revealed that learners also retain more of what they have learned when they devote more study time to difficult material rather than easy material (Tullis & Benjamin, 2011), explain texts to themselves (Chi et al., 1989), or employ repeated practice (Kornell & Bjork, 2008b). At a more general level, high school students who devote more time to math practice rather than leisure achieve higher GPA and standardized test scores and high school graduation and college enrollment rates (Galla et al., 2014).

1.2. *Making self-regulated learning decisions*

Why is it that students do not always choose the strategies that would seem to maximize their learning (Hacker, Bol, Horgan, & Rakow, 2000)? Here, we discuss how several existing lines of research together yield an emerging view of how learners choose their self-regulated learning strategies—and why they sometimes choose poorly.

First, there is general evidence that learners do attempt to make self-regulated learning decisions that reflect their metacognitive judgments. That is, in the influential framework of Nelson and Narens (1990), they exert metacognitive *control*. For instance, when learners judge material as more difficult, they often spend more study time on it (Dunlosky & Hertzog, 1998; for review, see Son & Metcalfe, 2000) and sometimes even apply different, more effective study strategies (Benjamin & Bird, 2006; c.f., Pyc & Dunlosky, 2010; Son, 2004, 2010). Learners also adjust study or retrieval strategies based on how they expect to be tested on the material later (Finley & Benjamin, 2012) or on how others have performed (Fraundorf & Benjamin, 2015).

However, translating metacognitive judgments into control of study would be ineffective if learners cannot accurately assess what they have learned or what study techniques will help them learn; that is, if their metacognitive *monitoring* is poor. Indeed, there is reason to think that metacognitive monitoring is challenging. Learners do not have direct access to the strength or accuracy of their memory traces (Koriat, 1997), nor do most students report having been taught normative information about what constitutes effective study (Geller et al., 2018; Hartwig & Dunlosky, 2012; Kornell & Bjork, 2007; McCabe, 2011). Instead, according to the cue-utilization framework (Koriat, 1997), listeners must *infer* what they know and how well they are learning on the basis of various cues, including both their internal state and their environment (Koriat, 1997; Schwartz, Benjamin, & Bjork, 1997). For example, consider a student preparing for a chemistry exam by re-reading the textbook and doing practice problems. Our student cannot definitively know which of these exercises will yield the best performance on an exam that has yet to occur. Future retention can only be inferred from the student's past and present perceptions, including both internal sources (Which chemical formulas come to mind quickly?) and external (Do responses to the practice problems match the answers in the back of the book? Has this strategy yielded good grades in the past?).

Because metacognitive monitoring is inferential, it may not always correspond to objective reality. Many laboratory experiments have assayed learners' judgments of future memory (*judgments of learning*; JOLs; Nelson & Dunlosky, 1991) for material learned under different study conditions. These studies have revealed that learners often assign *lower* JOLs to study conditions that characteristically result in *better* learning, such as retrieval practice (Karpicke & Roediger, 2008; Karpicke, 2009; Tullis, Finley, & Benjamin, 2013), repeated study (Koriat, Sheffer, & Ma'ayan, 2002; Kornell & Bjork, 2009; Kornell, Rhodes, Castel, & Tauber, 2011), distributed practice (Logan, Castel, Haber, & Viehman, 2012), generation from a prompt (Besken, 2016), and interleaving of categories (Wahlheim, Finn, & Jacoby, 2012). Further, JOLs are unduly influenced by variables that are largely irrelevant to long-term learning, such as font size or speech fluency (Carpenter, Wilford, Kornell, & Mullaney, 2013; Kornell et al., 2011; Rawson & Dunlosky, 2002; Rhodes & Castel, 2008, 2009; Serra & Dunlosky, 2010). Unfortunately, then, judgments of learning are not always an accurate basis for the choice between study strategies.

What, specifically, leads to errors in judgments of learning? Other work assessing the bases of metacognitive judgments and decisions have found that they are influenced in part by global metacognitive *beliefs* about the normative efficacy of different strategies (Fraundorf & Benjamin, 2014; Koriat, 1997; Koriat, Bjork, Sheffer, & Bar, 2004; Kornell & Bjork, 2009). And, these beliefs may be inaccurate because the properties of human memory are not always intuitive (Bjork et al., 2013; Kornell & Bjork, 2007; Simons & Chabris, 2011, 2012), even for advanced psychology students (McCabe, 2011; Simons & Chabris, 2011). But, an even more crucial influence on metacognitive monitoring may be learners' own subjective experience with the to-be-learned material (Kelley & Jacoby, 1996; Koriat, 1997; Koriat et al., 2004). Many learners appear to use in-the-moment processing fluency[1] to monitor their learning, a strategy known as the *ease-of-processing heuristic* (Finn & Tauber, 2015; Kornell et al., 2011; Benjamin, Bjork, & Hirshman, 1998; Benjamin et al., 1998; Rhodes & Castel, 2008; Undorf & Erdfelder, 2011; see also the closely related notion of *easily learned, easily remembered*; Koriat, 2008). This heuristic may be bolstered by the broader beliefs in Western cultures that achievement is a marker of intelligence and that intelligence cannot be changed (Heine et al., 2001; Plaut & Markus, 2005). So, many learners who encounter difficulty or disfluency in their learning may interpret these struggles as a sign of failure or inability to learn (Yeager & Dweck, 2012). By contrast, learners who instead view intelligence as less fixed and subject to change do not necessarily interpret difficulty this way (Miele, Finn, & Molden, 2011).

Unfortunately, initial or current ease of processing may be misleading in choosing self-regulated learning strategies (Bjork, 1994; Koriat & Bjork, 2005; Schmidt & Bjork, 1992; Soderstrom & Bjork, 2015; Sungkhasettee, Friedman, & Castel, 2011). Other work assessing the objective efficacy of different learning strategies has found that the strategies that produce better long-term retention tend to require more difficult cognitive processing in the short-term—the principle of *desirable difficulties* (Bjork, 1994; Bjork & Bjork, 1992; Schmidt & Bjork, 1992). Thus, students who rely on initial or current ease of processing to monitor their learning are very likely to be misled about which materials and study strategies yield the most enduring learning.

1.3. The misinterpreted-effort hypothesis

Synthesizing the bodies of evidence reviewed above suggests what we term the *misinterpreted-effort hypothesis* about how learners choose self-regulated learning strategies—and why those choices may be ineffective. Specifically, we hypothesize that learners' self-regulated learning decisions are often guided by which strategies learners perceive as more effective for learning. These perceptions of learning effectiveness, in turn, are often guided—or misguided—by the mental effort or difficulty experienced during study. Greater perceived effort, even when resulting in better objective learning, leads learners to *feel* that they are learning *less*, especially because many learners hold the mindset that effort is a sign of poor ability rather than growth (Yeager & Dweck, 2012). In turn, this perception leads individuals to eschew the strategy that gave rise to the experience of effort. That is, the experience of mental effort is hypothesized to have an indirect effect on self-regulated learning decisions, mediated by perceived learning (panel A of Fig. 1).

For instance, consider retrieval practice. Retrieval practice typically leads to better long-term retention than restudying (Rowland, 2014). But whereas restudy requires only passive reading, practicing retrieval requires an active memory search—a cognitively laborious task. Learners who judge learning on the basis of current ease of processing are thus likely to perceive retrieval practice as resulting in less learning than restudy. Thus, they are likely to—perhaps unwisely—choose restudy over retrieval practice for future study.

Although this hypothesis is *suggested* by a growing body of evidence on self-regulated learning, there are other plausible pathways by which learners may come to their decisions. A second, contrasting model of self-regulated learning decisions is what we term the *effort-avoidance hypothesis* (panel B of Fig. 1). Work in cognitive science and behavioral economics on cognitive control (e.g., Chevalier, 2018; Westbrook, Kester, & Braver, 2013) has produced robust evidence that individuals forego mentally effortful tasks (e.g., complex working memory tasks), even when doing so reduces rewards (e.g., monetary payments). Thus, it is possible that greater mental effort, in and of itself, may have a *direct* effect on self-regulated learning decisions, regardless of perceived learning (Duckworth, Kirby, Tsukayama, Berstein, & Ericsson, 2011; Guo et al., 2016; Perez, Cromley, & Kaplan, 2014). Learners may "take the easy way out" and avoid choosing a study strategy they feel takes more effort, even if they are aware it is effective for learning. Such a pattern would be analogous to how many people know that physical fitness and eating healthy are important (Hofmann, Friese, & Wiers, 2008; Mann, De Ridder, & Fujita, 2013), but do not engage in these behaviors because the costs of doing so at any given time are too high relative to alternative activities, such as sleeping in or watching TV (Kurzban, Duckworth, Kable, & Myers, 2013; Rothman et al., 2015). The effort-avoidance hypothesis is also broadly consistent with expectancy-value theory (Eccles &

[1] Some researchers have queried the effects of different types of processing fluency, such as fluency in retrieving from memory versus fluency in visual perception. We do not attempt to titrate the effects of particular kinds of fluency in the present study, especially because they all appear to contribute to judgments; for further discussion, see Alter and Oppenheimer (2009) and Undorf and Erdfelder (2011).

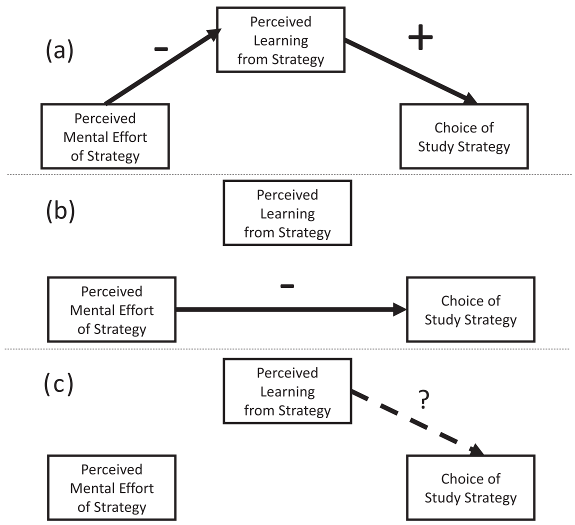

Fig. 1. Three theoretical models of self-regulated learning decisions.

Wigfield, 2002), in which students' choices are thought to reflect not just the perceived benefits of the choice (i.e., potential learning) but also the perceived *costs* of doing so (Inzlicht, Shenhav, & Olivola, 2018; Shenhav et al., 2017), including psychological costs, such as effort, boredom, or fatigue (Barron & Hulleman, 2015; Eccles & Wigfield, 2002; Flake, Barron, Hulleman, McCoach, & Welsh, 2015).

A third possibility is what we term the *effort-disregard hypothesis* (panel C): Experienced mental effort does not exert a robust influence on study strategy decisions, either directly or indirectly. Many metacognitive theories posit that subjective experiences, such as mental effort, are only one possible basis for metacognitive judgments (Fraundorf & Benjamin, 2014; Kelley & Jacoby, 1996; Koriat, 1997), so it possible that choices of study strategies are largely made on some other basis. One such basis may be already-established habits (Ariel & Dunlosky, 2013; Ariel, Al-Harthy, Was, & Dunlosky, 2011). Another is normative beliefs about how to study, acquired either explicitly from instruction or implicitly from common practices (e.g., in formal education, examples of different categories are often presented in separate blocks rather than interleaved; Roediger & Pyc, 2012). In this case, we might expect to see an influence of perceived learning effectiveness on study strategy decisions, but no influence of mental effort on either of these constructs.

1.4. The needed evidence

As we reviewed above, our misinterpreted-effort hypothesis is supported by several individual strands of evidence: (a) Broadly speaking, learners' study decisions reflect their judgments of learning (Son & Metcalfe, 2000), (b) judgments of learning are often influenced by perceived effort or fluency (Benjamin et al., 1998; Kornell et al., 2011), and (c) study strategies more effective for long-term learning are often more effortful or disfluent in the moment (Duckworth et al., 2011; Bjork & Bjork, 1992; Schmidt & Bjork, 1992).

However, several critical links in this hypothesized chain have not been fully established. First, there is surprisingly little direct evidence that judgments of learning predict *how* people study. There is fairly substantial evidence that learners' JOLs predict—and, indeed, causally determine (Metcalfe & Finn, 2008)—*which* materials they choose to study (for review, see Son & Metcalfe, 2000). And, it is clear that learners assign higher JOLs to materials studied under some conditions than others (e.g., Karpicke, 2009; Karpicke & Roediger, 2008; Koriat et al., 2002; Kornell et al., 2011; Kornell & Bjork, 2009; Logan et al., 2012; Tullis et al., 2013; Yan et al., 2016). What is less clear is whether those differences in perceived learning predict or even determine whether learners choose to *employ* those strategies in the future. That is, are learners' choices among different study strategies dictated by the perceived efficacy of those strategies? This association cannot be taken for granted; for instance, learners might instead choose strategies they consider less effortful (as we discuss above), or they might choose study activities to assess their current understanding rather than to acquire new knowledge (Kornell & Son, 2009).

Second, as a corollary to the above, the cue-utilization framework (Koriat, 1997) predicts that it should specifically be the *perceived* learning under these various strategies, not objective performance, that determines strategy choice. That is, because learners do not have direct access to their objective level of learning or knowledge (Koriat, 1997; Schwartz et al., 1997), they must make decisions based on their perceptions of learning; if perceived and objective learning differ, it should ultimately be perceived learning that predicts choices. The cue-utilization framework has received support insofar as there is ample evidence that JOLs rely on cues, such as fluency of processing, that do not always perfectly correspond to objective learning (Benjamin et al., 1998; Benjamin et al., 1998; Koriat, 1997; Kornell et al., 2011) and that JOLs, rather than objective performance, causally determine which items are chosen for study (Metcalfe & Finn, 2008). But, again, there is comparatively little data on learners' choices at the level of entire strategies, and, in particular, contrasting objective versus perceived learning as a predictor for this choice.

Third, relatively few studies have assessed subjective effort. The principle of desirable difficulties is derived from evidence that conditions that result in more enduring learning or retention often yield lower initial performance in objective accuracy. However, our misinterpreted-effort hypothesis predicts that learners should also find these conditions *subjectively* more difficult, and it is these subjective perceptions that may account for poor self-regulated learning decisions. Indeed, in direct comparisons, measures of subjective effort more accurately predict JOLs than objective effort (Koriat, Ackerman et al., 2014). Nevertheless, relatively few studies have directly assessed learners' experienced mental effort and task perceptions; rather, conclusions have often based on measures of objective fluency (e.g., Besken, 2016; Besken & Mulligan, 2013) or simply the researchers' assumptions about which conditions are likely to be more difficult or effortful. (For further discussion, see Mueller, Dunlosky, Tauber, & Rhodes, 2014.)

Lastly, and most critically, although the work reviewed above suggests a mediating role of perceived learning in relating perceived mental effort and self-regulated learning decisions, direct tests of this hypothesis are rare. The links between perceived mental effort and perceived learning (i.e., the ease-of-processing heuristic) and between perceived learning and metacognitive control of study have often been established and tested separately; rarely have these constructs all been measured together in the same study with the same subjects. (But see Susser, Jin, & Mulligan, 2016, for evidence that that perceived mental effort mediates the relationship of fluency to JOLs, albeit without a link to study strategy decisions.) It is even less common to investigate these relations while *also* controlling for objective performance even though, as we note above, the cue-utilization framework (Koriat, 1997) predicts that it should specifically be perceived rather than objective learning that governs metacognitive behavior.

1.5. Present work

Here, we test the misinterpreted-effort hypothesis by measuring the mediating role of perceived learning in self-regulated learning decisions. In each of the studies reported below, participants experienced two different self-regulated strategies. In Studies 1 through 3, we explicitly assessed the perceived mental effort and perceived learning associated with the strategies using self-report questionnaires. (In Study 4, we eliminate these ratings to evaluate the effects of the rating procedure itself.) Because our experimental paradigm was new and it was not clear *a priori* when best to assess these perceptions, we queried them twice with the goal of obtaining converging evidence: (a) immediately after participants had completed each individual strategy (as is often done to measure visceral experience; Loewenstein, 1996) and (b) retrospectively as a semantic-differential comparison of the two strategies (paralleling other metacognitive studies; Kornell & Bjork, 2008b; Kornell, Castel, Eich, & Bjork, 2010; Yan et al., 2016; Zulkiply, McLean, Burt, & Bath, 2012). Lastly, after experiencing and rating both strategies, participants chose one of the two strategies to use to learn new material in the future.

This method allowed us to assess (a) perceived mental effort, (b) perceived learning, and (c) study strategy choices all within the same participants within the same study. Further, it allowed us to *relate* these constructs; that is, does an individual's judgment of the effectiveness of (for instance) restudying material predict his or her intention to use that strategy in the future? In particular, we tested the mediation model predicted by our misinterpreted-effort hypothesis: The perception of greater mental effort leads learners to feel that they are actually learning *less*, which in turn is associated with lesser odds of choosing that strategy in the future. That is, perceived learning mediates the relationship between perceived difficulty and study strategy choice, yielding an indirect relation of perceived difficulty and choice. By contrast, the effort-avoidance hypothesis, in which learners "take the easy way out" and avoid effortful strategies *because* they are effortful, predicts that perceived effort should show a *direct* effect on strategy choice, independent of perceived learning. Lastly, the effort-disregard hypothesis, in which the choice of study strategies is primarily made on some other metacognitive basis (e.g., normative beliefs about which strategies are most effective) predicts we should observe neither a direct nor indirect effect of perceived effort on strategy choice (though perhaps a link between perceived learning and strategy choice).

In Studies 2 and 3, we also obtained an additional measure of participants' *objective* performance during learning. This allowed us to test two additional predictions of the cue-utilization framework: Because metacognitive judgments are inferential in nature, (a) it is *inferred* or perceived learning that ultimately predicts self-regulated learning choices, and (b) inferential cues, such as mental effort, thus influence perceptions and behavior over and above objective learning.

2. Study 1

In Study 1, we tested these hypotheses in the context of one of the many decisions that learners must make: how to *order* the material they plan to study. For instance, a statistics student practicing examples of *t*-tests, correlations, and ANOVAs to prepare for a statistics exam must choose the order in which to tackle those examples. Or—as in the present study—a student of biology might learn to identify birds one family at a time or all families at once. Specifically, we contrast two different strategies for ordering category exemplars. A *blocked schedule* involves studying exemplars grouped together by category (e.g., all of the finches, then all of the jays, then all of the orioles) whereas an *interleaved schedule* involves studying exemplars of multiple categories mixed together (e.g., a mix of finches, jays, and orioles).

We target this contrast for two reasons: First, it has received substantial attention in past studies of self-regulated learning as a case of apparent metacognitive failure: Learners persist in perceiving blocked schedules as more effective even when those same learners do better with interleaved schedules (Kornell & Bjork, 2008a; Kornell et al., 2010; Wahlheim et al., 2012; Yan et al., 2016; Zulkiply et al., 2012). Second, the superiority of interleaved schedules for category learning in many situations (e.g., Birnbaum, Kornell, Bjork, & Bjork, 2013; Hall, Domingues, & Cavazos, 1994; Kang & Pashler, 2012; Rohrer, Dedrick, & Burgess, 2014; Rohrer & Taylor, 2007; Taylor & Rohrer, 2010; Wahlheim, Dunlosky, & Jacoby, 2011; c.f., Kurtz & Hovland, 1956) has led the principle of interleaving to be regarded as an important example of the applicability of cognitive psychology to education (Roediger & Pyc, 2012; Rohrer & Pashler, 2010). In particular, we tasked participants with learning to identify families of birds because similar stimuli have been used in a number of past studies of the ordering of exemplars, which have found robust benefits of interleaved schedules (Birnbaum et al., 2013; Tullis, Benjamin, & Ross, 2011; Wahlheim et al., 2012), and because learning to identify taxonomic categories is an educationally relevant task (Dunlosky et al., 2013, p. 40).

