



Self-regulated use of retrieval practice: associations with individual differences in non-cognitive and cognitive factors

Frida Bertilsson¹ · Tova Stenlund¹ · Anna Sundström¹ · Bert Jonsson²

Received: 12 November 2023 / Revised: 16 April 2024 / Accepted: 24 April 2024
© The Author(s) 2024

Abstract

Retrieval practice is a learning strategy that has repeatedly been found to have positive effects on memory and learning. However, studies indicate that students rarely use retrieval practice on a voluntary basis. The objective of the present study was to examine students' self-regulated use of retrieval practice, and to determine whether sex and individual differences in cognitive and non-cognitive aspects are related to optional use of practice testing. A classroom study was conducted with 146 upper-secondary school students taking courses in mathematics and Swedish. An ABAB design was used to compare students' optional and non-optional use of retrieval practice (i.e., repeated online quizzing). Students performed cognitive tasks to assess working memory capacity and fluid intelligence and completed self-reports of non-cognitive factors related to school achievement, such as grit, need for cognition (NFC), conscientiousness and openness. Quiz use was then compared using paired- and independent-samples t-tests, and hierarchical linear regression analyses explored relations to individual differences. The results showed that students completed significantly fewer quizzes in the optional sections than in the non-optional sections, and that females completed significantly more optional quizzes than males in Swedish, but not in mathematics. Further, the results showed that conscientiousness predicted optional quiz use in mathematics, whereas sex, NFC, conscientiousness, and openness predicted quiz use in Swedish. To conclude, although the findings show a relatively low optional/self-regulated use of practice testing, in line with earlier research, they suggest that sex and non-cognitive factors, such as personality characteristics, can predict optional use of practice testing.

Keywords Retrieval practice · Self-regulated learning · Individual differences · Cognitive factors · Non-cognitive factors · Sex-related differences

Introduction

Retrieval practice is a learning strategy that is based on the well-established finding that repeated retrieval of to-be-remembered information from long-term memory, usually through repeated testing, is highly beneficial for learning and retention when compared to other learning strategies (Carpenter et al., 2008; Dunlosky et al., 2013; Karpicke & Blunt, 2011; Rawson et al., 2013; Roediger & Karpicke, 2006; Rowland, 2014). In recent years, the research field of retrieval practice (also called practice testing, test-enhanced learning, etc.) has shifted from more experimental settings to classroom-based research using educationally relevant material. A number of reviews and meta-analyses have found that the benefits of retrieval practice can be transferred to classroom activities (Agarwal et al., 2021; Lamotte et al., 2021; Moreira et al., 2019; Schwieren et al., 2017; Sotola & Credé, 2021; Yang et al., 2021). In a review by Agarwal et al. (2021), the aim was to suggest practical recommendations for when retrieval practice is beneficial for learning. The findings indicated that all retrieval practice conditions in the included studies resulted in positive effects, which suggests that as long as students participate in some form of retrieval practice activity, they will experience a learning benefit. However, in all studies included in the review by Agarwal et al. (2021), participants were required to take part in the retrieval practice activities, which raises the question of how transferable the results are to classroom settings where students' actions are less controlled. With the effectiveness of retrieval practice in mind, it is crucial that students self-regulate their learning and actually use retrieval practice on a voluntary basis. The overall focus in the present study is to explore how students use the effective learning strategy of retrieval practice when it is optional, and whether individual differences might be related to self-regulated use of retrieval practice.

Self-regulated learning (SRL) concerns the processes students use to direct their own study behaviors in order to achieve their goals (Pintrich, 2000; Pintrich & Zusho, 2007; Zimmerman, 2001, 2013). More specifically, Zimmerman (2002) suggests that self-regulation processes include three phases. The first phase includes setting learning goals and planning for how to achieve them. The second phase involves action to achieve the goals set, which includes the use of learning strategies. In the final phase, learning is evaluated in relation to the goals set in the first phase. The ability to regulate one's own learning has been found to be an important skill for achieving academic success (Broadbent & Poon, 2015), and has also been suggested to contribute to or explain gender differences in school achievement (Weis et al., 2013). A common finding regarding gender differences in school achievement is that female students tend to have higher course grades regardless of school subject, whereas males tend to score higher on achievement tests (see, for example, Voyer & Voyer, 2014). Studies examining gender differences in SRL indicate that females have a greater tendency than males to employ strategies of SRL (Panadero et al., 2017; Zimmerman & Martinez-Pons, 1990), but there are mixed results regarding gender and the different components of SRL (Martinez-Lopez et al., 2017; Stanikzai, 2019). Note that the cited sources have used the term gender, however, the present study recorded the participants biological sex, which is why different terms are used.

Self-regulated use of retrieval practice

Although there are many studies about self-regulated learning in general, self-regulated use of retrieval practice has not received much attention. A few recent studies suggest that

although there is vast support for the effectiveness of retrieval practice, students tend to use the learning strategy to a low extent. Instead, students in both undergraduate (Blasiman et al., 2017) and secondary school (Dirkx et al., 2019) contexts tend to favor study strategies less beneficial for learning, such as restudy, and this appears to be a common flaw in students' self-regulated learning. For example, Tullis and Maddox (2020) investigated self-reported use of retrieval practice among middle- and high-school students, and found that while both age groups used retrieval practice, it was used to a lesser extent than re-reading. In a study investigating the optional use of retrieval practice, undergraduate students were provided optional online reviews, either in test format (a quiz) or read format (students were provided question and answer) (Corral et al., 2020). The results indicated that students underutilize retrieval practice, as only 12 percent of the total reviews were completed. Moreover, only 55 percent of the participants completed at least one review (either test or read format) during the semester (Corral et al., 2020). Trumbo et al. (2016) also investigated the optional use of quizzes but compared the effect of optional vs. required quizzes on test scores and final course grade. Participants in the optional quiz group spent less time on each quiz and completed fewer quizzes in comparison with participants in the required quiz group. This difference was reflected in final grade performance such that participants in the required quiz group achieved better final grade performance than participants in the optional quiz group. The authors suggest that students may need extrinsic motivation in order to utilize quizzes as a study method (Trumbo et al., 2016). To sum up, a vast amount of research shows that retrieval practice is an effective learning strategy. However, getting students to employ this strategy in their own studying seems to be a challenge, as findings indicate that students use retrieval practice to a low extent when the use is optional. One interesting question regarding the use of retrieval practice is whether the employment of this strategy might be related to individual differences.

As research indicates that there are sex-related differences in self-regulated learning (Marrs & Sigler, 2012), it appears worthwhile to examine sex differences with respect to use of retrieval practice. Studies on the use of retrieval practice have often neglected this question. However, a study by Gagnon and Cormier (2019) examined Canadian college students' use of self-testing and distributed practice. Results indicated that females (63%) reported that they used self-testing as a study strategy to a greater extent than males (57%).

Self-regulated learning, achievement, and individual differences

The framework of SRL involves cognitive, motivational and emotional aspects of learning (Panadero, 2017). Recently, it has been highlighted that individual differences need to be incorporated into SRL theory (Azevedo, 2020) and researchers argue that such an integration, including both cognitive and non-cognitive factors, might provide a better understanding of students' academic achievement (Wolters & Hussain, 2015). Cognitive factors, such as fluid intelligence, crystallized intelligence, working memory and previous academic performance, as well as non-cognitive factors, such as personality, procrastination and emotional intelligence, have been found to be associated with SRL (see, e.g., Pérez-González et al., 2022; Richardson et al., 2012). Further, it has been suggested that SRL mediates the relationship between personality characteristics and academic achievement (Pintrich, 2000). Thus, it would be of interest to investigate possible relationships between non-cognitive factors such as personality and self-regulated use of retrieval practice.