2.1. Method

2.1.1. Participants

We recruited participants through Amazon Mechanical Turk and paid them for their one-time survey completion. In this and all following studies, participants were limited to individuals using a computer who were 18 years of age or older and reported speaking English fluently.

To determine an appropriate sample size, we first estimated effect sizes from a pilot study ($N = 10$), which resembled Study 2 in design except that it used a shorter, pilot version (including only questions 1–5 in Appendix C) of our questionnaire on participants' immediate perceptions of each study strategy. To ensure that our sample size was sufficient to detect differences even in more subtle measures, we used the effect size from the questionnaire item that showed the *smallest* difference between study-strategy conditions ("How boring was the last exercise?", $d = 0.32$). Analysis using G*Power 3.1.9.3 (Faul, Erdfelder, Lang, & Buchner, 2007) indicated we would have 80% power to detect differences between study strategies in this question with minimum $N = 77$.

Our data collection on Mechanical Turk was run in batches of participants completing the experiment in parallel; we continued running batches until we reached this minimum sample size, yielding 90 participants for Study 1. One participant was then excluded because of technical issues with the online survey, leaving 89 participants for analysis (44% female; 83% White, 6% Black/African American, 11% Asian, 1% American Indian or Alaskan Native; 6% Hispanic or Latino; 1% age 18–24, 36% age 25–34, 36% age 35–44, 18% age 45–54, 7.9% age 55–64, and 1.1% age 65–74; 1% some high school, 16% high school diploma, 35% some college, 35% bachelor's degree, 1% some graduate school, 10% master's degree, 2% doctoral degree).

2.1.2. Materials

2.1.2.1. Learning materials. In Study 1, participants learned to distinguish three different families of birds in two groups (i.e., six bird families total) on the basis of photographs. There were two study phases; in one, participants learned to distinguish finches, jays, and orioles, and in the other, participants learned to distinguish sparrows, tyrant flycatchers, and wood warblers.

We used digital photographs from a previous study (Tullis, Benjamin, & Ross, 2011) on the benefits of interleaved practice for category learning; the photographs were originally obtained via the Internet. Each photograph consisted of a color image of a bird measuring 322 × 403 pixels. One photograph of each bird family was presented during a demonstration phase, 15 photographs of each bird family were presented during each study phase (45 birds total per phase), and one photograph of each bird family was used in the final wrap-up phase.

An additional four bird families—penguins, flamingos, ducks, and parrots—were used in the initial two practice phases. For each of these families, three photographs were obtained from a stock photo website; these photographs varied in size somewhat but were generally of similar size as the study phase stimuli.

2.1.2.2. Immediate perception questionnaire. We created a questionnaire to assess participants' experience after completing each of the blocked and interleaved study strategies. Four questionnaire items assessed the perceived mental effort of the just-completed strategy (e.g., *How tiring was the last exercise?* 1 = *Not at all* to 6 = *A lot*), and four assessed participants' perceived learning from the strategy (e.g., *How likely are you to be able to distinguish between the types of birds?* 1 = *Not very likely* to 6 = *Extremely likely*). All items were on a 6-point Likert scale, with higher scores indicating greater endorsement. (See Appendix A for survey items.) As described in the Procedure, participants completed the questionnaire twice, once regarding the blocked schedule and once regarding the interleaved schedule.

We averaged the perception questionnaire items to create four composite scores: perceived mental effort of interleaved practice ($\alpha = 0.83$), perceived mental effort of blocked practice ($\alpha = 0.95$), perceived learning from interleaved practice ($\alpha = 0.76$), and perceived learning from blocked practice ($\alpha = 0.93$). The composites were created such that a higher score indicated greater endorsement of the construct (i.e., a more effective strategy and a more mentally effortful strategy).

2.1.2.3. Retrospective semantic-differential questionnaire. As a second source of converging evidence, and for consistency with other studies that had used end-of-experiment judgments (Kornell & Bjork, 2008b; Kornell et al., 2010; Yan et al., 2016; Zulkiply et al., 2012), we also created a second questionnaire querying participants' retrospective comparisons of the two strategies. This questionnaire consisted of eight semantic-differential items assessing participants' comparative perceptions of the two activities. In each item, participants judged between the two strategies on a continuum (e.g., *Which exercise required more mental effort?* 1 = *Grouped together to* 6 = *Not grouped together*). Four items assessed perceived mental effort ($\alpha = 0.86$) and four assessed perceived learning ($\alpha = 0.92$). (See Appendix B for survey items.)

2.1.3. Procedure

Participants completed the study on the Qualtrics online survey platform from a location of their choice. The overall task consisted of a demonstration phase, two practice phases, two study phases with associated immediate perception questionnaires, the

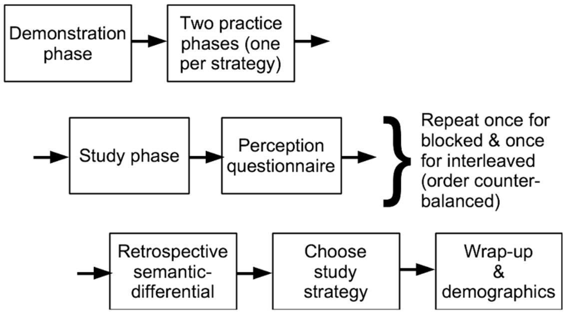

Fig. 2. Flow of study activities in Study 1.

retrospective-semantic differential questionnaire, a choice between the two strategies, a wrap-up phase, and demographics items. (See Fig. 2.) Participants' progress through the experiment was tracked by a progress bar at the top of the screen.

The demonstration phase was intended to present to participants, prior to learning, the type of knowledge they would need to acquire during the study phases. Participants were shown one bird from each of the six bird families included in the two study phases (one at a time) and were asked "Which type of bird is this?" Participants selected by choosing among the three bird families that would later appear together in the same study phase (i.e., either *finch, jay, or oriole?* or *sparrow, tyrant flycatcher, or wood warbler?*). After the participant made their choice, the correct answer was displayed (e.g., "The correct answer is **Oriole.**"), and the experiment advanced to the next exemplar. We expected that participants would be unlikely to know the correct answers to all of the demonstration items, thereby indicating to them that they would need to pay close attention to the study phases. This presupposition was supported: Mean performance on the demonstration phase (63%) was far from ceiling, with only 7% of participants answering all six demonstration items correctly.

In each subsequent study phase, participants learned to distinguish bird families by studying labeled exemplar pictures from each family, presented either (a) grouped by family (a blocked schedule, described to participants as "Grouped together by type" to avoid the presentation of jargon), or (b) randomly intermixed (an interleaved schedule, described as "Not grouped together by type"). In both conditions, each photo was presented one at a time for three seconds each with the bird family name shown above each image.

Two initial practice phases—one with a blocked schedule and one with an interleaved schedule—introduced the two study strategies to participants. In each practice phase, participants received only two bird families with three images each (penguins vs. flamingos in the first block, parrots vs. ducks in the second). When the practice phases were complete, participants were shown the prompt, "Next, you will begin the actual studying task. You will study the birds you will be tested on later. Remember, you will not be able to go back."

The practice phases were followed by the main study phases. In each of these phases, participants studied three families with 15 exemplars each. Immediately after completing one study phase, participants completed the immediate perception questionnaire regarding the just-completed strategy. Participants then used the other study strategy on the remaining trio of bird families and completed the perception questionnaire regarding that just-completed strategy. The order of strategies was counterbalanced across participants whereas the order of bird families was invariant; thus, across participants, each condition appeared equally with each stimulus set and in each serial position.

After experiencing and rating both strategies, participants completed the retrospective semantic-differential questionnaire to assess the perceptions of the strategies in relation to one another.

Next, participants chose which strategy they would use to study a new set of bird families. Participants were asked, "Imagine that you had to study more birds like you did today. Which strategy would you use to study the types of birds, so you would be able to take a test on them later?" Participants answered the two-choice question item by selecting either "Grouped together" or "Not grouped together."

Finally, in a wrap-up phase, one photograph from each bird family was presented; participants were asked "Which type of bird is this?" with answer choices corresponding to the correct bird family and the two other bird families that had appeared in the same study block. This wrap-up phase was not of primary theoretical interest (because our interest was in participants' perceptions and choices), but was included to fulfill the promise in the instructions that participants would take a test on the bird families. Following the wrap-up phase, participants complete a demographic questionnaire.

2.2. Results

2.2.1. Participant study strategy choice

When given the opportunity to block or interleave a third set of bird categories, 68% of participants chose a blocked schedule and 32% chose interleaved.

2.2.2. Immediate study-strategy perceptions

Why did participants favor a blocked schedule over interleaved? Participants judged the interleaved schedule ($M = 3.53$; $SD = 1.15$) as, on average, 0.21 points *less* effective (95% CI: $[-0.44, 0.02]$) for learning than the blocked schedule ($M = 3.74$;

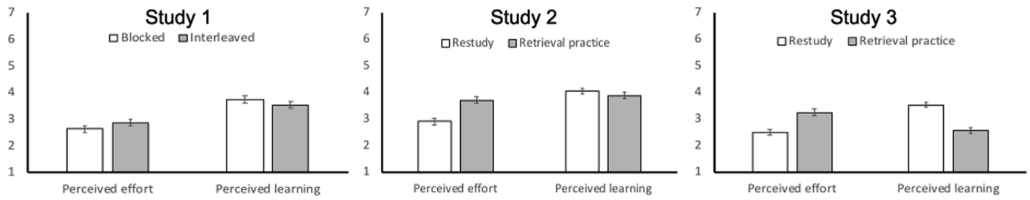

Fig. 3. Mean immediate perceived mental effort and immediate perceived learning effectiveness for different study strategies in each study.

SD = 1.19), $t(88) = -1.82$, $p = .073$, $d = 0.19$; this difference was marginally significant in a paired-samples *t*-test. They also rated the interleaved schedule ($M = 2.86$; SD = 1.17) as, on average, 0.23 points *more* mentally effortful (95% CI: [0.06, 9.38]) than the blocked schedule ($M = 2.63$; $SD = 1.19$), $t(88) = 2.79$, $p = .006$, $d = 0.30$, a significant difference. Fig. 3 displays the mean perceptions of the strategies in this and all following studies.

2.2.3. Relation of immediate perceptions to strategy choice

Next, we examined whether the perceptions of interleaved study as both less effective for learning and more mentally effortful explained participants' strategy selections. Since our critical dependent measure was participants' choice between a blocked and an interleaved schedule, the most relevant predictors were the differences between participants' perceptions of blocked vs. interleaved schedules, rather than the absolute levels. Thus, we created two difference scores: one for the difference in immediate perceptions of learning effectiveness for the interleaved schedule relative to the blocked schedule, and one for the difference in immediate perceptions of mental effort for the interleaved schedule relative to the blocked schedule. Positive values on these scales would indicate participants perceiving the interleaved schedule as more effortful or more effective for learning.

Separate logistic regression analyses showed that, for every point that participants perceived the interleaved schedule as more mentally effortful (relative to a blocked schedule), their odds of choosing it for future study declined 0.73 times (95% CI: [038, 0.88]), Wald $z = -3.09$, $p = .002$; Nagelkerke $R^2 = 0.18$. By contrast, for every point that participants perceived the interleaved schedule as more effective for learning (relative to blocked), their odds of choosing it increased 3.00 times (95% CI: [1.71, 5.24]) to choose it for future study, Wald $z = 3.84$, $p < .001$, Nagelkerke $R^2 = 0.28$. Given these independent direct effects of perceptions on study choice, we next tested the mediation hypothesis that perceived mental effort affects strategy choice by way of perceived learning.

We conducted a bootstrapped indirect effect analysis with 5000 samples using the PROCESS macro version 2.16.3 for SPSS (Hayes, 2013). In this model, the perceptions of mental effort difference score (interleaved minus blocked) was used as the independent variable, and the perceptions of learning effectiveness difference score (interleaved minus blocked) was the mediator. The dependent variable was strategy choice (0 = blocked, 1 = interleaved). Mediation was inferred if the 95% confidence interval for the crucial indirect effect parameter (i.e., perceptions of mental effort on strategy choice through the mediator) did not include zero.

Fig. 4 displays the results of this mediation analysis (overall Nagelkerke $R^2 = 0.30$). As hypothesized, participants who perceived the interleaved schedule as more mentally effortful (relative to blocked) also perceived interleaved material as less well learned (relative to blocked), $t = -6.83$, $p < .001$, $R^2 = 0.35$. Perceptions of the effectiveness of interleaving for learning in turn predicted greater odds of choosing an interleaved schedule for future study, $z = 2.83$, $p = .005$, partial $R^2 = 0.15$, indirect effect $= -0.76$, 95% CI $[-1.78, -0.10]$. The direct effect of perceived mental effort on strategy choice was not significant, $z = -1.30$, $p = .197$, partial $R^2 = 0.03$, suggesting full mediation.

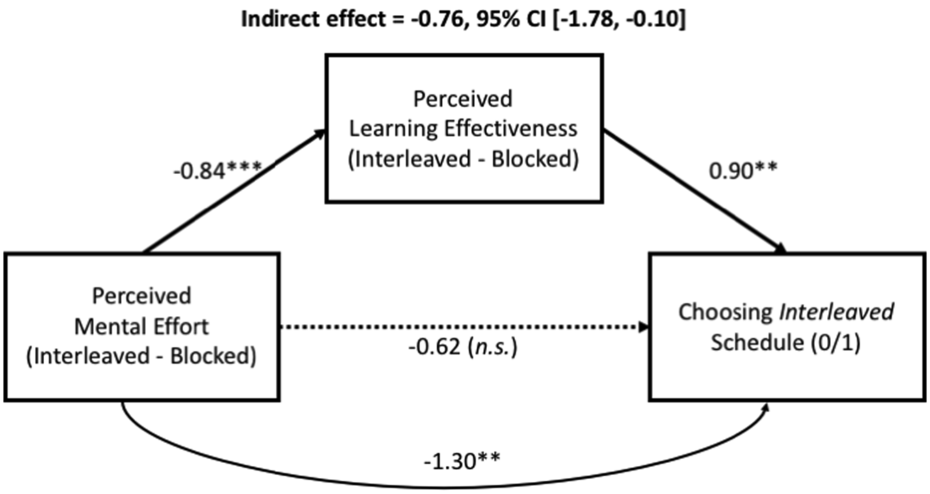

Fig. 4. Mediation analysis of immediate study strategy perceptions and choices in Study 1. $^{**}p < .01$, $^{***}p < .001$.

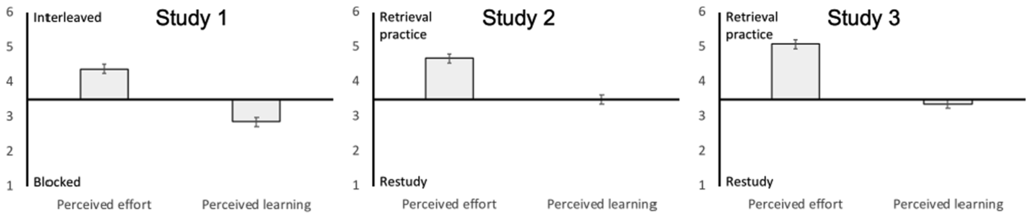

Fig. 5. Mean retrospective perceived mental effort and retrospective perceived learning effectiveness for different study strategies in each study. The solid line indicates the value on the scale that corresponds to no preference for one strategy over another.

2.2.4. Retrospective semantic differential

We also examined whether the same relations obtained in the retrospective comparisons of blocked and interleaved practice (displayed in Fig. 5). Because these judgments involved a single semantic-differential scale, with blocked practice on one end (1) and interleaved practice on the other (6), we compared participants' responses to the midpoint of the scale (3.5). These comparisons revealed similar a pattern as immediate perceptions: Learners judged an interleaved schedule as significantly more effortful than a blocked schedule ($M = 4.38$, $SD = 1.26$, $t(88) = 6.58$, $p < .001$, 95% CI: [4.11, 4.64], $d = 0.70$) and as significantly less effective for learning ($M = 2.85$, $SD = 1.51$, $t(88) = -4.03$, $p < .001$, 95% CI: [2.53, 3.17], $d = 0.43$).

2.2.5. Relation of retrospective comparisons to strategy choice

The retrospective comparisons predicted study strategy choices similarly to immediate perceptions: For every point that participants rated as the interleaved schedule as comparatively more effortful, their odds of choosing it for future study declined 0.87 times (95% CI: [0.70, 0.94]), Wald $z = -4.76$, $p < .001$, Nagelkerke $R^2 = 0.62$. And, for every point that participants who rated the interleaved schedule as more effective for learning, their odds of choosing it for future study increased 7.07 times (95% CI: [3.23, 15.48]), Wald $z = 4.90$, $p < .001$, Nagelkerke $R^2 = 0.73$.

We then applied the same mediation model (Fig. 6) to the retrospective semantic-differential items (overall Nagelkerke $R^2 = 0.78$). As hypothesized, participants who perceived the interleaved schedule as the more mentally effortful also perceived interleaved material as less well learned (relative to blocked), $t = -13.08$, $p < .001$, $R^2 = 0.66$, and this in turn predicted greater odds of choosing an interleaved schedule for future study, $z = 3.59$, $p < .001$, partial $R^2 = 0.39$, indirect effect $= -1.48$, 95% CI [-3.63, -0.39]. For the retrospective perceptions, the mediation was only partial, and there remained a significant direct effect of perceived mental effort on strategy choice, $z = -2.40$, $p = .016$, partial $R^2 = 0.17$.

2.2.6. Category learning

Lastly, although our primary interest in Experiment 1 was in how learners choose study strategies, we also considered the degree to which participants learned from the study materials. Recall that participants completed six categorization trials during the initial demonstration phase and another six in the final wrap-up phase; although primarily intended as part of our cover story, these also essentially constitute a pre-test and post-test. We compared categorization accuracy across the two time points. Overall categorization accuracy improved by 24 percentage points (95% CI: [19, 28]) from pre-test ($M = 47\%$; $SD = 17\%$) to post-test ($M = 71\%$; $SD = 21\%$), $t(88) = 9.79$, $p < .001$, $d = 1.04$, confirming that participants had learned from the study activities. Categorization accuracy did not significantly differ between the bird families studied with the interleaved schedule (M change score = 24%,

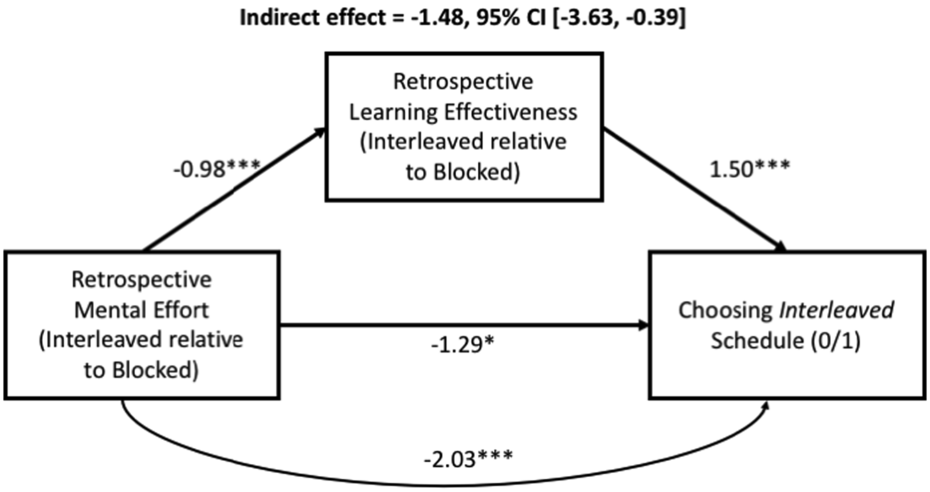

Fig. 6. Mediation analysis of retrospective study strategy comparisons and choices in Study 1. *$p < .05$, ***$p < .001$.