The five-factor model (FFM) of personality has been suggested as a comprehensive indicator of non-cognitive factors (Borghans et al., 2008). The FFM consists of five broad

personality factors: openness to experience, conscientiousness, extraversion, agreeableness and neuroticism (Costa & McCrae, 1992; Goldberg, 1993). Research on the relationship between FFM and academic achievement has shown conscientiousness to be the most consistent predictor of performance (Hakimi et al., 2011; Poropat, 2009; Sorić et al., 2017). Conscientiousness is defined as willingness to comply with conventional rules, and comprises facets of striving for achievement and self-discipline (Borghans et al., 2008). As such, the construct of conscientiousness is closely linked to persistence and motivation, and thus related to concepts such as *grit* and possibly also *need for cognition* (NFC). *Grit* is defined as “an individual’s perseverance and passion for long-term goals” (Duckworth et al., 2007, p. 1087), and might explain differences in individuals’ persistence or determination to succeed in learning situations. NFC is defined as the need to engage in and enjoy thinking and cognitively demanding tasks (Cacioppo & Petty, 1982), and might explain differences in individuals’ motivation when engaging in learning activities. The constructs of conscientiousness, *grit* and NFC have been found to be positively associated with academic performance (see, for example, Colling et al., 2022; Poropat, 2009; Wolters & Hussain, 2015). On a theoretical level, it also seems that NFC is highly related to the openness factor in the FFM, which has been defined as “the degree to which a person needs intellectual stimulation, change, and variety” (Borghans et al., 2008, p. 983). As with conscientiousness, openness has also been found to be positively associated with learning and academic achievement, while the other factors of the FFM seem to have weaker associations with performance (Bidjerano & Dai, 2007; Poropat, 2009).

Individual differences and retrieval practice

As regards retrieval practice, rather few studies have examined individual differences, and those studies conducted have mainly focused on relationships to the *effect of* retrieval practice (i.e., the so-called testing effect) rather than the *use of* this learning strategy. Studies have investigated the relationship between retrieval practice and cognitive functioning known to be associated with academic achievement. However, the findings are mixed. Regarding working memory capacity (WMC), some studies show no association between WMC and effect of retrieval practice (Bertilsson et al., 2017, 2021; Brewer & Unsworth, 2012; Wiklund-Hörnqvist et al., 2014), whereas others indicate that individuals with lower levels of WMC benefit more from retrieval practice (Agarwal et al., 2017). Regarding studies on episodic memory and fluid intelligence, Jonsson et al. (2021) found no association between general cognitive ability and the testing effect, whereas Brewer and Unsworth (2012) found that retrieval practice is more beneficial for individuals with poorer episodic memory and lower fluid intelligence. Studies on the testing effect have also investigated the relationships with non-cognitive factors such as NFC and *grit*. Previous research has not found any evidence that benefit from retrieval practice is influenced by individual differences in *grit* or NFC (Bertilsson et al., 2017, 2021; Stenlund et al., 2017; Wiklund-Hörnqvist et al., 2022). One possible explanation for these results could be that because the constructs of *grit* and NFC both include aspects of motivation, they may not be as relevant for explaining differences in performance when the researchers or teachers, and not the participants themselves, initiate the use of retrieval practice, as is most often the case in studies investigating retrieval practice. However, it might be the case that both cognitive and non-cognitive aspects are related to *self-regulated use of* retrieval practice. To our knowledge, there are only two studies (Fellman et al., 2020a, b) where individual differences have been investigated in relation to how students use retrieval practice. In these studies, medical students’ optional use of online quizzes was examined in

relation to individual differences, specifically in relation to reasoning or fluid intelligence (as measured by Raven's Advanced Progressive Matrices), NFC and grit (Fellman et al., 2020b) and working memory (Fellman et al., 2020a). The results showed that quiz use was related to reasoning and verbal working memory, but not to any of the non-cognitive variables. However, these two studies have limitations in the respect that homogenous samples of high-performing students were used, and one lacked demographical data, such as sex and age (Fellman et al., 2020a), restricting the possibility to generalize the results. In order to extend our knowledge on self-regulated use of retrieval practice, broader samples from other school settings need to be studied. In line with this, the present study examines optional use of retrieval practice in an upper-secondary school setting using educationally relevant material and taking cognitive and non-cognitive individual differences into account.

Purpose of the study

The overall aim of the study was to investigate self-regulated use of retrieval practice when integrated in courses in mathematics and Swedish, and to examine whether individual differences are related to this use. More specifically, the following research questions were addressed:

1. Does the use of retrieval practice differ depending on optional (i.e., outside the classroom) and non-optional (i.e., inside the classroom) retrieval practice conditions and are there sex-related differences?
2. Is the optional use of retrieval practice related to cognitive and non-cognitive aspects?

Method

Participants

One hundred forty-six upper-secondary school students ($M_{age} = 16.23$ years, $SD = 0.49$, 27% female) at science and technical programs in the northern part of Sweden were included in the study. The students were enrolled in classes in mathematics (96 students), and Swedish (26 students), or both mathematics and Swedish (24 students), resulting in a total of 120 mathematics students (31 female) and 50 Swedish students (16 female). The proportion of males and females in the study was relatively representative of the proportions of males/females in science and technical programs in upper-secondary schools on a national level (27% females in science programs, 19% females in technical programs; Statistics Sweden, 2023). The classes were led by teachers who were selected by the school to take part in the research project. Data from the mathematics classes were collected in two cohorts during the school years of 2018–2019 and 2019–2020. Data from four students were excluded from the study because they did not take part in the intervention or due to missing data in the measures of cognitive and non-cognitive aspects.

Materials and measures

Retrieval practice material

The material used for the retrieval practice intervention consisted of eight quizzes containing a total of 20 items in mathematics and Swedish, respectively. In mathematics, the content of each quiz corresponded to the content of a chapter in the course book (e.g., algebra & equations, geometry, etc.), with a focus on mathematical terms. In Swedish, the content of each quiz corresponded to different subtopics of the course (e.g., grammar, literature, etc.). The short-answer questions consisted of a definition of a mathematical term or concept related to Swedish and the name of the term or concept was typed in as an answer. The quizzes were made available to the students on their online school platform. Pre- and post-tests were conducted before and after each chapter/topic, but results from these were not included in the current study.

Non-cognitive measures

Short Grit Scale (GRIT-S) Grit was measured using a Swedish version of the Short Grit Scale (GRIT-S; Duckworth & Quinn, 2009). GRIT-S is a self-report instrument consisting of eight items, and is an adaption of the original Grit Scale (Duckworth et al., 2007). Half of the items assess consistency of interest (e.g., “I often set a goal but later choose to pursue a different one”) and the other half assess perseverance of effort (e.g., “I have achieved a goal that took years of work”). Four of the items are phrased negatively and are reversely scored. The items are responded to on a five-point scale ranging from “not like me at all” (1) to “very much like me” (5). The total score is generated by adding up the points awarded to each item and then dividing by the number of items. Higher scores indicate higher levels of grit. To ensure a high-quality and accurate translation of the questionnaire, back-translation by a professional translator was utilized. GRIT-S has demonstrated acceptable validity and reliability, with an internal consistency ranging between $\alpha=0.73$ and $\alpha=0.84$ (Duckworth & Quinn, 2009). In the current study $\alpha=0.61$.

The Mental Effort Tolerance Questionnaire (METQ) NFC was measured using a Swedish adaption of the original NFC scale (Cacioppo & Petty, 1982), the Mental Effort Tolerance Questionnaire (METQ; Dornic et al., 1991; Stenlund & Jonsson, 2017). METQ is a self-report scale consisting of 30 items, which represent both positive (e.g., “I really enjoy a task that involves coming up with new solutions”) and negative (e.g., “I only think as hard as I have to”) attitudes toward engaging in and enjoying thinking. Responses are given on a five-point scale, ranging from “strongly disagree” (1) to “strongly agree” (5). The items that capture negative attitudes are reversely scored. Responses are summed to a total score and high scores on the METQ indicate a high need for cognition. METQ has demonstrated good psychometric properties (Dornic et al., 1991; Stenlund & Jonsson, 2017), including an internal consistency of 0.80 in the present study.

Mini-IPIP The Mini International Personality Item Pool (Mini-IPIP; Donnellan et al., 2006) is a brief measure of the five-factor model of personality, which includes the dimensions of: neuroticism, extraversion, openness to experience, conscientiousness, and agreeableness. The instrument assesses each of these five personality factors using only four

items per factor. In the present study, a Swedish translation of the Mini-IPIP was used and only the subscales of conscientiousness and openness to experience were included. The subscale of conscientiousness comprised four items about the tendency to be organized and self-disciplined, written as short statements: “Get chores done right away,” “Often forget to put things back in their proper place” (scored reversely), “Like order,” and “Make a mess of things” (reversely scored). The subscale of openness to experience comprised four items about imagination and the tendency to enjoy abstract thinking: “Have a vivid imagination,” “Have difficulty understanding abstract ideas” (scored reversely), “Am not interested in abstract ideas” (scored reversely), and “Do not have a good imagination” (scored reversely). Responses are given on a five-point scale ranging from very inaccurate (1) to very accurate (5). The Mini-IPIP is well-validated in general population samples (Donnellan et al., 2006), and has demonstrated acceptable reliability in college samples (Baladasaro et al., 2013) as well as in the current study ($\alpha=0.58$ and $\alpha=0.75$ respectively for the included scales).