SD = 41%) and those studied with the blocked schedule (M change score = 23%, SD = 35%), $t(88) = 0.23$, $p = .82$, $d = 0.02$.

2.3. Discussion

In Study 1, we examined learners' perceptions of blocked and interleaved schedules and their choice between these strategies for future use. Consistent with prior research (Yan et al., 2016), we found that a sizable majority of people chose a blocked schedule over an interleaved schedule. This preference for a blocked schedule has been found even for tasks and materials for which an interleaved schedule results in superior learning (Kornell & Bjork, 2008b; Kornell et al., 2010; Wahlheim et al., 2012; Yan et al., 2016; Zulkiply et al., 2012), indicating it can sometimes be maladaptive. Note that we did not find such a significant benefit of interleaved practice within our present task, but that may be unsurprising given that our paradigm was designed to primarily assess differences in perception and was not ideal for testing ultimate learning (only three "test" trials per strategy).

Why do many learners choose a blocked schedule? In the present study, we directly queried participants' perceptions of each of the two strategies. We found that participants perceived the interleaved schedule as less effective for learning compared to the blocked schedule, replicating past results (Kornell & Bjork, 2008b; Kornell et al., 2010; Wahlheim et al., 2012; Yan et al., 2016; Zulkiply et al., 2012), even though there was no evidence of this in our (admittedly limited) measure of objective learning. In addition, we also observed that participants perceived an interleaved schedule as more mentally effortful. Further, the *more* that participants perceived the interleaved schedule as mentally effortful relative to a blocked schedule, the *less* effective for learning they judged it, and in turn, the less likely they were to choose an interleaved schedule for future study.

These results are noteworthy for several reasons. First, they provide evidence directly linking participants' perceptions of learning to their choice between interleaved and blocked schedules. Past work has found that learners typically judge an interleaved schedule as less effective for learning, but here we directly link those judgments to behavioral intentions by demonstrating that variation in the degree to which individuals perceive an interleaved schedule as effective for learning predicts their propensity to choose it over a blocked schedule for future study. More broadly, they extend the wealth of results demonstrating that metacognitive control of selecting individual *items* for study is governed by judgments of learning (Metcalfe & Finn, 2008; Son & Metcalfe, 2000); here, we observe that selections of entire study *strategies* are predicted by whether learners perceive those strategies as more or less effective for learning.

Second, the results reveal that learners also perceive an interleaved schedule as more mentally *effortful*. Past work has proposed that the reason learners disfavor interleaved schedules is ultimately rooted in a subjective experience of difficulty that arises when using an interleaved schedule (Kornell & Bjork, 2008b; Kornell et al., 2010; Wahlheim et al., 2012; Yan et al., 2016); however, those studies did not directly query participants' subjective experiences of effort or difficulty during study. Here, we confirmed that participants experience an interleaved schedule as more effortful than a blocked schedule.

Lastly, and most critically, our mediation model results provide support for our misinterpreted-effort hypothesis that learners' perceptions of greater mental effort using an interleaved schedule leads them to feel that they are actually learning *less* (in line with the ease-of-processing heuristic; Kornell et al., 2011), which in turn would lead them to *not choose* that strategy. This relationship was especially stark for participants' immediate perceptions of the strategies, where our results suggested that perceived learning fully mediated any influence of perceived effort on strategy choice, but it obtained via partial mediation in the retrospective judgments as well.

By contrast, there was only weaker evidence for the effort-avoidance hypothesis. There was not significant evidence for a direct effect of participants' immediate perceptions of mental effort on strategy choice; that is, it appeared learners did not avoid strategies simply because they were mentally taxing, but rather because they appeared to misinterpret that difficulty as conveying information about the strategy's effectiveness. When considering retrospective perceptions, mental effort did have a direct effect on strategy choice, but the indirect effect was still larger. The results also contradict the ignored-effort hypothesis, in which the choice of study strategies is based on factors other than the experience of mental effort.

Nevertheless, Study 1 contained at least one major omission: Our task provided us with no real measure of participants' *objective* performance during the study phase under these two competing schedules. This is a concern because individuals vary somewhat in the efficacy of these two schedules (e.g., in Yan et al., 2016, 10–30% of participants performed at least numerically better with a blocked schedule), and it is plausible that participants who are learning better with a strategy experience less mental effort. Thus, it is possible that the correlation we observed between perceptions of mental effort and perceptions of poor learning is wholly spurious and arose simply because of a confound with objective performance.

To rule out this confound in Study 2, we examine a different decision in self-regulated learning that permits us to examine initial, objective performance during learning *and* that tests the generality of our misperceived-effort hypothesis.

3. Study 2

In Study 1, we had examined the effects of perceived mental effort and perceived learning effectiveness in the context of choosing between blocked and interleaved schedules—a decision about the *ordering* of to-be-learned materials. However, these constructs are likely also relevant to numerous other decisions that learners face in the process of self-regulated learning. One such choice is how to engage with to-be-studied material. For instance, having read a chapter of a biology textbook, learners might choose to further engage with it either by re-reading the text (restudying) or by using practice questions to quiz themselves (retrieval practice). In Study 2, we sought to validate and extend our misinterpreted-effort hypothesis by testing whether it can also explain decisions about the *activities* performed during study.

An additional benefit of this design is that the use of retrieval practice allows provides us with a ready measure of objective performance during learning—namely, the degree of accuracy with which participants answer retrieval-practice trials. Because the benefits of retrieval practice are typically greater the more accurate that learners are during practice (Rowland, 2014), learners who perform more accurately during practice should in principle more strongly favor retrieval practice over restudy. Nevertheless, the cue-utilization framework (Koriat, 1997) posits that learners do not have direct access to their objective level of knowledge or learning and that metacognitive judgments are consequentially inferential (see also Schwartz et al., 1997). Thus, it should be ultimately be *perceived* learning, rather than *objective* performance, that determines the choice of study strategy. Indeed, it is precisely the lack of direct access to one's objective learning that necessitates heuristics such as interpreting effort as poor learning. Thus, our misinterpreted-effort hypothesis predicts an effect of mental effort on perceived learning even when controlling for objective learning.

3.1. Method

3.1.1. Participants

We again targeted minimum $N = 77$ based on our power analysis from Study 1 and ran batches of participants through Amazon Mechanical Turk until we reached this number, yielding 90 participants. One participant was excluded for technical reasons, leaving 89 participants for analysis (43% female; 74% White, 9% Black/African American, 18% Asian, 2% American Indian or Alaskan Native; 8% Hispanic or Latino; 11% high school diploma, 26% some college, 45% bachelor's degree, 1% some graduate school, 16% master's degree, 1% doctoral degree). The mean age of the 88 participants reporting their age was 35.7 years ($SD = 9.22$).

3.1.2. Materials

3.1.2.1. Learning materials. In Study 2, participants learned about various science topics. Two texts—one about the Great Barrier Reef (ReadWorks, 2013b) and one about pterosaurs (American Museum of Natural History, 2015)—were used in the main study phases in which participants experienced two different study strategies; each was approximately 1300–1400 words. A third text, about semiconductors (Weekly Reader Corporation, 2009), was used in the final phase and was shorter at approximately 1000 words. Each was written at an 8th to 12th grade level.

For the Great Barrier Reef and pterosaur texts, we constructed 15 sentences that each present one key fact from the text. Each sentence was constructed so that it could be presented as a fill-in-the-blank item in the retrieval practice condition (e.g., "Humpback whales use a method called ______ feeding for hunting.") or as a statement in the restudy condition (e.g., "Humpback whales use a method called <u>bubble net</u> feeding for hunting.").

Two additional texts—one on the history of the University of Pittsburgh Cathedral of Learning and the other on the Pittsburgh Panthers football team—were used in the initial two practice phases (University of Pittsburgh, 2016, 2017). These texts were much shorter than the main texts (82 and 77 words, respectively); one sentence was created for each practice text in the same manner as the 15 review phase items.

3.1.2.2. Immediate perception questionnaire. The questionnaire items assessing mental effort were identical to Study 1. The items assessing perceived learning effectiveness were adapted to reflect the characteristics of the prose materials (e.g., *How likely are you to remember the answers to the questions?* 1 = *Not very likely* to 6 = *Extremely likely*). (See Appendix C for survey items.) We averaged the respective perception questionnaire items to create composite scores in the same manner as Study 1. The alphas for all questionnaires were excellent, $\alpha s = 0.80$ to 0.93.

3.1.2.3. Retrospective semantic-differential questionnaire. The semantic-differential items assessing retrospective comparisons were identical to Study 1, except one item was modified with the word *birds* changed to *facts* (*Which facts do you think you'll remember better?*). The scale ranged from 1 = *Fill-in-the-blank* to 6 = *Review the fact* with $\alpha = 0.78$ for perceived effort and $\alpha = 0.91$ for perceived learning. (See Appendix D for survey items.)

3.1.3. Procedure

The sequence of activities was largely the same in Study 2 except that there was no additional demonstration phase showing what type of knowledge participants needed to acquire because the fill-in-the-blank and restudy items in the practice phase already demonstrated this. Thus, the sequence of activities consisted of two practice phases, two study phases with associated immediate perception questionnaire, the retrospective semantic-differential questionnaire, a choice between the two strategies, an open-ended wrap-up phase that asked about what participants learned from the three passages, and demographics items (See Fig. 7.) Unlike Study 1, participants did not see a progress bar at the top of the screen.

Prior to the first main study phase, two initial practice phases, one with retrieval practice and one with restudy (presented in the same order as in the critical study phases), introduced the study strategies to participants. These passages were presented for only one minute, had only a single review item each, and did not include the perception questionnaire.

In each critical study phase, learners read one text. The entire text was presented on the screen for seven minutes, after which the screen automatically advanced. Participants then reviewed 15 key facts from the passage using one of two randomly assigned review activities. In *retrieval practice*, participants filled in a blank in each fact with the missing word or phrase using their memory (as in Hinze & Wiley, 2011); in *restudy*, participants reviewed each fact in intact form (i.e., with the blank filled for them in with the correct word or phrase). The facts were presented one at a time; after restudying or completing the fact, participants clicked a "Next" button

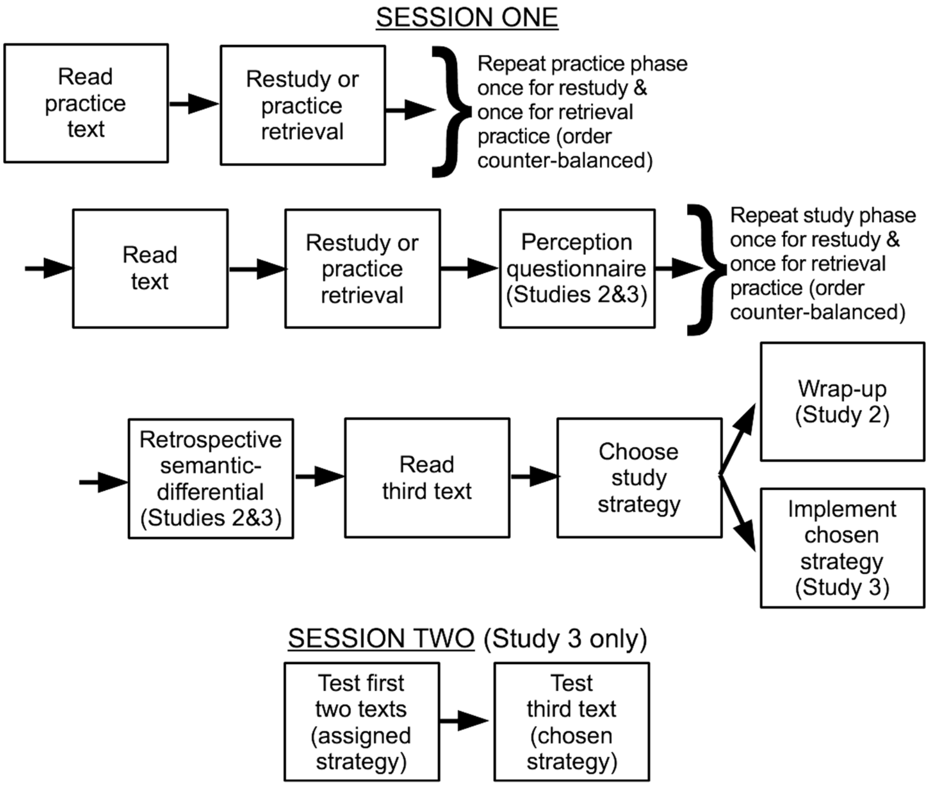

Fig. 7. Flow of study activities in Studies 2, 3, and 4.

to advance to the next item. Participants were not able to return to previous items. The facts were not presented in the same order as in the text but were presented in a single randomized order that was consistent across participants. Again, all cued-recall responses were scored as correct or incorrect by two trained research assistants ($\kappa = 0.96$); a third research assistant resolved any disagreements.

Immediately following the completion of the review activity, participants completed the immediate perception questionnaire regarding the just-completed strategy. To verify that the seven minutes of study provided sufficient time to participants to read the passage, the immediate-perception questionnaire additionally probed whether participants had had enough time to finish reading the associated passage; participants indicated they had finished the passage 90% of the time. After completing the immediate-perception questionnaire, participants then read and reviewed a second passage using the other study strategy and then answered the questionnaire about it. The Great Barrier Reef passage was always presented first and the pterosaurs passage second; however, the order of the study strategies was counterbalanced across participants. Thus, across participants, each condition appeared equally with each text and in each serial position.

As in Study 1, after experiencing and rating both strategies, participants completed the retrospective semantic-differential questionnaire. Participants then read the third text (on semiconductors), which was also displayed for seven minutes. After reading this text, participants were asked to choose between retrieval practice and restudying to study facts from it. Specifically, participants were asked, "Before you complete an exercise about this final passage, please select which you would like to use to help you learn the passage (so that you can take a test on it)." They were reminded of the two types of activities and were then asked, "Which type of exercise would you like to receive?" Participants answered the two-choice question item by selecting either "Review the fact" or "Fill-in-the-blank."

Since this strategy choice was our primary outcome measure, we did not require participants to implement the chosen strategy. Instead, participants were told they did not actually need to complete the strategy session for the third text. (In Study 3, participants actually implemented the chosen strategy to prepare for a test.) Rather, in the wrap-up phase, participants were simply asked, "What was the most interesting thing you learned from the three passages?" in order to fulfill the promise that they would be tested on the content they studied.

3.2. Results

3.2.1. Performance in retrieval practice

Participants answered 69% of the retrieval practice items correctly (*SD* = 22%). Meta-analysis (Rowland, 2014) suggests that retrieval practice robustly outperforms restudy even without feedback so long as participants are more than 50% accurate in their

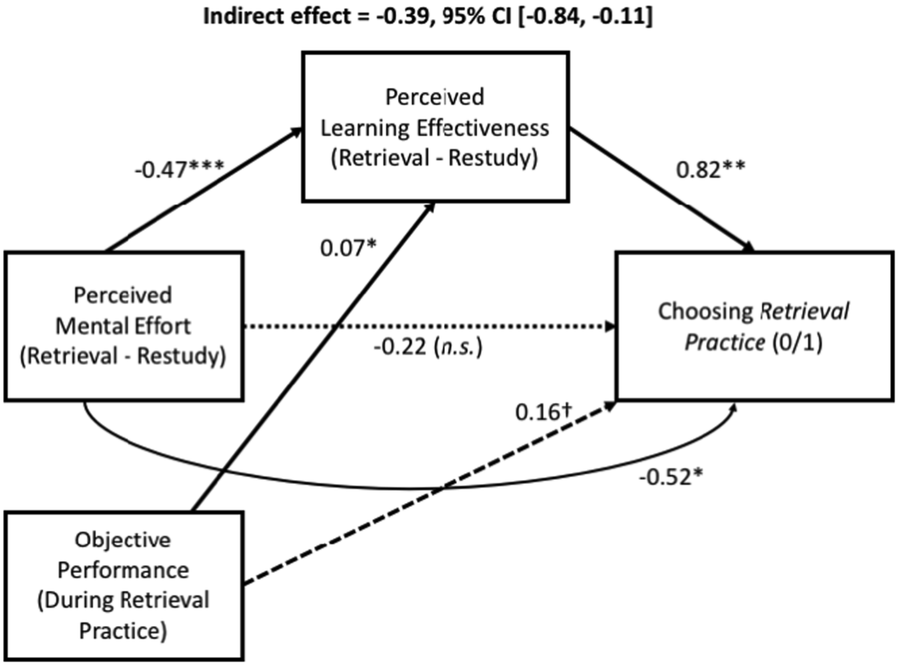

Fig. 8. Mediation analysis of immediate study strategy perceptions and choices in Study 2 while controlling for objective performance. †$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$.

practice attempts; thus, the level of performance in Study 2 is well within the range at which retrieval practice should be beneficial.

3.2.2. Participant study strategy choice

Nevertheless, when given the opportunity to choose retrieval practice or restudying for a third text, 31% of participants chose retrieval practice and 69% chose restudying.

3.2.3. Immediate study-strategy perceptions

Participants judged retrieval practice ($M = 3.87$; $SD = 1.15$) as, on average, 0.17 points *less* effective for learning (95% CI: [−0.44, 0.93]) than restudying ($M = 4.04$; SD = 1.10); this difference was not significant in a paired-sample *t*-test, $t(88) = -1.29$, $p = .202$, $d = 0.14$. They also rated retrieval practice ($M = 3.69$; SD = 1.25) as, on average, 0.79 points *more* mentally effortful (95% CI: [0.50, 1.08]) than restudying ($M = 2.90$; SD = 1.20), a significant difference, $t(88) = 5.392$, $p < .001$, $d = 0.57$.

3.2.4. Relation of perceptions to strategy choice

As we did in Study 1, we fit separate logistic regression analyses to test the direct-effect relationships between perceptions of mental effort and perceived learning on strategy choice (0 = restudy, 1 = retrieval practice). We again created difference scores by subtracting participants' scores on the mental effort and learning effectiveness scales taken immediately following restudy from these same scores taken immediately following retrieval practice; thus, positive values indicate when retrieval practice was viewed as more effortful or more effective for learning than restudy.

For every point that participants perceived retrieval practice as more mentally effortful (relative to restudy), their odds of choosing it for future study declined 0.44 times (95% CI [0.17, 0.62]) to choose it for future study, $z = -2.93$, $p = .003$, Nagelkerke $R^2 = 0.15$. Moreover, for every point that participants perceived retrieval practice as more effective for learning (relative to restudy), their odds of choosing it were 2.80 times greater (95% CI [1.59, 4.95]), $z = 3.56$, $p < .001$, Nagelkerke $R^2 = 0.27$. Given these bivariate relations with study choice, we next tested the hypothesis that perceived mental effort predicts strategy choice by way of perceived learning effectiveness.

A critical difference in Study 2 is that we could also include objective performance with retrieval practice (i.e., proportion accuracy during the practice trials) as a predictor of perceived learning and of strategy choice. Fig. 8 displays this expanded model[2] of immediate perceptions in Study 2. First, we consider influences on perceived learning (overall $R^2 = 0.35$). Objective performance did predict perceived learning; the better participants performed during retrieval practice, the more they perceived themselves as learning from retrieval practice (relative to restudy), $t = 2.22$, $p = .029$, partial $R^2 = 0.05$. Critically, mental effort incrementally

[2] Throughout all analyses, the same conclusions about the direct and indirect effects of mental effort were obtained if we excluded objective performance from the model. We focus here on the model that includes objective performance to test the prediction that perceived learning is a better predictor of study behavior than objective learning.

predicted perceived learning over and above objective learning, and, indeed, was a *larger* influence than objective performance; as hypothesized, participants who perceived retrieval practice as more mentally effortful (relative to restudy) also perceived material reviewed with retrieval practice as less well learned, $t = -5.89$, $p < .001$, partial $R^2 = 0.29$.

How did these constructs in turn influence study strategy choice (overall Nagelkerke $R^2 = 0.34$)? As hypothesized, perceptions of retrieval practice as comparatively effective for learning predicted greater odds of choosing retrieval practice for future study, $z = 2.63$, $p = .009$, partial $R^2 = 0.13$, indirect effect = -0.39, 95% CI [-0.84, -0.11]. The direct effect of perceived mental effort did not significantly predict strategy choice, $z = -0.94$, $p = .345$, partial $R^2 = 0.02$, suggesting full mediation, as in Study 1. Additionally, objective performance during retrieval practice had only a marginal effect on strategy choice that was smaller than that of perceived learning, $z = 1.73$, $p = .083$, partial $R^2 = 0.06$.

3.2.5. Retrospective semantic differential

We again turned to participants' retrospective comparisons, which were on a single semantic-differential scale[3] from restudy (1) to retrieval practice (6). Comparisons of participants' responses to the midpoint of the scale (3.5) again closely resembled the results from immediate perceptions: Learners judged retrieval practice as significantly more effortful than restudy ($M = 4.67$, $SD = 1.24$, $t(88) = -8.95$, $p < .001$, 95% CI: [4.41, 4.93], $d = 0.95$), but did not regard one strategy as significantly more effective for learning than the other ($M = 3.48$, $SD = 1.56$, $t(88) = 0.14$, $p = .892$, 95% CI: [3.15, 3.81], $d = 0.01$).

3.2.6. Relation of retrospective comparisons to strategy choice

As with the immediate perceptions, separate logistic regressions supported both of the bivariate relationships with study choice. For every point that participants rated retrieval practice as more mentally effortful relative to restudy, their odds of choosing it for future study were 0.59 times lower (95% CI: [0.37, 0.74]), $z = -4.02$, $p < .001$, Nagelkerke $R^2 = 0.29$. And, for every point that participants rated retrieval practice as more effective for learning, their odds of choosing it were 2.77 times greater (95% CI: [1.74, 44.3]), $z = 4.28$, $p < .001$, Nagelkerke $R^2 = 0.40$.

Given these bivariate relationships, we then applied our mediation model (Fig. 9) to the retrospective semantic-differential items. We first consider predictors of perceived learning (overall $R^2 = 0.35$). As with the immediate perceptions, there was a significant effect of objective performance during retrieval practice, $t = 2.22$, $p = .029$, partial $R^2 = 0.05$, but there was a much larger effect of mental effort such that the more participants perceived retrieval practice as comparatively more mentally effortful, the less they perceived it as comparatively effective for learning, $t = -5.89$, $p < .001$, partial $R^2 = 0.29$.

Turning to predictors of study strategy choice (overall Nagelkerke $R^2 = 0.52$), we found that perceptions of retrieval practice as effective for learning in turn predicted greater odds of choosing it for future study, $z = 3.41$, $p < .001$, partial $R^2 = 0.27$. Thus, there was an indirect effect of mental effort on study strategy choice (indirect effect = -0.41, 95% CI [-0.98, -0.12] . As in Study 1, this mediation was only partial when examining retrospective comparisons, and there remained a significant direct effect of perceived mental effort on strategy choice, $z = -2.73$, $p = .006$, partial $R^2 = 0.16$. Objective performance during retrieval practice did not significantly predict strategy choice over and above perceived learning and perceived effort, $z = 0.11$, $p = .254$, partial $R^2 = 0.03$.

3.2.7. Retention of information

As with Study 1, our primary interest was in learners' choices, but we also considered how the retrieval practice versus restudy strategies influenced participants' retention of the material. Recall that, at the end of the experiment, participants were prompted to report the most interesting fact they learned from the experiment. This task was included primarily as part of the cover story, but participants' choice of which fact to report may in part reflect their ability to recall the texts. Thus, we scored[4] whether this fact referred to the text reviewed with restudy ($M = 26\%$, SD = 45%), the text reviewed with retrieval practice ($M = 38\%$, $SD = 49\%$), the final text for which participants chose (but did not implement) a strategy ($M = 47\%$, $SD = 50\%$), or could not be coded ($M = 17\%$, $SD = 17\%$). Overall, participants were 11 percentage points (95% CI: [-1, 24]) more likely to report an interesting fact from the practiced text than the restudied text; this difference was marginally significant, $t(88) = 1.79$, $p = .077$, $d = 0.19$.

3.3. Discussion

In Study 2, we modified the procedure of Study 1 to test our theoretical model in the context of a different self-regulated learning decision: whether to practice retrieving learned material (in this case, in the form of fill-in-the-blank items) or restudy it in intact form. Most participants chose restudying over retrieval practice, and *why* our participants favored restudy replicated the same pattern for why they favored a blocked schedule in Study 1: Restudying was judged to be both more effective for learning and less mentally effortful than retrieval practice, and these variables predicted intentions for future study, such that the *more* mentally effortful participants perceived retrieval practice (relative to restudying), the *less* effective they felt retrieval practice was for learning, and in turn, the less likely they were to choose it over restudy.

An additional contribution of Study 2 is that we assessed participants' objective accuracy during retrieval practice, which meta-

[3] Due to researcher oversight, the semantic-differential scale in Studies 2 and 3 was in the opposite direction from Study 1 such that the normatively more effective strategy (retrieval practice) constituted the lower endpoint of the scale. We reverse the scale in reporting the results here to facilitate comparison across studies.

[4] Some responses referred to more than one text, which we scored as each of the texts referred to; thus, percentages sum to more than 100%.

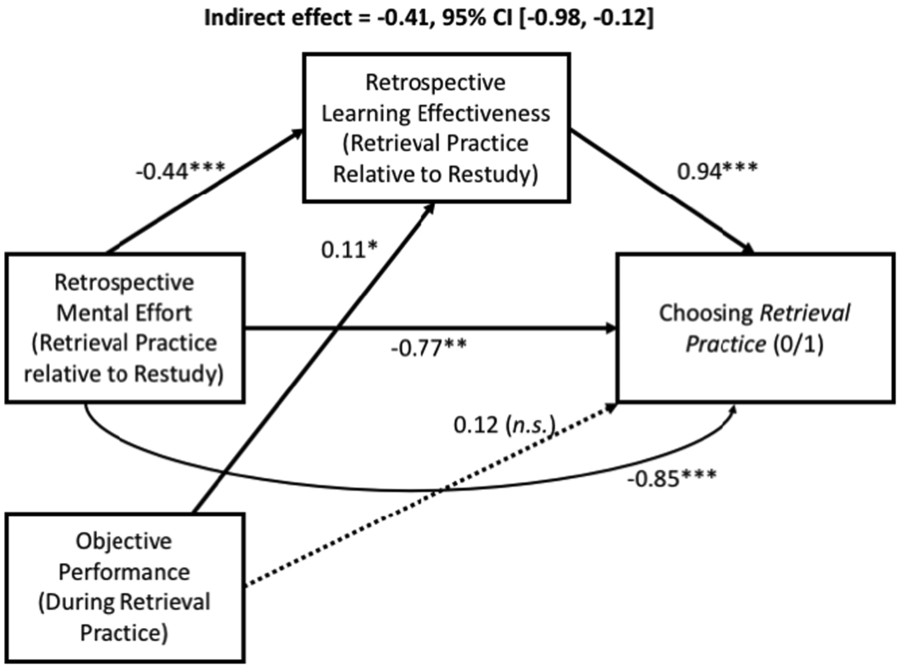

Fig. 9. Mediation analysis of retrospective study strategy comparisons and choices in Study 2 while controlling for objective performance. *$p < .05$, **$p < .01$, ***$p < .001$.

analysis has indicated modulates the magnitude of the retrieval practice benefit (Rowland, 2014). Objective performance did, indeed, predict whether participants considered retrieval practice comparatively effective for learning—but it was not as strong a predictor as perceptions of mental effort. Further, objective retrieval practice performance related only very weakly to the subsequent choice of whether to employ retrieval practice; this relationship was marginal in immediate perceptions, non-significant in retrospective comparisons, and of trivial effect size in both cases. Together, these results support the claims that metacognition is inferential and that learners do not have direct access to their level of learning or knowledge (Koriat, 1997; Schwartz et al., 1997): It was ultimately *perceived* learning, rather than objective performance, that predicted study intentions. This perceived learning must be inferred from other cues, such as mental effort; thus, the prominent cue of mental effort had a larger effect size in predicting perceived learning than did objective performance.

Combined, the results of Study 1 and Study 2 suggest that our misinterpreted-effort hypothesis may characterize a broad principle of self-regulated learning: Learners infer how well a study strategy contributes to their learning based in part on the experience of mental effort that it engenders. Strategies experienced as more mentally effortful are interpreted as less effective for learning and, consequently, not chosen. Across the two studies, this mediation model was observed to hold both across different kinds of self-regulated learning decisions (choosing the order of study materials versus choosing the activity performed on them) and across different kinds of to-be-learned materials (categorizing visual exemplars versus reading prose scientific content). In both studies, this relationship was particularly strong when participants' perceptions of a strategy were assessed immediately afterwards, where our results suggested that the misinterpreted-effort effect fully mediated the relationship of perceived effort to self-regulated learning decisions. For retrospective comparisons of the strategies, the mediation was only partial (though still statistically significant), and there was also a direct negative effect of a strategy's perceived effort on participants' choices to deploy it.

The studies presented to this point have shown that learners' perceptions of mental effort and, consequently, of learning effectiveness matter for their choice of study strategies. These choices are likely to be, in turn, consequential for learning outcomes insofar as some learning strategies are more effective for long-term learning than others. For example, meta-analysis (Rowland, 2014) indicates that retrieval practice outperforms restudy when participants are given feedback or even without feedback when initial performance is more than 50% accurate. Given that our participants performed well above this threshold, choosing retrieval practive over restudy would likely have contributed to their long-term learning of the texts; indeed, we saw some of this evidence insofar as participants were marginally more likely to recall an interesting fact from the practiced text than the restudied text. Further, given that the assignment of text topics to conditions was counterbalanced across participants, any such difference must be attributed to the study strategies and not to the content of one text being intrinsically more interesting than the other.

But, the benefits of retrieval practice were not generally reflected in our Study 2 participants' choices and perceptions: Participants did not judge retrieval practice as more effective for learning than restudy, they rated it as more effortful, and they were less likely to choose it for future study. Indeed, a general challenge is that strategies are often more effective for long-term learning are often initially more difficult or effortful (Schmidt & Bjork, 1992). A further implication of our results, then, is that misperceiving difficulty as poor learning may lead learners to choose less effective strategies and consequently learn and retain less than they could have. However, Study 2 was largely not designed to assess participants' long-term retention: We did not directly query whether participants accurately remembered the texts (only what they found most interesting), and we did not have participants review or attempt to recall the final text for which they choose their own strategy. In Study 3, we directly test the hypothesis that choosing retrieval practice over restudy predicts better retention—and, by extension, reveals consequences of perceiving those strategies as differentially effortful or effective.

4. Study 3

In order to test the consequences of learners' study strategy choices for long-term retention, we adapted the procedure of Study 2 to include two sessions separated by 48 h. In the initial session, participants experienced each of the restudy and retrieval practice strategies with different texts, and then chose a strategy to apply to a final text, as in Study 2; in addition, participants now actually applied that chosen strategy to the final text. Then, they returned for a second session 48 h later, in which their long-term memory for all of the texts was assessed via a cued recall test, identical in format to the retrieval practice items. This test allowed us to not only relate perceived effort and perceived learning to study-strategy choice, but to then relate study-strategy choice to retention; that is, did participants who studied the final text with retrieval practice indeed remember it better? In addition, by comparing memory for the two texts for which participants were assigned a strategy by the experimenter, we could carry out an experimental test of the standard retrieval practice effect (Karpicke & Roediger, 2008); that is, does retrieval practice produce greater long-term retention than restudy?

We also further tested the generality of our mediation model by extending it to (a) additional texts on other science topics and (b) to increase the relevance to theories of learning, a new population and setting: university students in a research lab, who more closely resemble classroom learners in both age and setting, rather than an online experiment that samples the general population.

4.1. Method

4.1.1. Participants

We again targeted minimum $N = 77$ based on our power analysis from Study 1. Because Study 3 used a two-session procedure and we expected some participant drop-out, we initially recruited 87 participants through the University of Pittsburgh undergraduate participant pool in exchange for partial credit towards a course requirement. Three participants were excluded who did not return for the second session, and four were excluded for technical issues that inhibited their completion of the study task, leaving 80 participants for analysis, sufficient for our desired level of power (54% female; 78% White, 8% Black/African American, 14% Asian, 1% American Indian or Alaskan Native, 4% other, 3% none of the options, 3% decline to state; 5% Hispanic or Latino; 68% first-years, 21% sophomores, 6% juniors, 4% seniors, 1% other). Their mean age was 18.9 years ($SD = 1.33$).

4.1.2. Materials

4.1.2.1. Learning materials. As in Study 2, participants learned about various science topics. To increase the generality of our findings, we added two additional texts (similar in length, reading level, and structure to those in Study 2) as an additional counterbalancing variable. Half the participants read the same texts as in Study 2; the other half instead read two new texts—one on the American Chestnut tree and one on water (ReadWorks, Inc., 2013a, 2013c)—and reviewed the associated 15 key passage facts. All participants still read the same final text on semiconductors. Because participants in this study now actually reviewed this final text using restudy or retrieval practice, we constructed 15 sentences of key facts from the final text as well. Similar to Study 2, participants could generally read the texts within the allotted seven minutes; participants indicated they finished reading the text 86% of the time.

4.1.2.2. Immediate perception questionnaire. Identical to Study 2. Alphas were adequate to excellent (αs = 0.73 to 0.88).

4.1.2.3. Retrospective semantic-differential questionnaire. Identical to Study 2. Consistency was excellent for retrospective comparisons of learning ($\alpha = 0.87$) but somewhat lower for effort ($\alpha = 0.67$).

4.1.2.4. Theory of intelligence questionnaire. Because at least one study suggests that the extent to which learners perceive difficulty as poor learning may vary as a function of their view of intelligence (Miele et al., 2011), we included on an exploratory basis a standard inventory of naïve theories of intelligence (Dweck, 1999). Participants responded to three items ($\alpha = 0.91$) on a 1–6 Likert scale ranging from *Strongly disagree* to *Strongly agree.* However, we observed no correlations between theories of intelligence on any of our critical measures,[5] so it will not be discussed further.

4.1.3. Procedure

Participants completed two study sessions in a research lab on the university's campus. Participants completed both sessions in a

[5] We note that this absence of a relation between naïve theories of intelligence and self-regulated learning perceptions or choices is consistent with at least two other investigations (Koriat, Nussinson et al., 2014; Yan, 2014) although a growth mindset may predict more nuanced appreciation of particular strategies (Yan, Thai, and Bjork, 2014).

lab room with up to two other participants; however, participants did not communicate with one another during the experiment.

The first session was identical to Study 2, except that (a) as noted above, half of the participants in Study 3 received two different texts and associated sentences in the initial phases of the study, (b) participants completed the theory of intelligence questionnaire, (c) participants did not receive an open-ended wrap up item (since they now received an actual test in the second session), and (d) most critically, participants actually performed their chosen study strategy (retrieval practice or restudy) on the final text. That is, after reading the final text, participants were asked, "Before you complete the exercise about the passage, please select which type of exercise you would like to use to help you learn the passage (so that you will be able to take a test on it during your next visit to the lab). Remember – For *review the fact*, you had to reread sentences that contained information from the passage. For *fill-in-the-blank*, you had to figure out the answers yourself based on your memory of the passage. Which type of exercise would you like to receive?" After selecting their strategy of choice, participants reviewed the 15 key passage facts in either the retrieval practice or restudy form.

Thus, the sequence of activities for the first session consisted of two practice phases, two study phases with associated immediate-perception questionnaires and questions about whether participants finished reading in the time allotted, retrospective semantic-differential questionnaire, the theory of intelligence questionnaire, a choice between the two review strategies, the review of the final passage facts using the chosen strategy, and demographics and background characteristics items. (See Fig. 7, above.)

Participants then returned to the lab 48 h after the start of their first session. In the second session, participants' memory was tested using the 15 fill-in-the-blank items for each of three passages texts studied during Session 1 (totaling 45 test questions). The texts were tested in the same order in which they were presented in Study 1; however, the 15 test items within each passage were presented in a new re-randomized order, held constant across participants. The test items were presented one at a time; participants typed their response in the blank and then clicked "Next" to move on to the next item. Just as in the first session, participants were not able to return to previous test items. Two trained research assistants scored each response to the cued-recall item as correct or incorrect ($\kappa = 0.95$); where they disagreed, a third research assistant broke the tie.

Lastly, at the end of the second session, participants answered two open-ended questions about their experience in the study; although not central to our hypothesis testing, these questions provided some additional, qualitative insight into participants' experiences. The first question asked participants to self-report their rationale for choosing their study strategy for the third text ("Why did you make that choice?"). We coded these responses for the presence of four themes suggested by our hypotheses and by the broader literature: choosing a strategy based on what is most effective for learning or memory (Nelson & Narens, 1990), choosing based on ease or avoidance of effort (Kurzban et al., 2013), choosing based on habit or typical study (Ariel & Dunlosky, 2013; Ariel et al., 2011), or choosing what was more interesting or engaging (Son & Metcalfe, 2000). The second question asked participants whether they thought their chosen study strategy was indeed more effective than the alternative ("Do you think there would have been a difference in your accuracy today if you would have chosen the other type of exercise?"). We coded these responses into three categories: participants who thought they would have learned more poorly with the other strategy, participants who thought they would have learned better, and participants who thought their choice made no difference in their learning.

4.2. Results

4.2.1. Performance in retrieval practice

All participants were assigned one text on which to use retrieval practice. Participants answered 54% of those retrieval practice items correctly (*SD* = 19%), which is notably less than in Study 2; we revisit this point in the Discussion below. For the participants who chose to use retrieval practice on the third text, practice performance on that text was slightly higher at 57% (SD = 18%), but still lower than in Study 2.

4.2.2. Participant study strategy choice

When given the opportunity to choose retrieval practice or restudying for the third text in Session 1, 47.5% of participants chose retrieval practice and 52.5% chose restudying.

4.2.3. Immediate study-strategy perceptions

Participants' perceptions of retrieval practice and restudy significantly differed on both dimensions: Participants judged retrieval practice ($M = 2.55$; SD = 0.91) as, on average, 0.98 points less effective for learning (95% CI [−1.27, −0.68]) than restudying ($M = 3.52$; SD = 0.96), $t(79) = -6.59$, $p < .001$, $d = 0.74$. They also rated retrieval practice ($M = 3.24$; SD = 1.11) as, on average, 0.76 points more mentally effortful (95% CI: [0.53, 1.00]) than restudying ($M = 2.48$; SD = 0.92), $t(79) = 6.40$, $p < .001$, $d = 0.72$.