Cognitive measures

Raven’s Advanced Progressive Matrices (RAPM) RAPM were used to capture participants’ general cognitive ability (Raven, 1990). RAPM is a non-verbal test, where items consist of a 3×3 matrix of geometric patterns with the bottom-right area missing. The participants are asked to complete the pattern by selecting one option among eight alternatives. The difficulty of the items advances progressively during the test. The original test includes 48 items, of which the first 12 are often used as practice items. However, in this study, a short version including 18 items and six practice items was used. The participants had 25 min to complete the tasks. The total number of correctly scored items was used as a dependent variable, where higher scores indicated higher general cognitive ability. Previous studies have indicated that RAPM has good construct validity (Schweizer et al., 2007) and demonstrates good psychometric properties in college samples (Arthur et al., 1999). Internal consistency in the current study was 0.77.

The Operation Span Task (Ospan) To measure working memory capacity (WMC), a standardized complex working memory task—an automated version of the Operation Span Task (Ospan; Unsworth et al., 2005)—was used. The Ospan is computer-administered and comprises two tasks: a letter span and a concurrent math task. The participant is asked to solve simple arithmetic tasks (processing demand), while simultaneously maintaining the presented letters in long-term memory (storage demand). These tasks are combined in sets that range from three to seven blocks, and each set size is performed for three trials, resulting in a total of 75 math problems and 75 letters. After each trial, the participant is shown a matrix of 12 letters, and is asked to recall the 3–7 letters in the order they were shown. Participants must have at least 85% correct on the math tasks, in order to ensure that they do not ignore these tasks in favor of rehearsing the letters. Ospan has demonstrated good test–retest reliability, $r=0.83$, acceptable internal consistency, $\alpha=0.78$, and good construct validity (Unsworth et al., 2005).

Procedure

The design of the retrieval practice implementation was conducted in collaboration with four teachers who were employed at the upper-secondary school. The teachers contributed

with valuable insight regarding the challenges associated with teaching in upper-secondary school and gave advice on how to design an intervention that could be a realistic part of everyday education for their students. The practical implementation was conducted by the teachers. At the start of each cohort, the students were introduced to the project and given an inspiring lecture about retrieval practice and how it should be used to improve retention. This was done to ensure that the students had knowledge about the benefits of retrieval practice, as previous studies have shown that students often tend to choose other, less-effective learning strategies.

At the start of each section of the courses, the quiz was made available to the students on their learning platform. The quiz was available for the duration of the corresponding section of the course, with an average time open of 11 days for the quizzes in the Swedish course, and 22 and 24 days, respectively, for the two cohorts in the mathematics course. For all sections, the quiz was available for the students to use of their own volition. In addition, for half of the sections, the quiz was also used once a week in the classroom (i.e., 3–4 times per chapter in mathematics and two times per section in Swedish). Sections/chapters for which quizzing was completely optional will be referred to as “optional” and sections/chapters for which quizzing was conducted in the classroom in addition to being available for voluntary use will be referred to as “non-optional”. It is important to note that while quizzing was an in-class activity during non-optional sections, students were sometimes absent from class for various reasons, which means that some students may not have completed any or only very few quizzes even in non-optional sections. Non-optional and optional quizzes were alternated between the sections, resulting in an ABAB design (A = non-optional/in the classroom and B = optional/outside the classroom). For practical reasons, the second cohort in mathematics was designed in the reverse order and started with an optional section (i.e., a BABA design). The purpose of the ABAB design was to investigate whether there were differences in quiz use between the two conditions.

The measures of non-cognitive aspects and cognitive abilities were collected in a group setting at the school, during two sessions of about 90 min each. The cognitive tasks were conducted on the students’ school-provided computers via an online platform. Non-cognitive aspects were measured through self-report questionnaires, using paper and pen. The students received two movie tickets for their participation in the study. The study was approved by the Regional Ethical Review Board, Sweden (2017/517–31), and written informed consent was obtained in accordance with the Declaration of Helsinki.

Statistical analyses

First, the data was subjected to descriptive analysis. In this analysis, two outliers (defined as z -score > 3.29 or < -3.29) were identified in the measures of NFC and RPM, respectively. Due to the sensitivity of the analysis used, these values were excluded. Secondly, paired-samples t -tests were used to examine differences in average quiz use between optional and non-optional conditions, and independent-samples t -tests examined differences between males and females as well as between cohorts in mathematics. Cohen’s d was used as a measure of effect size, for which values of 0.02 are considered a small, 0.05 a medium, and 0.08 a large effect (Cohen, 1992). Bivariate correlations between optional quiz use and cognitive and non-cognitive variables were calculated in order to investigate the strengths and directions of the relationships, as well as to detect potential problems with multicollinearity as indicated by elevated VIF values (> 10) (Tabachnick & Fidell, 2019). No multicollinearity was found. In order to determine which of the included independent variables

were most predictive of optional quizzing, as well as the potential influence of sex, the correlation analyses were followed up with hierarchical linear regression analyses. The dependent variable for the analyses was the total number of quizzes completed during the optional sections. Independent variables were the measures of fluid intelligence, WMC, NFC, grit, conscientiousness, and openness. Sex was also added as a predictor in the first step, fluid intelligence and WMC were added in the second step to control for cognitive abilities, and the non-cognitive variables NFC, grit, conscientiousness, and openness were added in the third step.

Results

Use of retrieval practice

First, a series of analyses were performed to investigate how the optional and non-optional quizzes had been used. Independent-samples t-tests showed that there was a clear difference in frequency of quiz use between optional and non-optional quizzing in both school subjects (see Figs. 1 and 2). The average number of quizzes completed was significantly higher in non-optional sections than in completely optional sections, in both mathematics, $t=20.02$, $p<0.001$, $d=1.8$, and Swedish, $t=7.11$, $p<0.001$, $d=1.0$ (see Table 1 for descriptive statistics). The variation in number of quizzes completed by each participant is also quite different between the two conditions.

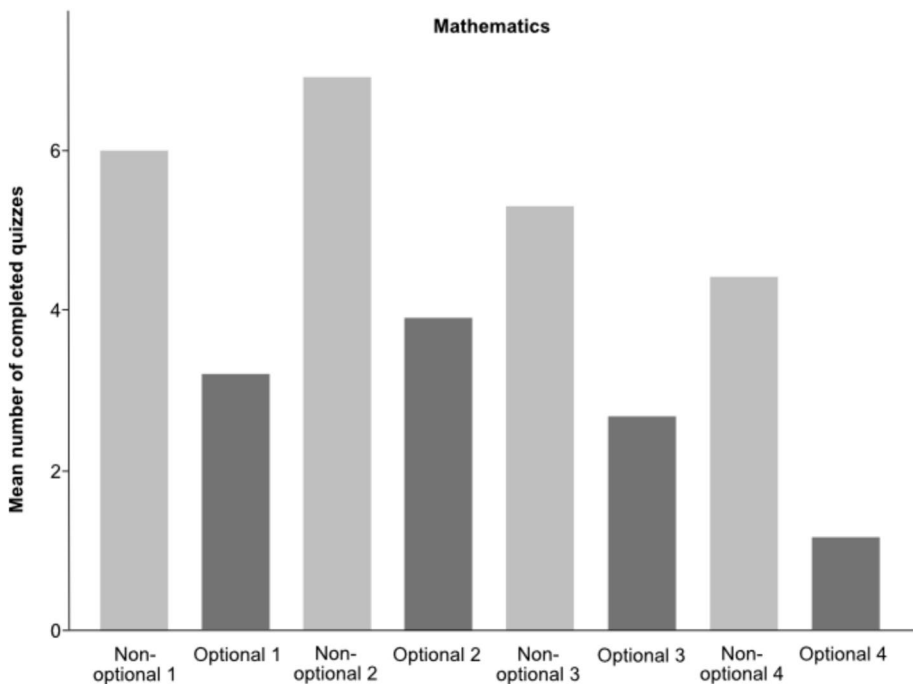


Fig. 1 Average number of completed quizzes for each chapter in mathematics

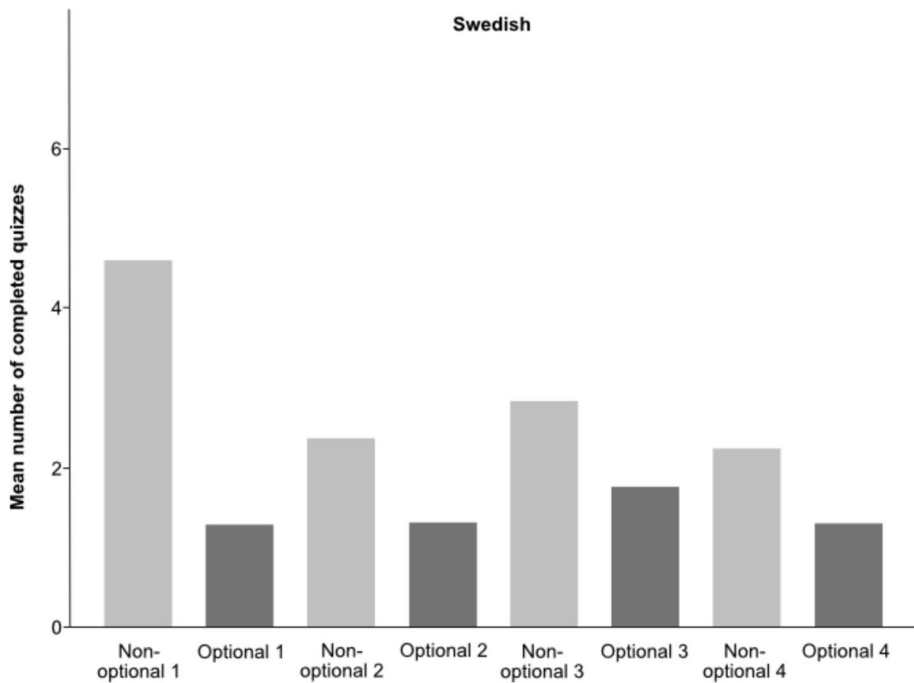


Fig. 2 Average number of completed quizzes for each section in Swedish

Table 1 Descriptive statistics for optional and non-optional quiz use in mathematics and Swedish, as well as means for male and female students separately

Chapter/section	Mathematics				Swedish			
	Mean	SD	M_{male}	M_{female}	Mean	SD	M_{male}	M_{female}
Optional								
1	3.25	3.18	3.08	3.74	1.28	2.00	0.91	2.06
2	3.92	3.83	3.67	4.65	1.32	2.47	0.59	2.88
3	2.67	3.25	2.85	2.13	1.76	2.33	1.18	3.00
4	1.15	1.76	0.90	1.87	1.30	2.13	0.79	2.38
Total	10.96	8.57	10.46	12.39	5.66	6.64	4.42	7.71
Non-optional								
1	6.07	3.77	5.53	7.61	4.60	3.20	3.76	6.38
2	6.95	2.36	6.70	7.68	2.36	2.68	1.74	3.69
3	5.30	3.48	5.47	4.81	2.84	3.32	1.53	5.63
4	4.41	2.62	4.20	5.00	2.24	2.38	1.71	3.38
Total	22.69	8.33	21.85	25.10	12.04	9.98	8.74	19.06

Note. For mathematics, data from both cohorts have been combined, i.e., Optional 1 includes data from the first optional quiz in both cohorts

As Table 1 illustrates, the quiz use is lower for the optional sections than for the non-optional sections. Moreover, the optional sections had a much larger share of students who had completed no or only very few quizzes. There was also a significant difference in number of completed quizzes between male and female students (see Table 1), with females completing more quizzes in Swedish in both the optional ($t=-3.30$, $p<0.01$, $d=0.9$) and non-optional sections ($t=-3.87$, $p<0.001$, $d=1.2$). In mathematics, there was a significant difference in completed quizzes between males and females for the non-optional quizzes ($t=-1.89$, $p=0.03$, $d=0.4$), but not for the optional quizzes. There was no significant difference between the two cohorts in mathematics in terms of completed quizzes ($t=1.40$, $p=0.08$, $d=0.3$).

Individual differences associated with the optional use of retrieval practice

In order to investigate whether students' individual differences in cognitive and non-cognitive factors were related to their self-regulated use of quizzes, correlational analyses and hierarchical regression analyses were performed. As the purpose was to examine self-regulated quiz use, the dependent variable used in these analyses was the total number of completed optional quizzes. The correlations suggested that grit and conscientiousness have weak to moderate positive associations with optional quiz use in mathematics and Swedish (see Table 2). In Swedish, there was also a positive relationship between NFC and optional quiz use.

In the next step, hierarchical linear regression analyses were performed to examine the predictive ability of the independent variables. The results showed that, controlling for sex in the first step, conscientiousness was the only statistically significant predictor for optional quiz use in mathematics, explaining 11 percent of the variance (see Table 3). For optional quiz use in Swedish, a somewhat different pattern emerged. Sex was a significant positive predictor in all three steps, and in the third step, NFC and conscientiousness were statistically significant predictors, with positive associations with quiz use, whereas openness was significantly and negatively related to quiz use (see Table 3). NFC and sex had the most predictive power, followed by conscientiousness and openness. Together, the significant non-cognitive predictors explained 28 percent of the variance in optional quiz use in Swedish.

Discussion

Although a large amount of research has shown convincing evidence that retrieval practice is a highly effective learning strategy, it seems to be an underutilized study strategy. The present study focused on the group of students that actually use retrieval practice on their own, and whether they differ from other students, which has thus far been a neglected research area. More precisely, the aim was to investigate the self-regulated use of retrieval practice (i.e., repeated online quizzing) in an intervention for upper-secondary school students, and to determine to what extent the use of optional quizzes outside the classroom is related to differences between males and females, as well as to individual differences in cognitive and non-cognitive factors related to academic success.

First, we compared optional (outside the classroom) with non-optional (in the classroom) use of retrieval practice in two school subjects, mathematics and Swedish, and examined whether sex-related differences could be found. As expected, and in line with

Table 2 Pearson correlations between the dependent variables optional quiz use in mathematics and Swedish, and the independent variables fluid intelligence, WMC, NFC, grit, conscientiousness, and openness

Variables	M	SD	N	1	2	3	4	5	6	7	8
1. Mathematics optional	10.96	8.57	120	-	0.38	-0.11	-0.02	0.06	0.21*	0.40**	-0.09
2. Swedish optional	5.66	6.46	50		-	0.27	-0.01	0.37**	0.37**	0.53**	-0.16
3. Fluid intelligence	15.48	4.00	137			-	0.17	0.23**	-0.05	-0.02	0.20*
4. WMC	36.52	15.48	139				-	0.07	-0.09	-0.04	0.05
5. NFC	100.56	12.08	140					-	0.16	0.18*	0.35**
6. Grit	3.17	0.56	143						-	0.50**	0.06
7. Conscientiousness	12.81	2.94	141							-	-0.07
8. Openness	13.52	3.23	141								-

Note. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Table 3 Hierarchical regression analyses using sex, fluid Intelligence, WMC, NFC, grit, conscientiousness, and openness as predictors of optional quiz use in mathematics and Swedish

Predictors	Mathematics				Swedish			
	B	SE	β	Adj. r^2	B	SE	β	Adj. r^2
Step 1				-0.01				0.21**
Sex	1.15	1.97	0.06		6.93	2.08	0.48**	
Step 2				-0.01				0.23**
Sex	1.36	1.99	0.07		7.02	2.06	0.49**	
Fluid intelligence	-0.23	0.22	-0.10		0.56	0.33	0.24	
WMC	-0.02	0.06	-0.04		-0.03	0.06	-0.08	
Step 3				0.12**				0.51***
Sex	-0.74	1.93	-0.04		4.98	1.77	0.35**	
Fluid intelligence	-0.20	0.21	-0.09		0.38	0.30	0.17	
WMC	-0.02	0.05	-0.03		0.02	0.05	0.04	
NFC	0.05	0.07	0.07		0.18	0.07	0.35*	
Grit	-0.01	1.66	-0.00		0.51	1.89	0.04	
Conscientiousness	1.20	0.36	0.39***		0.71	0.35	0.34*	
Openness	-0.18	0.27	-0.07		-0.54	0.25	-0.26*	