4.2.4. Relation of perceptions to strategy choice

As in Study 2, separate logistic regression analyses using difference scores showed that for every point that participants perceived retrieval practice as more mentally effortful relative to restudy, their odds of choosing retrieval practice for future study declined 0.48 times (95% CI [0.14, 0.68]) to, Wald $z = -2.54p = .011$, Nagelkerke $R^2 = 0.12$. By contrast, for every point that participants perceived retrieval practice as relatively more effective for learning, their odds of choosing it were 2.22 times greater (95% CI [1.44, 3.43]), Wald $z = 3.60$, $p < .001$, Nagelkerke $R^2 = 0.25$.

Fig. 10 depicts the mediation model for immediate perceptions in Study 3, controlling for objective performance during retrieval practice. We first examine what predicted perceptions of learning (overall $R^2 = 0.46$). As in Study 2, objective performance during retrieval predicted perceptions of material studied under retrieval practice as comparatively well-learned, $t = 4.37$, $p < .001$, partial $R^2 = 0.20$. But, the stronger predictor was mental effort, such that participants who perceived retrieval practice as relatively more

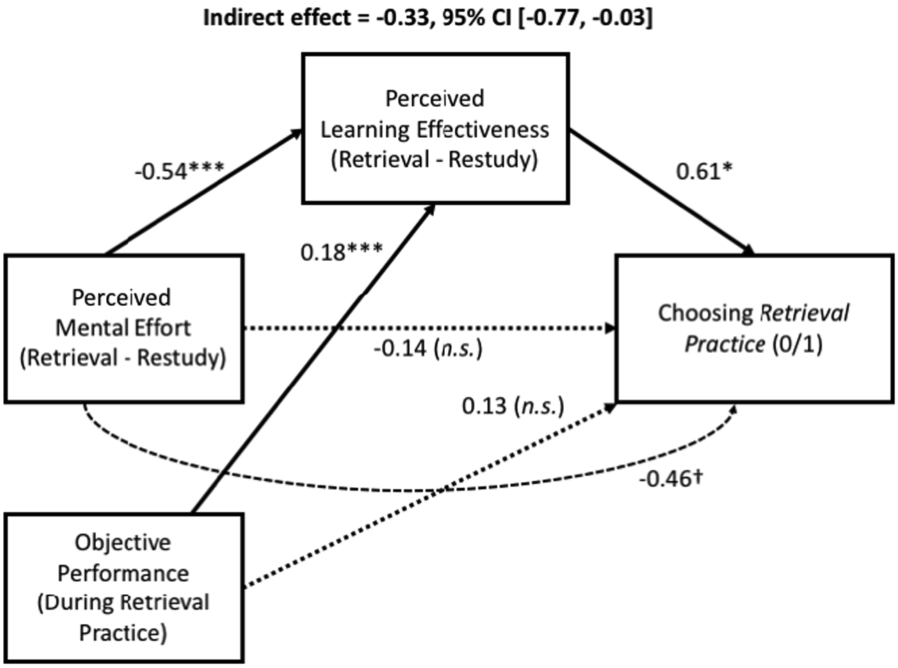

Fig. 10. Mediation analysis of immediate study strategy perceptions and choices in Study 3, controlling for objective performance. †$p < .10$, *$p < .05$, ***$p < .001$.

mentally effortful also perceived it as relatively ineffective for learning, $t = -4.85$, $p < .001$, $R^2 = 0.23$.

Next, we consider what predicted participants' choice of study strategy for the third text (overall Nagelkerke $R^2 = 0.27$). Perceptions of the effectiveness of retrieval practice for learning in turn predicted greater odds of choosing retrieval practice for future study, $z = 2.23$, $p = .026$, partial $R^2 = 0.09$, indirect effect of mental effort $= -0.33$, 95% CI $[-0.77, -0.03]$. The direct effect of perceived mental effort did not significantly predict strategy choice, $z = -0.45$, $p = .651$, partial $R^2 < 0.01$, suggesting full mediation. Nor did objective performance during practice significantly predict study strategy choice when controlling for perceived learning, $z = 1.16$, $p = .246$, partial $R^2 = 0.02$.

4.2.5. Retrospective semantic differential

Comparisons of the semantic-differential scale to its midpoint (3.5) revealed that learners judged retrieval practice as significantly more effortful than restudy ($M = 5.08$, $SD = 0.93$, $t(79) = -15.16$, $p < .001$, 95% CI: [4.87, 5.29], $d = 1.70$), but, contrary to the immediate perceptions, did not regard one strategy as significantly more effective for learning than the other ($M = 3.36$, $SD = 1.55$, $t(79) = -0.84$, $p = .402$, 95% CI: [3.03, 3.69], $d = 0.09$).

4.2.6. Relation of retrospective comparisons to strategy choice

As in Study 2, separate logistic regression analyses supported each of the bivariate relations with strategy choice. For every 1-unit increase in participants' rating of the mental effort of retrieval practice relative to restudy, their odds of choosing it for future study declined 0.58 times (95% CI: [0.01, 0.66]), $z = -1.97$, $p = .049$, Nagelkerke $R^2 = 0.07$. And, for every 1-unit increase in participants' rating of retrieval practice as effective for learning relative to restudy, their odds of choosing it 2.81 times increase (95% CI: [1.82, 4.36]) to choose it for future study, Wald $z = 4.64$, $p < .001$, Nagelkerke $R^2 = 0.44$.

We applied our mediation model (Fig. 11) to the retrospective semantic-differential comparisons in Study 3. Again, we find that perceptions of retrieval practice as comparatively effective for learning (overall $R^2 = 0.31$) were predicted positively by objective performance during retrieval practice, $t = 4.89$, $p < .001$, partial $R^2 = 0.24$, and negatively by perceptions of retrieval practice as comparatively effortful, $t = -2.32$, $p = .023$, partial $R^2 = 0.07$.

When we turn to the choice of strategy for the third text (overall Nagelkerke $R^2 = 0.44$), perceptions of retrieval practice as comparatively more effective for learning in turn predicted greater odds of choosing it for future study, $z = 3.88$, $p < .001$, partial $R^2 = 0.31$, indirect effect of mental effort $= -0.34$, 95% CI $[-0.87, -0.01]$. Unlike in the retrospective analyses in Study 1 and 2 (but consistent with all of the analyses of immediate perceptions), the direct effect of perceived mental effort on strategy choice was not significant, $z = -0.64$, $p = .522$, partial $R^2 = 0.01$, suggesting full mediation, nor was there a significant effect of objective performance, $z = 0.05$, $p = .457$, partial $R^2 < 0.01$

4.2.7. Post-test for assigned study activities

The most critical addition to Study 3 was a second session to examine how the study strategies affected participants' actual long-

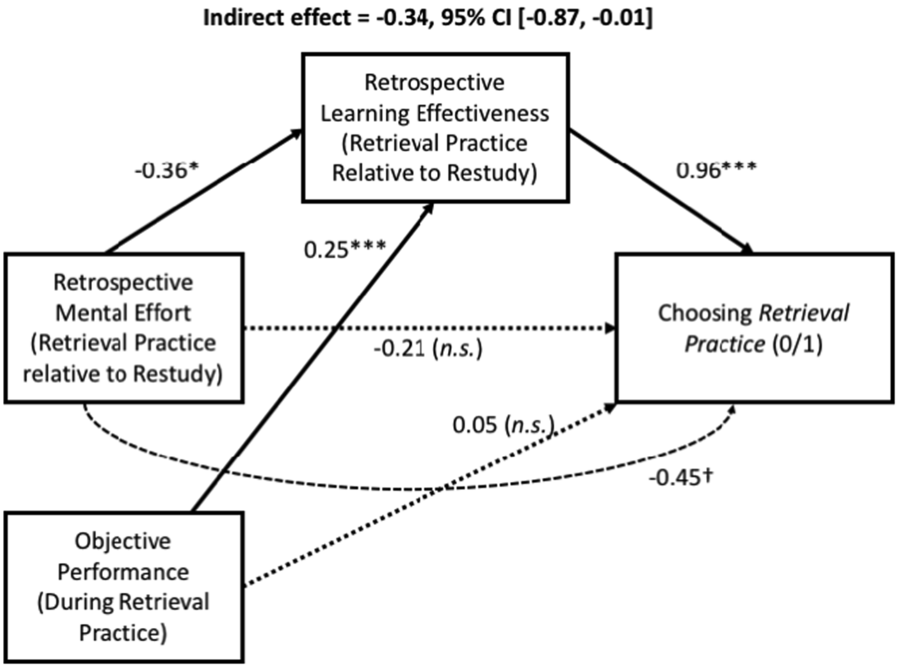

Fig. 11. Mediation analysis of retrospective study strategy comparisons and choices in Study 3, controlling for objective performance. †$p < .10$, *$p < .05$, ***$p < .001$.

term memory for the texts. We first compared performance on the initial texts, for which participants were assigned a study strategy. A paired-samples t-test revealed that participants were marginally better at remembering texts for which they were assigned to use restudy ($M = 8.31$ items correct out of 15; SD = 2.79) than assigned to use retrieval practice ($M = 7.69$; SD = 2.80), M of the difference = 0.63 (95% CI: [−0.057, 1.31]), $t(79) = 1.82$, $p = .072$, $d = 0.20$. This is counter to the typical testing effect, wherein material reviewed using retrieval practice strategy yield long-term retention than material reviewed using restudy.

4.2.8. Post-test for chosen study activity

Recall that for the final text, learners made their own choice during the first session whether to use restudy or retrieval practice. An independent samples t-test revealed that participants who selected retrieval practice ($M = 7.95$ items correct of 15, $SD = 2.42$) remembered significantly more of the final passage than participants who selected restudy ($M = 6.40$, $SD = 3.12$), M of the difference = 1.54 (95% CI: [0.29, 2.79]), $t(78) = 2.46$, $p = .016$, $d = 0.55$.

4.2.9. Open-ended items

Although our primary interest was in the quantitative measures, we also examined responses to the two open-ended questions at the end of the experiment.

The first question asked participants to freely report why they had chosen the study strategy they selected. Overall, 68% of participants referred to choosing a strategy better for learning, 19% referred to the ease of the strategy, 6% referred to habit, and 1% referred to choosing the strategy that was more interesting[6]. The rates at which participants self-reported choosing a strategy based on learning effectiveness did not significantly vary between restudy ($M = 64\%$, $SD = 49\%$) and retrieval practice ($M = 71\%$, $SD = 46\%$), $t(78) = -0.64$, $p = .525$, $d = 0.14$. This result is consistent with our mediation analyses, in which the primary predictor of strategy choice was always perceived effectiveness for learning. Although ease of processing was cited less frequently by participants as a reason for their choice, participants were more likely to mention this factor when they chose restudy ($M = 29\%$, $SD = 46\%$) than when they chose retrieval practice ($M = 8\%$, $SD = 27\%$), $t(78) = 2.42$, $p = .018$, $d = 0.54$. This response also accords with our mediation analysis, which suggests a secondary direct path whereby some participants favor restudy simply because it is less effortful.

The second open-ended question queried whether participants thought their choice of strategy ended up being preferable for learning the material. Although less relevant to our primary research question of *why* participants choose the study strategies they do, it is interesting to note that participants who chose retrieval practice were marginally more likely to feel that their choice helped them learn ($M = 58\%$, $SD = 50\%$) than participants who chose restudy ($M = 38\%$, $SD = 49\%$), $t(78) = 1.78$, $p = .078$, $d = 0.40$. This is consistent with the claim that, at least under some conditions, participants may come to recognize the value of such strategies

[6] These percentages do not sum to 100% because some responses referred to more than one theme and some referred to none of these themes or were uncodeable.

with experience (Tullis et al., 2013).

4.3. Discussion

In Study 3, more participants again chose restudy to learned facts rather than practice retrieving them, although this preference was not as strong as in Study 2. Nonetheless, our primary research question was not about the base rate at which any particular strategy is chosen, but rather about *why* participants who choose a particular strategy favor that strategy. Here, Study 3 closely replicated Study 2: Retrieval practice was perceived to be significantly more mentally effortful than restudying, and restudying was judged as significantly more effective for learning than retrieval practice. Most critically, our hypothesized mediation model was again supported such that the *more* participants perceived a strategy to be mentally effortful, the *less* effective they felt it was for their learning, and in turn the less apt they were to choose that strategy for future study. In Study 3, full mediation was supported for both measures of participants' immediate perceptions of each strategy and their retrospective comparison of the two strategies; the perceived effort required by the study strategies did not have a significant direct effect on which strategy participants chose. These results further extend and validate our model with a different subject population (students in an undergraduate participant pool), a different study environment (a university research lab), and additional to-be-learned science content.

The most important addition to Study 3 was a second session two days later to assess participants' long-term memory for the materials. Because self-regulated learning strategies differ in their effectiveness for long-term retention, participants' perceptions and choices are likely to have consequences for how effectively they learn. Our results were consistent with that claim: Participants who chose to review a science passage using retrieval practice better remembered the passage two days later than participants who merely restudied facts from the passage. We caution that because participants self-selected into choosing retrieval practice or choosing restudy, we cannot say on the basis of the present evidence alone that retrieval practice *caused* the superior learning, but past work (Rowland, 2014) does suggest the effect of retrieval practice is indeed causal. (We consider this issue further in Section 6.5.) In summary, perceiving retrieval practice as more effortful was associated with perceiving it as less effective for learning, and in turn not choosing it—a decision that was likely unfortunate because choosing retrieval practice was in truth associated with *better* long-term retention.

One caveat is that, for the two passages for which participants were randomly assigned a strategy, we did not replicate the typical benefits of retrieval practice over restudy. One reason for this may be that our study materials were comparatively difficult even during the first session: Meta-analysis (Rowland, 2014) has demonstrated that the benefits of retrieval practice are greater the better that learners perform during that retrieval practice, with an effect of retrieval practice potentially absent entirely if initial accuracy is less than 50%. In the case of the present study, average retrieval practice performance for these texts (55%) was only barely above that threshold. Thus, these passages may have simply been too difficult to elicit a reliable benefit of retrieval practice. Some evidence supports this link between initial retrieval practice performance and retention: The better that participants performed on the initial retrieval practice items, the better their retention for that practiced text relative to the restudied text (i.e., they showed a larger testing effect), $r = 0.49$, $p < .001$.

Nevertheless, the primary goal of Study 3 was to examine the consequences of study strategy *choice*, and for the critical text for which participants made their own choice of strategy, we did indeed see a typical testing effect whereby choosing retrieval practice was associated with more enduring learning than choosing restudy—and this choice, in turn, was predicted by participants' perceptions of the mental effort and learning effectiveness of retrieval practice.

5. Study 4

In our first three studies, we found that measures of participants' perceived mental effort and perceived learning predicted their choice of study strategy. This pattern is in line with our hypothesis that participants' choices of study strategies are typically influenced by such factors. However, an alternate explanation is that this relationship is an artifact of our experimental procedure, which required participants to rate the strategies on those two dimensions. It is possible that the process of rating study strategies on these dimensions itself induced participants to choose strategies on that basis even if they would not normally; for instance, the questionnaire items themselves may alter participants' mindset (Duckworth & Yeager, 2015; Job, Dweck, & Walton, 2010), or participants may desire to appear consistent with their previous questionnaire responses[7] (Peer & Gamliel, 2011). In particular, having rated restudy favorably compared to retrieval practice, participants may have then felt obligated to choose it as their study strategy.

We conducted Study 4 to test for this possibility. We used the same procedure as our one-day Study 2 except that participants did *not* provide any sort of ratings of the study strategies. Instead, they simply experienced both strategies and then made a final choice of strategy for the third text. If the act of rating strategies for perceived effort and learning had inflated participants' preference for the strategy they preferred on those dimensions (which, in this case, had been restudy), then we should see that preference reduced when participants did not make any such ratings.

[7] We thank an anonymous reviewer for suggesting this possibility.

5.1. Method

5.1.1. Participants

We recruited participants through the University of Pittsburgh undergraduate participant pool in exchange for partial credit towards a course requirement ($N = 57$ based on availability; 39% female; 68% White, 7% Black/African American, 16% Asian, 7% other race; 2% declined to report; 2% Hispanic or Latino; 19% high school diploma, 74% some college, 5% bachelor's degree, 1% some graduate school, 16% master's degree, 1% doctoral degree). The mean age of the 56 participants reporting their age was 19.75 years ($SD = 1.52$).

5.1.2. Materials

The materials were identical to Study 2.

5.1.3. Procedure

We adapted the procedure from Study 2 by eliminating both the immediate-perception and retrospective semantic-differential questionnaires. Participants experienced each of the strategies without rating, then made a final choice of strategy for the final third text. (See Fig. 7, above.)

5.2. Results

The sole measure collected in Study 4 was participants' final choice of study strategy. We compared these choices to those in Study 2, which used an identical one-day procedure but with the inclusion of ratings. In Study 4, 23% of participants chose retrieval practice and 77% chose restudying. A chi-square test for independence showed that these proportions did not significantly differ from the preferences in Study 2 (31% and 69%, respectively), $\chi^2_{(1)} = 0.90$, $p = .344$.

5.3. Discussion

Eliminating the ratings of perceived mental effort and perceived learning in Study 4 did not significantly change participants' ultimate preference for restudy versus retrieval practice. Indeed, participants were numerically *more* likely to choose restudy in Study 4 than in Study 2. Thus, there was no evidence that the act of providing favorable ratings to one strategy (i.e., restudy) in prior studies had artificially inflated participants' propensity to choose that strategy.

6. General discussion

In four studies, we examined the reasons and processes behind learners' study strategy choices and (in Study 3) the consequences of those choices. In the first three studies, participants experienced two different (and theoretically contrasting) study strategies, rated each for the mental effort experienced and its perceived effectiveness for learning, and made a choice of study strategy for future use. Although each of these constructs has been viewed as relevant to self-regulated learning, few studies have assessed all of them at the same time in the same participants to determine their relations.

Across these first three studies, a consistent pattern emerged: The more learners perceived a study strategy as mentally effortful, the less they judged it to be effective for learning; this, in turn, predicted their choice of study strategies such that the more effective that participants thought a strategy was for learning, the more likely they were to choose it for future study. That is, we observed an indirect effect of perceived effort on study strategy choices, mediated by perceived learning. This pattern held across differences in study strategies (the choice between blocked vs. interleaved schedules and the choice between restudy vs. retrieval practice), to-be-learned materials (learning to classify category exemplars and learning of science texts), subject populations (students and the general adult population), and locations (the Web or a laboratory). This indirect effect also always held across both participants' measures of immediate perceptions of each strategy and retrospective comparisons of the two strategies. It also held even when controlling for participants' *objective* performance during study.

Further, this mediation effect generally outweighed other influences on study choice that we measured: When we considered immediate perceptions of each strategy, there was no significant direct effect of perceived effort on study choices, suggesting the influence of perceived mental effort was fully mediated by how it influenced by perceptions of study. When we instead considered retrospective comparisons of each strategy, an additional direct effect emerged in two of the three studies whereby participants were less likely to choose strategies they regarded as mentally effortful, but the mediation effect remained even in those cases. Moreover, Study 3 suggested full mediation even in the retrospective comparisons.

Further, Study 3 suggests that these perceptions and choices have consequences for long-term retention of science texts: Not only did perceptions predict whether or not participants chose to employ retrieval practice, that choice in turn predicted performance on a test of retention 48 h later.

Lastly, in Study 4, we tested the possibility that providing the act of favorable ratings for certain study strategies, such as restudy, itself drove participants' subsequent choice of those strategies (e.g., if participants tried to conform their choice to their ratings). We adapted Study 2 to eliminate the ratings of perceived mental effort and perceived learning, but this alteration did not decrease (and, in fact, numerically *increased*) participants' ultimate preference for restudy relative to the original Study 2. Thus, there was no evidence that participants' preference for certain strategies was an artifact of providing favorable ratings to them in our procedure.