Note. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

previous research (Corral et al., 2020; Trumbo et al., 2016), the number of quizzes completed differed between the optional condition and the non-optional condition, with significantly more quizzes being completed during non-optional sections for both mathematics and Swedish. What makes this finding especially interesting is the fact that all students had been informed about the benefits of retrieval practice at the start of the intervention. Knowing that retrieval practice can be an unintuitive learning technique, starting off the intervention with an inspirational lecture about the method was an attempt to inspire the students to stick with the technique until they could see its effectiveness for themselves. While it is impossible to know what the use of quizzes would have been without this introduction, the generally low usage suggests that informing the participants of the observed benefits of retrieval practice did not have a strong inspirational effect (but see also Ariel & Karpicke, 2017). Thus, it seems important that teachers encourage students to use retrieval practice activities inside the classroom, and maybe more importantly, support students' development of the skill to self-regulate their learning so that they choose effective learning strategies (Zimmerman, 2002, 2013).

From a self-regulated learning (SRL) perspective, it is probable that there are sex differences with respect to the use of retrieval practice (see, for example, Panadero et al., 2017). While this has been a neglected research topic, one previous study does suggest that females use self-testing as a study strategy to a greater extent than males (Gagnon & Cormier, 2019). In line with this, the present study found that females completed more quizzes than males overall. With respect to optional quizzing, there was a significant difference between the sexes in Swedish, but not in math. One explanation for this finding might be that more quizzes were completed in mathematics overall (in both the optional and non-optional conditions) in comparison with Swedish (see Figs. 1 and 2). It is possible that the use of quizzes is a more natural and traditional part of teaching, and way of learning, in the mathematics classroom than in the Swedish classroom. For example, in a study examining

Canadian teachers' assessment practices in mathematics ($N = 1096$), quizzes were used to a large extent, both as formative assessments in order to get a sense of students' understanding (89%), and as summative assessments (79%; Suurtamm et al., 2010).

Second, besides examining non-optional and optional use of retrieval practice and sex-related differences in this use, we also examined whether there are individual differences with respect to cognitive and non-cognitive factors related to the optional use of retrieval practice. Previous studies have indicated that cognitive abilities such as fluid intelligence and working memory capacity (Brewer & Unsworth, 2012; Minear et al., 2018), and non-cognitive factors such as grit, conscientiousness and openness are related to self-regulated learning and academic achievement (Pérez-González et al., 2022; Richardson et al., 2012). To our knowledge, however, there are only two studies (Fellman et al., 2020a, b) that have investigated individual differences in relation to optional use of retrieval practice. The findings from these studies indicated that higher cognitive ability and verbal working memory were related to greater optional use of retrieval practice, whereas no differences were observed in the non-cognitive variables grit and NFC. In contrast, the findings from the present study showed that NFC (Swedish), as well as conscientiousness (mathematics, Swedish), were positively related to optional quiz use, which is in line with earlier research in the SRL area (Pérez-González et al., 2022; Richardson et al., 2012), although no relationships to cognitive abilities were found. From a theoretical perspective, it has been suggested that both cognitive ability and personality characteristics, such as conscientiousness, might serve important functions in self-regulated learning with respect to employing study strategies (Pérez-González et al., 2022), and the different student populations (undergraduate vs. upper-secondary school) and different subject areas examined in the present study and by Fellman et al. (2020b) as well as Fellman et al. (2020a) might explain the different results observed in the different studies. In addition, as regards the results related to intelligence, Fellman et al. (2020b) found that a one-unit increase in RAPM score increased the odds ratio for belonging to the high-retrieval practice group by 1.18, but additional analyses on high-retrieval practice users revealed that intelligence did not predict quiz use per session. Thus, the practical relevance of the observed association should be considered.

Another interesting finding from the present study was that in mathematics, conscientiousness was the only statistically significant predictor of optional quiz use, whereas in Swedish, several statistically significant predictors emerged. Sex, conscientiousness, and NFC were positively related to optional quiz use, while openness to experience was negatively related to quiz use. Thus, in mathematics, students who are more conscientious completed more quizzes, while quiz use in Swedish was related to being conscientious and having a higher NFC. The negative relationship to openness can be understood in the sense that those who score low on openness, that is, those who are conventional and traditional in their behavior and prefer familiar routines to a greater extent (Aitken Harris, 2004), use optional quizzing more frequently as it can become a form of routine. The differences between the two school subjects, regarding non-cognitive factors relating to the optional use of retrieval practice, might be explained by differences in how these subjects are traditionally taught and in familiarity with quizzes, as mentioned above.

Overall, the finding that self-regulated use of retrieval practice is related to non-cognitive factors that have been found to be strong predictors of academic success implies that students who already are more likely to succeed in their academic endeavors are those using retrieval practice outside the classroom. Importantly, this was found despite giving all students information about retrieval practice that could have inspired them to use the strategy in a different way than their usual manner of engaging in educational tasks. Therefore, the results emphasize the important role of the educator in ensuring that the

students actively engage with effective learning strategies by including retrieval practice in the classroom activities.

Limitations

A strength of the present study is the inclusion of both cognitive and non-cognitive factors to examine individual differences in self-regulated use of retrieval practice, which contributes to the research area. Nevertheless, the study has some limitations that need to be mentioned. Considering that the implementation of retrieval practice was conducted with a focus on ecological validity, rather than experimental rigor, some limitations are to be expected. Unforeseen circumstances necessitated a slight adjustment to the original design, resulting in the change from an ABAB design in cohort 1 to a BABA design in cohort 2. However, since the purpose of the study was to investigate voluntary quiz use, this change should not impact the results in any significant way. This is supported by a *t*-test confirming that the two cohorts did not differ in terms of completed quizzes. In addition, using a quasi-experimental design inherently incorporates some general limitations. In this case, the inclusion of two subjects and four separate teachers may have included differences between the groups or impacted how the quizzing was practically implemented in the different classrooms. The effects of teacher- and group characteristics on students' study behaviors are important to consider and should be further explored in future research. Another limitation is the large difference in sample size in the mathematics and Swedish courses which makes it difficult to make comparisons between the subjects. Finally, although the proportion of females in the sample was comparable to the proportion of female students in science and technical programs in general, the relatively low number of female students implies that the sex-related differences found should be interpreted with caution. Hence, it is important to examine individual differences in self-regulated use of effective learning strategies, such as retrieval practice, in larger studies using more representative samples.

Conclusions and further studies

Not surprisingly, the results in the present study show that use of retrieval practice outside the classroom is low when compared to the use of retrieval practice inside the classroom. The results also suggest that sex-related and individual differences in non-cognitive factors can predict self-regulated use of retrieval practice. Female students seem to use retrieval practice as an optional strategy to a greater extent than males. Moreover, students who are more persistent (more conscientious) and who have a higher motivation to engage in learning activities (NFC) as well as a preference for routines (low on openness) use retrieval practice as an optional study strategy to a greater extent. To make sure that all students, regardless of their sex and non-cognitive aspects, benefit from effective learning strategies such as practice testing, it is evident that the strategy needs to be implemented in the classroom under non-optional conditions. On the other hand, we also need to learn more about students' motivations and what might inspire them to use practice testing as a study strategy.

Acknowledgements We thank Tony Quillbard for conducting the computer programming required for this study.

Funding Open access funding provided by Umeå University. Funding was received from Umeå School of Education to Bert Jonsson and Frida Bertilsson (project memory and learning) and the Swedish Research Council (grant number 721-2014-2099) to Bert Jonsson.

Data Availability The data that support the findings of this study are available from the corresponding author, FB, upon reasonable request.

Declarations

Conflict of interest The authors declare no conflict of interest.

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

References

- Agarwal, P. K., Finley, J. R., Rose, N. S., & Roediger, H. L., III. (2017). Benefits from retrieval practice are greater for students with lower working memory capacity. *Memory*, 25(6), 764–771. <https://doi.org/10.1080/09658211.2016.1220579>
- Agarwal, P. K., Nunes, L. D., & Blunt, J. R. (2021). Retrieval practice consistently benefits student learning: A systematic review of applied research in schools and classrooms. *Educational Psychology Review*, 33(4), 1409–1453. <https://doi.org/10.1007/s10648-021-09595-9>
- Aitken Harris, J. (2004). Measured intelligence, achievement, openness to experience, and creativity. *Personality and Individual Differences*, 36(4), 913–929. [https://doi.org/10.1016/S0191-8869\(03\)00161-2](https://doi.org/10.1016/S0191-8869(03)00161-2)
- Ariel, R., & Karpicke, J. D. (2017). Improving self-regulated learning with a retrieval practice intervention. *Journal of Experimental Psychology: Applied*. <https://doi.org/10.1037/xap0000133>
- Arthur, W., Tubre, T. C., Paul, D. S., & Sanchez-Ku, M. L. (1999). College-sample psychometric and normative data on a short form of the Raven Advanced Progressive Matrices Test. *Journal of Psychoeducational Assessment*, 17(4), 354–361. <https://doi.org/10.1177/073428299901700405>
- Azevedo, R. (2020). Reflections on the field of metacognition: Issues, challenges, and opportunities. *Metacognition and Learning*, 15(2), 91–98. <https://doi.org/10.1007/s11409-020-09231-x>
- Baldasaro, R. E., Shanahan, M. J., & Bauer, D. J. (2013). Psychometric properties of the mini-IPIP in a large, nationally representative sample of young adults. *Journal of Personality Assessment*, 95(1), 74–84. <https://doi.org/10.1080/00223891.2012.700466>
- Bertilsson, F., Stenlund, T., Wiklund-Hörnqvist, C., & Jonsson, B. (2021). Retrieval Practice: Beneficial for All Students or Moderated by Individual Differences? *Psychology Learning & Teaching*, 20(1), 21–39. <https://doi.org/10.1177/1475725720973494>
- Bertilsson, F., Wiklund-Hörnqvist, C., Stenlund, T., & Jonsson, B. (2017). The testing effect and its relation to working memory capacity and personality characteristics. *Journal of Cognitive Education and Psychology*, 16(3), 241–259. <https://doi.org/10.1891/1945-8959.16.3.241>
- Bidjerano, T., & Dai, D. Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences*, 17(1), 69–81. <https://doi.org/10.1016/j.lindif.2007.02.001>
- Blasiman, R. N., Dunlosky, J., & Rawson, K. A. (2017). The what, how much, and when of study strategies: Comparing intended versus actual study behaviour. *Memory*, 25(6), 784–792. <https://doi.org/10.1080/09658211.2016.1221974>
- Borghans, L., Duckworth, A. L., Heckman, J. J., & ter Weel, B. (2008). The economics and psychology of personality traits. *Journal of Human Resources*, 43(4), 972–1059. <https://doi.org/10.3368/jhr.43.4.972>
- Brewer, G. A., & Unsworth, N. (2012). Individual differences in the effects of retrieval from long-term memory. *Journal of Memory and Language*, 66(3), 407–415. <https://doi.org/10.1016/j.jml.2011.12.009>

- Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, 27, 1–13. <https://doi.org/10.1016/j.iheduc.2015.04.007>
- Cacioppo, J. T., & Petty, R. E. (1982). The need for cognition. *Journal of Personality and Social Psychology*, 42(1), 116–131. <https://doi.org/10.1037/0022-3514.42.1.116>
- Carpenter, S. K., Pashler, H., Wixted, J. T., & Vul, E. (2008). The effects of tests on learning and forgetting. *Memory and Cognition*, 36(2), 438–448. <https://doi.org/10.3758/MC.36.2.438>
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112(1), 155–159. <https://doi.org/10.1037/0033-2909.112.1.155>
- Colling, J., Wollschläger, R., Keller, U., Preckel, F., & Fischbach, A. (2022). Need for Cognition and its relation to academic achievement in different learning environments. *Learning and Individual Differences*, 93, 102110. <https://doi.org/10.1016/j.lindif.2021.102110>
- Corral, D., Carpenter, S. K., Perkins, K., & Gentile, D. A. (2020). Assessing students' use of optional online lecture reviews. *Applied Cognitive Psychology*, 34(2), 318–329. <https://doi.org/10.1002/acp.3618>
- Costa, P. T., & McCrae, R. R. (1992). The five-factor model of personality and its relevance to personality disorders. *Journal of Personality Disorders*, 6(4), 343–359. <https://doi.org/10.1521/pedi.1992.6.4.343>
- Dirkx, K. J. H., Camp, G., Kester, L., & Kirschner, P. A. (2019). Do secondary school students make use of effective study strategies when they study on their own? *Applied Cognitive Psychology*, 33(5), 952–957. <https://doi.org/10.1002/acp.3584>
- Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment*, 18(2), 192–203. <https://doi.org/10.1037/1040-3590.18.2.192>
- Dornic, S., Ekehammar, B., & Laaksonen, T. (1991). Tolerance for mental effort: Self-ratings related to perception, performance and personality. *Personality and Individual Differences*, 12(3), 313–319. [https://doi.org/10.1016/0191-8869\(91\)90118-U](https://doi.org/10.1016/0191-8869(91)90118-U)
- Duckworth, A. L., & Quinn, P. D. (2009). Development and Validation of the Short Grit Scale (Grit-S). *Journal of Personality Assessment*, 91(2), 166–174. <https://doi.org/10.1080/00223890802634290>
- Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: Perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, 92(6), 1087–1101. <https://doi.org/10.1037/0022-3514.92.6.1087>
- Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques. *Psychological Science in the Public Interest*, 14(1), 4–58. <https://doi.org/10.1177/1529100612453266>
- Fellman, D., Lincke, A., Berge, E., & Jonsson, B. (2020a). Predicting Visuospatial and Verbal Working Memory by Individual Differences in E-Learning Activities. *Frontiers in Education*, 5, 1–11. <https://doi.org/10.3389/educ.2020.00022>
- Fellman, D., Lincke, A., & Jonsson, B. (2020b). Do Individual Differences in Cognition and Personality Predict Retrieval Practice Activities on MOOCs? *Frontiers in Psychology*, 11(2076), 1–10. <https://doi.org/10.3389/fpsyg.2020.02076>
- Gagnon, M., & Cormier, S. (2019). Retrieval practice and distributed practice: The case of French Canadian Students. *Canadian Journal of School Psychology*, 34(2), 83–97. <https://doi.org/10.1177/0829573518773225>
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(1), 26–34. <https://doi.org/10.1037/0003-066X.48.1.26>
- Hakimi, S., Hejazi, E., & Lavasani, M. G. (2011). The relationships between personality traits and students' academic achievement. *Procedia - Social and Behavioral Sciences*, 29, 836–845. <https://doi.org/10.1016/j.sbspro.2011.11.312>
- Jonsson, B., Wiklund-Hörnqvist, C., Stenlund, T., Andersson, M., & Nyberg, L. (2021). A Learning Method for All: The Testing Effect Is Independent of Cognitive Ability. *Journal of Educational Psychology*, 113(5), 972–985. <https://doi.org/10.1037/edu0000627>
- Karpicke, J. D., & Blunt, J. R. (2011). Retrieval practice produces more learning than elaborative studying with concept mapping. *Science*, 331(6018), 772–775. <https://doi.org/10.1126/science.1199327>
- Lamotte, M., Izaute, M., & Darnon, C. (2021). Can tests improve learning in real university classrooms? *Journal of Cognitive Psychology*, 33(8), 974–992. <https://doi.org/10.1080/20445911.2021.1956939>
- Marrs, H., & Sigler, E. A. (2012). Male academic performance in college: The possible role of study strategies. *Psychology of Men & Masculinity*, 13(2), 227–241. <https://doi.org/10.1037/a0022247>
- Martinez-Lopez, R., Yot, C., Tuovila, I., & Perera-Rodríguez, V.-H. (2017). Online self-regulated learning questionnaire in a Russian MOOC. *Computers in Human Behavior*, 75, 966–974. <https://doi.org/10.1016/j.chb.2017.06.015>