6.1. Metacognitive monitoring and control

Together, this series of studies makes several contributions to our understanding of the metacognitive processes that underlie self-regulated learning. First, they demonstrate that learners' choices about how to study are predicted by their judgments of which strategies result in the best learning. Prior work has demonstrated that learners often elect *which* individual items to study based on their judgments of learning (Son & Metcalfe, 2000). Here, we show that decisions about *how* to study (i.e., which study strategies to apply) are also predicted by perceived learning, at the level of a strategy as a whole rather than individual items. This result furthers the claim that learners can exert effective metacognitive control (in the framework of Nelson & Narens, 1990) insofar as they can translate their metacognitive judgments into corresponding study decisions (even if those judgments themselves may be erroneous). More generally, this work supports the idea that perceptions experienced while studying have consequences for future study choices.

Second, our results reinforce the long-held notion that metamnemonic processes are *inferential* (Koriat, 1997; Schwartz et al., 1997): Learners do not have direct access to their state of knowledge or rate at which they are learning. Rather, they must *infer* this from cues in their mental and/or external environment (Jacoby & Whitehouse, 1989; Koriat, 1997), using heuristics that are often accurate but that are not infallible (Benjamin, Bjork, & Schwartz, 1998), such as using the ease of processing to infer learning (Kornell et al., 2011). These inferential processes may outweigh even objective evidence about their own learning: for instance, participants who have been shown the empirical benefits of an interleaved schedule for their own learning nevertheless often persist in judging a blocked schedule as more effective (Yan et al., 2016). In the present study, this principle was realized in the difference between measures of objective performance and perceived learning. In both Studies 2 and 3, participants' objective performance under retrieval practice predicted their perceptions of that strategy's effectiveness for learning, suggesting some ability to accurately monitor their learning. But, it was ultimately those *perceptions* of learning effectiveness that predicted which strategies people chose; when controlling for perceived learning, there was no direct effect of objective performance on study strategy choice. That is, although objective performance with a particular study strategy no doubt informs perceptions of how effectiveness it is for learning, it is ultimately learners' *perceptions* that determine their behavior (especially since learners often do not have access to objective data on how well they learn under different strategies).

Third, our results also provide a new form of evidence supporting the principle of desirable difficulties. It has long been observed that ease of initial learning is not always an indicator of long-term retention (e.g., Benjamin et al., 1998; Soderstrom & Bjork, 2015) and that the conditions most effective for long-term retention often require *more* difficulty and effort during initial learning (Bjork, 1994; Bjork & Bjork, 1992; Schmidt & Bjork, 1992). But, past work on these desirable difficulties has typically characterized conditions as *difficult* based either on objective performance or fluency during learning (e.g., variable practice conditions result in lower performance initially but better performance later; Schmidt & Bjork, 1992) or on intuitions about which conditions are likely to be experienced as difficult (e.g., words in small fonts are assumed to be more disfluent for participants; for further discussion, see Mueller et al., 2014). Here, we directly queried learners' own perceptions of their experience with various study strategies. We found that two strategies that have been found to result in more enduring learning—interleaving categories (Birnbaum et al., 2013; Hall et al., 1994; Kang & Pashler, 2012; Kornell & Bjork, 2008a; Kornell et al., 2010; Rohrer & Taylor, 2007; Rohrer et al., 2014; Taylor & Rohrer, 2010; Wahlheim et al., 2011, 2012; Yan et al., 2016; Zulkiply et al., 2012) and retrieval practice (Carpenter et al., 2008; Karpicke & Roediger, 2008; Karpicke, 2012; Roediger & Karpicke, 2006; Rowland, 2014)—were indeed perceived by learners as comparatively effortful, a perception that may influence their study strategy choices.

One point of divergence between our series of studies and the broader self-regulated learning literature is that we found only partial support for one particular study strategy argued to be normatively more effective for learning: retrieval practice (for meta-analysis, see Rowland, 2014). In Study 3, participants randomly assigned to study a text using retrieval practice actually performed *worse* than those using restudy. (It must be noted, though, that when given the opportunity to choose between the two, participants who chose retrieval practice did have better retention—and that an ancillary measure in Study 2 that asked participants to recall a single interesting fact did show a marginal benefit for randomly-assigned retrieval practice.) It is not necessarily clear what accounts for this discrepancy, although low performance during the initial retrieval practice may be a factor; the present studies were primarily intended to assess *why* learners are making the choices they are, rather than *which* choices are normatively more effective. Indeed, these studies (and in particular, Study 3) point to the need to better understand what strategies work for whom, how, and under what conditions (for further discussion, see, e.g., Agarwal, Finley, Rose, & Roediger, 2017; Karpicke & Aue, 2015; van Gog & Sweller, 2015), and if normatively less effective strategies could actually be a better choice for some learners.

6.2. Making self-regulated learning decisions

The most critical outcome of the present studies is that they support our misinterpreted-effort hypothesis of why learners frequently make ineffective self-regulated learning decisions. Specifically, when learners experience a study strategy, such as simple restudy of a text, as fluent and not effortful, they perceive that strategy to be more effective for learning (i.e., the *ease of processing heuristic;* Kornell et al., 2011), and they are consequently more likely to choose it for future study. Conversely, when learners experience a study strategy, such as retrieval practice, as difficult and effortful (and thus do not have the feeling of fluency associated with a less effortful strategy), they perceive that strategy to be less effective for learning, and they are less apt to employ it. However, learners' perception that difficult processing results in poor learning is at odds with the finding that study strategies that are effective for long-term retention often require more effort or difficulty during initial learning (i.e., desirable difficulties). Indeed, we observed in Study 3 that choosing the effortful retrieval-practice strategy over restudy was associated with better retention 48 h later. Thus, although the decision not to employ effortful strategies may *feel* logical to the learner (because they have judged those strategies as ineffective for learning), this heuristic is likely to prove suboptimal for long-term learning. In sum, learners experience effective learning strategies like retrieval practice as effortful, and they misperceive that effort as indicative of poor learning, so they do not make full use of these effective strategies. We found this hypothesis supported in all three studies, both in measures of learners' immediate perceptions of each study strategy and in their retrospective comparisons.

We found less support for two alternative accounts of self-regulated learning decisions. Under an effort-avoidance hypothesis, learners seek to minimize mental effort and avoid study strategies they perceive as effortful even if they know those strategies are effective for learning (for further discussion, see Duckworth et al., 2011; Guo et al., 2016, Perez et al., 2014). This hypothesis predicts a direct relation between perceived effort and strategy choice, independent of perceived learning. We did find some support for this prediction—but only in participants' retrospective comparisons of the strategies (not in their immediate perceptions), and only in Studies 1 and 2. Together, the results suggest that a particularly powerful influence on learners' self-regulated learning decisions is the interpretation of mentally effortful strategies as yielding poor learning (and hence being undesirable), but that learners to some degree may also avoid strategies simply because they are taxing to implement. Indeed, these two influences may both be active in a single decision, as learners judge whether a seemingly effective strategy is "worth the effort."

Finally, a third possible hypothesis is the effort-disregard hypothesis. Metacognitive judgments can be made on multiple bases (Kelley & Jacoby, 1996; Koriat, 1997), so it is possible that choices of study strategies are *not* guided by the experience of mental effort and are made on some other basis (or bases). For example, if learners generally defer to already-established habitual strategies (Ariel & Dunlosky, 2013; Ariel et al., 2011), or they choose strategies based on a normative belief about how to study, then their experience of mental effort would be largely irrelevant to their own decisions. We found a hypothesis of complete disregard of mental effort was not supported insofar as perceived mental effort *was* closely linked to perceived learning, and, consequently, to learners' choices.

Of course, it is plausible that some of these other bases—such as habits or beliefs—also influence self-regulated learning decisions *in addition to* experienced effort. Indeed, a valuable future direction would be to additionally assess these constructs to determine how they additionally influence decisions, or even interact with perceived mental effort, in a paradigm such as this one. For example, general beliefs about study can be assessed separately from in-the-moment experience by asking participants to plan future study without engaging it (Yan, Soderstrom, Seneviratna, Bjork, & Bjork, 2017) or to predict the outcome of an experiment that they did not participate in (Koriat et al., 2004; Kornell et al., 2011). Nevertheless, we argue that the effects of mental effort that we observe here are likely to be relevant in many ecologically valid study situations in that nearly any time learners attempt actual study or learning activities, they will be confronted with the mental effort (or lack thereof) associated with those activities. That is, even if participants have a normative belief they can use to plan future study, when it comes time to carry out those study activities, they will experience mental effort or mental ease, which could then influence their perceptions of learning and choice of study strategies.

6.3. Integrating cognitive and motivation science

More broadly, the present work points to the importance of considering—and synthesizing—both cognitive and motivational perspectives in understanding self-regulated learning. For instance, work in cognitive psychology on the comparative effectiveness of particular study strategies has inspired efforts to teach learners those strategies (e.g., Ariel & Karpicke, 2018; Bielaczyc, Pirolli, & Brown, 1995; Broeren, Verkoeijen, Heijltjes, Smeets, & Arends, 2018; Chi, De Leeuw, Chiu, & Lavancher, 1994; Einstein, Mullet, & Harrison, 2012; McNamara, 2004; McNamara, Levinstein, & Boonthum, 2004; McNamara, O'Reilly, Best, & Ozuru, 2006). Although these interventions have had some success, one challenge is even students trained on these strategies do not always use them and often quickly revert to old strategies (Schunn, Moss, Huppert, & Schneider, 2009; Susser & McCabe, 2013) and do not always show benefits in learning (Broeren et al., 2018). Indeed, explicit instruction may not even change learners' *perceptions* of which strategies are effective for learning (Yan et al., 2016).

These prior outcomes raise the question of why normative beliefs and knowledge appear to exert only a limited influence on self-regulated learning decisions. Our results suggest that this pattern may be explained by an appeal to motivation science: Verbal knowledge about ideal behavior is often overridden by the visceral experience of effort (Hofmann et al., 2008; Kurzban et al., 2013; Mann et al., 2013; Rothman et al., 2015). In particular, explicit instruction on the effectiveness of particular strategies may be outweighed by the *experience* of mental effort, which many learners appear to interpret as an indicator of poor learning.

Conversely, for motivation science, our results highlight the importance of cognitive interpretations of learners' experience. Substantial work in motivation science has studied the factors modulating mental effort and its effect on learning (Barron & Hulleman, 2015; Eccles & Wigfield, 2002; Flake et al., 2015; Inzlicht et al., 2018; Westbrook & Braver, 2015) as well as how students and other learners might be induced to exert more effort (Autin & Croizet, 2012; Eskreis-Winkler et al., 2016; Job et al., 2010; Job, Friese, & Bernecker, 2015; Mrazek et al., 2018; Smith & Oyserman, 2015). However, our results suggest that it is ultimately the cognitive *interpretation* of that effort that is the most important influence on behavior (see also Koriat, Nussinson et al., 2014; Susser et al., 2016; Unkelbach, 2006); this effect reliably emerged every time we tested for it, whereas a direct effect of effort on study strategy choice appeared less robustly. That is, participants generally did not avoid study strategies simply because they were perceived as effortful, but rather because they interpreted their experience of effort as indicating that they were not learning from those strategies. Our misinterpreted-effort hypothesis implies that motivation and effort cannot be considered independent of the cognitive interpretation of those experiences; conversely, cognitive judgments of learning are likely powerfully influenced by visceral motivational experiences. An interdisciplinary approach may be most fruitful in understanding the interplay of cognitive and motivational factors in self-regulated learning.

6.4. Implications for practice

This interaction of cognitive and motivational factors also has implications for how to facilitate effective self-regulated learning. Specifically, targeting either cognition or motivation alone may be insufficient to change learners' behavior. For instance, if the most common reason that learners eschew effective strategies is not because they want to avoid the required mental effort but because they interpret that effort as indicative of poor learning, then attempts to enhance learning by targeting the level of student effort alone may be misguided. The problem may not be that learners are not trying hard enough, but rather that they do not understand the relationship between expended effort and learning.

Past work has also found that simply telling or showing learners which strategy is best for learning does not always change their judgments (e.g., Yan et al., 2016). Our misinterpreted-effort hypothesis suggests that instruction on cognitive principles alone may be insufficient to change judgments of learning or study behavior given the powerful visceral *experience* of mental effort and its influence on perceived learning.

Rather, our findings suggest the *link* between perceived effort and perceived learning may be a critical locus for intervention. In reality, many strategies effective for long-term retention, such as retrieval practice, require initial, "desirable" difficulty, which our current results suggest gives rise to the experience of mental effort. However, for learners, more effort is often interpreted as a sign of *less* learning (Heine et al., 2001; Plaut & Markus, 2005; Yeager & Dweck, 2012). This interpretation can be harmful to learning because the way learners perceive a strategy's effectiveness for learning is directly related to whether or not they choose to employ it. For this reason, learners—and those teaching learners—may need to take a different perspective of the experience of effort itself; that is, to develop a new *metamotivational belief* (Scholer, Miele, Murayama, & Fujita, 2018). Specifically, it may be beneficial to reframe feelings of mental effort as a sign of learning and growth rather than of failure, which might weaken or reverse the link between mental effort and (lack of) perceived learning. (For a review of similar approaches, see Yeager & Walton, 2011). A further benefit of such reframing is that it could support students' metacognitive understanding of effective learning in a way that generalizes across decisions, rather than having to individually learn which strategy is better in each case (e.g., learning whether to practice retrieval or restudy, then learning whether to use a blocked or interleaved schedule, etc.).

Of course, there may be some situations in which additional effort does *not* prove fruitful (e.g., the *labor-in-vain effect*; Nelson & Leonesio, 1988). Although effective study strategies often entail initial effort, the converse is not true: adding arbitrary initial difficulties does not necessarily enhance learning (Yue, Castel, & Bjork, 2013). Indeed, one of the reasons that learners may interpret difficulty or disfluency as a sign of poor learning is that, in many cases, difficulty *does* augur poor retention (Benjamin et al., 1998; Koriat, 2008). The challenge for students is to overcome this preconception when it comes to effective study strategies and, ultimately, to learn when difficulty predicts learning (e.g., when the difficulty during learning resembles the difficulties that will be present when the information must be used later; Blaxton, 1989; Schmidt & Bjork, 1992) and when it does not.

6.5. A causal relation?

One limitation of our present study is that our data were primarily correlational in nature. Although in each study we assigned learners to experience both of the strategies in question (either a blocked schedule vs. an interleaved schedule, or restudy vs. retrieval practice), we did not directly manipulate their perceptions. Thus, although our misinterpreted-effort hypothesis proposes that the experience of mental effort has a *causal* role in the perception of poor learning, and that perceptions of learning efficacy in turn influence study strategy decisions, the present data taken alone cannot speak to the direction of causation.

However, other studies provide evidence that each link in this indirect effect is indeed causal. First, numerous studies have experimentally manipulated ease of processing and found that it influences perceptions of learning. For example, Carpenter et al. (2013) varied the fluency of the speaker in a video-recorded lecture while holding the lecture content constant. Participants randomly assigned to the fluent lecture perceived greater learning, even though this variable was unrelated to actual recall (see also Fiechter, Fealing, Gerrard, & Kornell, 2018). Similarly, when the volume of words is experimentally manipulated, louder (and thus easier to perceive) words are perceived as better learned even though volume is unrelated to actual recall (Rhodes & Castel, 2009). Similar effects have also been observed with experimental manipulations of superfluous diagrams (Serra & Dunlosky, 2010), font size (Kornell et al., 2011; Rhodes & Castel, 2008), and the coherence of text (Rawson & Dunlosky, 2002). Indeed, what we regard as a contribution of the present study is to generalize these experimental manipulations of ease of processing to the type of mental effort and difficulty encountered as learners choose among authentic study strategies.

The second link in our mediation model was between perceived learning and study choices. Is this relationship also causal? Prior researchers have argued so, insofar as study decisions track with experimental manipulations that spuriously influence JOLs, rather than with objective learning (Metcalfe & Finn, 2008). We observed an analogous pattern here in that the final choice of study strategy was predicted by *perceived* learning rather than objective performance.

Lastly, a question of causality arises in interpreting the superior retention of learners who chose retrieval practice in Study 3: Does choosing retrieval practice help people learn better, or do learners choose retrieval practice because they are already high-performing students (who perhaps view all study strategies as less effortful)? Here, again, our present study is ambiguous taken alone, but a large number of experimental studies have provided robust evidence for a causal benefit of retrieval practice (for meta-analysis, Rowland, 2014), including in authentic educational settings (e.g., Agarwal, Bain, & Chamberlain, 2012; Carpenter, Pashler, & Cepeda, 2009; Einstein et al., 2012; McDaniel, Anderson, Derbish, & Morrisette, 2007; McDermott, Agarwal, D'Antonio, Roediger, & McDaniel, 2014; Nungester & Duchastel, 1982). Further, we also note that the indirect effect of mental effort on study strategy choice remained even after controlling for participants' actual learning performance; thus, it cannot be attributed to high-performing learners simply experiencing less effort. Nevertheless, causal conclusions could be further strengthened with future experimental work, such as with the reframing manipulations we discuss above.

6.6. Limitations and future directions

The present set of studies also have several other limitations. We observed our misinterpreted-effort hypothesis to explain decisions both between interleaving versus blocking categories and between restudying versus practicing retrieval; we examined these particular decisions because of the attention given to them in both the theoretical and applied literature. However, learners must make many other sorts of decisions in self-regulated learning, such as deciding whether to distribute practice over time or mass it into a single study episode (Baddeley & Longman, 1978; Benjamin & Bird, 2006; Pyc & Dunlosky, 2010; Son, 2004, 2010; Toppino, Cohen, Davis, & Moors, 2009) and deciding how long or how many times to study (Karpicke & Roediger, 2007; Karpicke, 2009; Kornell & Bjork, 2008; Postman, 1965; Pyc & Rawson, 2011; Rawson & Dunlosky, 2011; Son & Metcalfe, 2000; Tullis & Benjamin, 2011; Vaughn & Rawson, 2011). Further, learners do not face only binary decisions between two study strategies but rather a broader panoply of options about how, when, what, how long, and how often to study. Although it seems likely that our hypothesis could also be applied to other decisions in self-regulated learning decisions, its generality across all of these decisions remains to be tested.

We also tested perceptions and self-regulated learning decisions only in an experimental context (although we chose our materials to be educationally relevant). However, given the importance of self-regulated learning for educational success, it would be valuable to confirm that our misinterpreted-effort hypothesis is also supported in the classroom context. Educational practice might also benefit from potential interventions on the pattern and process of study strategy choices we observed, such as with the reframing procedures discussed above.

A particular advantage of testing the misinterpreted-effort hypothesis in authentic educational contexts is that it would assay learners' self-regulated learning choices when their learning outcomes are more consequential. In the present studies, participants had no particular incentive to master the study materials (beyond social acceptability or a desire not to "look stupid"; Hawkins, Brown, Steyvers, Wagenmakers, 2011, 2012). Could this account for the influence of perceived effort we observed? We argue it is unlikely for three reasons. First, there was evidence in all three studies that participants *did* learn the material: Classification performance improved from pre-test to post-test in Study 1, retrieval practice performance was more than 50% accurate in both Studies 2 and 3, and participants remembered more than half of the material even 48 h later in Study 3. Second, the free-response items in Study 3 (although subject to the limitations of self-reporting) also suggested that participants were generally attempting to learn the material; a majority of participants (68%) self-reported choosing a study strategy based on what would be most effective for learning the material, and only 19% reported choosing based on what would minimize the effort required in the experiment. Lastly, and most importantly, the pattern of results we observed was *not* what would be expected if participants simply chose the least effortful strategy because they did not really need to learn the material: The primary way that perceived mental effort influenced study strategies was that it predicted perceptions of learning, and it was perceived learning effectiveness that proximally predicted strategy choice; that is, participants were making a genuine attempt to learn the material by choosing the strategies that they regarded as more effective for learning. Evidence for a direct effect of perceived effort on choices (i.e., avoiding strategies simply because they require more effort than participants are willing to exert) was comparatively weak. Indeed, we think it is remarkable that we observed weaker support for the effort-avoidance hypothesis despite the fact that our paradigm presented what was in many ways the environment *most* likely to reveal such an effect (i.e., there was no real incentive to learn). Nevertheless, it will be important to validate the misperceived-effort hypothesis in more authentic educational contexts in the future.

Another critical future direction will be to determine whether the perception of mental effort as poor learning is ever absent or reversed. While we observed general support for our misinterpreted-effort hypothesis across participants, it is possible that individual learners have different conceptualizations of mental effort. For example, learners with a more incremental views of intelligence (i.e., people who believe that intelligence can be improved with effort) do not necessarily rely on an ease-of-processing heuristic in judging learning (Miele et al., 2011). In our mediation model, this would be realized as a weaker, or even reversed, link between mental effort and perceived (lack of) learning. We did not observe any link between naïve theories of intelligence and study-strategy perceptions or choices in our own data (nor have several other studies; Koriat et al., 2014; Yan, 2014), but individual differences in these heuristics merit further study.

This goal could also be aided by taking a developmental perspective (Koriat, Ackerman et al., 2014). We observed evidence that learners use an ease-of-processing heuristic in translating mental effort into perceived learning, but it is not clear how or when this heuristic develops. It would be interesting to examine if the pattern of perceptions and choice observed in our studies for adults holds for children across developmental stages. Is a high level of mental effort perceived to be indicative of poor learning effectiveness for all stages of development, or do children view effort differently? If we can understand when in development learners begin to misperceive effort as poor learning, interventions at that point could support a better and more positive view of mental effort as it relates to learning effectiveness.

7. Conclusion

Learners are frequently tasked with choosing how, when, what, and how long to study. These self-regulated learning decisions have consequences for long-term learning and retention because not all study strategies are equally effective. Unfortunately, learners often make suboptimal self-regulated learning decisions, employing normatively less-effective strategies rather than more effective ones. Why? In four studies spanning a variety of study strategies, to-be-learned material, and participant populations, we obtained support for a misinterpreted-effort hypothesis of self-regulated learning decisions: Learners perceive the mental effort that is associated with many normatively effectively strategies as a sign of poor learning, and hence they perceive these strategies as ineffective and do not choose them.

Appendix A. Interleaved vs. Blocked Perception Items

Perceptions of Mental Effort

1. How tiring was the last exercise?
   1 = *Not at all* to 6 = *A lot*
2. How mentally exhausting was the last exercise?
   1 = *Not at all* to 6 = *A lot*
3. How difficult was the last exercise?
   1 = *Not at all* to 6 = *A lot*
4. How boring was the last exercise?
   1 = *Not at all* to 6 = *A lot*

Perceptions of Learning Effectiveness

1. How likely are you to be able to distinguish between the types of birds?
   1 = *Not very likely* to 6 = *Extremely likely*
2. How good do you think your memory for the different types of birds will be?
   1 = *Not very good* to 6 = *Extremely good*
3. How effective was this exercise in helping you to distinguish between the types of birds?
   1 = *Not very effective* to 6 = *Extremely effective*
4. How well did you learn do distinguish the types of birds?
   1 = *Not very well* to 6 = *Extremely well*

Appendix B. Interleaved vs. Blocked Semantic Differential Items

Perceptions of Mental Effort

1. Which exercise required more mental effort?
   1 = *Grouped together* to 6 = *Not grouped together*
2. Which strategy would be harder for other people?
   1 = *Grouped together* to 6 = *Not grouped together*
3. Which strategy was more enjoyable? (reverse-scored)
   1 = *Grouped together* to 6 = *Not grouped together*
4. Which strategy was harder for you?
   1 = *Grouped together* to 6 = *Not grouped together*

Perceptions of Learning Effectiveness

1. Which birds do you think you'll remember better?
   1 = *Birds grouped together* to 6 = *Birds not grouped together*
2. Which do you think is a more effective learning strategy for you?
   1 = *Grouped together* to 6 = *Not grouped together*
3. Which do you think is a more effective learning strategy for the average person?
   1 = *Grouped together* to 6 = *Not grouped together*
4. Which strategy would you use to study in the future?
   1 = *Grouped together* to 6 = *Not grouped together*

Appendix C. Retrieval Practice vs. Restudy Perception Items

Perceptions of Mental Effort

1. How tiring was the last exercise?
   1 = *Not at all* to 6 = *A lot*
2. How mentally exhausting was the last exercise?
   1 = *Not at all* to 6 = *A lot*
3. How difficult was the last exercise?

1 = *Not at all* to 6 = *A lot*
4. How boring was the last exercise?
   1 = *Not at all* to 6 = *A lot*

Perceptions of Learning Effectiveness

1. How likely are you to remember the answers to the questions?
   1 = *Not very likely* to 6 = *Extremely likely*
2. How good do you think your memory for the answers will be?
   1 = *Not very good* to 6 = *Extremely good*
3. How effective was this exercise in helping you to learn the answers to the questions?
   1 = *Not very effective* to 6 = *Extremely effective*
4. How well did you learn the answers to the questions?
   1 = *Not very well* to 6 = *Extremely well*

Appendix D. Retrieval Practice vs. Restudy Semantic Differential Items

Perceptions of Mental Effort

1. Which exercise required more mental effort?
   1 = *Fill-in-the-blank* to 6 = *Review the fact*
2. Which strategy would be harder for other people?
   1 = *Fill-in-the-blank* to 6 = *Review the fact*
3. Which strategy was more enjoyable? (reverse-scored)
   1 = *Fill-in-the-blank* to 6 = *Review the fact*
4. Which strategy was harder for you?
   1 = *Fill-in-the-blank* to 6 = *Review the fact*

Perceptions of Learning Effectiveness

1. Which facts do you think you'll remember better?
   1 = *Fill-in-the-blank* to 6 = *Review the fact*
2. Which do you think is a more effective learning strategy for you?
   1 = *Fill-in-the-blank* to 6 = *Review the fact*
3. Which do you think is a more effective learning strategy for the average person?
   1 = *Fill-in-the-blank* to 6 = *Review the fact*
4. Which strategy would you use to study in the future?
   1 = *Fill-in-the-blank* to 6 = *Review the fact*

References

Agarwal, P. K., Bain, P. M., & Chamberlain, R. W. (2012). The value of applied research: Retrieval practice improves classroom learning and recommendations from a teacher, a principal, and a scientist. *Educational Psychology Review, 24*, 437–448. https://doi.org/10.1007/s10648-012-9210-2.

Agarwal, P. K., Finley, J. R., Rose, N. S., & Roediger, H. L. (2017). Benefits from retrieval practice are greater for students with lower working memory capacity. *Memory, 25*, 764–771. https://doi.org/10.1080/09658211.2016.1220579.

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

American Museum of Natural History (2015). Unearthing pterosaurs. Retrieved from https://www.readworks.org/article/Unearthing-Pterosaurs/bb70e3b3-391d-4fa8-aec5-298afcddcb55.

Ariel, R., Al-Harthy, I. S., Was, C. A., & Dunlosky, J. (2011). Habitual reading biases in the allocation of study time. *Psychonomic Bulletin & Review, 18*, 1015–1021. https://doi.org/10.3758/s13423-011-0128-3.

Ariel, R., & Dunlosky, J. (2013). When do learners shift from habitual-based to agenda-based processes when selecting items for study? *Memory & Cognition, 41*, 416–428. https://doi.org/10.3758/s13421-012-0267-4.

Ariel, R., & Karpicke, J. D. (2018). Improving self-regulated learning with a retrieval practice intervention. *Journal of Experimental Psychology: Applied, 24*, 43–56. https://doi.org/10.1037/xap0000133.

Autin, F., & Croizet, J.-C. (2012). Improving working memory efficiency by reframing metacognitive interpretation of task difficulty. *Journal of Experimental Psychology: General, 141*, 610–618. https://doi.org/10.1037/a0027478.

Baddeley, A. D., & Longman, D. J. A. (1978). The influence of length and frequency of training session on the rate of learning to type. *Ergonomics, 21*, 627–635.

Barron, K. E., & Hulleman, C. S. (2015). Expectancy-value-cost model of motivation. In J. S. Eccles, & K. Salmelo-Aro (Eds.). *International encyclopedia of social & behavioral sciences, 2nd edition: Motivational psychology* (pp. 503–509). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.26099-6.

Benjamin, A. S., & Bird, R. D. (2006). Metacognitive control of the spacing of study repetitions. *Journal of Memory and Language, 55*, 126–137. https://doi.org/10.1016/j.jml.2006.02.003.

Benjamin, A. S., Bjork, R. A., & Hirshman, E. (1998). Predicting the future and reconstructing the past: A Bayesian characterization of the utility of subjective fluency. *Acta Psychologica, 98*, 267–290. https://doi.org/10.1016/S0001-6918(97)00046-2.

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

Besken, M. (2016). Picture-perfect is not perfect for metamamory: Testing the perceptual fluency hypothesis with degraded images. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 42*, 1417–1433. https://doi.org/10.1037/xlm0000246.

Besken, M., & Mulligan, N. W. (2013). Easily perceived, easily remembered? Perceptual interference produces a double dissociation between metamemory and memory performance. *Memory & Cognition, 41*, 897–903. https://doi.org/10.3758/s13421-013-0307-8.

Bielaczyc, K., Pirolli, P. L., & Brown, A. L. (1995). Training in self-explanation and self-regulation strategies: Investigating the effects of knowledge acquisition activities on problem solving. *Cognition and Instruction, 13*, 221–252. https://doi.org/10.1207/s1532690xci1302_3.

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*, 392–402. https://doi.org/10.3758/s13421-012-0272-7.

Bjork, R. A. (1994). Memory and metamemory considerations in the training of human beings. In J. Metcalfe, & A. P. Shimamura (Eds.). *Metacognition: Knowing about knowing* (pp. 185–206). Boston, MA: The MIT Press.

Bjork, R. A., & Bjork, E. L. (1992). A new theory of disuse and an old theory of stimulus fluctuation. In A. Healy, S. Kosslyn, & R. Shiffrin (Vol. Eds.), *Learning processes to cognitive processes: Essays in honors of William Estes: Vol. 2*, (pp. 35–67). Hillsdale, NJ: Erlbaum.

Bjork, R. A., Dunlosky, J., & Kornell, N. (2013). Self-regulated learning: Beliefs, techniques, and illusions. *Annual Review of Psychology, 64*, 417–444. https://doi.org/10.1146/annurev-psych-113011-143823.

Blaxton, T. A. (1989). Investigating dissociations among memory measures: Support for a transfer-appropriate processing framework. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 657–668. https://doi.org/10.1037/0278-7393.15.4.657.

Broeren, M., Verkoeijen, P., Heijltjes, A., Smeets, G., & Arends, L. (2018). *Can instruction support the self-regulated use of retrieval practice in higher education? Poster presented at the 59th annual meeting of the psychonomic society, New Orleans, LA*.

Carpenter, S. K., Pashler, H., & Cepeda, N. J. (2009). Using tests to enhance 8th graders' retention of U.S. history facts. *Applied Cognitive Psychology, 23*, 760–771. https://doi.org/10.1002/acp.1507.

Carpenter, S. K., Pashler, H., Wixted, J. T., & Vul, E. (2008). The effects of tests on learning and forgetting. *Memory & Cognition, 36*, 438–448. https://doi.org/10.3758/MC.36.2.438.

Carpenter, S. K., Wilford, M. M., Kornell, N., & Mullaney, K. M. (2013). Appearances can be deceiving: Instructor fluency increases perceptions of learning without increasing actual learning. *Psychonomic Bulletin & Review, 20*, 1350–1356. https://doi.org/10.3758/s13423-013-0442-z.

Chevalier, N. (2018). Willing to think hard? The subjective value of cognitive effort in children. *Child Development, 89*, 1283–1295. https://doi.org/10.1111/cdev.12805.

Chi, M. T. H., Bassok, M., Lewis, M. W., Reimann, P., & Glaser, R. (1989). Self-explanations: How students study and use examples in learning to solve problems. *Cognitive Science, 13*, 145–182. https://doi.org/10.1016/0364-0213(89)90002-5.

Chi, M. T. H., De Leeuw, N., Chiu, M. H., & Lavancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive Science, 18*, 439–477. https://doi.org/10.1016/0364-0213(94)90016-7.

Duckworth, A. L., Kirby, T. A., Tsukayama, E., Berstein, H., & Ericsson, K. A. (2011). Deliberate practice spells scucess: Why grittier competitors triumph at the National Spelling Bee. *Social Psychological and Personality Science, 2*, 174–181. https://doi.org/10.1177/1948550610385872.

Duckworth, A. L., & Yeager, D. S. (2015). Measurement matters: Assessing personal qualities other than cognitive ability for educational purposes. *Educational Researcher, 44*, 237–251. https://doi.org/10.3102/0013189x15584327.

Dunlosky, J., & Hertzog, C. (1997). Older and younger adults use a functionally identical algorithm to select items for restudy during multitrial learning. *Journal of Gerontology B: Psychological Sciences, 52*, P178–P186.

Dunlosky, J., & Metcalfe, J. (2009). *Metacognition.* Thousand Oaks, CA: Sage Publications.

Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*, 4–58. https://doi.org/10.1177/1529100612453266.

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development.* New York: Psychology Press.

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*, 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153.

Einstein, G. O., Mullet, H. G., & Harrison, T. L. (2012). The testing effect: Illustrating a fundamental concept and changing study strategies. *Teaching of Psychology, 39*, 190–193. https://doi.org/10.1177/0098628312450432.

Eskreis-Winkler, L., Shulman, E. P., Young, V., Tsukayama, E., Brunwasser, S. M., & Duckworth, A. L. (2016). Using wise interventions to motivate deliberate practice. *Journal of Personality and Social Psychology, 111*, 728–744. https://doi.org/10.1037/pspp0000074.

Faul, F., Erdfelder, E., Lang, A., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175–191. https://doi.org/10.3758/bf03193146.

Fiechter, J. L., Fealing, C., Gerrard, R., & Kornell, N. (2018). Audiovisual quality impacts assessments of job candidates in video interviews: Evidence for an AV quality bias. *Cognitive Research: Principles and Implications, 3*, 47. https://doi.org/10.1186/s41235-018-0139-y.

Finley, J. R., & Benjamin, A. S. (2012). Adaptive and qualitative changes in encoding strategy with experience: Evidence from the test-expectancy paradigm. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 38*, 632–652. https://doi.org/10.1037/a0026215.

Finley, J. R., Tullis, J. G., & Benjamin, A. S. (2010). Metacognitive Control of Learning and Remembering. In M. S. Khine, & I. M. Saleh (Eds.). *New science of learning: cognition, computers and collaboration in education* (pp. 109–131). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4419-5716-0_6.

Finn, B., & Tauber, S. K. (2015). When confidence is not a signal of knowing: How students' experiences and beliefs about processing fluency can lead to miscalibrated confidence. *Educational Psychology Review, 27*, 567–586. https://doi.org/10.1007/s10648-015-9313-7.

Flake, J. K., Barron, K. E., Hulleman, C., McCoach, B. D., & Welsh, M. E. (2015). Measuring cost: The forgotten component of expectancy-value theory. *Contemporary Educational Psychology, 41*, 232–244. https://doi.org/10.1016/j.cedpsych.2015.03.002.

Fraundorf, S. H., & Benjamin, A. S. (2014). Knowing the crowd within: Metacognitive limits on combining multiple judgments. *Journal of Memory and Language, 71*, 17–38. https://doi.org/10.1016/j.jml.2013.10.002.

Fraundorf, S. H., & Benjamin, A. S. (2015). Conflict and metacognitive control: The mismatch-monitoring hypothesis of how others' knowledge states affect recall. *Memory, 24*, 1108–1122. https://doi.org/10.1080/09658211.2015.1069853.

Galla, B. M., Plummer, B. D., White, R. E., Meketon, D., D'Mello, S. K., & Duckworth, A. L. (2014). The Academic Diligence Task (ADT): Assessing individual differences in effort on tedious but important schoolwork. *Contemporary Educational Psychology, 39*, 314–325. https://doi.org/10.1016/j.cedpsych.2014.08.001.

Geller, J., Toftness, A. R., Armstrong, P. I., Carpenter, S. K., Manz, C. L., Coffman, C. R., & Lamm, M. H. (2018). Study strategies and beliefs about learning as a function of academic achievement and achievement goals. *Memory, 26*, 683–690. https://doi.org/10.1080/09658211.2017.1397175.

Guo, J., Nagengast, B., Marsh, H. W., Kelava, A., Gaspard, H., Brandt, H., ... Trautwein, U. (2016). Probing the unique contributions of self-concept, task values, and their interactions using multiple value facets and multiple academic outcomes. *AERA Open, 2*, 1–20. https://doi.org/10.1177/2332858415626884.

Hacker, D. J., Bol, L., Horgan, D. D., & Rakow, E. A. (2000). Test prediction and performance in a classroom context. *Journal of Educational Psychology, 92*, 160–170. https://doi.org/10.1037/0022-0663.92.1.160.

Hall, K. G., Domingues, D. A., & Cavazos, R. (1994). Contextual interference effects with skilled baseball players. *Perceptual and Motor Skills, 78*, 835–841. https://doi.org/10.2466/pms.1994.78.3.835.

Hartwig, M. K., & Dunlosky, J. (2012). Study strategies of college students: Are self-testing and scheduling related to achievement? *Psychonomic Bulletin & Review, 19*, 126–134. https://doi.org/10.3758/s13423-011-0181-y.

Hawkins, G. E., Brown, S. D., Steyvers, M., & Wagenmakers, E.-J. (2012). An optimal adjustment procedure to minimize experiment time in decisions with multiple alternatives. *Psychonomic Bulletin & Review, 19*, 339–348. https://doi.org/10.3758/s13423-012-0216-z.

Hawkins, G. E., Brown, S. D., Steyvers, M., & Wagenmakers, E.-J. (2011). *Leave the experiment as quickly as possible, without looking stupid: An optimal adjustment procedure to explain context effects in multi-alternative choice. Poster presented at the 52nd annual meeting of the psychonomic society, Seattle, WA*.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach.* New York, NY: Guilford Press.

Heine, S. J., Lehman, D. R., Ide, E., Leung, C., Kitayama, S., Takata, T., & Matsumoto, H. (2001). Divergent consequences of success and failure in Japan and North America: An investigation of self-improving motivations and malleable selves. *Journal of Personality and Social Psychology, 81*, 599–615. https://doi.org/10.1037/0022-3514.81.4.599.

Hinze, S. R., & Wiley, J. (2011). Testing the limits of testing effects using completion tests. *Memory, 19*, 290–304. https://doi.org/10.1080/09658211.2011.560121.

Hofmann, W., Friese, M., & Wiers, R. W. (2008). Impulsive versus reflective influences on health behavior: A theoretical framework and empirical review. *Health Psychology Review, 2*, 111–137. https://doi.org/10.1080/17437190802617668.

Inzlicht, M., Shenhav, A., & Olivola, C. Y. (2018). The effort paradox: Effort is both costly and valued. *Trends in Cognitive Sciences, 22*, 337–349. https://doi.org/10.1016/j.tics.2018.01.007.