- Minear, M., Coane, J. H., Boland, S. C., Cooney, L. H., & Albat, M. (2018). The benefits of retrieval practice depend on item difficulty and intelligence. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(9), 1474–1486. <https://doi.org/10.1037/xlm0000486>
- Moreira, B. F. T., Pinto, T. S. S., Starling, D. S. V., & Jaeger, A. (2019). Retrieval practice in classroom settings: A review of applied research. *Frontiers in Education*, *4*(February), 1–16. <https://doi.org/10.3389/educ.2019.00005>
- Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, *22*, 74–98. <https://doi.org/10.1016/j.edurev.2017.08.004>
- Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*. <https://doi.org/10.3389/fpsyg.2017.00422>
- Pérez-González, J.-C., Filella, G., Soldevila, A., Faiad, Y., & Sanchez-Ruiz, M.-J. (2022). Integrating self-regulated learning and individual differences in the prediction of university academic achievement across a three-year-long degree. *Metacognition and Learning*, *17*(3), 1141–1165. <https://doi.org/10.1007/s11409-022-09315-w>
- Pintrich, P. R., & Zusho, A. (2007). Student motivation and self-regulated learning in the college classroom. In *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Perspective* (pp. 731–810). Springer Netherlands. https://doi.org/10.1007/1-4020-5742-3_16
- Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). Elsevier. <https://doi.org/10.1016/B978-012109890-2/50043-3>
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, *135*(2), 322–338. <https://doi.org/10.1037/a0014996>
- Raven, J. C. (1990). *Advances Progressive Matrices: Sets I*. Oxford Univ Press.
- Rawson, K. A., Dunlosky, J., & Sciarrelli, S. M. (2013). The power of successive relearning: Improving performance on course exams and long-term retention. *Educational Psychology Review*, *25*(4), 523–548. <https://doi.org/10.1007/s10648-013-9240-4>
- Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353–387. <https://doi.org/10.1037/a0026838>
- Roediger, H. L., & Karpicke, J. D. (2006). Test-enhanced learning: Taking memory tests improves long-term retention. *Psychological Science*, *17*(3), 249–255. <https://doi.org/10.1111/j.1467-9280.2006.01693.x>
- Rowland, C. A. (2014). The effect of testing versus restudy on retention: A meta-analytic review of the testing effect. *Psychological Bulletin*, *140*(6), 1432–1463. <https://doi.org/10.1037/a0037559>
- Schweizer, K., Goldhammer, F., Rauch, W., & Moosbrugger, H. (2007). On the validity of Raven's matrices test: Does spatial ability contribute to performance? *Personality and Individual Differences*, *43*(8), 1998–2010. <https://doi.org/10.1016/j.paid.2007.06.008>
- Schwieren, J., Barenberg, J., & Dutke, S. (2017). The testing effect in the psychology classroom: A meta-analytic perspective. *Psychology Learning & Teaching*, *16*(2), 179–196. <https://doi.org/10.1177/1475725717695149>
- Sorić, I., Penezić, Z., & Burić, I. (2017). The Big Five personality traits, goal orientations, and academic achievement. *Learning and Individual Differences*, *54*, 126–134. <https://doi.org/10.1016/j.lindif.2017.01.024>
- Sotola, L. K., & Créde, M. (2021). Regarding class quizzes: A meta-analytic synthesis of studies on the relationship between frequent low-stakes testing and class performance. *Educational Psychology Review*, *33*(2), 407–426. <https://doi.org/10.1007/s10648-020-09563-9>
- Stanikzai, M. (2019). Self-regulated learning: An exploratory study (Level and gender difference). *Journal of Multidisciplinary*, *4*, 76–81. <https://doi.org/10.5281/zenodo.2595252>
- Statistics Sweden. (2023). <https://www.statistikdatabasen.scb.se/pxweb/sv/ssd/>
- Stenlund, T., & Jonsson, B. (2017). Assessing the Willingness to Elaborate among Young Students: Psychometric Evaluation of a Swedish Need for Cognition Scale. *Frontiers in Education*, *2*(2). <https://doi.org/10.3389/educ.2017.00002>
- Stenlund, T., Jönsson, F. U., & Jonsson, B. (2017). Group discussions and test-enhanced learning: Individual learning outcomes and personality characteristics. *Educational Psychology*, *37*(2), 145–156. <https://doi.org/10.1080/01443410.2016.1143087>
- Suurtamm, C., Koch, M., & Arden, A. (2010). Teachers' assessment practices in mathematics: Classrooms in the context of reform. *Assessment in Education: Principles, Policy & Practice*, *17*(4), 399–417. <https://doi.org/10.1080/0969594X.2010.497469>
- Tabachnick, B. G., & Fidell, L. S. (2019). *Using Multivariate Statistics* (7th ed.). Pearson.

- Trumbo, M. C., Leiting, K. A., McDaniel, M. A., & Hodge, G. K. (2016). Effects of reinforcement on test-enhanced learning in a large, diverse introductory college psychology course. *Journal of Experimental Psychology: Applied*, 22(2), 148–160. <https://doi.org/10.1037/xap0000082>
- Tullis, J. G., & Maddox, G. B. (2020). Self-reported use of retrieval practice varies across age and domain. *Metacognition and Learning*, 15(2), 129–154. <https://doi.org/10.1007/s11409-020-09223-x>
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, 37(3), 498–505. <https://doi.org/10.3758/BF03192720>
- Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, 140(4), 1174–1204. <https://doi.org/10.1037/a0036620>
- Weis, M., Heikamp, T., & Trommsdorff, G. (2013). Gender differences in school achievement: The role of self-regulation. *Frontiers in Psychology*, 4. <https://doi.org/10.3389/fpsyg.2013.00442>
- Wiklund-Hörnqvist, C., Jonsson, B., & Nyberg, L. (2014). Strengthening concept learning by repeated testing. *Scandinavian Journal of Psychology*, 55(1), 10–16. <https://doi.org/10.1111/sjop.12093>
- Wiklund-Hörnqvist, C., Stillesjö, S., Andersson, M., Jonsson, B., & Nyberg, L. (2022). Retrieval Practice Is Effective Regardless of Self-Reported Need for Cognition—Behavioral and Brain Imaging Evidence. *Frontiers in Psychology*, 12, 1–10. <https://doi.org/10.3389/fpsyg.2021.797395>
- Wolters, C. A., & Hussain, M. (2015). Investigating grit and its relations with college students' self-regulated learning and academic achievement. *Metacognition and Learning*, 10(3), 293–311. <https://doi.org/10.1007/s11409-014-9128-9>
- Yang, C., Luo, L., Vadillo, M. A., Yu, R., & Shanks, D. R. (2021). Testing (quizzing) boosts classroom learning: A systematic and meta-analytic review. *Psychological Bulletin*, 147(4), 399–435. <https://doi.org/10.1037/bul0000309>
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, 41(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2
- Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist*, 48(3), 135–147. <https://doi.org/10.1080/00461520.2013.794676>
- Zimmerman, B. J., & Martinez-Pons, M. (1990). Student differences in self-regulated learning: Relating grade, sex, and giftedness to self-efficacy and strategy use. *Journal of Educational Psychology*, 82(1), 51–59. <https://doi.org/10.1037/0022-0663.82.1.51>
- Zimmerman, B. J. (2001). *Self-Regulated Learning and Academic Achievement: Theoretical Perspective*. In B. J. Zimmerman & D. H. Schunk (Eds.) (pp. 1–37). Routledge.

Frida Bertilsson Department of Psychology, Umeå University, SE-901 87, Umeå, Sweden. E-mail: frida.bertilsson@umu.se. Web site of department: <https://www.umu.se/en/departement-of-psychology/>
Most relevant publications in the field of Psychology of Education: Bertilsson, F., Stenlund, T., Wiklund-Hörnqvist, C., & Jonsson, B. (2021). Retrieval practice: Beneficial for all students or moderated by individual differences? *Psychology Learning & Teaching*, 20(1), 21–39. <https://doi.org/10.1177/1475725720973494>
 Bertilsson, F., Wiklund-Hörnqvist, C., Stenlund, T., & Jonsson, B. (2017). The testing effect and its relation to working memory capacity and personality characteristics. *Journal of Cognitive Education and Psychology*, 16(3), 241–259. <https://doi.org/10.1891/1945-8959.16.3.241>

Tova Stenlund Department of Psychology, Umeå University, SE-901 87, Umeå, Sweden. E-mail: tova.stenlund@umu.se. Web site of department: <https://www.umu.se/en/departement-of-psychology/>
Most relevant publications in the field of Psychology of Education: Stenlund, T., Lyrén, P.-E., & Eklöf, H. (2018). The successful test taker: exploring test-taking behavior profiles through cluster analysis. *European Journal of Psychology of Education*, 33(2), 403–417. <https://doi.org/10.1007/s10212-017-0332-2>
 Stenlund, T., Eklöf, H., & Lyrén, P.-E. (2017). Group differences in test-taking behaviour: An example from a high-stakes testing program. *Assessment in education: Principles, Policy & Practice*, 24(1), 4–20. <https://doi.org/10.1080/0969594X.2016.1142935>
 Stenlund, T., & Jonsson, B. (2017). Assessing the willingness to elaborate among young students: Psychometric evaluation of a Swedish need for cognition scale. *Frontiers in Education*, 2, 2. <https://doi.org/10.3389/feduc.2017.00002>
 Stenlund, T., Jönsson, F., & Jonsson, B. (2017). Group discussions and test-enhanced learning: individual learning outcomes and personality characteristics. *Educational Psychology*, 37(2), 145–156. <https://doi.org/10.1080/01443410.2016.1143087>
 Jonsson, B., Wiklund-Hörnqvist, C., Stenlund, T., Andersson, M., & Nyberg, L. (2021). A learning method for all: The testing effect is independent of cognitive ability. *Journal of Educational Psychology*, 113(5), 972–985. <https://doi.org/10.1037/edu0000627>
 Inzunza, M., Brown, G.T.L., Stenlund, T., & Wikström, C. (2022). The relationship between subconstructs of empathy and general cognitive ability in the context of policing. *Frontiers in*

Psychology, 13, 907610. <https://doi.org/10.3389/fpsyg.2022.907610>

Inzunza, M., Stenlund, T., & Wikström, C. (2019). Measuring perspective taking among police recruits: A comparison of self-reported and objective measures. *Policing: An International Journal*. <https://doi.org/10.1108/PIJPSM-09-2018-0129>

Anna Sundström Department of Psychology, Umeå University, SE-901 87, Umeå, Sweden. E-mail: anna.e.sundstrom@umu.se. Web site of department: <https://www.umu.se/en/departement-of-psychology/>

Most relevant publications in the field of Psychology of Education: Sundström, A.E., & Hakelind, C. (2023). Examining clinical skills and abilities in psychology – implementation and evaluation of an objective structured clinical examination in psychology. *The Journal of Mental Health Training, Education and Practice*, 18(2), 97–110. <https://doi.org/10.1108/JMHTEP-10-2021-0124>

Hakelind, C., & Sundström, A.E. (2022). Examining Skills and Abilities During the Pandemic – Psychology Students' and Examiners' Perceptions of a Digital OSCE. *Psychology Learning & Teaching*, 21(3), 278–295. <https://doi.org/10.1177/14757257221114038>

Knehta, E., & Sundström, A. (2019). 'It was, perhaps, the most important one' students' perceptions of national tests in terms of test-taking motivation. *Assessment in Education: Principles, Policy & Practice*, 26(2), 202–221. <https://doi.org/10.1080/0969594X.2017.1323725>

Stenlund, T., Sundström, A., & Jonsson, B. (2016). Effects of repeated testing on short- and long-term memory performance across different test formats. *Educational Psychology*, 36(10), 1710–1727. <https://doi.org/10.1080/01443410.2014.953037>

Bert Jonsson Department of Applied Educational Science, Umeå University, SE-901 87, Umeå, Sweden. E-mail: bert.jonsson@umu.se. Web site of department: <https://www.umu.se/en/departement-of-applied-educational-science/>

Most relevant publications in the field of Psychology of Education (from 2020): Bertilsson, F., Stenlund, T., Wiklund-Hörnqvist, C., & Jonsson, B. (2021). Retrieval practice: Beneficial for all students or moderated by individual differences? *Psychology Learning and Teaching-Plat*, 20(1), 21–39. <https://doi.org/10.1177/1475725720973494>

Fellman, D., Lincke, A., & Jonsson, B. (2020). Do individual differences in cognition and personality predict retrieval practice activities on MOOCs? *Frontiers in Psychology*, 11, 2076. <https://doi.org/10.3389/fpsyg.2020.02076>

Jonsson, B., Granberg, C., & Lithner, J. (2020). Gaining mathematical understanding: The effects of creative mathematical reasoning and cognitive proficiency. *Frontiers in Psychology*, 11(3666). <https://doi.org/10.3389/fpsyg.2020.574366>

Jonsson, B., Wiklund-Hörnqvist, C., Stenlund, T., Andersson, M., & Nyberg, L. (2020). A learning method for all: The testing effect is independent of cognitive ability. *Journal of Educational Psychology*, 113(5), 972–985. <https://doi.org/10.1037/edu0000627>

Levlin, M., Wiklund-Hörnqvist, C., Sandgren, O., Karlsson, S., & Jonsson, B. (2022). Evaluating the effect of rich vocabulary instruction and retrieval practice on the classroom vocabulary skills of children with (developmental) language disorder. *Language, Speech, and Hearing Services in Schools*, 53(2), 542–560. https://doi.org/10.1044/2021_LSHSS-21-00101

Sammallahti, E., Finell, J., Jonsson, B., & Korhonen, J. (2023). A meta-analysis of math anxiety interventions. *Journal of Numerical Cognition*, 9(2), 346–362. <https://doi.org/10.5964/jnc.8401>

Stillesjö, S., Karlsson Wirebring, L., Andersson, M., Granberg, C., Lithner, J., Jonsson, B., . . . Wiklund-Hörnqvist, C. (2021). Active math and grammar learning engages overlapping brain networks. *Proceedings of the National Academy of Sciences*, 118(46), e2106520118. <https://doi.org/10.1073/pnas.2106520118>

Wiklund-Hörnqvist, C., Stillesjö, S., Andersson, M., Jonsson, B., & Nyberg, L. (2022). Retrieval Practice Is Effective Regardless of Self-Reported Need for Cognition - Behavioral and Brain Imaging Evidence. *Frontiers in Psychology*, 12. Retrieved from: <https://www.frontiersin.org/article/https://doi.org/10.3389/fpsyg.2021.797395>

Wirebring, L. K., Wiklund-Hörnqvist, C., Stillesjö, S., Granberg, C., Lithner, J., Andersson, M., . . . Jonsson, B. (2021). Effects on brain activity after creative mathematical reasoning when considering individual differences in cognitive ability. *bioRxiv*. <https://doi.org/10.1101/2021.03.30.437492>

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

Frida Bertilsson is PhD student with a master's degree in cognitive science. She is part of a research project denoted as "memory and learning" which focuses on studying learning strategies in relation to cognitive and non-cognitive factors for learning.

Tova Stenlund is an associate professor at the Department of Psychology, Umeå University, Sweden. She has a PhD degree in Educational Measurement and her research and teaching interest lie in the field of educational psychology and cognitive psychology, with a particular focus on validity of test and assessment results, test-taking behaviour, and effects of repeated testing on memory and learning.

Anna Sundström is an Associate Professor in psychology at the Department of Psychology, Umeå University Sweden. She has a PhD degree in Educational Measurement and her research and teaching interest concerns applied psychometrics with a special focus on evaluating effects of psychological treatment, as well as and teaching and learning in psychology.

Bert Jonsson is a Professor at the Department of Applied Educational Science. He is the principal investigator of the project "the learning brain" financed by the Swedish research council. The project investigates fundamental questions arising in educational science and pertaining to the cognitive neuroscience of children's learning. The project focuses mainly on learning strategies in relation to cognition.

Authors and Affiliations

Frida Bertilsson¹  · **Tova Stenlund**¹  · **Anna Sundström**¹  · **Bert Jonsson**² 

✉ Frida Bertilsson
frida.bertilsson@umu.se

Tova Stenlund
tova.stenlund@umu.se

Anna Sundström
anna.e.sundstrom@umu.se

Bert Jonsson
bert.jonsson@umu.se

¹ Department of Psychology, Umeå University, SE-901 87 Umeå, Sweden

² Department of Applied Educational Science, Umeå University, SE-901 87 Umeå, Sweden