Jacoby, L. L., & Whitehouse, K. (1989). An illusion of memory: False recognition influenced by unconscious perception. *Journal of Experimetnal Psychology: General, 118*, 126–135.

Job, V., Dweck, C. S., & Walton, G. M. (2010). Ego depletion-is it all in your head? Implicit theories about willpower affect self-regulation. *Psychological Science, 21*, 1686–1693. https://doi.org/10.1177/0956797610384745.

Job, V., Friese, M., & Bernecker, K. (2015). Effects of practicing self-control on academic performance. *Motivation Science, 1*, 219–232. https://doi.org/10.1037/mot0000024.

Kang, S. H. K., & Pashler, H. (2012). Learning painting styles: Spacing is advantageous when it promotes discriminative contrast. *Applied Cognitive Psychology, 26*, 97–103. https://doi.org/10.1002/acp.1801.

Karpicke, J. D. (2009). Metacognitive control and strategy selection: Deciding to practice retrieval during learning. *Journal of Experimental Psychology: General, 138*, 469–486. https://doi.org/10.1037/a0017341.

Karpicke, J. D. (2012). Retrieval-based learning: Active retrieval promotes meaningful learning. *Current Directions in Psychological Science, 21*, 157–163. https://doi.org/10.1177/0963721412443552.

Karpicke, J. D., & Aue, W. R. (2015). The testing effect is alive and well with complex materials. *Educational Psychology Review, 27*, 317–326. https://doi.org/10.1007/s10648-015-9309-3.

Karpicke, J. D., & Roediger, H. L. (2007). Repeated retrieval during learning is the key to long-term retention. *Journal of Memory and Language, 57*, 151–162. https://doi.org/10.1016/j.jml.2006.09.004.

Karpicke, J. D., & Roediger, H. L. (2008). The critical importance of retrieval for learning. *Science, 319*, 966–968. https://doi.org/10.1126/science.1152408.

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

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

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

Koriat, A., Ackerman, R., Adiv, S., Lockl, K., & Schneider, W. (2014). The effects of goal-driven and data-driven regulation of metacognitive monitoring during learning: A developmental perspective. *Journal of Experimental Psychology: General, 143*, 386–403. https://doi.org/10.1037/a0031768.

Koriat, A., & Bjork, R. A. (2005). Illusions of competence in monitoring one's knowledge during study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 187–194. https://doi.org/10.1037/0278-7393.31.2.187.

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

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*, 1624–1637. https://doi.org/10.1037/xlm0000009.

Koriat, A., Sheffer, L., & Ma'ayan, H. (2002). Comparing objective and subjective learning curves: Judgments of learning exhibit increased underconfidence with practice. *Journal of Experimental Psychology. General, 131*, 147–162. https://doi.org/10.1037/0096-3445.131.2.147.

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. (2008b). 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., & Bjork, R. A. (2008a). Optimising self-regulated study: The benefits—and costs—of dropping flashcards. *Memory, 16*, 125–136. https://doi.org/10.1080/09658210701763899.

Kornell, N., & Bjork, R. A. (2009). A stability bias in human memory: Overestimating remembering and underestimating learning. *Journal of Experimental Psychology. General, 138*, 449–468. https://doi.org/10.1037/a0017350.

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. https://doi.org/10.1037/a0017807.

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

Kornell, N., & Son, L. K. (2009). Learners' choices and beliefs about self-testing. *Memory, 17*, 493–501. https://doi.org/10.1080/09658210902832915.

Kurtz, K. H., & Hovland, C. I. (1956). Concept learning with differing sequences of instances. *Journal of Experimental Psychology, 51*(4), 239–243. https://doi.org/10.1037/h0040295.

Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences, 36*, 661–679. https://doi.org/10.1017/S0140525X12003196.

Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes, 65*, 272–292.

Logan, J. M., Castel, A. D., Haber, S., & Viehman, E. J. (2012). Metacognition and the spacing effect: The role of repetition, feedback, and instruction on judgments of learning for massed and spaced rehearsal. *Metacognition and Learning, 7*, 175–195. https://doi.org/10.1007/s11409-012-9090-3.

Mann, T., De Ridder, D., & Fujita, K. (2013). Self-regulation of health behavior: Social psychological approaches to goal setting and goal striving. *Health Psychology, 32*, 487–498. https://doi.org/10.1037/a0028533.

McCabe, J. (2011). Metacognitive awareness of learning strategies in undergraduates. *Memory & Cognition, 39*, 462–476. https://doi.org/10.3758/s13421-010-0035-2.

McDaniel, M. A., Anderson, J. L., Derbish, M. H., & Morrisette, N. (2007). Testing the testing effect in the classroom. *European Journal of Cognitive Psychology, 19*, 494–513. https://doi.org/10.1080/09541440701326154.

McDermott, K. B., Agarwal, P. K., D'Antonio, L., Roediger, H. L., & McDaniel, M. A. (2014). Both multiple-choice and short-answer quizzes enhance later exam performance in middle and high school classes. *Journal of Experimental Psychology: Applied, 20*, 3–21. https://doi.org/10.1037/xap0000004.

McNamara, D. S. (2004). SERT: Self-explanation reading training. *Discourse Processes, 38*, 1–30.

McNamara, D. S., Levinstein, I. B., & Boonthum, C. (2004). iSTART: Interactive strategy training for active reading and thinking. *Behavior Research Methods, Instruments, & Computers, 36*, 222–233. https://doi.org/10.3758/BF03195567.

McNamara, D. S., O'Reilly, T. P., Best, R. M., & Ozuru, Y. (2006). Improving adolescent students' reading comprehension with iSTART. *Journal of Educational Computing Research, 34*, 147–171. https://doi.org/10.2190/1RU5-HDTJ-A5C8-JVWE.

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

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

Mrazek, A. J., Ihm, E. D., Molden, D. C., Mrazek, M. D., Zedelius, C. M., & Schooer, J. W. (2018). Expanding minds: Growth mindsets of self-regulation and the influences on effort and perseverance. *Journal of Experimental Social Psychology, 79*, 164–180. https://doi.org/10.1016/j.jesp.2018.07.003.

Mueller, M. L., Dunlosky, J., Tauber, S. K., & Rhodes, M. G. (2014). The font-size effect on judgments of learning: Does it exemplify fluency effects or reflect people's

beliefs about memory? *Journal of Memory and Language, 70*, 1–12. https://doi.org/10.1016/j.jml.2013.09.007.
Nelson, T. O., & Dunlosky, J. (1991). When people's judgments of learning (JOLs) are extremely accurate at predicting subsequent recall: The "delayed-JOL effect". *Psychological Science, 2*, 267–270. https://doi.org/10.1111/j.1467-9280.1991.tb00147.x.
Nelson, T. O., & Leonesio, R. J. (1988). Allocation of self-paced study time and the "labor-in-vain effect". *Journal of Experimental Psychology: Learning, Memory, and Cognition, 14*, 676–686.
Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. In G. Bower (Ed.). *The psychology of learning and motivation* (pp. 125–173). New York: Academic Press.
Nungester, R. J., & Duchastel, P. C. (1982). Testing versus review: Effects on retention. *Journal of Education Psychology, 74*, 18–22. https://doi.org/10.1037/0022-0663.74.1.18.
Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review, 14*, 187–193. https://doi.org/10.3758/bf03194050.
Peer, E., & Gamliel, E. (2011). Too reliable to be true? Response bias as a potential source of inflation in paper-and-pencil questionnaire reliability. *Practical Assessment, Research & Evaluation, 16*.
Perez, T., Cromley, J. G., & Kaplan, A. (2014). The role of identity development, values, and costs in college STEM retention. *Journal of Educational Psychology, 106*, 315–329. https://doi.org/10.1037/a0034027.
Plaut, V. C., & Markus, H. R. (2005). The "inside" story: A cultural-historical analysis of being smart and motivated, American style. In C. Dweck, & A. Elliott (Eds.). *Handbook of competence and motivation* (pp. 457–488). New York, NY: Guilford Press.
Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin, 135*, 322–338. https://doi.org/10.1037/a0014996.
Postman, L. (1965). Retention as a function of degree of overlearning. *Science, 135*, 666–667. https://doi.org/10.1126/science.135.3504.666.
Pyc, M. A., & Dunlosky, J. (2010). Toward an understanding of students' allocation of study time: Why do they decide to mass or space their practice? *Memory & Cognition, 38*, 431–440. https://doi.org/10.3758/MC.38.4.431.
Pyc, M. A., & Rawson, K. A. (2011). Costs and benefits of dropout schedules of test-restudy practice: Implications for student learning. *Applied Cognitive Psychology, 25*, 87–95. https://doi.org/10.1002/acp.1646.
Rawson, K. A., & Dunlosky, J. (2002). Are performance predictions for text based on ease of processing? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 28*, 69–80. https://doi.org/10.1037//0278-7393.28.1.6.
Rawson, K. A., & Dunlosky, J. (2011). Optimizing schedules of retrieval practice for durable and efficient learning; How much is enough? *Journal of Experimental Psychology: General, 140*, 283–302. https://doi.org/10.1037/a0023956.
ReadWorks, Inc. (2013a). The American chestnut. Retrieved from https://www.readworks.org/article/The-American-Chestnut/94810bf5-e192-4c2a-a1fe-35ba9af41cd3.
ReadWorks, Inc. (2013b). The Great Barrier Reef. Retrieved from https://www.readworks.org/article/The-Great-Barrier-Reef/0e48bb0e-8759-403f-ad6c-0493101ea727.
ReadWorks, Inc. (2013c). Water on the Earth. Retrieved from https://www.readworks.org/article/Water-on-the-Earth/c9935af1-b930-438e-b5ca-c1a4beb3f554.
Rhodes, M. G., & Castel, A. D. (2008). Memory predictions are influenced by perceptual information: Evidence for metacognitive illusions. *Journal of Experimental Psychology. General, 137*, 615–625. https://doi.org/10.1037/a0013684.
Rhodes, M. G., & Castel, A. D. (2009). Metacognitive illusions for auditory information: Effects on monitoring and control. *Psychonomic Bulletin & Review, 16*, 550–554. https://doi.org/10.3758/PBR.16.3.550.
Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests imporves long-term retention. *Psychological Science, 17*, 249–255. https://doi.org/10.1111/j.1467-9280.2006.01693.x.
Roediger, H. L., & Pyc, M. A. (2012). Inexpensive techniques to improve education: Applying cognitive psychology to enhance educational practice. *Journal of Applied Research in Memory and Cognition, 1*, 242–248 doi: j.jarmac.2012.09.002.
Rohrer, D., Dedrick, R. F., & Burgess, K. (2014). The benefit of interleaved mathematics practice is not limited to superficially similar kinds of problems. *Psychonomic Bulletin & Review, 21*, 1323–1330. https://doi.org/10.3758/s13423-014-0588-3.
Rohrer, D., & Pashler, H. (2010). Recent research on human learning challenges conventional instructional strategies. *Educational Researcher, 39*(5), 406–412. https://doi.org/10.3102/0013189x10374770.
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.
Rothman, A. J., Gollwitzer, P. M., Grant, A. M., Neal, D. T., Sheeran, P., & Wood, W. (2015). Hale and hearty policies: How psychological science can create and maintain healthy habits. *Perspectives on Psychological Science, 10*, 701–705. https://doi.org/10.1177/1745691615598515.
Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin, 140*, 1–32. https://doi.org/10.1037/a0037559.
Schmidt, R. A., & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3*, 207–217. https://doi.org/10.1111/j.1467-9280.1992.tb00029.x.
Scholer, A. A., Miele, D. B., Murayama, K., & Fujita, K. (2018). New directions in self-regulation: The role of metamotivational beliefs. *Current Directions in Psychological Science, 27*, 437–442. https://doi.org/10.1177/0963721418790549.
Schunn, C. D., Moss, J., Huppert, T., & Schneider, W. (2009). *Real-time NIRS feedback may help improve self-explanation learning strategy use. Paper presented at the society for neuroscience, Chicago, IL*.
Schwartz, B. L., Benjamin, A. S., & Bjork, R. A. (1997). The inferential and experiental basis of metamemory. *Current Directions in Psychological Science, 6*, 132–137. https://doi.org/10.1111/1467-8721.ep10772899.
Serra, M. J., & Dunlosky, J. (2010). Metacomprehension judgements reflect the belief that diagrams improve learning from text. *Memory, 18*, 698–711. https://doi.org/10.1080/09658211.2010.506441.
Shenhav, A., Musslick, S., Lieder, F., Kool, W., Griffiths, T. L., Cohen, J. D., & Botvinick, M. M. (2017). Toward a rational and mechanistic account of mental effort. *Annual Review of Neuroscience, 40*, 99–124.
Simons, D. J., & Chabris, C. F. (2011). What people believe about how memory works: A representative sample of the U.S. population. *PLoS ONE, 6*, e22757. https://doi.org/10.1371/journal.pone.0022757.
Simons, D. J., & Chabris, C. F. (2012). Common (mis)beliefs about memory: A replication and comparison of telephone and Mechanical Turk survey methods. *PLoS ONE, 7*, e51876. https://doi.org/10.1371/journal.pone.0051876.
Smith, G. C., & Oyserman, D. (2015). Just not worth my time? Experienced difficulty and time investment. *Social Cognition, 33*, 1–18 doi: 101521soco201533285.
Soderstrom, N. C., & Bjork, R. A. (2015). Learning versus performance: An integrative view. *Perspectives on Psychological Science, 10*, 176–199. https://doi.org/10.1177/1745691615569000.
Son, L. K. (2004). Spacing one's study: Evidence for a metacognitive control strategy. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 601–604. https://doi.org/10.1037/0278-7393.30.3.601.
Son, L. K. (2010). Metacognitive control and the spacing effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 36*, 255–262. https://doi.org/10.1037/a0017892.
Son, L. K., & Metcalfe, J. (2000). Metacognitive and control strategies in study-time allocation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*, 204–221. https://doi.org/10.1037/0278-7393.26.1.204.
Sungkhasettee, V., Friedman, M. C., & Castel, A. D. (2011). Memory and metamemory for inverted words: Illusions of competency and desirable difficulties. *Psychonomic Bulletin & Review, 18*, 973–978. https://doi.org/10.3758/s13423-011-0114-9.
Susser, J. A., Jin, A., & Mulligan, N. W. (2016). Identity priming consistently affects perceptual fluency but only affects metamemory when primes are obvious. *Journal*

of Experimental Psychology: Learning, Memory, and Cognition, 42, 657–662. https://doi.org/10.1037/xlm0000189.
Susser, J. A., & McCabe, J. (2013). From the lab to the dorm room: Metacognitive awareness and use of spaced study. *Instructional Science, 41*, 345–363. https://doi.org/10.1007/s11251-012-9231-8.
Taylor, K., & Rohrer, D. (2010). The effects of interleaved practice. *Applied Cognitive Psychology, 24*, 837–848. https://doi.org/10.1002/acp.1598.
Toppino, T. C., Cohen, M. S., Davis, M. L., & Moors, A. C. (2009). Metacognitive control over the distribution of practice: When is spacing preferred? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 1352–1358. https://doi.org/10.1037/a0016371.
Tullis, J. G., & Benjamin, A. S. (2011). On the effectiveness of self-paced learning. *Journal of Memory and Language, 64*(2), 109–118. https://doi.org/10.1016/j.jml.2010.11.002.
Tullis, J. G., Benjamin, A. S., & Ross, B. H. (2011). *A metacognitive illusion in category learning. Poster presented at the 52th annual meeting of the psychonomic society, Seattle, WA*.
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.
Undorf, M., & Erdfelder, E. (2011). Judgments of learning reflect encoding fluency: Conclusive effort in favor of the ease-of-processing hypothesis. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37*, 1264–1269. https://doi.org/10.1037/a0023719.
University of Pittsburgh (2016). Pitt football through the years. Retrieved from http://exhibit.library.pitt.edu/football-through-years/.
University of Pittsburgh (2017). Preserving our architectural treasures. Retrieved from http://www.treasures.pitt.edu/history/.
Unkelbach, C. (2006). The learned interpretation of cognitive fluency. *Psychological Science, 17*, 339–345.
van Gog, T., & Sweller, J. (2015). Not new, but nearly forgotten: The testing effect decreases or even disappears as the complexity of learning materials increases. *Educational Psychology Review, 27*, 247–264. https://doi.org/10.1007/s10648-015-9310-x.
Vaughn, K. E., & Rawson, K. A. (2011). Diagnosing criterion-level effects on memory: What aspects of memory are enhanced by repeated retrieval? *Psychological Science, 22*, 1127–1131. https://doi.org/10.1177/0956797611417724.
Wahlheim, C. N., Dunlosky, J., & Jacoby, L. L. (2011). Spacing enhances the learning of natural concepts: An investigation of mechanisms, metacognition, and aging. *Memory & Cognition, 39*(5), 750–763. https://doi.org/10.3758/s13421-010-0063-y.
Wahlheim, C. N., Finn, B., & Jacoby, L. L. (2012). Metacognitive judgments of repetition and variability effects in natural concept learning: Evidence for variability neglect. *Memory & Cognition, 40*(5), 703–716. https://doi.org/10.3758/s13421-011-0180-2.
Weekly Reader Corporation (2009). Short circuits. Retrieved from https://www.readworks.org/article/Short-Circuits/ff97e001-bccc-41d5-b53a-58015dfe8995#!.
Westbrook, A., & Braver, T. (2015). Cognitive effort: A neuroeconomic approach. *Cognitive, Affective, & Behavioral Neuroscience, 15*, 395–415. https://doi.org/10.3758/s13415-015-0334-y.
Westbrook, A., Kester, D., & Braver, T. S. (2013). What is the subjective cost of cognitive effort? Load, trait, and aging effects revealed by economic preference. *PLoS ONE, 8*, e68210. https://doi.org/10.1371/journal.pone.0068210.
Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.). *The educational psychology series. Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ: Lawrence Erlbaum Associates. https://doi.org/10.1016/j.chb.2007.09.009.
Yan, V. X. (2014). *Learning concepts and categories from examples: How learners' beliefs match and mismatch the empirical evidence.* [unpublished doctoral dissertation]Los Angeles: University of California.
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. https://doi.org/10.1037/xge0000177.
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. https://doi.org/10.1037/xap0000139.
Yan, V. X., Thai, K.-P., & Bjork, R. A. (2014). Habits and beliefs that guide self-regulated learning: Do they vary with mindset? *Journal of Applied Research in Memory and Cognition, 3*, 140–152 10.106/j.jarmac.2014.04.003.
Yeager, D. S., & Dweck, C. S. (2012). Mindsets that promote resilience: When students believe that personal characteristics can be developed. *Educational Psychologist, 47*, 302–314. https://doi.org/10.1080/00461520.2012.722805.
Yeager, D. S., Henderson, M. D., Paunesku, D., Walton, G. M., D'Mello, S., Spitzer, B. J., & Duckworth, A. L. (2014). Boring but important: A self-transcendent purpose for learning fosters academic self-regulation. *Journal of Personality and Social Psychology, 107*, 559–580. https://doi.org/10.1037/a0037637.supp.
Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*, 267–301. https://doi.org/10.3102/0034654311405999.
Yue, C. L., Castel, A. D., & Bjork, R. A. (2013). When disfluency is—and is not—a desirable difficulty: The influence of typeface clarity on metacognitive judgments and memory. *Memory & Cognition, 41*, 229–241. https://doi.org/10.3758/s13421-012-0255-8.
Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*, 82–91. https://doi.org/10.1006/ceps.1999.1016.
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. https://doi.org/10.1016/j.learninstruc.2011.11.002.